Nonlinear Dynamics and Chaos

Prof. George Haller
Transcription: Trevor Winstral

2022

Contents

Part I

Nonlinear Dynamics and Chaos 1

Chapter 1

Introduction

First we shall introduce the most important characters in our following exploration. The ideas and definitions here will be recurring regularly as we examine them from different perspectives and using different tools. The content covered by this course can be found in the following books. For further details on some of the results, we recommend consulting these.

- J. Guckenheimer & P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields [Guckenheimer and Holmes, 1990];
- F. Verhulst, Nonlinear Differential Equations and Dynamical Systems [Verhulst, 1989];
- V. I. Arnold, Ordinary Differential Equations [Arnold, 1992];
- S. Strogatz, Nonlinear Dynamics and Chaos [Strogatz, 2000].

Definition 1.1 (Dynamical System (DS))**.** A triple $(P, E, \mathcal{F})$, with

- P : the phase space for the dynamical variable $x \in P$,
- E : base space of the evolutionary variable (e.g. time) $t \in E$,
- $\mathcal{F}$: the evolution rule (deterministic) which defines the transition from one state to the next.

The two main types of evolutionary variable spaces are

(i) Discrete dynamical systems (DDS) $t \in E = \mathbb{Z}$ with trajectory $\{x_0, x_1, \ldots\}$,

(ii) Continuous dynamical systems (CDS) $t \in E = \mathbb{R}$ with trajectory $\{x_t\}_{t\in\mathbb{R}}$.

Corresponding to these there are various types of evolution rules

(i) In a DDS we have iterated mappings

$$\boxed{x_{n+1} = F(x_n, n).}$$

If there is no explicit dependence on n, i.e. $\frac{\partial F}{\partial n} = 0$, then

$$\boxed{x_{n+1} F(x_n) = F(F(x_{n-1})) = \underbrace{F \circ \ldots \circ F}_{n+1 \text{ times}}(x_0) = F^{n+1}(x_0).}$$

Example 1.1 (Cobweb diagram of a one-dimensional DDSs). In such cases and in one-dimensional problems, a simple way to analyze the behavior of the system is the so-called *cobweb* diagram. We may plot x_{n+1} as a function of x_n, as demonstrated in Fig. 1.1. The image of an initial condition x_0 lies on the graph at $x_{n+1} = F(x_0)$. We can also compute the next iterate by horizontally projecting the point $(x_0, F(x_0))$ to the diagonal line defined by $x_{n+1} = x_n$. Following the porjection of this point to the horizontal axis (x_n) we find the intersection with the graph at the point $(x_1, F(x_1))$. It follows that fixed points on the cobweb diagram correspond to the intersection of the graph of F with the diagonal line $x_{n+1} = x_n$.

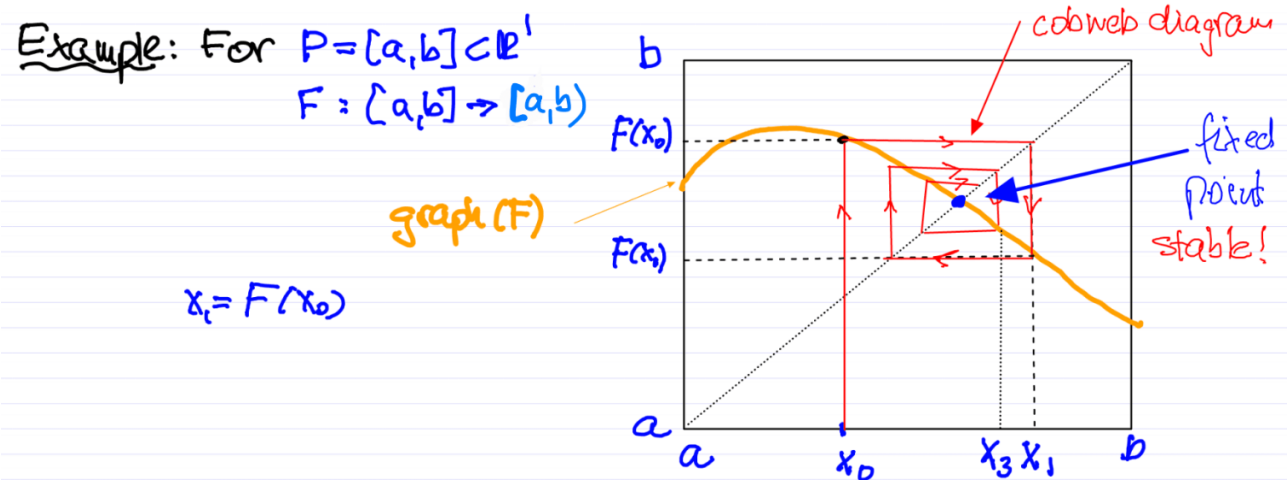


Figure 1.1: Analysis of a one-dimensional system defined on the interval $x \in [a, b]$ using the cobweb diagram

(ii) In a CDS we have a first order system of ordinary differential equations (ODE)

$$\boxed{\dot{x} = f(x, t)}$$

for $x \in P$ and $t \in E$. This yields the initial value problem (IVP):

$$\begin{cases} \dot{x} = f(x,t) \\ x(t_0) = x_0 \end{cases}$$

Assuming there exists a unique solution $\varphi(t; t_0, x_0)$ with $\dot{\varphi} = f(\phi, t)$ and $\varphi(t_0) = x_0$, then the following flow map is well defined

$$\boxed{F_{t_0}^t(x_0) := \varphi(t; t_0, x_0).}$$

Geometrically, this solution can be viewed as a trajectory in phase space (cf. Fig. 1.2).

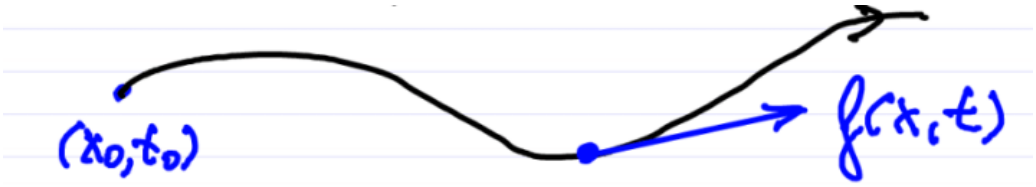


Figure 1.2: Trajectory of a continuous dynamical system. The RHS is given by f(x,t), which is the tangent vector to this curve at the point x at time t.

Such an $F_{t_0}^t$ has the properties

(a) $F_{t_0}^t$ is as smooth as $f(x,t)$,

(b) $F_{t_0}^{t_0} = I$ and $F_{t_0}^{t_2} = F_{t_1}^{t_2} \circ F_{t_0}^{t_1}$,

(c) $\left(F_{t_0}^t\right)^{-1} = F_t^{t_0}$ exists and is smooth.

Properties (a) and (b) together are called the group property. A special case of continuous dynamical systems is the autonomous system.

$$\boxed{\dot{x} = f(x).}$$

The autonomy of a system implies

$$x(s, t_0, x_0) = x(\underbrace{s - t_0}_{t}, 0, x_0) \overset{!}{=} x(t, x_0).$$

The induced flow map in this case is the one-parameter family of maps

$$\boxed{F^t = F_0^t : x_0 \mapsto x(t, x_0).}$$

Example 1.2 (Logisitic Equation). For a resource-limited population, we have the following dynamical system for $a > 0$, $b > 0$, and the population $x \in \mathbb{R}_+ \cup \{0\}$

$$\dot{x} = ax(b - x).$$

In this case we have $E = \mathbb{R}$ and $\mathcal{F} = \{F^t\}_{t=-\infty}^{+\infty}$. This system has globally existing unique solutions (see later). We may analyze the behavior of this system by plotting $\dot{x}$ as a function of x, analogously to the cobweb diagram. This is demonstrated in Fig. 1.3. At x values, where $\dot{x}$ is positive $x(t)$ is growing, while at negative values it is decreasing. This means, that fixed points, at which $x(t) = \text{const.}$ correspond to intersections of the graphs with the horizontal axis.

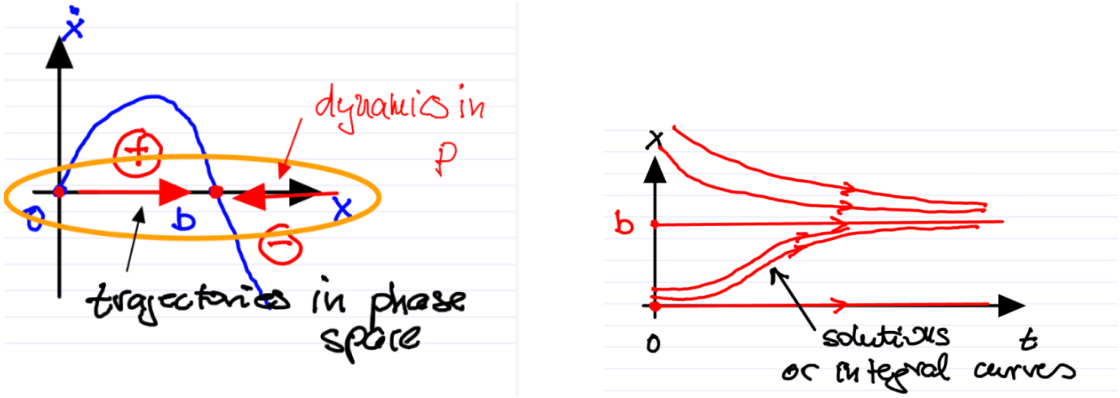


Figure 1.3: Left: Analysis of the right hand side. Right: Evolution in the extended phase space $P \times \mathbb{R}$.

Example 1.3 (Pendulum). Given the equation of motion

$$ml^2\ddot{\varphi} = -mgl\sin(\varphi).$$

We let $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ to transform into the first-order ODE form

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\dfrac{g}{l}\sin(x_1). \end{cases}$$

Thus we have

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; \quad f(x) = \begin{pmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) \end{pmatrix}.$$

Qualitative analysis gives the following facts

- $(x_1, x_2) = (0, 0)$ and $(x_1, x_2) = (\pi, 0)$ are zeros of f.
- Energy is conserved, hence both small and large amplitude oscillations are expected.
- The function $f(x)$ has symmetries: it is invariant under the transformation $(x_1, x_2, t) \mapsto (x_1, -x_2, -t)$ and $(x_1, x_2, t) \mapsto (-x_1, x_2, -t)$. See the left panel of Fig. 1.4.

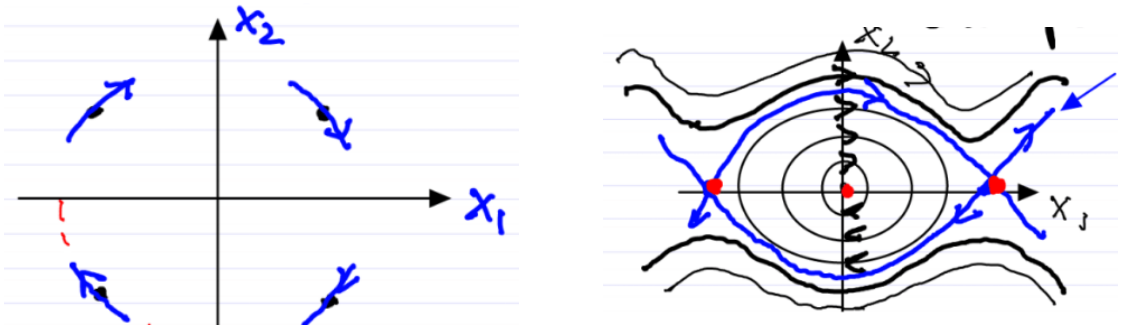


Figure 1.4: Left: The symmetries of the dynamical system. Right: Phase portrait of the pendulum. Red dots show the fixed points, while the blue trajectories make up the separatrix.

Definition 1.2. A separatrix is a boundary (i.e. a codimension-1 surface) in phase space which separates regions of qualitatively different behaviors. In practice, it is unobservable by itself and connects different fixed points. The separatrix of the pendulum is shown in the right panel of Fig. 4.

Example 1.4 (Exploit geometry of phase space for analysis). Consider two cities, A and B. The two cities are connected by two roads, denoted by the blue and green curves of the left panel of Fig. 1.5. We assume that travelling on the two roads, it is possible for two bikes to make it from A to B without ever being further away from each other than a distance $d < D$.

Assume two trucks are trying to make it between A and B, on different roads in the opposite direction, carrying a load of width D. Given this information, can the trucks make it without hitting each other? We can view this problem as a continuous dynamical system with two coordinates x_1 and x_2 that parameterize the two routes between A and B. This dynamical system is, in general, non autonomous.

The right panel of Fig 1.5 shows the trajectories of the two trucks and the two bikes in phase space. The two trajectories must intersect by continuity, thus at that point the trucks must be at the same positions as the bikes, implying they are within distance D. Therefore the trucks must crash!

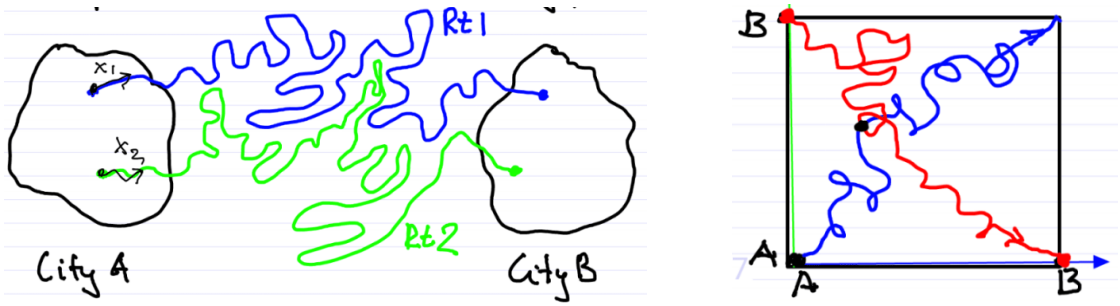


Figure 1.5: Left: An example of the two bike routes. Right: Blue represents the trajectory of the two bikes, red represents the trajectory of the two trucks.

Chapter 2

Fundamentals

In this chapter, we first review some fundamental properties of continuous dynamical systems that will be used heavily in later chapters. As we will see, these technical results are interesting in their own right. They can help in interpreting or cross-checking numerical results or physical models for self-consistency or accuracy.

2.1 Existence and uniqueness of solutions

Consider

$$\begin{cases} \dot{x} = f(x,t); \quad x \in \mathbb{R}^n \\ x(t_0) = x_0 \end{cases}.$$

Does this initial value problem have a unique solution? We have the following theorems to help us answer that question.

Theorem 2.1 (Peano)**.** *If $f \in \mathcal{C}^0$ near (x_0, t_0), then there exists a local solution $\varphi(t)$, i.e.,*

$$\dot{\varphi}(t) = f(\varphi(t), t), \varphi(t_0) = x_0;\ \forall t \in (t_0 - \varepsilon, t_0 + \varepsilon);\ 0 < \varepsilon \ll 1.$$

Example 2.1 (Free falling mass). Consider a point mass of mass m at position x. The acceleration due to gravity is denoted by g. Measuring the potential energy from the reference point $x = x_0$, we have the total energy is conserved.

$$\frac{1}{2}m\dot{x}^2 = mg(x - x_0).$$

This implies that

$$\begin{cases} \dot{x} = \sqrt{2g(x - x_0)} \\ x(0) = x_0 \end{cases}$$

on the set $P = \{x \in \mathbb{R} : \ x \geq x_0\}$. Therefore we have that $f \in \mathcal{C}^0$ in phase space, so by Peano's theorem (cf. Theorem 2.1), there exists a local solution. A schematic diagram is shown in Fig. 2.1. The solution is actually $x(t) = x_0 + \frac{g}{2}(t - t_0)^2$, however $x(t) = x_0$ is also a solution to

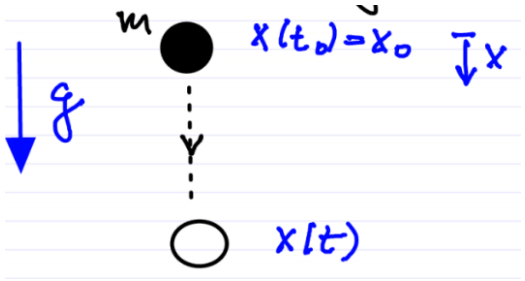


Figure 2.1: Schematic diagram of the point mass in free fall.

the IVP, therefore we do not have a unique solution. Physically there exists a solution, but this IVP was derived from a heuristic energy-principle, not from Newton's laws, which are not equivalent.

Definition 2.1. A function f is called locally Lipschitz around x_0 if there exists an open set U_{x_0} and $L > 0$ such that for all $x, y \in U_{x_0}$

$$\boxed{\|f(y,t) - f(x,t)\| \leq L\|y - x\|.}$$

Example 2.2 (Lipschitz functions). Fig. 2.2 shows an example of a Lipschitz and a non-Lipschitz function around x_0.

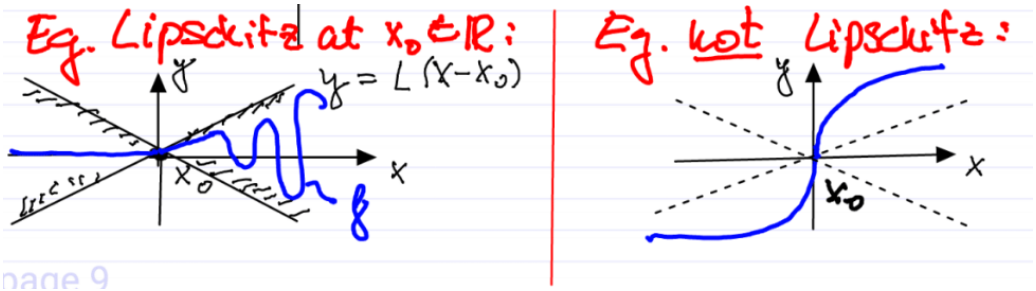


Figure 2.2: Interpretation of the Lipschitz property.

Theorem 2.2 (Picard). *Assume*

(i) $f \in \mathcal{C}^0$ *in* t *near* (t_0, x_0),

(ii) f *is locally Lipschitz in* x *near* (t_0, x_0).

Then there exists a unique local solution to the IVP. The proof can be found in [Arnold, 1992].

Note the following relations. If f is $\mathcal{C}^1 \implies f$ is Lipschitz $\implies f$ is $\mathcal{C}^0$.

Example 2.3 (Free falling mass revisted). We check if f is Lipschitz.

$$\frac{|f(x) - f(x_0)|}{|x - x_0|} = \frac{\sqrt{2g}}{\sqrt{|x - x_0|}} \geq L|x - x_0|.$$

Thus f is not Lipschitz near x_0.

2.2 Geometric consequences of uniqueness

If the solution is unique, we have a few facts that can be derived from the geometric point of view.

(i) The trajectories of autonomous systems cannot intersect. Note that fixed points do not violate this (e.g. pendulum equations). See Fig. 2.3 which shows the phase portrait of the pendulum.

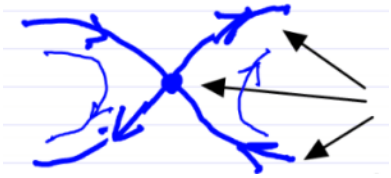


Figure 2.3: The phase portrait of the pendulum. Trajectories do not intersect since each arrow is pointing at separate trajectories.

(ii) For non-autonomous systems, intersections in phase space are possible: a trajectory may occupy the same point x at a different time instants (see the left panel of Fig. 2.4. In this case we can extend the phase space in order to get an autonomous system where there cannot be any intersections.

$$X = \begin{pmatrix} x \\ t \end{pmatrix}, \; F(X) = \begin{pmatrix} f(x,t) \\ 1 \end{pmatrix}; \; \dot{X} = F(X).$$

2.3 Local vs global existence

Example 2.4 (Exploding solution).

$$\begin{cases} \dot{x} = x^2 \\ x(t_0) = 1. \end{cases}$$

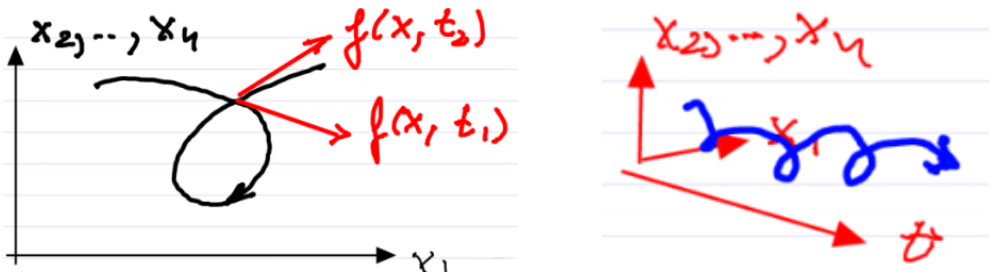


Figure 2.4: Left: Intersecting trajectories in phase space for a non-autonomous system. Right: The same trajectory in the extended phase space, without intersections.

Integrating yields the solution $x(t) = \frac{1}{1-(t-t_0)}$. This solution blows up at $t_\infty = t_0 + 1$, therefore the solution is only local. This is demonstrated in Fig. 2.5.

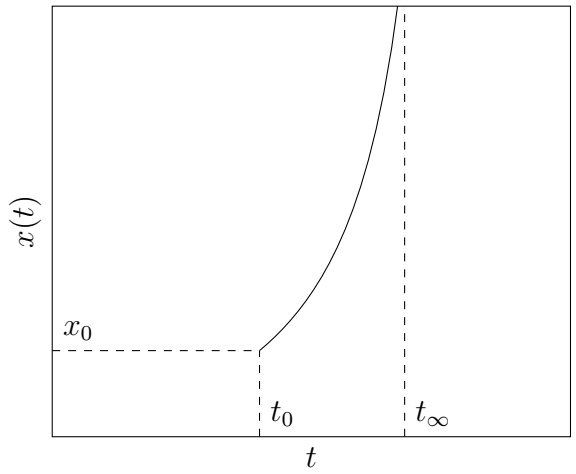


Figure 2.5: Solution to the ODE $\dot{x} = x^2$ started from $x(t_0) = 1$.

To address this problem of local solutions not being able to be continued into global solution, we have the following theorem.

Theorem 2.3 (Continuation of solution)**.** *If a local solutions cannot be continued to a time $t = T$, then we must have*

$$\boxed{\lim_{t \to T} \|x(t)\| = \infty.}$$

The proof can be found in [Arnold, 1992].

Example 2.5 (Coupled Pendulum System). Consider two pendula of masses m_1 and m_2. They both have length l. The angles of these pendula are denoted by φ_1 and φ_2. Let us assume

that they are coupled by a nonlinear spring, which can be described by a potential $V(\varphi_1, \varphi_2)$. This setup is illustrated in Fig. 2.6. We set $x_1 = \varphi_1,\ x_2 = \dot{\varphi}_1,\ x_3 = \varphi_2,\ x_4 = \dot{\varphi}_2$ and get the following equation of motion

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = \dots \\ \dot{x}_3 = x_4 \\ \dot{x}_4 = \dots \end{cases}$$

The RHS is smooth, therefore there exists a unique local solution to any IVP. The phase space

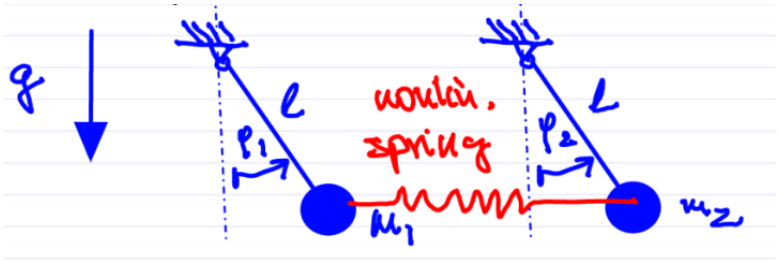


Figure 2.6: Physical setup of the coupled pendulum with a nonlinear spring.

is given by

$$P = \{x :\ x_1 \in S^1,\ x_2 \in \mathbb{R},\ x_3 \in S^1,\ x_4 \in \mathbb{R}\} = S^1 \times \mathbb{R} \times S^1 \times \mathbb{R}.$$

Where S^1 is the 1 dimensional sphere (i.e. a circle). With this space we know that $\|x_1\|$ and $\|x_3\|$ are bounded. Due to energy being conserved we have

$$E = T + V = \frac{1}{2} m_1 l_1 x_2^2 + \frac{1}{2} m_2 l_2 x_4^2 + \underbrace{V(x_1, x_3)}_{\geq 0}$$

$$E = E_0 = \text{constant} \geq 0.$$

Hence $\|x_2\|$ and $\|x_4\|$ are also bounded, therefore all solutions exist globally.

Definition 2.2. A linear system is one such that for $x \in \mathbb{R}^n,\ A(t) \in \mathbb{R}^{n \times n}$ and $A \in \mathcal{C}^0$

$$\boxed{\dot{x} = A(t)x.}$$

Remark 2.4. Note that A can be written as $A = S + \Omega$ where $S = \frac{1}{2}(A + A^T)$ is symmetric (i.e. $S = S^T$) and $\Omega = \frac{1}{2}(A - A^T)$ is skew symmetric (i.e. $\Omega = -\Omega^T$). Furthermore the eigenvalues of S, λ_i, are all real and their respective eigenvectors, e_i, are orthogonal.

Example 2.6 (Global existence in linear systems).

$$\begin{aligned}\langle x, \dot{x}\rangle &= \frac{1}{2}\frac{d}{dt}\|x(t)\|^2 = \langle x, A(t)x\rangle = \langle x, (S(t)+\Omega(t))x\rangle \\ &= \langle x, S(t)x\rangle + \underbrace{\langle x, \Omega(t)x\rangle}_{=0} \overset{(*)}{=} \sum_{i=1}^{n} \lambda_i(t)x_i^2 \\ &\leq \lambda_{\max}(t)\sum_{i=1}^{n} x_i^2 = \lambda_{\max}(t)\|x(t)\|^2.\end{aligned}$$

Where in $(*)$ we used that $x = \sum_{i=1}^{n} x_i e_i$ with $\|e_i\| = 1$ and $e_i \perp e_j$ for all $i \neq j$. Thus we get

$$\frac{\frac{1}{2}\frac{d}{dt}\|x(t)\|^2}{\|x(t)\|^2} \leq \lambda_{\max}(t) \implies \int_{t_0}^{t} \log\left(\frac{\|x(s)\|^2}{\|x(t_0)\|^2}\right) ds \leq \lambda_{\max}(s)ds.$$

By exponentiating both sides, we obtain

$$\boxed{\|x(t)\| \leq \|x(t_0)\| \exp\left(\int_{t_0}^{t} \lambda_{\max}(s)ds\right).}$$

Therefore, by the continuation theorem, global solutions exist as long as $\int_{t_0}^{t} \lambda_{\max}(s)ds < \infty$.

2.4 Dependence on initial conditions

Given the IVP

$$\begin{cases}\dot{x} = f(x,t) \\ x(t_0) = x_0.\end{cases}$$

With $x \in \mathbb{R}^n$ and $f \in \mathcal{C}^r$ for some $r \geq 1$, we have the solution $x(t; t_0, x_0)$.

The dependence of the solution on initial data is of interest to us. This is due to us wanting the solution to be robust with respect to errors and uncertainties in the initial data. To address this, we have Theorem 2.5.

Theorem 2.5. *If $f \in \mathcal{C}^r$ for $r \geq 1$ then $x(t; t_0, x_0)$ is $\mathcal{C}^r$ in (t_0, x_0). Proof in [Arnold, 1992].*

The geometric meaning of this is that for $U \subset P \subset \mathbb{R}^n$ we have that $F_{t_0}^t(U)$ is a smooth deformation of U (cf. Fig. 2.7). It turns out $\left(F_{t_0}^t\right)^{-1} = F_t^{t_0}$ is also $\mathcal{C}^r$, hence we have that $F_{t_0}^t$ is a diffeomorphism.

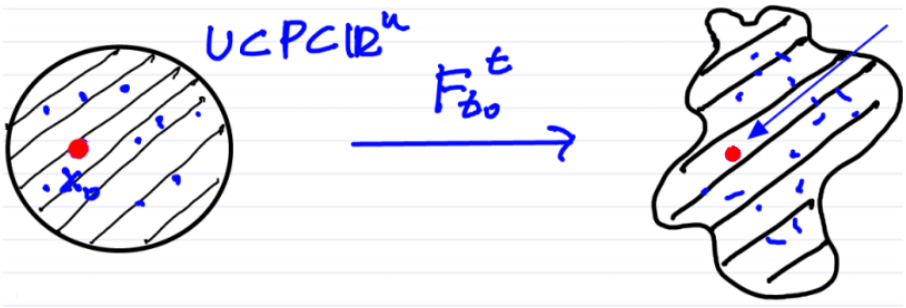


Figure 2.7: The smooth transformation of U. The red point on the right it $F_{t_0}^t(x_0)$, i.e. the image of x_0 under the evolution operator.

Remark 2.6 (The total differential). We denote the total differential of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ as Df. The total differential is a function which takes a location x as the argument and returns the derivative of f at the point x, i.e. the Jacobian. This implies evaluating the Jacobian at the point x. For a function $f(x, y) = f(x_1, \ldots, x_n, y_1, \ldots, y_m) : \mathbb{R}^{n+m} \to \mathbb{R}^k$ the total differential Df means with respect to all of the variables and the total differential with respect to x, written $D_x f$ is the total differential only taken with respect to the x variables. Thus for $f(x, y) : \mathbb{R}^{n+m} \to \mathbb{R}^k$ we have the total differential

$$Df = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} & \frac{\partial f_k}{\partial y_1} & \cdots & \frac{\partial f_k}{\partial y_m} \end{pmatrix};$$

$$Df(x_0, y_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0, y_0) & \frac{\partial f_1}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial y_m}(x_0, y_0) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial x_n}(x_0, y_0) & \frac{\partial f_k}{\partial y_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial y_m}(x_0, y_0) \end{pmatrix},$$

and the total differential with respect to x

$$D_x f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix}; \quad D_x f(x_0, y_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0, y_0) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(x_0, y_0) & \cdots & \frac{\partial f_k}{\partial x_n}(x_0, y_0) \end{pmatrix}.$$

Now, how can we compute the Jacobian of the flow map $\frac{\partial x(t;t_0,x_0)}{\partial x_0} = DF_{t_0}^t(x_0)$? We start from the IVP and take the gradient (with respect to x_0) of both sides. On the left hand side we can exchange order of the time derivative and the gradient and on the right hand side we

use the chain rule. We end up with the equation

$$\frac{d}{dt}\frac{\partial x}{\partial x_0} = D_x f(x(t;t_0,x_0),t)\frac{\partial x}{\partial x_0}; \quad \frac{\partial x}{\partial x_0} \in \mathbb{R}^{n\times n}.$$

This means, that the flow map gradient satisfies the IVP

$$\begin{aligned}\frac{d}{dt}\left[DF_{t_0}^t(x_0)\right] &= D_x f(F_{t_0}^t(x_0),t)DF_{t_0}^t(x_0)\\ DF_{t_0}^{t_0}(x_0) &= I.\end{aligned}$$

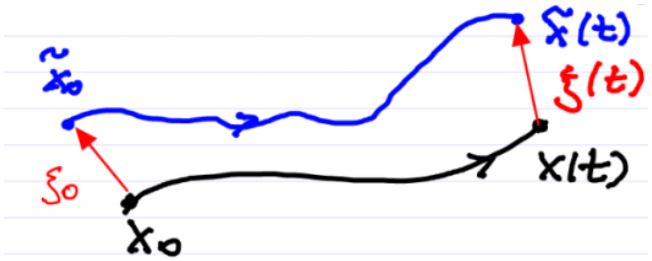
This is called the equation of variations, which is a linear, non-autonomous ODE for the matrix $M = DF_{t_0}^t(x_0)$

$$\begin{cases}\dot{M} = D_x f(x(t;t_0,x_0))M\\ M(t_0) = I.\end{cases}$$

Example 2.7 (Locations of extreme deformation in phase space). We define

$$\begin{aligned}\xi(t) &:= \tilde{x}(t) - x(t) = x(t;t_0,\tilde{x}_0) - x(t;t_0,x_0)\\ &= x(t;t_0,x_0) + \frac{\partial x}{\partial x_0}(t;t_0,x_0)\xi_0 + \mathcal{O}(\|\xi_0\|^2) - x(t;t_0,x_0)\\ &= DF_{t_0}^t(x_0)\xi_0 + \mathcal{O}(\|\xi_0\|^2).\end{aligned}$$

Where we used the Taylor expansion and assume the perturbation to x_0 is small, i.e. $\|\xi_0\| \ll 1$. Therefore we have



$$\begin{aligned}\|\xi(t)\|^2 &= \langle DF_{t_0}^t(x_0)\xi_0, DF_{t_0}^t(x_0)\xi_0\rangle + \mathcal{O}(\|\xi_0\|^3)\\ &= \langle \xi_0, \underbrace{\left[DF_{t_0}^t(x_0)\right]^T DF_{t_0}^t(x_0)}_{=:C_{t_0}^t(x_0)}\xi_0\rangle + \mathcal{O}(\|\xi_0\|^3).\end{aligned}$$

$C^t_{t_0}(x_0)$ is known as the Cauchy-Green strain tensor. It is positive definite and symmetric and due to its dependence on the initial condition, $C^t_{t_0}(x_0)$ actually defines a *tensor field*. Therefore the largest possible deformation is

$$\max_{x_0,\xi_0} \frac{\|\xi(t)\|^2}{\|\xi_0\|^2} = \max_{x_0,\xi_0} \frac{\langle \xi_0, C^t_{t_0}(x_0)\xi_0\rangle}{\|\xi_0\|^2} = \max_{x_0} \lambda_n(x_0).$$

Where we used that $C^t_{t_0}$ is positive definite in the last equality, and that $\lambda_n(x_0)$ is the largest eigenvalue of $C^t_{t_0}(x_0)$. Because we typically have exponential growth we introduce the following quantity.

Definition 2.3. The finite-time Lyapunov exponent is defined as

$$\boxed{\mathrm{FTLE}^t_{t_0}(x_0) := \frac{1}{2|t-t_0|}\log(\lambda_n(x_0)).}$$

The FTLE is a diagnostic quantity for Lagrangian Coherent Structures (LCS), i.e. influential surfaces governing the evolution in P.

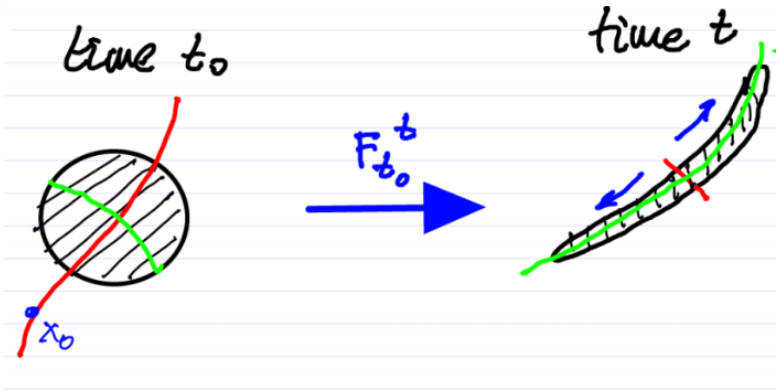


Figure 2.8: On the left the red ridge represents large values of $\mathrm{FTLE}^t_{t_0}$, on the right the green ridge the high values of $\mathrm{FTLE}^{t_0}_t$.

The ridges of $\mathrm{FTLE}^t_{t_0}$ are the repelling LCS, meanwhile the ridges of $\mathrm{FTLE}^{t_0}_t$ are the attracting LCS as depicted in Fig. 2.8. Now we are left with the problem of computing $F^t_{t_0}(x_0)$. Recall that analytically we start with $F^t_{t_0}(x_0)$ and use this to calculate $DF^t_{t_0}(x_0)$. From here we can find $C^t_{t_0}(x_0)$, giving us $\lambda_n(x_0)$ and thereby the FTLE. We now outline a process to compute the FTLE numerically.

(i) Define an initial $M \times N$ grid of initial data $x_0(i,j) \in \mathbb{R}^2$.

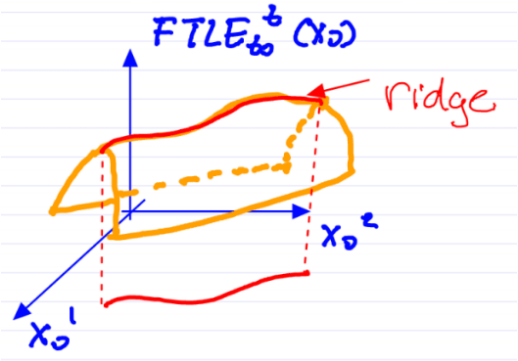


Figure 2.9: The projection of the FTLE ridge onto the initial value space.

(ii) Launch trajectories numerically from grid points to obtain a discrete approximation of $F^t_{t_0}(x_0)$ as $F^t_{t_0}(x_0(i,j))$.

(iii) Use finite differencing to approximate

$$DF^t_{t_0}(x_0(i,j)) \approx \begin{pmatrix} \frac{x(t;t_0,x_0(i,j)+\delta e_1)_1 - x(t;t_0,x_0(i,j)-\delta e_1)_1}{2\delta} & \cdots & \frac{x(t;t_0,x_0(i,j)+\delta e_n)_1 - x(t;t_0,x_0(i,j)-\delta e_n)_1}{2\delta} \\ \vdots & & \vdots \\ \frac{x(t;t_0,x_0(i,j)+\delta e_1)_n - x(t;t_0,x_0(i,j)-\delta e_1)_n}{2\delta} & \cdots & \frac{x(t;t_0,x_0(i,j)+\delta e_n)_n - x(t;t_0,x_0(i,j)-\delta e_n)_n}{2\delta} \end{pmatrix}.$$

This process then yields the surface we see in Fig. 2.9.

Example 2.8 (Calculating the FTLE for the double gyre). Due to incompressibility, we can define the two dimensional flow using a single scalar function called the stream function.

$$\Psi(x,y) = -\sin(\pi x)\sin(\pi y).$$

The components (u,v) of the fluid velocity $(v = (u,v))$ are obtained as partial derivatives of the stream function, according to the formulas

$$\begin{cases} u = \dfrac{\partial \Psi}{\partial y} \\ v = -\dfrac{\partial \Psi}{\partial x}. \end{cases}$$

The Lagrangian trajectories of fluid particles obey the differential equations (i.e. we have the fluid velocity field)

$$\begin{cases} \dot{x} = u = \dfrac{\partial \Psi}{\partial y} \\ \dot{y} = v = -\dfrac{\partial \Psi}{\partial x}. \end{cases}$$

Interestingly, in this case, the phase space coincides with the physical space spanned by the coordinates (x, y).

Remark 2.7. This is an example of a Hamiltonian system, where Ψ is the Hamiltonian (usually denoted as H).

For any autonomous Hamiltonian system we have that the Hamiltonian is constant along trajectories. We can verify this as follows

$$\frac{d}{dt}\Psi(x(t), y(t)) = \frac{\partial \Psi}{\partial x}\dot{x} + \frac{\partial \Psi}{\partial y}\dot{y} = 0.$$

So we have that trajectories are level curves of $\Psi(x, y)$. We can then derive the phase portrait from the level curves of Ψ. Further, we have that $\dot{x} = \frac{\partial \Psi}{\partial y} = -\pi \sin(\pi x)\cos(\pi y)$ which yields that $\text{sign}(\dot{x}) = -\text{sign}(\sin(\pi x))\text{sign}(\cos(\pi y))$. Putting these together we can construct the contour plot with arrows. The contour plot, and FTLE approximation are shown in Fig. 2.10.

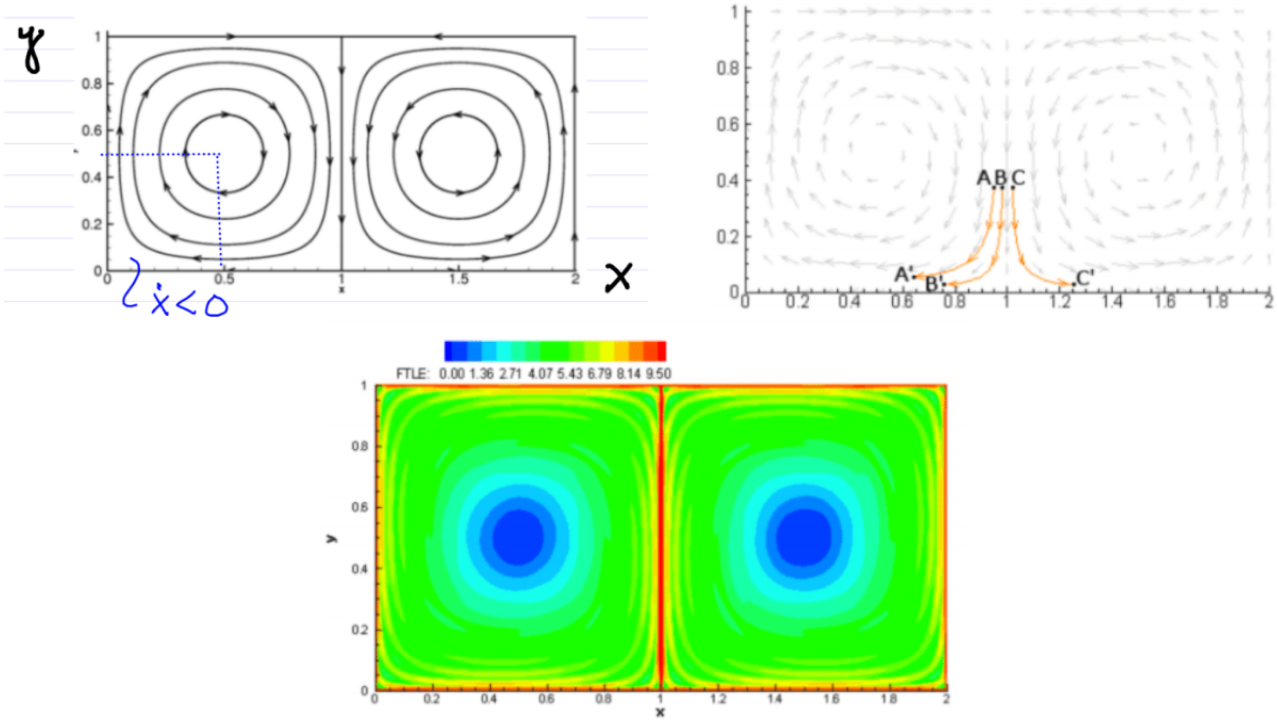


Figure 2.10: Top left: The analytic phase plot. Top right: The exploration done to calculate FTLE. Bottom: The FTLE plot. Figures here were taken from Shawn Shadden of UC Berkeley.

Example 2.9 (ABC flow). Let our dynamical system be defined as follows with $A, B, C \in \mathbb{R}$

$$\begin{cases} \dot{x} = A\sin(z) + C\cos(y) \\ \dot{y} = B\sin(x) + A\cos(z) \\ \dot{z} = C\sin(y) + B\cos(x). \end{cases}$$

This is an exact solution to Euler's equations. We have an autonomous velocity field. Depending on parameters it can even generate chaotic fluid trajectories. The numerical approximation of the FTLE for the ABC flow is depicted in Fig. 2.11.

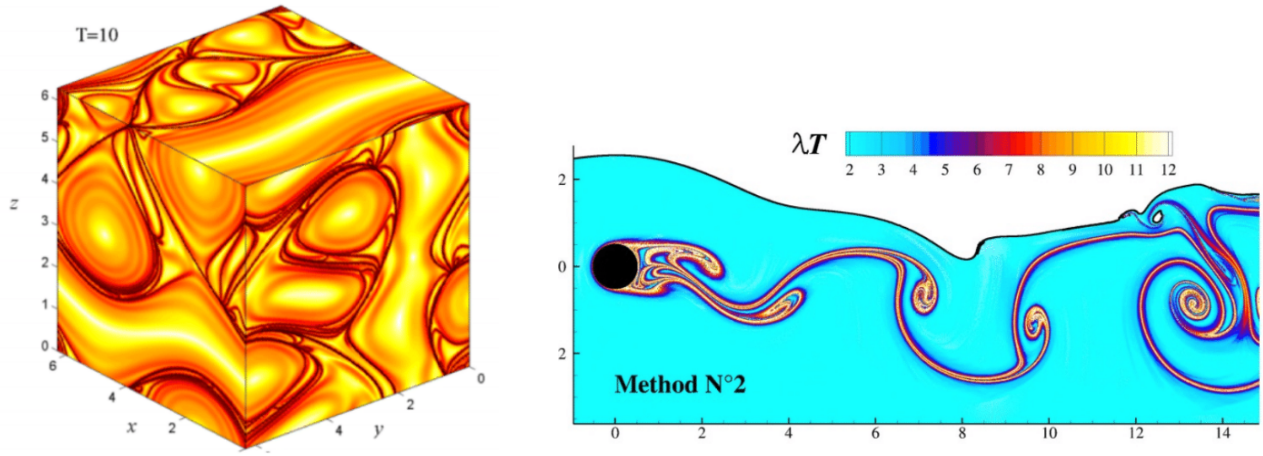


Figure 2.11: Left: numerically calculated FTLE field of the ABC flow. Darker colors signify higher FTLE values [Haller, 2001]. Right: Again the FTLE is plotted, for vortex shedding behind a cylinder under a free surface [Sun et al., 2016].

2.5 Dependence on parameters

We now have the IVP

$$\begin{cases} \dot{x} = f(x, t, \mu) \\ x(t_0) = x_0. \end{cases}$$

With $x \in \mathbb{R}^n,\ f \in \mathcal{C}^r,\ r \geq 1$, therefore we have a solution $x(t; t_0, x_0, \mu) \in \mathcal{C}^r_{x_0}$.

We now examine how solutions depend μ. This is critical as solutions should be robust to changes or uncertainties in the model.

Example 2.10 (Perturbation Theory). Given a weakly nonlinear oscillator

$$m\ddot{x} + c\dot{x} + kx = \varepsilon f(x, \dot{x}, t),\ 0 \leq \varepsilon \ll 1,\ x \in \mathbb{R}.$$

The usual approach is to seek solutions by expanding from the known solution of the linear limit $\varepsilon = 0$, i.e.

$$x_\varepsilon(t) = \varphi_0(t) + \varepsilon\varphi_1(t) + \varepsilon^2\varphi_2(t) + \ldots + \mathcal{O}(\varepsilon^r).$$

If $x_\varepsilon(t)$ is in $\mathcal{C}^r_\varepsilon$, we have $\varphi_1(t) = \frac{\partial x_\varepsilon(t)}{\partial \varepsilon}\Big|_{\varepsilon=0}$ and $\varphi_2(t) = \frac{\partial^2 x_\varepsilon(t)}{\partial \varepsilon^2}\Big|_{\varepsilon=0}$

Regularity with respect to the parameter μ actually follows from regularity with respect to the initial condition x_0. We can use the following trick to extend the IVP with a dummy variable μ

$$\begin{cases} \dot{x} = f(x, t, u) \\ \dot{\mu} = 0 \\ x(t_0) = x_0 \\ \mu(t_0) = \mu_0. \end{cases}$$

Thus with $X = \begin{pmatrix} x \\ \mu \end{pmatrix} \in \mathbb{R}^{n+p}$ and $F(X_0) = \begin{pmatrix} f \\ 0 \end{pmatrix};\ X_0 = \begin{pmatrix} x_0 \\ \mu_0 \end{pmatrix}$. We have the extended IVP

$$\begin{cases} \dot{X} = F(X) \\ X(t_0) = X_0. \end{cases} \tag{2.1}$$

Applying the previous result on regularity with respect to x_0 to (2.1), we have that $f \in \mathcal{C}^r_{x,\mu}$ implies that $X(t) \in \mathcal{C}^r_{X_0}$ in turn implying that $x(t; t_0, x_0, \cdot) \in \mathcal{C}^r_\mu$. The solution is as smooth in parameters as the RHS of the dynamical system.

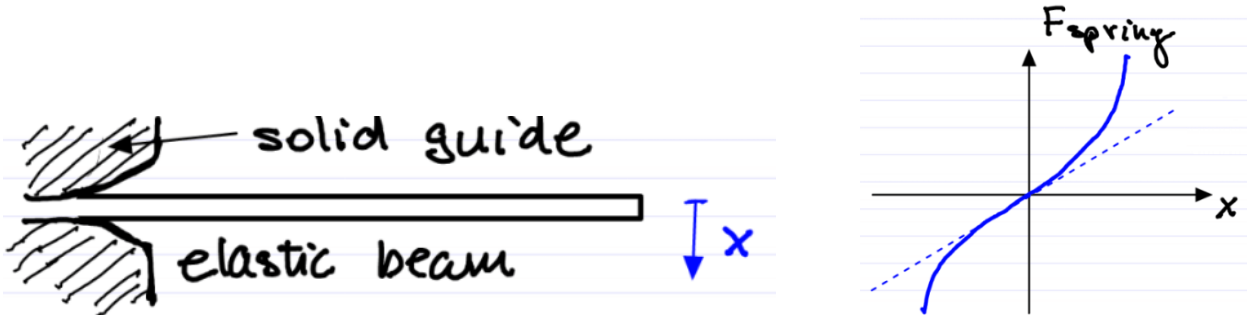


Figure 2.12: Setup for the nonlinear springboard.

Example 2.11 (Periodic Oscillations of a nonlinear springboard). Given an elastic beam extending from a solid guide, we measure the deflection of this beam with the variable x. This system is illustrated in the left panel of Fig. 2.12. By increasing x, the effective free length of the beam is shortened, thereby stiffening the spring nonlinearly. The effect of this nonlinearity on the force exerted on the spring is illustrated in the right panel of Fig. 2.12. This setup yields the following equations of motion

$$\begin{cases} \ddot{x} + x + \varepsilon x^3 = 0; \quad 0 \leq \varepsilon \ll 1 \\ x(0) = a_0; \quad \dot{x}(0) = 0. \end{cases}$$

So we have weak nonlinearity with no known explicit solution. Although weak, this nonlinearity is still significant, as can be seen in Fig. 2.13. Rewriting this as a first order ODE ($x_1 = x$; $x_2 = \dot{x}$), and note that the RHS is $\mathcal{C}^r_{x,\mu}$, therefore there exists a unique local solution that is also $\mathcal{C}^r_\mu$. Thus the expansion is justified

$$x_\varepsilon(t) = \varphi_0(t) + \varepsilon\varphi_1(t) + \ldots + \mathcal{O}(\varepsilon^r). \tag{2.2}$$

We can see, by substitution, that for $\varepsilon = 0$ we find that $\varphi_0(t) = a_0 \cos(t)$.

Now we look specifically for T-periodic solutions, as we would expect such a solution physically, therefore we have

$$\varphi_i(t) = \varphi_i(t + T).$$

The period T still has to be determined. Plugging this power series into (2.2) into the IVP to get

$$\begin{aligned} \mathcal{O}(1): &\quad \ddot{\varphi}_0 + \varphi_0 &&= 0 \\ \mathcal{O}(\varepsilon): &\quad \ddot{\varphi}_1 + \underbrace{\varphi_1}_{\omega=1} &&= -\varphi_0^3 = -a_0^3 \cos^3(t) = -a_0^3 \left[\frac{1}{4}\cos(3t) + \frac{3}{4} \underbrace{\cos(t)}_{\text{resonance}} \right]. \end{aligned} \tag{2.3}$$

We can see that (2.3) is a linear oscillator with a forcing coming from the zeroth order solution. Since the zeroth order solution $\varphi_0 = a_0 \cos(t)$ already solves the IVP we have the following initial conditions

$$\varphi_1(0) = 0; \quad \dot{\varphi}_1(t) = 0.$$

This holds as $\varphi_0 = a_0 \cos(t)$ already solves the IVP. The general solution to this equation is the sum of two terms. We add the general solution of the homogeneous part and a particular solution to the inhomogeneous part. We can write this solution to (2.3) as

$$\begin{aligned}\varphi_1(t) &= \varphi_1^{\text{hom}}(t) + \varphi_1^{\text{part}}(t) \\ &= \underbrace{A\cos(t) + B\sin(t)}_{\text{TBD from initial conditions}} + \underbrace{C\cos(3t) + Dt\cos(t) + Et\sin(t)}_{\text{TBD from (2.3)}}.\end{aligned}$$

Observe that due to a resonance between the natural frequency of the oscillator and the forcing secular terms, $t\cos(t)$ and $t\sin(t)$ appear. Thus it cannot be periodic, so our Ansatz already fails for $i = 1$. We conclude that no solution of this type exists. Our Ansatz was too restrictive and T should depend on ε.

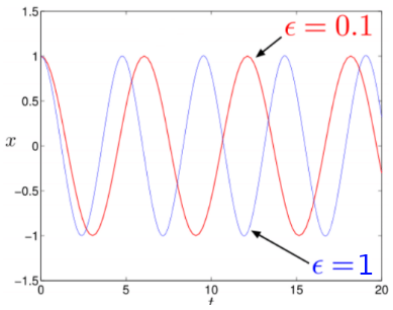


Figure 2.13: Numerical integration of x for $a_0 = 1$ and different values of ε.

Lindstedt's idea We should seek a solution of the form

$$x_\varepsilon(t) = \varphi_0(t;\varepsilon) + \varepsilon\varphi_1(t;\varepsilon) + \varepsilon^2\varphi_2(t;\varepsilon) + \mathcal{O}(\varepsilon^3).$$

Furthermore φ_i should be T_ε periodic, i.e. the period should depend on the strength of the nonlinearity ε.

$$\varphi_i(t + T_\varepsilon;\varepsilon) = \varphi_i(t;\varepsilon).$$

Rewriting the period as

$$T_\varepsilon = \frac{2\pi}{\omega(\varepsilon)}; \quad \omega(\varepsilon) = 1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \mathcal{O}(\varepsilon^3).$$

We then rescale time according to $\tau = \omega(\varepsilon)t$ to find

$$\frac{d}{d\tau} = \frac{1}{\omega(\varepsilon)}\frac{d}{dt} \implies \boxed{[\omega(\varepsilon)]^2 x'' + x + \varepsilon x^3 = 0.}$$

Where we have taken x' to represent $\frac{dx}{d\tau}$. Plugging our expression into the rescaled ODE yields

$$\left(1 + 2\varepsilon\omega_1 + \mathcal{O}(\varepsilon^2)\right)\left[\varphi_0'' + \varepsilon\varphi'' + \mathcal{O}(\varepsilon^2)\right] + \left[\varphi_0 + \varepsilon\varphi_1 + \mathcal{O}(\varepsilon^2)\right] + \varepsilon\left[\varphi_0^3 + \mathcal{O}(\varepsilon)\right] = 0.$$

Matching equal powers of ε yields

$$\begin{aligned}
&\mathcal{O}(1): \ \varphi_0'' + \varphi_0 = 0 \implies \varphi_0(\tau) = a_0\cos(\tau); \quad \varphi_0(0) = a_0; \quad \dot{\varphi}_0(0) = 0 \\
&\mathcal{O}(\varepsilon): \ \varphi_1'' + \varphi_1 = -\phi_0^3 - 2\omega_1\varphi_0'' = \left(2\omega_1 a_0 - \frac{3}{4}a_0^3\right)\underbrace{\cos(\tau)}_{\text{resonance}} - \frac{a_0^3}{4}\cos(3\tau); \\
&\qquad \varphi_1(0) = 0; \quad \dot{\varphi}_1(0) = 0.
\end{aligned}$$

From the first line, we can see the initial conditions are fulfilled. In this step we used that $\dot{\varphi}(t=0) = 0$ if and only if $\omega(\varepsilon)\varphi'(0) = 0$. We get the solution

$$\varphi_1(t) = A\cos(\tau) + B\sin(\tau) + C\cos(3\tau) + D\tau\cos(\tau) + E\tau\sin(\tau).$$

The presence of resonance again excludes periodic solutions, but now we can select ω_1 to eliminate these terms.

$$2\omega_1 a_0 - \frac{3}{4}a_0^3 = 0 \implies \boxed{\omega_1 = \frac{3}{8}a_0^2.}$$

This successfully eliminates the resonance and determines the missing frequency term at $\mathcal{O}(\varepsilon)$. Thus we find

$$x_\varepsilon(\tau) = a_0\cos(\tau) - \frac{\varepsilon}{32}a_0^3\left(\cos(\tau) - \cos(3\tau)\right) + \mathcal{O}(\varepsilon^2).$$

In the original time scaling this is

$$x_\varepsilon(t) = a_0\cos(\omega t) - \frac{\varepsilon}{32}a_0^3\left(\cos(\omega t) - \cos(3\omega t)\right) + \mathcal{O}(\varepsilon^2); \quad \omega = 1 + \frac{3}{8}\varepsilon a_0^2 + \mathcal{O}(\varepsilon^2).$$

This procedure can be continued to higher order terms, where we select ω_2 so that the $\mathcal{O}(\varepsilon^2)$ terms cancel.

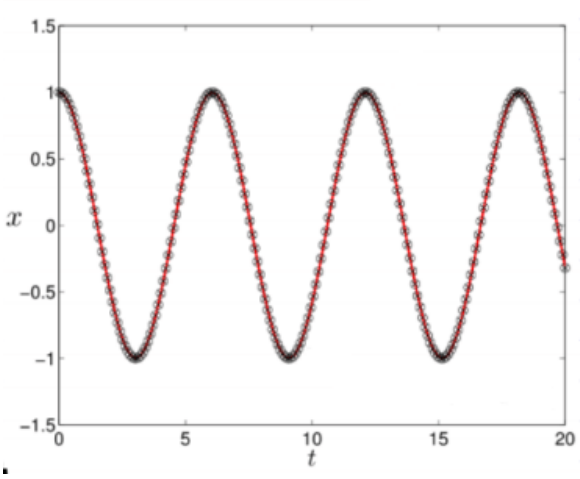


Figure 2.14: Approximation (dots) vs analytic solution (solid line) of x on the time interval $[0, 20]$.

Chapter 3

Stability of fixed points

Now we would like to begin to explore the behaviour of dynamical systems around fixed points. This will allow us to find out if we should expect to observe a fixed state, and to understand what happens if we perturb the system away from this fixed state.

3.1 Basic definitions

Consider

$$\dot{x} = f(x,t),\ x \in \mathbb{R}^n,\ f \in \mathcal{C}^1.$$

Assume that $x = 0$ is a fixed point, i.e. $f(0,t) = 0$ for all $t \in \mathbb{R}$. If the fixed point is originally at $x_0 \neq 0$, shift it to zero by letting $\tilde{x} := x - x_0$, therefore

$$\dot{\tilde{x}} = \dot{x} = f(x_0 + \tilde{x}, t) = \tilde{f}(\tilde{x}, t).$$

We would like to understand how the dynamical system behaves near its equilibrium state. To this end we introduce the following definitions.

Definition 3.1 (Lyapupnov Stability)**.** The fixed point $x = 0$ is stable if for all t_0, for all $\varepsilon > 0$ small enough, there exists a $\delta = \delta(t_0, \varepsilon)$, such that for all $x_0 \in \mathbb{R}^n$ with $\|x_0\| \leq \delta$, we have

$$\boxed{\|x(t; t_0, x_0)\| \leq \varepsilon \quad \forall t \geq t_0.}$$

Remark 3.1 (N-dimensional ball). When writing $\mathcal{B}(r)$ we refer to the ball of radius r in $\mathbb{R}^n$, i.e. the set $\{x :\ \|x\| < r\}$.

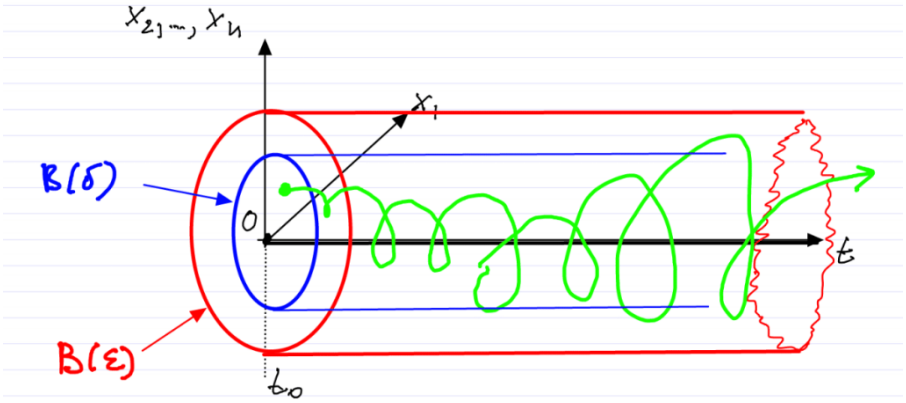


Figure 3.1: An example such a δ for a given Lyapunov stable fixed point.

Example 3.1 (Stability of the lower equilibrium of the pendulum). Recall the equation of motion of the pendulum $\ddot{\varphi} + \sin(\varphi) = 0$, that we transform into a first order ODE by setting $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ to obtain

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\sin(x_1). \end{cases}$$

For small $\varepsilon > 0$, this geometric procedure gives a $\delta(\varepsilon) > 0$ such that the definition of stability is satisfied for $x = 0$. We can see in Fig. 3.2 that for any initial point chosen within the blue circle, it's trajectory remains within the red circle for all time (cf. Fig. 3.1). Therefore $x = 0$ is (Lyapunov) stable.

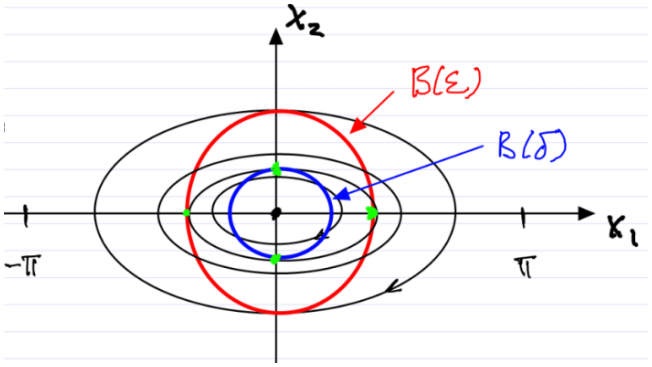


Figure 3.2: Stability of lower equilibrium for the pendulum, here $0 < \varepsilon < \pi$.

Definition 3.2 (Asymptotic stability)**.** The fixed point $x = 0$ is *asymptotically stable* if

(i) it is stable,

(ii) for all t_0, there exists $\delta_0(t_0, \varepsilon)$ such that for every x_0 with $\|x_0\| \leq \delta_0$ we have

$$\boxed{\lim_{t\to\infty} x(t; t_0, x_0) = 0.}$$

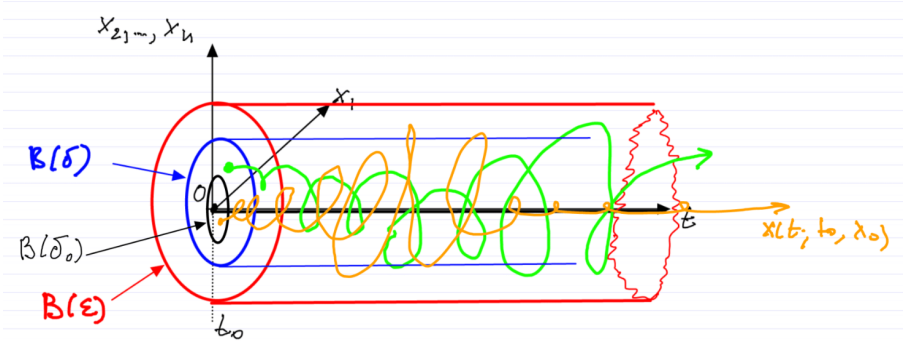


Figure 3.3: An example for an asymptotically stable fixed point (black trajectory).

Definition 3.3 (Domain of attraction)**.** The *domain of attraction* of the fixed point $x = 0$ is the set of all x_0's for which

$$\boxed{\lim_{t\to\infty} x(t; t_0, x_0) = 0.}$$

Example 3.2 (Damped pendulum). We have the equation of motion with the linear damping coefficient c

$$\ddot{\varphi} + c\dot{\varphi} + \sin(\varphi) = 0, \quad c > 0.$$

Transforming into a first-order ODE with $x_1 = \varphi$ and $x_2 = \dot{\varphi}$ gives

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -cx_2 - \sin(x_1). \end{cases}$$

The total energy is given by

$$E = \frac{1}{2}x_2^2 + (1 - \cos(x_1)).$$

Further we have the rate of energy change

$$\frac{d}{dt}E(x_1(t), x_2(t)) = x_2\left(\dot{x}_2 + \sin(x_1)\right) = -cx_2^2.$$

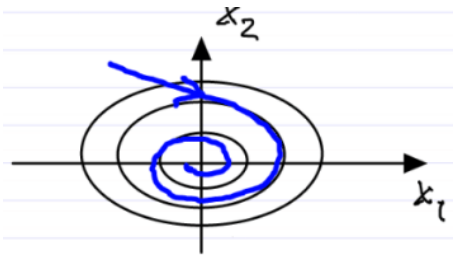


Figure 3.4: An example of a trajectory which loses energy, in this case due to damping.

Therefore, along trajectories energy decreases monotonically as shown in Fig 3.4. By the $\mathcal{C}^0$ dependence of the trajectory on initial conditions, the trajectories remain close to the undamped oscillations for small $c > 0$. We conclude that trajectories are inward spirals for a small dissapation $c > 0$. The fixed point $x = 0$ is still Lyapunov stable, but asymptotic stability does not yet follow (is the limit of $x(t)$ equal to 0?).

Remark 3.2 (LaSalle's invariance principle). This conclusion follows rigorously from LaSalle's invariance principle, namely if we assume that $\dot{x} = f(x)$, $f \in \mathcal{C}^1$, and that there exists a $V \in \mathcal{C}^1$ with

$$\dot{V} = \frac{dV(x(t))}{dt} \leq 0.$$

Then the set of accumulation points for any trajectory is contained in the set of trajectories that stay within the set $I = \{x \in \mathbb{R}^n : \dot{V}(x) = 0\}$.

Example 3.3. Consider the following dynamical system in polar coordinates, i.e. $r\cos(\theta) = x$ and $r\sin(\theta) = y$,

$$\begin{cases} \dot{r} = r(1-r) \\ \dot{\theta} = \sin^2\left(\dfrac{\theta}{2}\right). \end{cases}$$

Note that $r = 0$ is a fixed point, the set $r = 1$ is an invariant circle, and the set $\theta = 0$ is an invariant set. An invariant set is a set such that if the dynamical system is started on the set, it remains in the set for all time. Examining the radial evolution reveals that the equation of motion decouples. We see that $\dot{\theta} \geq 0$, so rotation is either positive or null.

From Fig. 3.5 we can see that both of the fixed points, $(0,0)$ and $(1,0)$, are not stable. However, inspecting Fig. 3.5 we see that that $p = (1,0)$ is an example of an attractor: a set with an open neighborhood of points that all approach the set as $t \to \infty$.

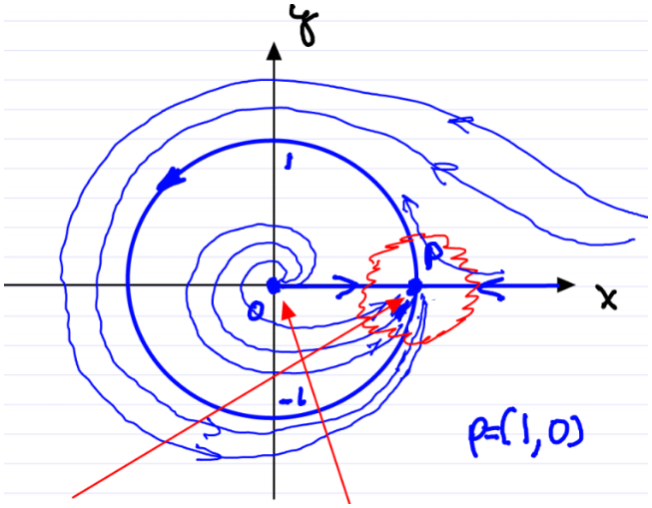


Figure 3.5: Phase portrait of the dynamical system in cartesian coordinates, with the red arrows pointing to the two unstable equilibria.

Definition 3.4 (Invariant set)**.** The set $S \subset P$ is an *invariant set* for the flow map $F^t : P \to P$ if $F^t(S) = S$ for all $t \in \mathbb{R}$.

Definition 3.5 (Unstable point)**.** A fixed point $x = 0$ is unstable if it is not stable.

Remark 3.3. We can negate a mathematical statement by using the reverse relational operators outside the statements involving these operators i.e. $\exists \to \forall$ and $\forall \to \exists$. For example we have for continuity $\forall\varepsilon\ \exists\delta :\ \|f(x) - f(y)\| < \varepsilon$ if $\|x - y\| < \delta$, meanwhile for discontinuity we have $\exists\varepsilon :\ \forall\delta :\ \|f(x) - f(y)\| \geq \varepsilon$ for $\|x - y\| < \delta$.

In our case for stability we have

$$\forall\varepsilon, t_0 : \quad \exists\delta > 0 : \quad \forall x_0 \text{ with } \|x_0\| < \delta : \quad \|x(t)\| \leq \varepsilon \quad \forall t \geq t_0.$$

Meanwhile for instability

$$\exists\varepsilon, t_0 : \quad \underbrace{\forall\delta > 0}_{\text{“for arbitrarily small”}} \quad : \quad \exists x_0 \text{ with } \|x_0\| < \delta : \quad \|x(t)\| > \varepsilon \quad \underbrace{\exists t \geq t_0}_{\text{“for some”}} .$$

This negation is demonstration in Fig. 3.6.

Remark 3.4. By $\mathcal{C}^0$ dependence of trajectories on initial conditions, if $x(t; t_0, x_0)$ leaves $\mathcal{B}(\varepsilon)$, then for $\tilde{x}_0$ close enough to x_0, $x(t; t_0, \tilde{x}_0)$ also leaves $\mathcal{B}(\varepsilon)$. Since this is true on an open set around x_0, the measure of such trajectories in nonzero, the instability is observable!

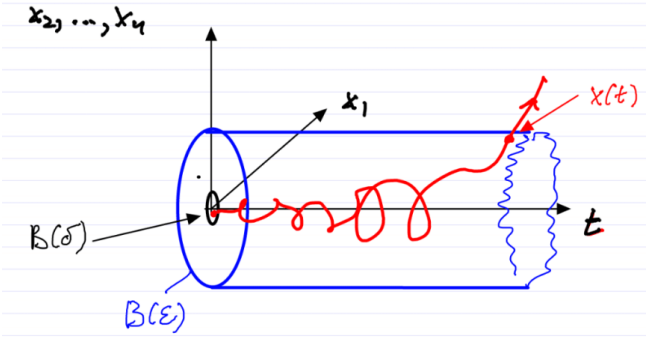


Figure 3.6: Example of an unstable fixed point, with the red trajectory representing a trajectory starting arbitrarily close to the fixed point, leaving a given ε-ball.

Example 3.4 (Unstable fixed point of pendulum). In contrast, we can have that infinitely many trajectories converge to the fixed point, yet it is still unstable, as illustrated in Fig. 3.7. In fact, the converging trajectories form a measure-zero set, thus the stability near the unstable equilibrium is unobservable.

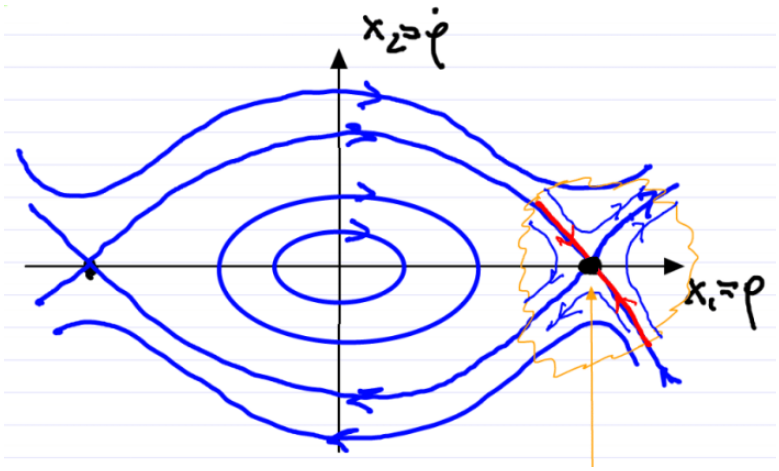


Figure 3.7: The phase portrait around the unstable fixed point of the pendulum, with the stable trajectories (red).

3.2 Stability based on linearization

We would like to derive a more general method to analyze the stability of fixed points, thus we try to simplify our system around the fixed point and discover what this can tell us about the full (unsimplified) system. In the following section we shall always assume that our system is autonomous. We will have the following setup

$$\dot{x} = f(x), \quad f \in \mathcal{C}^1, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{R}^n. \tag{3.1}$$

If $f(p) = 0$, then p is a fixed point. By transforming using $y = x - p$, we have that in the transformed system $y = 0$ is a fixed point. Furthermore, we have that around $y = 0$ the ODE is

$$\dot{y} = f(p + y) = \underbrace{f(p)}_{=0} + Df(p)y + o(\|y\|) = Df(p)y + o(\|y\|).$$

Remark 3.5. Since we only assumed one continuous derivative for the function f the remainder term in the Taylor approximation is $o(\|y\|)$. The little o notation means that

$$\lim_{\|y\| \to 0} \frac{o(\|y\|)}{\|y\|} = 0.$$

Definition 3.6 (Linearized ODE)**.** We define the *linearization* of (3.1) at the fixed point p as

$$\boxed{\dot{y} = Ay; \quad y \in \mathbb{R}^n, \quad A := Df(p) \in \mathbb{R}_{n \times n}; \quad Df(p) = \left(\begin{matrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{matrix} \right) \Bigg|_{x=p}.} \tag{3.2}$$

Now we would like to study the stability of the fixed point $y = 0$ in (3.2). From this analysis, we want to know the relevance of our results for the full nonlinear system (3.1).

3.3 Review of linear dynamical systems

Recall the setup

$$\dot{y} = A(t)y, \quad y \in \mathbb{R}^n, \quad A \in \mathbb{R}^{n \times n}, \quad A \in \mathcal{C}^0_t.$$

The following facts have already been established

- We know that the global existence and uniqueness of solutions is guaranteed.
- The superposition principle holds; namely the linear combination of solutions is also a solution.
- There exists a set of n linearly independent solutions: $\varphi_1(t), \ldots, \varphi_n(t) \in \mathbb{R}_n$.
- The general solution is

$$y(t) = \sum_{i=1}^{n} c_i \varphi(t) = \underbrace{\begin{bmatrix} \varphi_1(t) & \ldots & \varphi_n(t) \end{bmatrix}}_{\Psi(t)\text{: fundamental matrix solution}} \underbrace{\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}}_{c} = \Psi(t)c; \quad \dot{\Psi} = A(t)\Psi.$$

- We have the initial value problem $y(t_0) = y_0$ which implies

$$\Psi(t_0)c = y_0 \implies y(t) = \underbrace{\Psi(t)\left[\Psi(t_0)\right]^{-1}}_{\phi(t) := F_{t_0}^t} y_0$$

  Where we used that the $\varphi_i(t)$ are linearly independent in the last equality. And we have the *normalized fundamental matrix* $\phi(t)$ equal to the flow map, with $\phi(t_0) = I$.

- In the autonomous case $\dot{x} = Ax$ solutions can be practically constructed.

  (i) **Explicit Solution**

$$\phi(t) = e^{At} := \sum_{j=0}^{\infty} \frac{1}{j!}(At)^j$$

  With $0! = 1$ and $0^0 = I$. We can verify that this is indeed a solution

$$\dot{\phi}(t) = \sum_{j=1}^{\infty} \frac{1}{(j-1)!}(At)^{j-1}A = A\sum_{j=0}^{\infty} \frac{1}{j!}(At)^j = Ae^{At} = A\phi(t).$$

  Where we used that A commutes with its powers in the second equality. We now have that each column of $\phi(t)$ satisfies $\dot{y} = Ay$.

  Remark 3.6. For a scalar ODE $\dot{y} = a(t)y$ for $y \in \mathbb{R}$, the solution is known $y(t) = e^{\int_{t_0}^t a(s)ds} y_0$. However, this does not extend to the higher dimensional $\dot{y} = A(t)y$. In fact, in general, $\phi(t) = e^{\int_{t_0}^t A(s)ds}$ is <u>not</u> a solution. We can check this

$$\dot{\phi} = \sum_{j=0}^{\infty} \frac{1}{j!}\frac{d}{dt}\left(\int_{t_0}^t A(s)ds\right)^j = \sum_{j=1}^{\infty} \frac{1}{(j-1)!}\left(\int_{t_0}^t A(s)ds\right)^{j-1} A(t) \neq A(t)\phi(t).$$

  The nonequality holds as $A(t)$ does not generally commute with $\int A(s)ds$.

(ii) **Solution from eigenfunctions** If we have an autonomous system, we can solve the ODE without an infinite series. We have

$$\dot{y} = Ay, \quad y \in \mathbb{R}^n, \quad y(0) = y_0. \tag{3.3}$$

Substituting $\varphi(t) = e^{\lambda t}s$ for $\lambda \in \mathbb{C}$ and $s \in \mathbb{C}^n$ into (3.3) yields

$$\lambda s = As \implies (A - \lambda I)s = 0 \iff \det(A - \lambda I) = 0.$$

Therefore λ must be an eigenvalue of A and s must be the corresponding eigenvector. We call $\det(A-\lambda I)$ the *characteristic equation* of A. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues and $s_1, \ldots, s_n$ be the corresponding eigenvectors. In the case that some eigenvalues are repeated, some of the s_i may be generalized eigenvectors. We then have two cases.

(a) A is semisimple, i.e. the eigenvectors are linearly independent (which is always the case if the λ_i all have algebraic multiplicity of one). Then we have the solution

$$y(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t} s_i = \sum_{j=1}^{n} c_j e^{(\mathrm{Re}\lambda_j)t} e^{i(\mathrm{Im}\lambda_j)t} s_j.$$

Where we used $\lambda_j = \mathrm{Re}\lambda_j + i\mathrm{Im}\lambda_j$.

(b) A is <u>not</u> semisimple, i.e. has repeated eigenvalues (but not enough linearly independent eigenvectors). Then we assume that λ_k has algebraic multiplicity $a_k > g_k$, where a_k measures the multiplicity of λ_k as a root of $\det(A - \lambda I) = 0$, and g_k is the number of linearly independent eigenvectors for λ_k, also called the *geometric muliplicity* of λ_k. Even in this case, λ_k gives rise to a_k linearly independent solutions of the form

$$\underbrace{P_0}_{=s_k} e^{\lambda_k t},\ P_1(t)e^{\lambda_k t},\ P_2(t)e^{\lambda_k t}, \ldots, P_{a_{k-1}}(t)e^{\lambda_k t}$$

where $P_j(t)$ is a vector polynomial of t of order j or less.

3.4 Stability of fixed points in autonomous linear systems

First we note that we can bound our solution

$$\|y(t)\| = \|\phi(t)y_0\| \leq \underbrace{\|\phi(t)\|}_{\text{Operator norm}} \|y_0\| \leq Ce^{\mu t}\|y_0\|. \tag{3.5}$$

Where $\mu = \max_j(\mathrm{Re}\lambda_j) + \nu$, with $\nu > 0$, as small as needed, provided we increase C appropriately. If A is semisimple, then $\nu = 0$ can be selected.

Theorem 3.7 (Stability of fixed points in linear systems)**.** *Given $y = 0$ a fixed point of the linear system $\dot{y} = Ay$ with $A \in \mathbb{R}^{n\times n}$ the following statements hold:*

(i) Assume that $Re\lambda_j < 0$ for all j. Then $y = 0$ is asymptotically stable.

(ii) Assume that $Re\lambda_j \leq 0$ for all j, and for all λ_k with $Re\lambda_k = 0$ we have $a_k = g_k$. Then $y = 0$ is stable.

(iii) Assume there exists a k such that $Re\lambda_k > 0$. Then $y = 0$ is unstable.

These scenarios are illustrated in Fig. 3.8.

Proof. (i) Pick $\varepsilon > 0$, and select $\nu > 0$ small, such that $\mu < 0$. Then pick $C > 0$ such that (3.5) holds, and let $\delta = \frac{\varepsilon}{C}$. This implies (since $\|y_0\| \leq \delta$) that

$$\|y(t)\| \leq \varepsilon e^{\mu t} \leq \varepsilon,$$

and

$$\|y(t)\| \leq \varepsilon e^{\mu t} \xrightarrow{t\to\infty} 0.$$

Where the limit holds as $\mu < 0$.

(ii) Again choose $\delta = \frac{\varepsilon}{C}$ and note that $\mu = \max_j(\mathrm{Re}\lambda_j) + \nu = 0 + \nu = 0$ ($\nu = 0$ as $a_k = g_k$). Then stability follows by (3.5). However, asymptotic stability does not hold, as $\varphi(t) = Ce^{i(\mathrm{Im}\lambda_j)t}$ solutions exist.

(iii) There exists a solution of the form

$$\varphi(t) = C_k e^{\lambda_k t} s_k = C_k e^{(\mathrm{Re}\lambda_k)t} e^{i(\mathrm{Im}\lambda_k)t} s_k.$$

In turn this implies

$$\|\varphi(t)\| = C_k e^{(\mathrm{Re}\lambda_k)t} \|s_k\| \xrightarrow{t\to\infty} \infty.$$

□

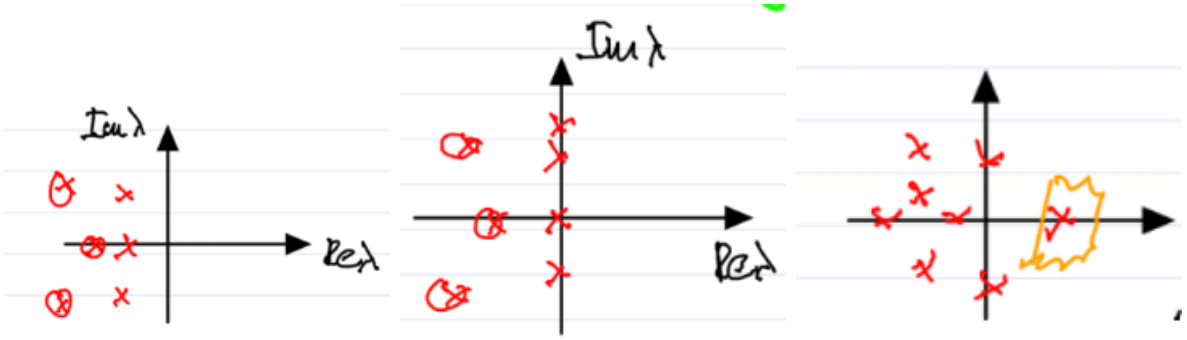


Figure 3.8: Eigenvalue arrangements for scenarios (i), (ii), and (iii) (from left to right) in Theorem 3.7

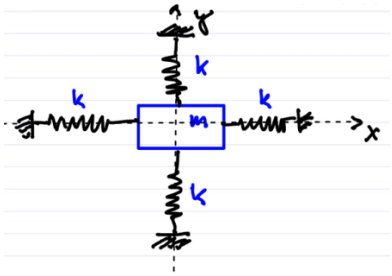


Figure 3.9: Arrangement of coupled oscillators with rectangular mass in the middle.

Example 3.5 (Stability analysis of 2 degrees of freedom coupled oscillators). Given a rectangular mass m with a spring of stiffness coefficient k attached to each side extending to fixed walls in each cardinal direction. We want to know the stability of the equilibrium where all of the springs are equally extended. This dynamical system is depicted in Fig. 3.9. First, note that this is a conservative system, i.e. $E = \text{const}$. Next we transform the coordinates so that the equations of motion can be brought into the form of an ODE for this dynamical system

$$x = \begin{pmatrix} x \\ \dot{x} \\ y \\ \dot{y} \end{pmatrix}.$$

Thus we have a 4-dimensional, nonlinear, system of ODEs. We now linearize this at the fixed point $(x, y) = (0, 0)$, i.e. $\dot{x} = Ax$ with $x \in \mathbb{R}^n$.

The system exhibits full spatial symmetry in x and y, hence the eigenmodes will be the same in the x and y directions. This means we have repeated pairs of purely imaginary eigenvalues for A: $\lambda_{1,2} = \lambda_{3,4} = \pm i\omega$. It is clear that scenarios (i) and (iii) of Theorem 3.7 do not apply to the linearized ODE. So we need to check if (ii) applies.

We have that $\mathrm{Re}\lambda_k = 0$ for $k = 1, 2, 3, 4$. Also $a_k = 2$ for $k = 1, 2, 3, 4$. Now assume $g_k < 2$. Then there would exists solutions of the form $te^{\pm i\omega t}s_k$, but this would contradict the conservation of energy, as either the (nonnegative) kinetic energy and/or the (nonnegative) potential energy would grow unbounded. Hence, the total energy could not be conserved. Therefore we know that $g_k = a_k$ and we can apply (ii) to find $x = y = 0$ is Lyapunov stable for the linearized system. What does this imply for the nonlinear system?

3.5 Stability of fixed points in nonlinear systems

Following the previous example, we would like to know what information about the stability of fixed points of nonlinear systems we can derive from the linearized system. The full nonlinear system is

$$\dot{x} = f(x), \quad f(x_0) = 0, \quad x \in \mathbb{R}^n, \quad f \in \mathcal{C}^1. \tag{3.6}$$

And its linearization at the fixed point x_0

$$\dot{y} = Df(x_0)y, \quad y \in \mathbb{R}^n, \quad Df(x_0) \in \mathbb{R}^{n\times n}. \tag{3.7}$$

We would like to conclude that the linearized dynamics are qualitatively similar to the nonlinear dynamics. In order to study if this is the case, we have to formalize *similar* mathematically.

Definition 3.7 ($\mathcal{C}^k$ equivalence of dynamical systems)**.** Consider two autonomous dynamical systems:

(i)

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \quad f \in \mathcal{C}^1; \quad F^t : x_0 \mapsto x(t; x_0). \tag{3.8}$$

(ii)

$$\dot{x} = g(x), \quad x \in \mathbb{R}^n, \quad g \in \mathcal{C}^1; \quad G^t : x_0 \mapsto x(t; x_0).$$

The two dynamical systems are $\mathcal{C}^k$ *equivalent*, for $k \in \mathbb{N}$, on an open set $U \subset \mathbb{R}^n$, if there exists a $\mathcal{C}^k$ diffeomorphism $h : U \to U$ that maps orbits of (i) into orbits of (ii), while preserving the orientation but not necessarily the exact parameterization of the orbit by time. Specifically for all $x \in U$, any $t_1 \in \mathbb{R}$ there exists a $t_2 \in \mathbb{R}$ such that

$$\boxed{h(F^{t_1}(x)) = G^{t_2}(h(x)).}$$

$h : U \to U$ does this for all $x \in U$ in a $\mathcal{C}^k$ fashion. This equivalence through a function h is demonstrated in Fig 3.10.

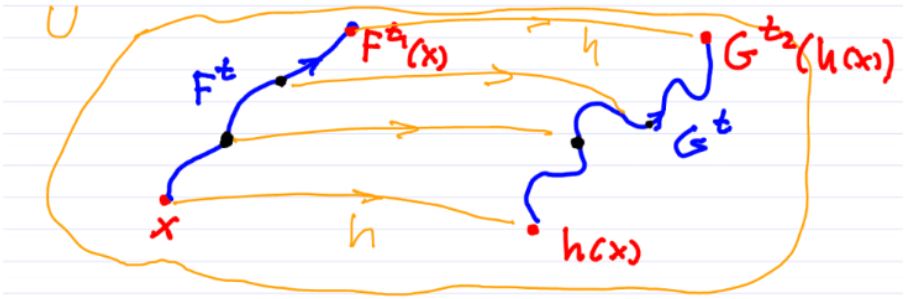


Figure 3.10: The function h mapping the orbits of the dynamical system describing F into the system describing G.

Proposition 3.8 (Restrictiveness of smooth equivalence)**.** *Consider two $\mathcal{C}^k$-equivalent dynamical systems for $k > 0$ with the right hand sides f and g, as in the definitions above. Let x_0 be a fixed point of F and y_0 a fixed point of G such that these correspond to each other under the $\mathcal{C}^k$ equivalence, i.e. $y_0 = h(x_0)$. The eigenvalues of $Df(x_0)$ must be constant, positive multiples of the eigenvalues of $Dg(y_0) = Dg(h(x_0))$.*

Proof. By the $\mathcal{C}^k$ equivalence we have that there exists a $\tau(x,t)$ with $\frac{d\tau}{dt} > 0$ such that $h(F^t(x)) = G^{\tau(x,t)}(h(x))$. Further assume that τ is at least continuously differentiable in its arguments. Differentiating the previous equality with respect to x yields

$$DhDF^t = \dot{G}^\tau \frac{d\tau}{dx} + DG^\tau Dh.$$

At a fixed point x_0 we have that $\dot{G}^\tau(x_0) = g(G^\tau(x_0)) = 0$. This implies that the first term on the right hand side of the differentiated equality is equal to 0. Now denote the linearizations of f and g at the fixed point by $A = Df(x_0)$ and $B = Dg(h(x_0))$ respectively. Then the matrices e^{At} and $e^{B\tau}$ must be similar matrices, in turn implying that At and $B\tau(x_0,t)$ must have the same spectrum, i.e.

$$\frac{\lambda_i(B)}{\lambda_i(A)} = \frac{t}{\tau(x_0,t)} = C = \text{const.} > 0.$$

Therefore the eigenvalues of the linearizations are constant, positive multiples of each other. □

Definition 3.8 ($\mathcal{C}^k$ conjugacy)**.** Consider the same two dynamical systems as before. The two dynamical systems are $\mathcal{C}^k$ *conjugate* if there exists a $\mathcal{C}^k$ diffeomorphism $h : U \to U$ that maps orbits of (i) into (ii), while preserving orientation and the parameterization of the orbit by time. Specifically for all $x \in U$ and $t \in \mathbb{R}$ we have

$$\boxed{h(F^t(x)) = G^t(h(x)).}$$

Proposition 3.9 (Restrictiveness of smooth conjugacy)**.** *Consider two $\mathcal{C}^k$ conjugate systems, as in the definition above. In addition to the previous proposition, now the linearizations of the two systems will have the same spectra at each fixed point.*

Proof. The $\mathcal{C}^k$ conjugacy requires preserving the parameterizations of the orbits, i.e. $\tau(x,t) = t$. Hence, this is just a specification of the previous proof. □

Definition 3.9 (Topological equivalence)**.** For $k = 0$, $\mathcal{C}^k$ equivalence is also called *topological equivalence*. In this case, a continuous, invertible deformation takes orbits of one system into the orbits of the other. Under these conditions, $h : U \to U$ is called a *homeomorhpism*.

Example 3.6 (Topologically equivalent linear systems for $n = 2$). To illustrate the meaning of topological equivalence, Fig. 3.11 shows three linear systems ($\dot{x} = Ax$ for $x \in \mathbb{R}^2$) which are topologically equivalent.

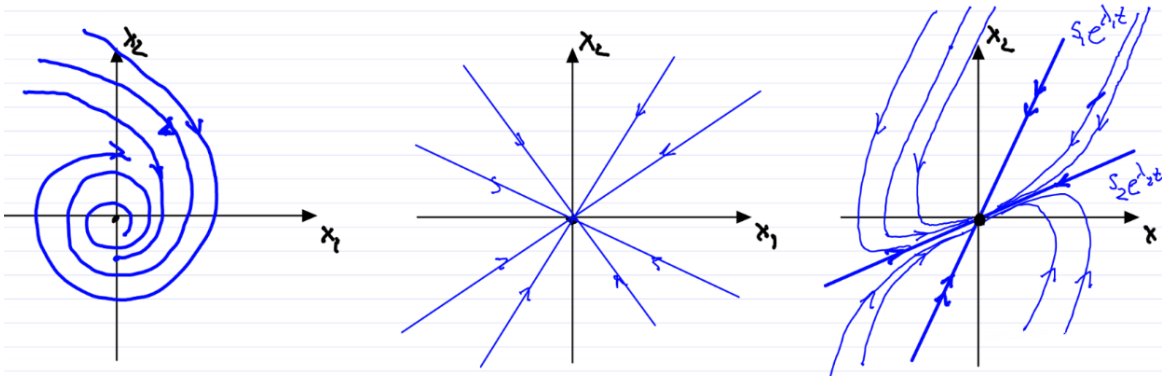


Figure 3.11: Three topologically equivalent 2-dimensional linear systems. Left: The stable spiral. Middle: The sink. Left: The stable node.

The stable spiral has the eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$ for $\alpha < 0$ and $\beta \neq 0$. The sink has the eigenvalues $\lambda_1 = \lambda_2 < 0$. and The stable node has the eigenvalues $\lambda_1 < \lambda_2 < 0$. Note here that the number of eigenvalues λ_i with $\mathrm{Re}\lambda_i < 0$, $\mathrm{Re}\lambda_i = 0$, and $\mathrm{Re}\lambda_i > 0$ is the same in all three of these cases, namely for each system the real part of both eigenvalues are less than 0.

Remark 3.10. The above systems are not $\mathcal{C}^k$ equivalent or conjugate, as the spectra differ by more than a multiplicative constant.

Example 3.7 (Topologically inequivalent linear systems for $n = 2$). As a counter example, we now present three linear systems which are not topologically equivalent. The stable spiral (from before), the unstable spiral, and the saddle. The unstable spiral has the eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$ for $\alpha > 0$ and $\beta \neq 0$ (note the different sign for α). The saddle has the eigenvalues $\lambda_1 < 0 < \lambda_2$. These systems are depicted in Fig 3.12.

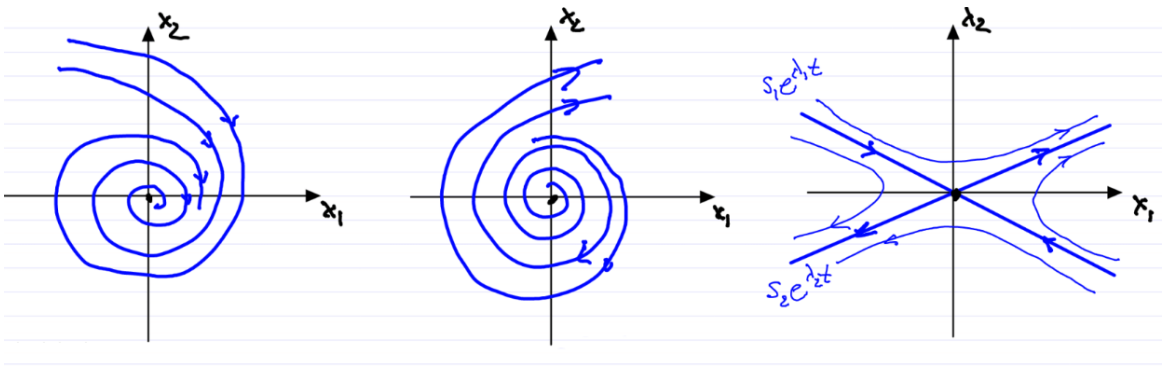


Figure 3.12: Three 2-dimensional linear systems which are not topologically equivalent. Left: The stable spiral. Middle: The unstable spiral. Right: The saddle.

Note here that the eigenvalue configurations in terms of the number of λ_i with real part less than 0 are different in each case. Building on the role of the eigenvalue configuration we noted in the previous examples, we introduce the concept of a hyperbolic fixed point.

Definition 3.10 (Hyperbolic fixed point)**.** We call the fixed point $x = x_0$ a *hyperbolic fixed point* of (3.6) if each of the eigenvalues λ_i of its linearization (3.7) satisfy

$$\boxed{\mathrm{Re}\lambda_i \neq 0.}$$

Geometrically the eigenvalue configuration of a hyperbolic fixed point is shown in Fig. 3.13.

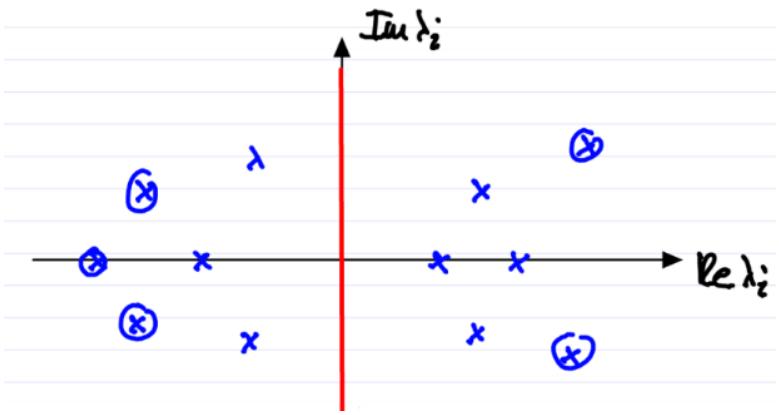


Figure 3.13: The eigenvalue configuration of a hyperbolic fixed point, i.e. no eigenvalues are on the imaginary axis (red).

Proposition 3.11. *Under small perturbations to the nonlinear system, the linearized stability type of the hyperbolic fixed point is preserved.*

Before proving this result, recall the Implicit Function Theorem (without proof).

Theorem 3.12 (Implicit Function Theorem ($n+1$ dimensional case))**.** *For a function $F : \mathbb{R}^{n+1} \to \mathbb{R}^n$ which is $\mathcal{C}^1$, if $F(x_0, y_0) = 0$ and the Jacobian $D_x F(x_0, y_0)$ is nonsingular (invertible), then there exists a nearby solution to $F(x, y) = 0$, for $x_y = x_0 + \mathcal{O}(|y - y_0|)$. Further x_y is as smooth in y as $F(x, y)$.*

Proof (Proposition). Add a small perturbation to (3.8) i.e.

$$\dot{x} = f(x) + \varepsilon g(x); \quad |\varepsilon| \ll 1, \quad f(x_0) = 0.$$

Now we ask if the perturbed system has a fixed point x_ε near x_0. We frame this in terms of the implicit function theorem

$$F(x, \varepsilon) = f(x) + \varepsilon g(x) \stackrel{?}{=} 0; \quad F(x_0, 0) = 0; \quad x \in \mathbb{R}^n, \quad F : \mathbb{R}^{n+1} \to \mathbb{R}^1.$$

We check that $D_x F(x_0, 0)$ is nonsingular exactly when $Df(x_0)$ is; this is fulfilled as we have no zero eigenvalues. The linearization at the perturbed fixed point takes the form

$$\dot{y} = D\left[f(x) + \varepsilon g(x)\right]\big|_{x=x_\varepsilon} y = \left[Df(x_0 + \mathcal{O}(\varepsilon)) + \varepsilon Dg(x_0 + \mathcal{O}(\varepsilon))\right] y = \underbrace{\left[Df(x_0) + \mathcal{O}(\varepsilon)\right]}_{=A_\varepsilon} y.$$

In the last equality we used the Taylor expansion in ε. We have that the roots of $\det(A_\varepsilon - \lambda I) = 0$ depend continuously on the parameter ε. These roots correspond to the eigenvalues of A_ε. Therefore, the roots stay within an $\mathcal{O}(\varepsilon)$ neighborhood of the eigenvalues of $Df(x_0)$ (see Fig. 3.14). Hence we have that the eigenvalue configuration is unchanged for small enough ε. □

Remark 3.13. In the above proof, not only does the hyperbolicity of fixed points remain preserved, but also the stability type.

Meanwhile, for nonhyperbolic fixed points, this is not the case, and the smallest perturbation may change their stability type. This is due to the fact, that no matter how small the scale of the perturbation (ε) the $\mathcal{O}(\varepsilon)$ ball around eigenvalues on the imaginary axis will always intersect with $\mathbb{C} - \{\mathrm{Im}\lambda_i = 0\}$ (i.e. points which are not on the imaginary axis).

Now we would like to connect the preservation of stability type under nonlinear perturbation to analyzing the stability type of fixed points of nonlinear dynamical systems based on their linearization.

Theorem 3.14 (Hartman-Grobman)**.** *If the fixed point x_0 of the nonlinear system (3.6) is hyperbolic, then the linearization (3.7) is topologically equivalent to the nonlinear system in a neighborhood of x_0.*

Consequence: *For hyperbolic fixed points, linearization predicts the correct stability type and orbit geometry near x_0.*

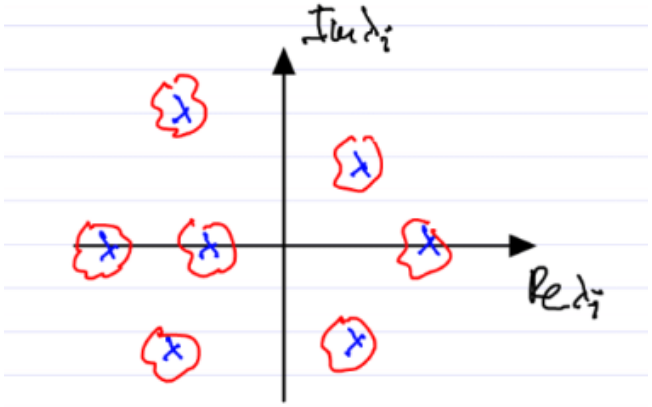


Figure 3.14: The eigenvalue configuration the $\mathcal{O}(\varepsilon)$ neighborhood (red) drawn around each eigenvalue (blue).

Now we would like to apply this to the pendulum to systematically derive the stability type of its fixed points.

Example 3.8 (Stability analysis of the pendulum via Hartman-Grobman). Recall the transformed ODE for the pendulum

$$\dot{x} = f(x) = \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\sin(x_1) \end{cases}; \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We have two fixed points $p = (\pi, 0)$ and $q = (0, 0)$. First we analyze the stability of the fixed point p. The differential of f at a point a is

$$Df(a) = \begin{pmatrix} 0 & 1 \\ -\cos(a_1) & 0 \end{pmatrix}.$$

Start by linearizing at p

$$\dot{y} = Ay; \quad A = Df(p) = \begin{pmatrix} 0 & 1 \\ -\cos(x_1) & 0 \end{pmatrix}_{x=p} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now we have to check if p is hyperbolic

$$\det(A - \lambda I) = \lambda^2 - 1 = 0 \implies \lambda_{1,2} = \pm 1; \; s_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; s_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Neither of the eigenvalues lie on the imaginary axis, so p is hyperbolic. This allows us to move between the nonlinear and linearized system for the stability analysis without compromising

our results. We find the linearized dynamics to be

$$y(t) = C_1 e^t s_1 + C_2 e^{-t} s_2.$$

Using the initial conditions y_0 the trajectory can be expressed as

$$y(t) = F^t y_0,$$

where F^t is the normalized fundamental matrix solution. We can now fully describe the phase portrait of the linearization. The *stable subspace* E^S is $\text{span}\{s_2\} = \{y_0 : \ F^t y_0 \xrightarrow{t\to\infty} 0\}$. The unstable subspace E^U is $\text{span}\{s_1\} = \{y_0 : \ F^t y_0 \xrightarrow{t\to-\infty} 0\}$. The phase portrait near the fixed point p is illustrated in Fig. 3.15

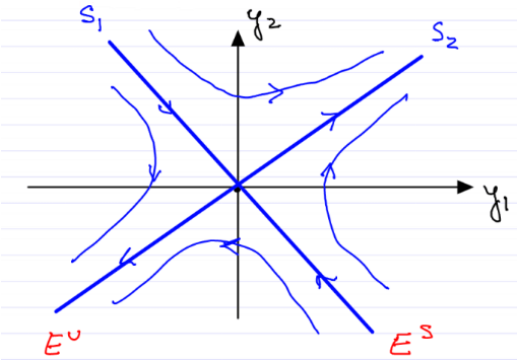


Figure 3.15: The phase portrait of the linearized pendulum in a neighborhood around p.

The nonlinear phase portrait is topologically equivalent to the linear one. Further we can define the stable and unstable manifolds of p for the nonlinear system. We designate $F^t(\cdot)$ to be the flow map for the nonlinear system after this point. The *stable manifold* of p is

$$\boxed{W^S = \{x_0 : \ F^t(x_0) \xrightarrow{t\to\infty} p\}.}$$

and the *unstable manifold* of p

$$\boxed{W^U = \{x_0 : \ F^t(x_0) \xrightarrow{t\to-\infty} p\}.}$$

Both of these are $\mathcal{C}^0$ curves through p and their existence follows from the Hartman-Grobman theorem. These manifolds are shown in the nonlinear phase portrait around p in Fig. 3.16.

Next we analyze the stability of the fixed point q. Once again, our first step is to linearize

$$\dot{y} = Ay, \quad A = Df(q) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad \det(A - \lambda I) = 0.$$

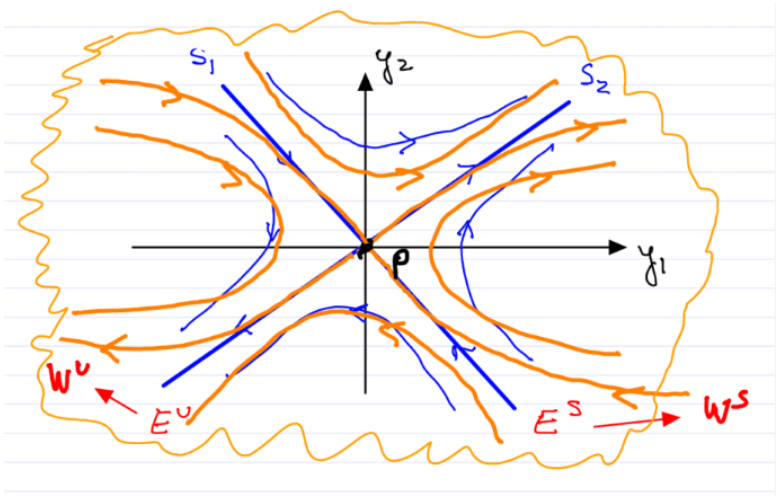


Figure 3.16: The phase portrait of the pendulum on a neighborhood around p with the stable and unstable manifolds of the nonlinear system as well as the stable and unstable spaces of the linearization.

From here, we see that the determinant is equal $\lambda^2 + 1 = 0$, yielding the roots $\lambda_{1,2} = \pm i$, i.e. the fixed point is not hyperbolic. Thus the linearized dynamics is inconclusive for the nonlinear system. In this particular case, q turns out to be stable by the definition of Lyapunov stability. Later we will use another approach to show this directly.

In the last example, the importance of hyperbolicity was not accentuated, as the latter fixed point had the same stability type in the linearized system as in the full system. This leads us to question if there are cases where the stability type between the linear and nonlinear systems is not preserved.

Example 3.9 (Criticality of hyperbolicity in Hartman-Grobman). Let the dynamical system be

$$\dot{x} = ax^3, \quad x \in \mathbb{R}, \quad a \neq 0.$$

This system has a fixed point at $x = 0$, linearizing here gives

$$A = 3ax^2\big|_{x=0} = 0 \implies \dot{y} = 0y = 0.$$

We have a single root $\lambda_1 = 0$, hence $x = 0$ is a nonhyperbolic fixed point. Disregarding this fact, we may be inclined to conclude that $x = 0$ is a stable fixed point, since $y = 0$ is trivially a fixed point of the linearization $\dot{y} = 0$. This is not the case, as we can see by analyzing the full nonlinear dynamics for $a > 0$ as in Fig. 3.17, where we observe that $x = 0$ is an unstable fixed point.

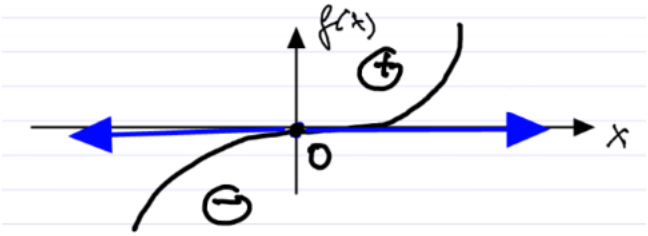


Figure 3.17: Nonlinear dynamics for the dynamical system $\dot{x} = ax^3$ with $a > 0$.

Now that we have understood how to use the Hartman-Grobman theorem, we would like to to be able to definitely conclude the stability type of a fixed point, once we have the linearization. To achieve this, we require a sufficient and necessary criterion for all of the eigenvalues of the linearized system to be left of the imaginary axis, i.e. $\mathrm{Re}\lambda_i < 0$ for all i.

Theorem 3.15 (Routh-Hurwitz)**.** *Consider the polynomial*

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0.$$

Without loss of generality assume $a_0 > 0$, if $a_0 < 0$ then multiply by -1 and if $a_0 = 0$ then $\lambda = 0$ is a root and therefore we cannot have asymptotic stability. Next, define the following series of subdeterminants

$$D_0 = a_0,\ D_1 = a_1,\ D_2 = \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix},\ D_3 = \begin{vmatrix} a_1 & a_0 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \ldots,\ D_n = \begin{vmatrix} a_1 & a_0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots & a_n & a_{n-1} & a_{n-2} \\ 0 & \ldots & & 0 & a_n \end{vmatrix}$$

Then we have that if and only if for all i $D_i > 0$ then $Re\lambda_i < 0$ for all i.

A weaker neccesary condition is that if for all i $Re\lambda_i < 0$ then $a_i > 0$ for all i. Therefore if there exists an $a_k < 0$, we know immediately that the fixed point cannot be asymptotically stable as not all of the $Re\lambda_i$ can be strictly negative.

Remark 3.16. For a given i we construct the matrix used for calculating D_i as follows: write the elements $a_1, \ldots, a_i$ along the diagonal, then in each row k write the a_j in descending index order such that a_k aligns with the placement inherited from us writing along the diagonal. The leftover spaces are filled with zeros.

Remark 3.17. Adolf Hurwitz discovered this criterion independently of Edward Routh in 1895 while holding a chair at the ETH.

Example 3.10 (Applying the Routh-Hurwitz criterion). Given the polynomial

$$a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0.$$

The Routh-Hurwitz criterion is

$$D_0 = a_0 > 0;\ D_1 = a_1 > 0;\ D_2 = a_1a_2 - a_0a_3 > 0;\ D_3 = a_3D_2 > 0.$$

Therefore

$$\boxed{a_0 > 0,\ a_1 > 0,\ a_1a_2 - a_0a_3 > 0,\ a_3 > 0}$$

forms a sufficient and necessary condition for asymptotic stability ($a_i > 0$ follows from here) for $n = 3$.

A more refined linearization result comes from Sternberg's Theorem [Chicone, 1999].

Theorem 3.18 (Sternberg)**.** *Assume the following*

(i) The fixed point is asymptotically stable, i.e. $Re\lambda_j < 0$ *for all* $j = 1, \dots, n$ *(this also implies hyperbolicity);*

(ii) $f \in \mathcal{C}^r$ *for* $r \in \mathbb{N} \cup \{\infty\}$ *and* $r > \frac{\max_j |Re\lambda_j|}{\min_j |Im\lambda_j|}$*;*

(iii) The system is nonresonant, i.e.

$$\sum_{j=1}^{n} m_j\lambda_j \neq \lambda_k,\ \forall k$$

for any sequence $m_j \in \mathbb{N}$ *with* $\sum_{j=1}^{n} m_j \geq 2$.

Then the following holds: $\dot{x} = f(x)$ *with* $f(0) = 0$ *is locally* $\mathcal{C}^r$ *conjugate to its linear part* $\dot{y} = Ay$ *for* $A = Df(0)$.

Example 3.11 (Watt's centrifugal governor for steam engines). Now we put together everything built until now in the example of Watt's centrifugal governor for steam engines. Originally this system was used in mills in the 1700's, then it was adapted by Watt to the steam engine in 1788. This adaptation has been credited as a major factor in the industrial revolution, and is a first example of feedback control. The system is outlined in Fig. 3.18. The two masses (of mass $\frac{m}{2}$) rotate counter clockwise, and their position in radians is given by θ, with their rotational velocity $\dot{\theta}$. The masses are attached by a rod of length L and their deflection from the vertical position is measured by φ. Smaller $\dot{\theta}$ allowed for an increase in steam supply.

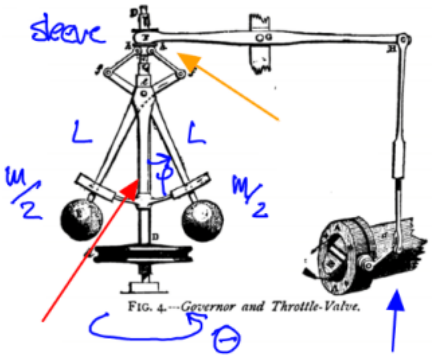


Figure 3.18: Schematic for Watt's centrifugal governor. The yellow arrow points towards a damper on the rotation about the spindle (red arrow). On the right the blue arrow designates a steam engine cylinder.

Following changes in the design, the systems suddenly became unstable. To address this Vishnegradky studied the root cause in 1877. We first derive the equation of motion. For the governor we use the equation of motion for a rotating hoop (with viscous damping coefficient b).

$$mL^2\ddot{\varphi} + bL^2\left(\frac{g}{L} - \dot{\theta}^2\cos(\varphi)\right)\sin(\varphi) = 0; \quad b > 0.$$

Next, let ω denote the angular velocity of the steam engine, i.e. with the gear ratio n we have $\dot{\theta} = n\omega$. Then we find

$$m\ddot{\varphi} = -b\dot{\varphi} - m\left(\frac{g}{L} - n^2w^2\cos(\varphi)\right)\sin(\varphi).$$

Now we derive the equation of motion for the steam engine. Denote the moment of inertia for the engine by J, the driving torque from the steam as P_1 and the constant load P, we obtain

$$J\dot{\omega} = P_1 - P.$$

In this case we have $P_1 = P^* + k\left(\cos(\varphi) - \cos(\varphi^*)\right)$ for the desired operation angle φ^*, the gain k, and $P*$ the value of P at φ^*. Putting this together yields

$$J\dot{\omega} = k\cos(\varphi) - P_0; \quad P_0 = P - P^* + k\cos(\varphi^*).$$

Let $\dot{\varphi} = \Psi$ to transform into a three-dimensional set of equations (ODE)

$$\begin{cases} \dot{\varphi} = \Psi \\ \dot{\Psi} = -\dfrac{b}{m}\Psi - \left(\dfrac{g}{L} - n^2\omega^2\cos(\varphi)\right)\sin(\varphi)\,; \\ \dot{\omega} = \dfrac{k}{J}\cos(\varphi) - \dfrac{P_0}{J}. \end{cases} \quad x = \begin{pmatrix} \varphi \\ \Psi \\ \omega \end{pmatrix}.$$

Then our operation point is the fixed point x_0 of this system

$$f(x_0) = 0 \implies \Psi_0 = 0;\ \omega_0^2 = \frac{g}{Ln^2\cos(\varphi_0)};\ \cos(\varphi_0) = \frac{P_0}{k}.$$

If $\sin(\varphi_0) = 0$, we have an unphysical state and ignore this case. For simplification set $L = 1$, this could be formally achieved by nondimensionalizing the length L. Now we linearize at the fixed point x_0

$$\dot{y} = Ay; \quad A = Df(x_0) = \begin{pmatrix} 0 & 1 & 0 \\ n^2\omega^2\cos(2\varphi_0) - g\cos(\varphi_0) & -\frac{b}{m} & n^2\omega_0\sin(2\varphi_0) \\ -\frac{k}{J}\sin(\varphi_0) & 0 & 0 \end{pmatrix}.$$

With this we obtain the characteristic equation $\det(A - \lambda I) = 0$

$$\underbrace{1}_{a_3}\lambda^3 + \underbrace{\frac{b}{m}}_{a_2}\lambda^2 + \underbrace{g\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)}}_{a_1}\lambda + \underbrace{g\frac{2k\sin^2(\varphi_0)}{J\omega_0}}_{a_0} = 0.$$

Now check the Routh-Hurwitz criterion for asymptotic stability:

(i) The necessary condition for $\mathrm{Re}\lambda_k < 0$ for all k: $a_j > 0$ for all j is fulfilled.

(ii) Next we check the subdeterminants

$$\begin{aligned} D_0 &= a_0 = g\frac{2k\sin^2(\varphi_0)}{J\omega_0} > 0; \\ D_1 &= a_1 = g\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)} > 0; \\ D_2 &= \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix} = a_1a_2 - a_0a_3 = g\frac{b}{m}\frac{\sin^2(\varphi_0)}{\cos(\varphi_0)} - g\frac{2k\sin^2(\varphi_0)}{J\omega_0} > 0; \\ D_3 &> 0 \iff D_2 > 0 \text{ and } a_3 = 1 > 0. \end{aligned}$$

The only actual condition for x_0 to be asymptotically stable is $D_2 > 0$

$$\frac{bJ}{m} > \frac{2P_0}{\omega_0}.$$

From the equation for the fixed points we know

$$P_0\omega_0^2 = \frac{gk}{n^2} = \text{const.} \tag{3.9}$$
$$\omega_0^2 + 2P_0\omega_0\frac{d\omega_0}{dP_0} = 0.$$

From the first equation, we realize that we must write $\omega_0 = \omega_0(P_0)$. To obtain the second equation, we derive the first equation (3.9) with respect to P_0. Therefore we find

$$\frac{d\omega_0}{dP_0} = -\frac{\omega_0}{2P_0}.$$

Next, define the *non-uniformity of performance*

$$\nu = \left|\frac{d\omega_0}{dP_0}\right| = \frac{\omega_0}{2P_0}.$$

Then our criterion for asymptotic stability becomes

$$\frac{bJ}{m}\nu > 1.$$

To conclude, the following have harmful effects of stability

- Bigger engines which increase m
- Better machining of surfaces decreasing b
- Increased operating speed decreasing J
- Versatility in operation decreasing ν.

These have the effect of pushing the eigenvalues to the right in the complex plane (towards the imaginary axis) as is illustrated in Fig 3.19.

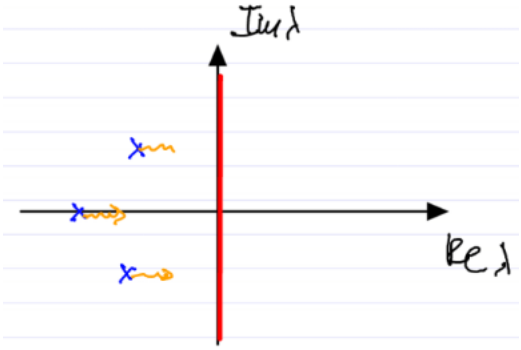


Figure 3.19: Change in the eigenvalue configuration as small modification are made to the Watt engine governor.

3.6 Data-driven linear modeling of dynamical systems

Given are $m+1$ vectors of n-dimensional observables $x_0, \ldots, x_m \in \mathbb{R}^n$ observed at Δt intervals. An example of such data would be many different trajectories or snapshots of a single trajectory. Typically $m \gg 1$, i.e. there are a high number of observations.

Our objective is to find the best fitting linear dynamical system

$$\dot{x} = Ax; \quad x \in \mathbb{R}^n; \quad A \in \mathbb{R}^{n \times n}.$$

Since the data is given as discrete observations, this is equivalent to seeking the best $B = e^{A\Delta t}$ such that $x_{k+1} = Bx_k$ for $k = 1, \ldots, m-1$. To solve this we synthesize the data into matrices

$$X = \begin{bmatrix} x_0 \ldots x_{m-1} \end{bmatrix} \in \mathbb{R}^{n \times m}; \quad Y = \begin{bmatrix} x_1 \ldots x_m \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Notice that the matrix Y is shifted forward compared to the matrix X by Δt. Next, we solve for a matrix B that minimizes the distance between Y and BX. That is, we find

$$B_* = \underset{B \in \mathbb{R}^{n \times n}}{\operatorname{argmin}}\, g(B); \quad g(B) = \|Y - BX\|^2,$$

where the matrix norm is the Frobenius norm. We denote the components of B by b_{ij}, the components of Y by y_{ij}, and the components of X by x_{ij}. Note that

$$\begin{aligned} \frac{\partial g}{\partial b_{ij}} &= \frac{\partial}{\partial b_{ij}} \sum_{l,m,p} \left(y_{lm} - b_{lp} x_{pm}\right)^2 = -2 \sum_{l,m,p} \left(y_{lm} - b_{lp} x_{pm}\right) \delta_{il} \delta_{jp} x_{pm} \\ &= -2 \sum_{m,p} \left(y_{im} - b_{ip} x_{pm}\right) x_{jm} = 2 \left(YX^T - BXX^T\right). \end{aligned}$$

In the third equality we have used the properties of the Kronecker-delta δ_{ij}. Therefore we must have that $D_B g = -2\left(YX^T - BXX^T\right) = 0$, so we must solve the linear system of equations

$$BXX^T = YX^T \in \mathbb{R}^{n\times n}.$$

In the case $n > m$, the matrix X^T has more columns than rows, thus X^T is singular. Hence XX^T has a nonempty kernel and is therefore not invertible. For $n < m$ this is not necessarily the case. For a solution, we can use the pseudo inverse $(XX^T)^\dagger$, the superscript here is called *dagger*.

Recall that if $A \in \mathbb{R}^{n\times n}$ is invertible then $A^{-1} = A^\dagger$. If A is not invertible, then calculate the singular value decomposition of $A = U\Sigma V^T$ for $U, V \in \mathrm{SO}(n)$ and Σ a diagonal matrix. The entries of Σ are the strictly positive singular values. The columns of U and V are the left and right singular vectors. We then define the pseudo inverse as

$$A^\dagger = V\Sigma^\dagger U^T; \quad \Sigma^\dagger_{ij} = \begin{cases} \frac{1}{\Sigma_{ij}} & \Sigma_{ij} \neq 0 \\ 0 & \Sigma_{ij} = 0. \end{cases}$$

This means that for $Ax = b$, the best solution in a least squares sense is given by $x = A^\dagger b$. In our case, the solution to 3.10 is

$$B = \left(YX^T\right)\left(XX^T\right)^\dagger.$$

This procedure is called *dynamic mode decomposition* (DMD) [Schmid, 2010, Kutz et al., 2015]. The eigenvectors of B are called *DMD modes* and the normalized exponents of those corresponding eigenvalues are called *DMD eigenvalues*.

The procedure has been extended in *extended DMD*, where DMD was performed on functions of the observables (typically polynomials) [Williams et al., 2015]. This effectively tries to construct the linearization from data. Note here that none of these procedures can describe intrinsically nonlinear phenomena (i.e. those with coexisting isolated compact invariant sets).

Remark 3.19. The numerical cost of (pseudo)inverting the $n \times n$ matrix, XX^T, can be reduced when $n \gg m$.

(i) We can calculate the singular value decomposition of the data matrix X instead.

(ii) By keeping only the leading r singular values in this decomposition, we obtain a reduced (rank-r) approximation of the matrix $B \in \mathbb{R}^{n\times n}$ as follows. Denote the singular value decomposition by $X = U\Sigma V^T$. This also means that $(XX^T) = U\Sigma^2 U^T$, i. e. we get the same U as before, but from an SVD of an $n \times m$ matrix, which is much less demanding numerically.

With the leading r singular vectors and singular values σ, X can be written as

$$X \approx U_r \Sigma_r V_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T, \qquad U = [u_1, ..., u_n] \quad V = [v_1, ..., v_m].$$

(iii) The matrix B can be approximately written as

$$\begin{aligned} B &= Y X^T (X X^T)^\dagger \\ &\approx Y V_r \Sigma_r U_r^T U_r (\Sigma_r^2)^\dagger U_r^T \\ &= Y V_r \Sigma_r^\dagger U_r^T = Y (X_r)^\dagger. \end{aligned}$$

We denote the truncated approximation of B as $B_r = Y(X_r)^\dagger$.

3.7 Lyapunov's direct (second) method for stability

Stability analysis via linearization is not perfect. For nonhyperbolic fixed points, the results obtained are inconclusive, and this case turns out to be somewhat common for conservative systems. Furthermore, linearization does not give any insight into the size of the domain of stability. Hence, a method for determining stability type without relying on linearization would be desirable. This is given to us by Lyapunov's direct method.

Theorem 3.20. *Consider*

$$\dot{x} = f(x), \quad f \in \mathcal{C}^1, \quad x \in \mathbb{R}^n; \quad f(x_0) = 0.$$

Assume that there exists a function $V : U \to \mathbb{R}$ with $V \in \mathcal{C}^1(U)$ for an open set $U \subset \mathbb{R}^n$ and $x_0 \in U$ which fulfills the following

(i) V is positive definite in a neighborhood of x_0, i.e.

$$V(x_0) = 0; \quad V(x) > 0, \ x \in U - \{x_0\}.$$

(ii) $\dot{V}$ is negative semidefinite in the same neighborhood, i.e.

$$\dot{V} = \frac{d}{dt} V(x(t)) = \langle DV(x(t)), \dot{x}(t) \rangle = \langle DV(x(t)), f(x(t)) \rangle \leq 0, \ x \in U.$$

Then $x = x_0$ is Lyapunov stable (cf. Def. 3.1). V is called a Lyapunov function. The hypotheses are illustrated in Fig. 3.20.

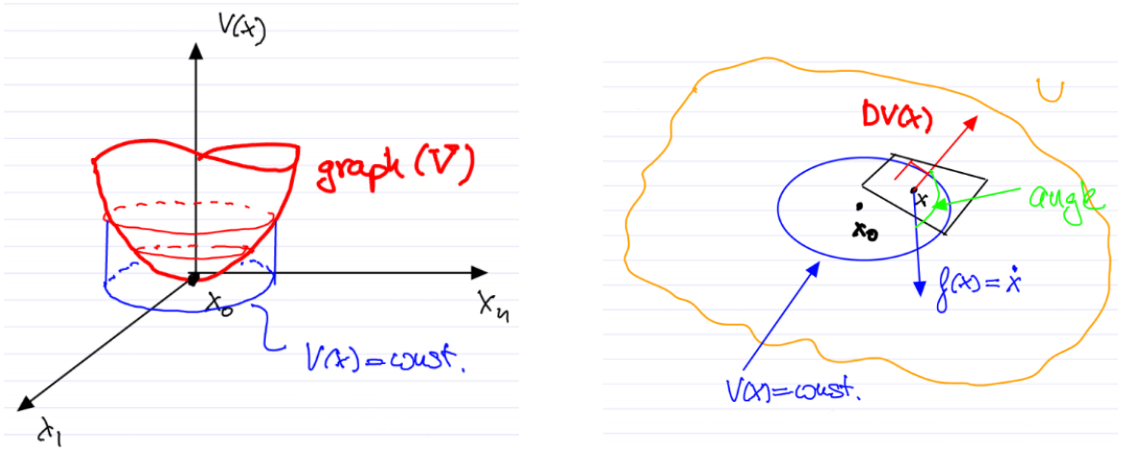


Figure 3.20: Geometric interpretation of hypotheses of Lyapunov's direct method. Left depicts hypothesis (i) and right hypothesis (ii). The angle between $DV(x)$ and $f(x)$ is at least $\frac{\pi}{2}$. In each image the blue $V(x)$ =const. is a level surface diffeomorphic to S^{n-1}.

Remark 3.21. To denote the boundary of a set A we write ∂A.

Proof. First choose $\varepsilon > 0$ and define $\alpha(\varepsilon) = \min_{x \in \partial B_\varepsilon(x_0)} V(x) > 0$. Note that $\alpha(\varepsilon)$ is well defined as $V \in \mathcal{C}^0(U)$ and $\partial B_\varepsilon(x)$ is compact and spherical for small enough ε. There exists an $x^* \in \partial B_\varepsilon(x)$ with $V(x^*) = \alpha(\varepsilon) \leq V(x)$ for all $x \in \partial B_\varepsilon(x_0)$. Next, define $U_\varepsilon = \{x \in B_\varepsilon(x_0) : V(x) < \alpha(\varepsilon)$. Notice that $x_0 \in U_\varepsilon$ because $V(x_0) = 0$ and $V(x) \geq 0$ on U. Further U_ε is open due to the continuity of V. We have that $U_\varepsilon \cap \partial B_\varepsilon(x_0)$ is empty by definintion, noting that for all $x \in \partial B_\varepsilon(x_0)$ we have $V(x) \geq \alpha(\varepsilon)$. Therefore there exists a ball $B_{\delta(\varepsilon)} \subset U_\varepsilon$ which contains x_0. This can be seen in Fig. 3.21.

Now observe that for every $\tilde{x}_0 \in B_{\delta(\varepsilon)}(x_0)$ we have that along trajectories $V(x(t; \tilde{x}_0)) \leq V(\tilde{x}_0) < \alpha(\varepsilon)$. The first inequality comes from hypothesis (ii) and the second inequality from the definition of U_ε. This implies that for $x(t; \tilde{x}_0) \in U_\varepsilon$ we have that $x(t; \tilde{x}_0)$ is not in $\partial B_\varepsilon(x_0)$ for any $t > 0$. The trajectory $x(t; \tilde{x}_0)$ is continuous, in order for it to leave the ball $B_\varepsilon(x_0)$ it must intersect the boundary $\partial B_\varepsilon(x_0)$. At this point $V(x)$ will atain a value of at least $\alpha(\varepsilon)$ by definition, however this is in contradiction to the fact that $V(x(t; \tilde{x}_0))$ is strictly smaller that $\alpha(\varepsilon)$. Therefore must stay in the ball $B_\varepsilon(x_0)$ for all times. □

Now we present some extensions to this theorem for various different stability types. These are necessary, as the previous theorem only provides a sufficient condition for stability (it is not "if and only if").

Theorem 3.22 (Theorem 2)**.** *Consider the same dynamical system. Assume*

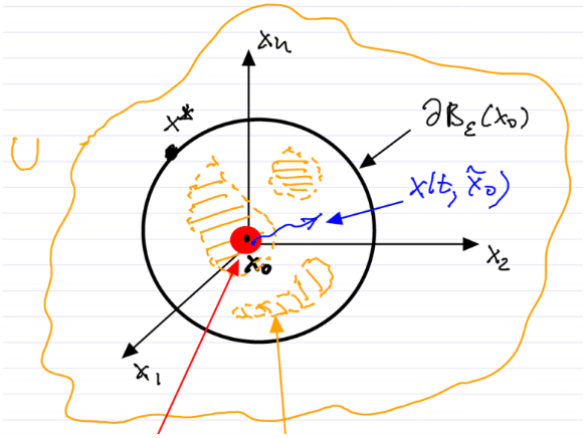


Figure 3.21: The constellation of U, U_ε, $\partial B_\varepsilon(x_0)$, and $B_{\delta(\varepsilon)}(x_0)$ from the proof of Lyapunov's direct method. The red arrow points at $B_{\delta(\varepsilon)}(x_0)$ and the yellow arrow at a connected component of U_ε.

(i) $V(x)$ *is positive definite,*

(ii) $\dot{V}(x)$ *is negative definite, i.e.*

$$\dot{V}(x) < 0, \ x \in U - \{x_0\}.$$

Then $x = x_0$ *is asymptotically stable. These hypotheses are illustrated in Fig. 3.22.*

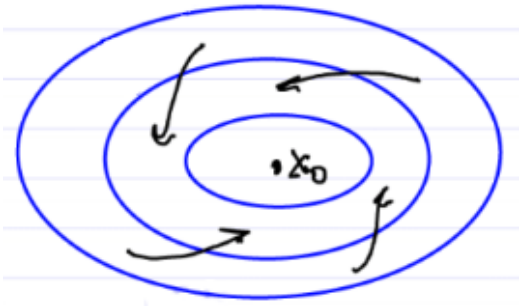


Figure 3.22: The hypotheses of Theorem 2 illustrated, the key difference being that the arrows (denoting the flow of the dynamical system) cross the level surfaces of V (blue rings) toward x_0.

Theorem 3.23 (Theorem 3)**.** *Consider the same dynamical system. Assume*

(i) $V(x)$ *is positive definite,*

(ii) $\dot{V}(x)$ *is positive definite, i.e.*

$$\dot{V}(x) > 0, \ x \in U - \{x_0\}.$$

Then $x = x_0$ *is unstable. The hypotheses are illustrated in Fig. 3.23.*

Figure 3.23: The hypotheses of Theorem 3 illustrated, the key difference being that the arrows (denoting the flow of the dynamical system) cross the level surfaces of V (blue rings) away from x_0.

Theorem 3.24 (Theorem 4)**.** *Consider the same dynamical system. Assume*

(i) $V(x)$ *is indefinite, i.e. arbitrarily close to* x_0 *there exists* $a, b \in U$ *such that* $V(x_1) \cdot V(x_2) < 0$ *(they have opposite signs and are not equal to 0) and* $V(x_0) = 0$*.*

(ii) $\dot{V}(x)$ *is definite near* x_0 *(either positive or negative).*

Then $x = x_0$ *is unstable. The geometry of the hypotheses are illustrated in Fig. 3.24.*

Remark 3.25. In each of these theorems, the definiteness of $\dot{V}$ can be replaced by semidefiniteness, if we add that the set $\{x \in U : \ \dot{V}(x) = 0\}$ does not contain full trajectories of the system. This is called Krasovsky's condition.

Now we would like to put these theorems into practice with a few examples.

Example 3.12 (Stability analysis of the pendulum with Lyapunov's direct method). Recall that using linearization, we were only able to conclude the stability type of one of the fixed points for the dynamical system of the pendulum

$$ml^2\ddot{\varphi} + mgl\sin(\varphi) = 0.$$

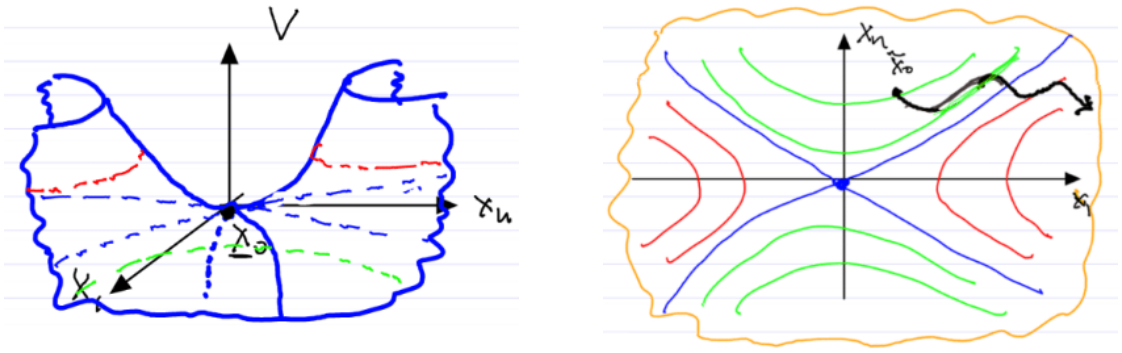


Figure 3.24: The geometry of hypotheses (i) (left) and (ii) (right) for $\dot{V}$ positive definite of Theorem 4 illustrated. On the right level surfaces are designated by lines, blue corresponds to $V = 0$, red $V > 0$, and green $V < 0$, which can be seen as the dotted lines of the same colors on the left.

Now we will use the energy as a Lyapunov function, which is often very useful. The energy is given by

$$E(x) = E(\varphi, \dot{\varphi}) = \frac{1}{2}ml^2\dot{\varphi}^2 + mgl(1 - \cos(\varphi)) = \frac{1}{2}ml^2\dot{x_2}^2 + mgl(1 - \cos(x_1)).$$

Transforming the dynamical system to be an system of first order ODEs we obtain

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varphi \\ \dot{\varphi} \end{pmatrix}; \quad \dot{x} = \begin{pmatrix} \dot{x_1} \\ \dot{x_2} \end{pmatrix} = \begin{pmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) \end{pmatrix} = f(x).$$

At the fixed point $x = (0, 0)$ we have

$$E(0,0) = 0; \quad DE(0,0) = 0 \in \mathbb{R}^{2\times 2}; \quad D^2E(0,0) = \begin{pmatrix} mgl & 0 \\ 0 & ml^2 \end{pmatrix}.$$

We have that the Hessian of E is positive definite. Therefore E is positive definite at $(0, 0)$. Further we have

$$\dot{E}(x) = \langle DE(x), f(x)\rangle = \begin{pmatrix} mgl\sin(x_1) & ml^2x_2 \end{pmatrix} \begin{pmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) \end{pmatrix} = 0.$$

Thus $\dot{E}$ is negative semidefinite. Now the hypotheses of Theorem 3.20 are fulfilled, hence $x = (0, 0)$ is Lyapunov stable. Importantly, this is a nonlinear result and we did not need to refer to the linearization of the system!

At the fixed point $x = (\pi, 0)$. We check again using the energy. First we realize that $E(\pi, 0) = 2mgl$, so we subtract the constant and redefine to obtain $\tilde{E} = E - 2mgl$. Now

$\tilde{E}(\pi, 0) = 0$, although this is not essential. Next we calculate

$$DE(\pi, 0) = 0 \in \mathbb{R}^{n \times n}; \quad D^2E(\pi, 0) = \begin{pmatrix} -mgl & 0 \\ 0 & ml^2 \end{pmatrix}.$$

Hence, E is indefinite at $(\pi, 0)$. From the previous calcuation we alread know that $\dot{E}(\pi, 0) = 0$, i.e. $\dot{E}$ is semidefinite. We cannot apply Theorem 4, but linearization already concluded that $(\pi, 0)$ was unstable.

Example 3.13 (Stability analysis of the friction pendulum). We have a shaft which is constantly rotating with angular speed Ω, around this shaft is a sleeve which rubs against the shaft creating friction. To this sleeve is a mass m attached at distance l, the deflection of this mass from its standard position (directly below the shaft) is measured by φ. The gravity constant is given by g. This setup is depicted in Fig. 3.25. The torque driving the pendulum is given by

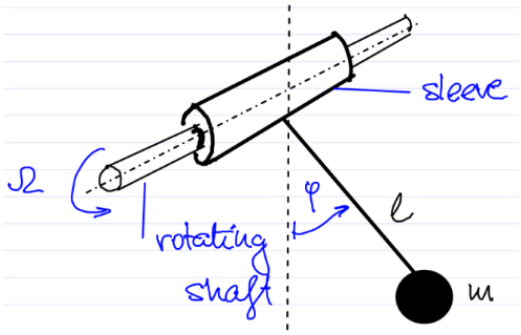


Figure 3.25: The setup of the friction pendulum.

$$T(\dot{\varphi}) = T_0 \text{sign}(\Omega - \dot{\varphi}).$$

From this we get the equation of motion

$$ml^2\ddot{\varphi} + mgl \sin(\varphi) = T(\dot{\varphi}) = T_0 \text{sign}(\Omega - \dot{\varphi}).$$

By assuming $\Omega \gg 1$, i.e. very fast rotation of the shaft, we find that $T(\dot{\varphi}) = T_0$. Next we transform the coordinates in order to get an ODE

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varphi \\ \dot{\varphi} \end{pmatrix}; \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) + \frac{T_0}{ml^2} \end{pmatrix}.$$

Thus the fixed point x_0 is at $(\overline{x}_1, \overline{x}_2)$ given by

$$\sin(\overline{x}_1) = \frac{T_0}{mgl}; \quad \overline{x}_2 = 0.$$

Since the system is forced, energy is not conserved and we cannot use it as a Lyapunov function. However, there may still exist a conserved quantity. Consider using

$V(x) = [\text{total energy at time } t] - [\text{work put in between time } t \text{ and } t_0] = [\text{initial energy}] = \text{const.}$

This is a conserved quantity in all mechanics problems, but generally cannot be calculated without already knowing the trajectories. Here we can, as

$$V(x(t)) = \frac{1}{2}mlx_0^2 + mgl(1-\cos(x_1)) - T_0(\underbrace{x_1}_{\varphi(t)} - \underbrace{x_1(0)}_{\varphi(0)}).$$

We can drop the constant term $T_0 x_1(0)$. Now verify that V is indeed constant

$$\dot{V}(x) = \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2 = (mgl\sin(x_1) - T_0)\,x_2 + ml^2 x_2\left(-\frac{g}{l}\sin(x_1) + \frac{T_0}{ml^2}\right) = 0.$$

Hence $\dot{V}(x)$ is semidefinite, and may be used to conclude stability or instability. We must check the Lyapunov conditions for $V(x)$

$$DV(x_0) = \begin{pmatrix} mgl\sin(\overline{x}_1) - T_0 & ml^2\overline{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$
$$D^2V(x_0) = \begin{pmatrix} mgl\cos(\overline{x}_1) & 0 \\ 0 & ml^2 \end{pmatrix}.$$

The Hessian is positive definite as long as $\overline{x}_1 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, and all fixed points in this region are stable by Theorem 3.20. The Hessian is indefinite if $\overline{x}_1 \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, but in this case Theorem 4 is not applicable as $\dot{V}$ is not definite.

Chapter 4

Bifurcations of fixed points

4.1 Local nonlinear dynamics near fixed points

We are interested in the local nonlinear dynamics around fixed points. Consider

$$\dot{x} = f(x); \quad f \in \mathcal{C}^r, r \geq 1; \quad f(p) = 0, \tag{4.1}$$

i.e. p is a fixed point of the dynamical system. The linearized system at p is

$$\dot{y} = Df(p)y, \ y \in \mathbb{R}^n, \ Df(p) \in \mathbb{R}^{n \times n}. \tag{4.2}$$

The linearization has the eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ with multiplicities counted. Corresponding to these eigenvalues are the eigenvectors $e_1, \ldots, e_n \in \mathbb{C}^n$, including generalized eigenvectors for when the algebraic multiplicity is greater than the geometric multiplicity. The eigenvector e_j is real when $\lambda_j \in \mathbb{R}$.

Definition 4.1. The following subspaces are invariant for the linearized dynamical system:

(i) The *stable subspace*

$$\boxed{E^S = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(e_j) < 0 \right\},}$$

(ii) The *unstable subspace*

$$\boxed{E^U = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(e_j) > 0 \right\},}$$

(iii) The *center subspace*

$$\boxed{E^C = \text{span}_j \left\{ \text{Re}(e_j), \text{Im}(e_j) : \ \text{Re}(e_j) = 0 \right\}.}$$

Remark 4.1. Note here that the following facts hold

(i) $E^C = \emptyset$ if and only if p is hyperbolic,

(ii) $E^{U,S}$ and E^C are invariant subspaces of (4.2) by construction,

(iii) Solutions of (4.2) in E^S (resp. E^U) decay to $y = 0$ as $t \to \infty$ (resp. $t \to -\infty$).

We now discuss what happens to these subspaces in the nonlinear system (4.1).

Theorem 4.2 (Center Manifold Theorem)**.** *The following hold:*

(i) There exists a unique stable manifold $W^S(p)$ *for* (4.1)*, such that*

- $W^S(p)$ *is a* $\mathcal{C}^r$ *manifold (surface), tangent to* E^S *at* p *with dim* $W^S(p) =$ *dim* E^S*,*
- $W^S(p)$ *is invariant for* (4.1) *and for* $x \in W^S(p)$ *we have*

$$\|F^t(x) - p\| \leq K_S \exp\left[t\left(\max_{Re(\lambda_j)<0}(Re(\lambda_j)) + \varepsilon_S\right)\right]$$

for $t \geq 0$*,* $0 < \varepsilon_S \ll 1$*, and* $\|x - p\|$ *small enough.*

(ii) There exists a unique unstable manifold $W^U(p)$ *for* (4.1)*, such that*

- $W^U(p)$ *is a* $\mathcal{C}^r$ *manifold (surface), tangent to* E^U *at* p *with dim* $W^U(p) =$ *dim* E^U*,*
- $W^U(p)$ *is invariant for* (4.1) *and for* $x \in W^U(p)$ *we have*

$$\|F^t(x) - p\| \leq K_U \exp\left[t\left(\max_{Re(\lambda_j)>0}(Re(\lambda_j)) + \varepsilon_U\right)\right]$$

for $t \leq 0$*,* $0 < \varepsilon_U \ll 1$*, and* $\|x - p\|$ *small enough.*

(iii) There exists a (not necessarily unique) center manifold $W^C(p)$ *for* (4.1)*, such that*

- $W^C(p)$ *is a* $\mathcal{C}^{r-1}$ *manifold (surface), tangent to* E^C *at* p *with dim* $W^C(p) =$ *dim* E^C*,*

The geometry of these manifolds and their corresponding subspaces are sketched in Fig. 4.1.

Notice that Theorem 4.2 does not state anything about the asymptotic dynamics on the center manifold, in contrast to that of the unstable and stable manifolds. This means that the overall dynamics depends crucially on the center manifold. In the case when $W^u = \emptyset$, even the stability type of the fixed point is determined by $W^C(p)$. This is why it will be the subject of further investigations.

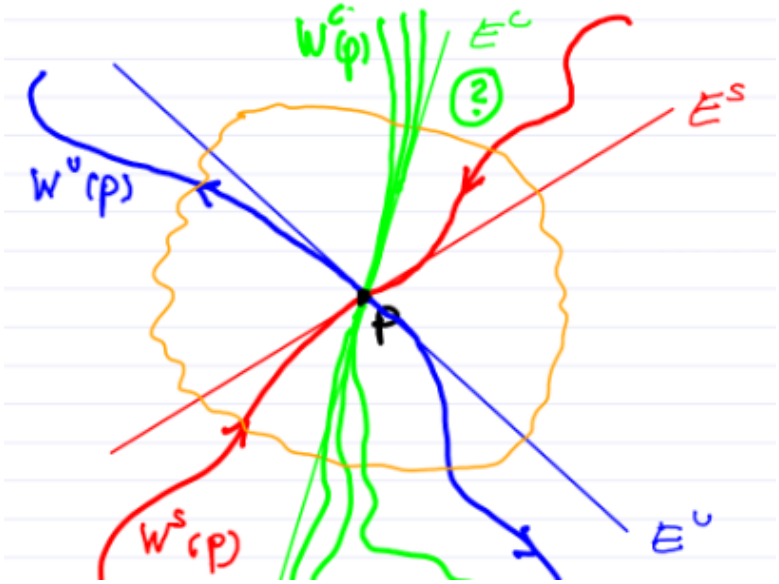


Figure 4.1: A sketch of the stable (red), unstable (blue), and center manifolds (green), along with their respective linear subspaces, that are invariant under the linearized dynamics. Note the existence of multiple center manifolds and the unique unstable/stable manifolds.

4.2 The center manifold

Example 4.1 (Uniqueness of the center manifold). Since Theorem 4.2 gave no guarantees on the uniqueness of the center manifold, now we explore this non-uniqueness in an example.

$$\begin{cases} \dot{x} = x^2 \\ \dot{y} = -y. \end{cases}$$

First, linearize at the origin to find the linearized dynamics

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

The linearized dynamics is illustrated in Fig. 4.2. We find the invariant subspaces

$$E^C = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}; \quad E^S = \text{span}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}; \quad E^U = \emptyset.$$

The nonlinear manifolds are illustrated with the invariant subspaces from the linearization in Fig. 4.2. Observe that there exist infinitely many center manifolds which are all invariant and all tangent to E^C at the origin. We also see that although the fixed point at the origin is stable in the linearized system, it is unstable in the nonlinear system.

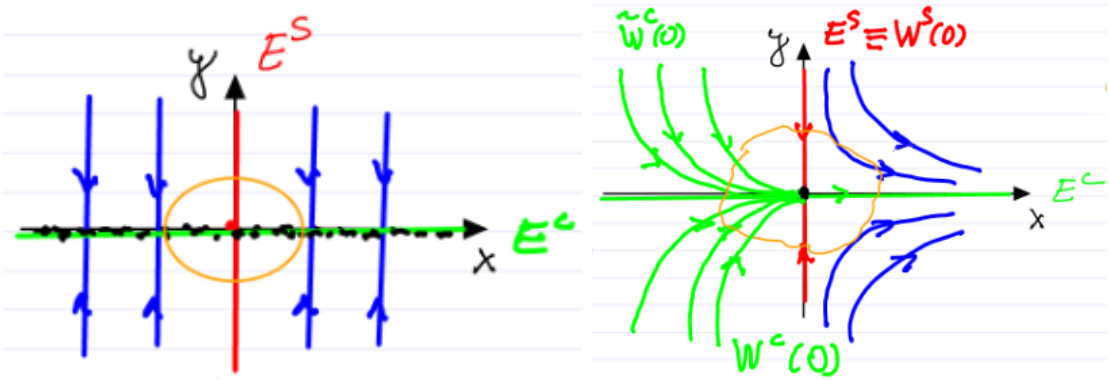


Figure 4.2: Left: The linearized dynamics around the origin. Right: The nonlinear phase portrait.

We now present a method to calculate $W^C(p)$ in general cases.

(i) Consider

$$\dot{z} = F(z); \quad F(0) = 0; \quad z \in \mathbb{R}^{c+d}; \quad F \in \mathcal{C}^r,$$

where c represents the number of center directions at the origin (dim E^C) and d denotes the remaining directions (dim E^U + dim E^S).

(ii) Now block-diagonalize the linearization. This consists of four steps

(a) First linearize the dynamics to find $\dot{z} = Mz + \mathcal{O}(\|z\|^2)$ with $M = DF(0) \in \mathbb{R}^{(c+d)\times(c+d)}$.

(b) Define the transformation matrix

$$T = \begin{bmatrix} a_1 & \dots & a_c & b_1 & \dots & b_d \end{bmatrix} = \begin{bmatrix} \text{basis in } E^C & \text{basis in } E^U \oplus E^S \end{bmatrix}.$$

(c) Pass to the basis from the transformation matrix by introducing a new set of coordinates as $z = T\xi$

$$\dot{\xi} = T^{-1}\dot{z} = T^{-1}MT\xi + T^{-1}\mathcal{O}(\|T\xi\|^2) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \xi + \mathcal{O}(\|\xi\|^2).$$

The matrices A and B are elements of $\mathbb{R}^{c\times c}$ and $\mathbb{R}^{d\times d}$ respectively.

(d) Let $\xi = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^c \times \mathbb{R}^d$, the x-coordinates are aligned with E^C and the y-coordinates are perpendicular to them. We then find

$$\dot{x} = Ax + f(x, y); \quad \dot{y} = By + g(x, y), \tag{4.3}$$

for $f, g \in \mathcal{C}^r$ and $f, g = \mathcal{O}(\|x\|^2, \|y\|^2, \|x\|\|y\|)$. The geometry in these coordinates is depicted in Fig. 4.3. The center manifold is given by

$$W^C(0) = \{(x, y) \in U : \ y = h(x)\}$$

for $h : \mathbb{R}^c \to \mathbb{R}^d$ and $h \in \mathcal{C}^{r-1}$ as in the theorem.

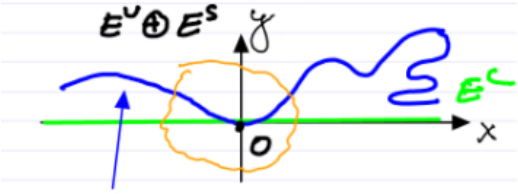


Figure 4.3: The geometry of the nonlinear system in the transformed coordinates aligned with the invariant subspaces of the linearization. The blue arrow points to the center manifold $W^C(0)$.

(iii) Now we use the invariance of $W^C(0)$, i.e. for all t, it holds that $y(t) = h(x(t))$. With this we find

$$\dot{y} = Dh(x(t))\dot{x}(t).$$

We can now substitute $\dot{x}$ and $\dot{y}$ from (4.3) to get the following nonlinear partial differential equation (PDE) for $h(x)$

$$\boxed{Bh(x) + g(x, h(x)) = Dh(x)\left[Ax + f(x, h(x))\right].} \tag{4.4}$$

In addition to (4.4) being nonlinear, the boundary conditions are also unknown. Therefore there is little hope for solving it analytically.

(iv) Instead take the Taylor expansion of (4.4) to approximate the solution

$$h(x) = \underbrace{h(0)}_{=0} + \underbrace{Dh(0)}_{=0} x + \frac{1}{2}\underbrace{D^2h(0)}_{3-\text{tensor}} \otimes x \otimes x + \mathcal{O}(\|x\|^3),$$

where the first two terms are 0 due to the tangency to E^C at 0. We therefore have that $h = O(x^2)$ and we are justified in seeking $W^C(0)$ in this form. We can then restrict the system to the center manifold by substituting $y = h(x)$ to get the reduced dynamics as

$$\boxed{\dot{x} = Ax + f(x, h(x)).}$$

Example 4.2 (Finding the center manifold). Consider the dynamical system

$$\begin{cases} \dot{x} = xy \\ \dot{y} = -y + \alpha x. \end{cases}$$

First we linearize at $(0,0)$ to get

$$M = \begin{pmatrix} [0] & [0] \\ [0] & [-1] \end{pmatrix}$$

which is already in block-matrix form. The dimensions of the stable, unstable, and center subspaces of the linearization are 1, 0, and 1 respectively. Hence the stability type depends on the dynamics on the center manifold $W^C(0)$. We now look for a parameterization of $W^C(0)$ in the form

$$h(x) = ax^2 + bx^3 + cx^4 + \mathcal{O}(x^5).$$

Since the right-hand side of the dynamical system was $\mathcal{C}^\infty$, this expansion could formally be continued to any order. However, the expansion will not converge as that would imply the uniqueness of the center manifold. In this example, we work with a finite, order 4 truncation. Now use the invariance (the PDE (4.4) we derived above) to find

$$\dot{y} = h'(x)\dot{x} = \left[2ax + 3bx^2 + 4cx^3 + \mathcal{O}(x^4)\right] x \left[ax^2 + bx^3 + cx^4 + \mathcal{O}(x^5)\right]. \tag{4.5}$$

On the other hand, from the dynamical system we know

$$\dot{y} = -h(x) + \alpha x^2 = (\alpha - a)x^2 - bx^3 - cx^4 + \mathcal{O}(x^5). \tag{4.6}$$

Equations (4.5) and (4.6) must hold for all x small enough, therefore the coefficients of equal powers of x in these two equations must match.

$$\begin{aligned} \mathcal{O}(x^2): &\ \alpha = a \\ \mathcal{O}(x^3): &\ b = 0 \\ \mathcal{O}(x^4): &\ 2a^2 = -c. \end{aligned}$$

Therefore we find

$$\boxed{h(x) = \alpha x^2 - 2\alpha^2 x^4 + \mathcal{O}(x^5).}$$

Then the dynamics on $W^C(0)$ becomes

$$\boxed{\dot{x} = xh(x) = \alpha x^3(1 - 2\alpha x^2) + \mathcal{O}(x^6).}$$

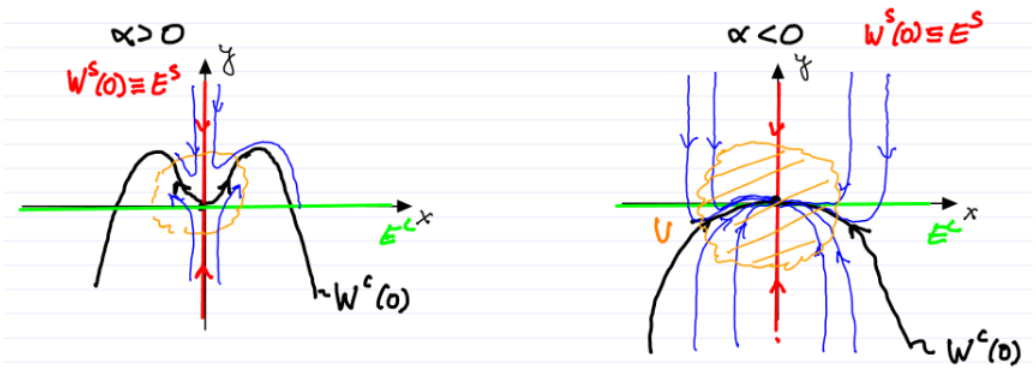


Figure 4.4: Left: The nonlinear dynamics on the center manifold for $\alpha > 0$. Right: The nonlinear dynamics on the center manifold for $\alpha < 0$.

The dynamics is depicted in Fig. 4.4. For $\alpha > 0$ the origin is unstable, meanwhile for $\alpha < 0$ the origin is asymptotically stable.

The full local stable manifold for $\alpha < 0$ is $\overline{W}^S(0) = U$ and it is of dimension 2. The difference between $\overline{W}^S(0)$ and $W^S(0)$ is that in general the decay rate in $\overline{W}^S(0) - W^S(0)$ is generally weaker than the rate guaranteed in the Center Manifold Theorem.

Remark 4.3. The $\mathcal{O}(x^5)$ truncation has two hyperbolic fixed points at $x = \pm\frac{1}{\sqrt{2\alpha}}$, however the full system has no such fixed points. The reason for this is that away from the origin, the $\mathcal{O}(x^6)$ terms are no longer guaranteed to be small relative to the $\mathcal{O}(x^5)$ terms, and the truncation this far away from 0 is not justified.

After this example we would like to explore if the concept of the center manifold is robust, as the existence of eigenvalues with $\mathrm{Re}(\lambda_i) = 0$ is not. We will explore this in an example.

Example 4.3 (Perturbing the previous example). Consider the following perturbed dynamical system

$$\begin{cases} \dot{x} = xy + \varepsilon x \\ \dot{y} = -y + \alpha x^2 \end{cases}; \quad |\varepsilon| \ll 1.$$

The linearization of this system yields

$$\begin{cases} \dot{x} = \varepsilon x \\ \dot{y} = -y. \end{cases}$$

Now the center manifold disappears as the center subspace E^C disappears for $\varepsilon > 0$!

4.3 Center manifolds depending on parameters

We begin with the setup

$$\begin{cases} \dot{x} = Ax + f(x, y, \varepsilon) \\ \dot{y} = By + g(x, y, \varepsilon) \end{cases}; \quad x \in \mathbb{R}^c,\ y \in \mathbb{R}^d,\ 0 \leq \varepsilon \ll 1;$$
$$f, g \in \mathcal{C}^r,\ f, g = \mathcal{O}(\|x\|^2, \|y\|^2, \|x\|\|x\|, \varepsilon\|x\|, \varepsilon\|y\|).$$

The order $\varepsilon\|x\|$ and $\varepsilon\|y\|$ terms are due to the perturbation of the linear part. Now assume $\mathrm{Re}(\lambda_j(A)) =$ for $j = 1, \ldots, c$ (the center directions) and $\mathrm{Re}(\lambda_j(B)) \neq$ for $j = 1, \ldots, d$ (the hyperbolic directions). Next, rewrite $\tilde{x} = \begin{pmatrix} x \\ \varepsilon \end{pmatrix}$ and $\tilde{y} = y$ to obtain the system

$$\begin{cases} \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{f}(\tilde{x}, \tilde{y}) \\ \dot{\tilde{y}} = \tilde{B}\tilde{y} + \tilde{g}(\tilde{x}, \tilde{y}) \end{cases}; \quad \tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{(c+1)\times(c+1)}; \quad \tilde{f} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{4.8}$$

Here, $\tilde{g} = g$ and $\tilde{B} = B$. Further, note that span $\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \right\}$ is an invariant subspace for $\tilde{A}$. The eigenvalues of $\tilde{A}$ are the same as those of A and an additional 0, thus there are $c + 1$ center directions and d hyperbolic directions. Applying the center manifold theorem to the fixed point $0 \in \mathbb{R}^{c+1+d}$ of (4.8) we obtain that there exists a $\tilde{W}^C(0)$ $\mathcal{C}^{r-1}$ manifold tangent to E^C at $\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ which is invariant and is of dimension $c+1$. The geometry of this manifold is illustrated in Fig. 4.5. Note in the figure that there is no dynamics in the ε direction, as $\dot{\varepsilon} = 0$ and that $(x, y) = (0, 0) \in \mathbb{R}^{c+d}$ remains a fixed point for $\varepsilon \neq 0$.

Computing $\tilde{W}^C(0)$ is done in a similar fashion as before. We use the center manifold theorem to get

$$\tilde{y} = y = \tilde{h}(\tilde{x}) = \tilde{h}(x, \varepsilon) = \mathcal{O}(\|x\|^2, \varepsilon\|x\|, \varepsilon^2) = \mathcal{O}(\|x\|^2, \varepsilon\|x\|).$$

The order ε^2 term was dropped as $x = 0$ must remain a fixed point. The function h describes the graph of $W^C_\varepsilon(0)$. Then the reduced dynamics on $W^C_\varepsilon(0)$ is

$$\boxed{\dot{x} = Ax + f(x, \tilde{h}(x, \varepsilon), \varepsilon).}$$

This can then be applied to the perturbed example from above.

Example 4.4 (Revisting the perturbation). Recall the dynamical system

$$\begin{cases} \dot{x} = xy + \varepsilon x \\ \dot{y} = -y + \alpha x^2. \end{cases}$$

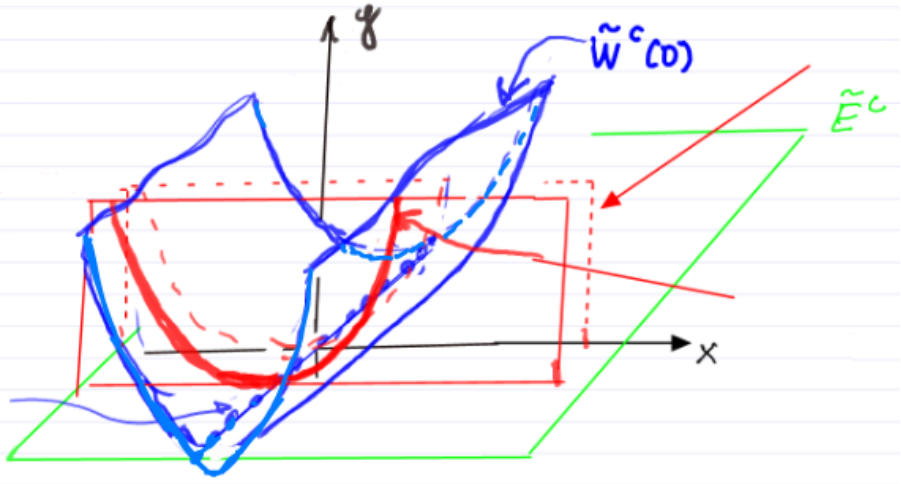


Figure 4.5: Geometry of the center manifold with the perturbation. The straight red arrow designates the cut at $\varepsilon = 0$ which is equal to $W^C(0)$, the squiggly red arrow points at the continuation of the center manifold from $\varepsilon = 0$ to $\varepsilon \neq 0$.

We have the persisting fixed point at $(x, y) = (0, 0) \in \mathbb{R}^{c+d}$ and the system is already in standard form with

$$A = 0; \quad B = -1; \quad f(x, y, \varepsilon) = xy + \varepsilon x; \quad g(x, y, \varepsilon) = -\alpha x^2.$$

We apply the center manifold theorem and get the existence of $W^C_\varepsilon(0)$ for $|\varepsilon| \ll 1$. This manifold satisfies

$$y = \tilde{h}(x, \varepsilon) = ax^2 + bx\varepsilon + c\varepsilon^2 + dx^3 + ex^2\varepsilon + jx\varepsilon^2 + k\varepsilon^3 + lx^4 \ldots. \tag{4.9}$$

The term $c\varepsilon^2$ must be equal to 0 for all ε such that the fixed point persists, therefore $c = 0$. Next the invariance $y(t) = \tilde{h}(x(t)), \varepsilon)$ is used, taking the time derivative on both sides yields

$$\dot{y} = [2ax + b\varepsilon + \mathcal{O}(2)] \underbrace{[\mathcal{O}(3) + \varepsilon x]}_{=\dot{x} \text{ from ODE and (4.9)}} = 2a\varepsilon x^2 + b\varepsilon^2 x + \mathcal{O}(4).$$

The $\mathcal{O}(n)$ designates terms of total degree n, for example x^n or $x^{n-k}\varepsilon^k$. From the ODE we find

$$\dot{y} = -y + \alpha x^2 = -ax^2 - bx\varepsilon - c\varepsilon^2 - dx^3 - ex^2\varepsilon - jx\varepsilon^2 - k\varepsilon^3 - \mathcal{O}(4) + \alpha x^2.$$

Comparing equal powers in these two equations we find

$$\begin{aligned}
&\mathcal{O}(x^2): \ 0 = \alpha - a; && \mathcal{O}(\varepsilon^2): \ 0 = -c; && \mathcal{O}(x\varepsilon): \ 0 = -b;\\
&\mathcal{O}(\varepsilon^3): \ 0 = -t; && \mathcal{O}(x^3): \ 0 = -d; && \mathcal{O}(x^2\varepsilon): \ 2a = -e;\\
&\mathcal{O}(x\varepsilon^2): \ b = -j.
\end{aligned}$$

Thus the shape of $W^C_\varepsilon(0)$ is given by

$$y = \tilde{h}(x, \varepsilon) = \alpha(1 - 2\varepsilon)x^2 + \mathcal{O}(4).$$

The dynamics on $W^C_\varepsilon(0)$ is

$$\dot{x} = \varepsilon x + \alpha(1 - 2\varepsilon)x^3 + \mathcal{O}(5).$$

We can see there is no substantial change in the shape of $W^C_\varepsilon(0)$ relative to the $\varepsilon = 0$ case. The stability type is determined by the sign of ε and a two time-scale dynamic persists.

From this example we may still wonder what effect the higher order terms have on the center manifold.

4.4 Normal forms

For a general treatment see [Guckenheimer and Holmes, 1990], here we will consider one example to illustrate the idea originally coming from Poincaré.

Example 4.5 (Reduced dynamics on 1-dimensional manifold). Consider the following 1-dimensional dynamical system

$$\dot{x} = x(\mu - x^2) + x^4; \quad 0 \leq |\mu| \ll 1.$$

The fixed points are at $x = 0$ and the roots of $g_\mu(x) = \mu - x^2 + x^3$. This function g_μ is illustrated in Fig. 4.6. By plotting x as a function of μ such that $g_\mu(x) = 0$ we get the *bifurcation diagram*

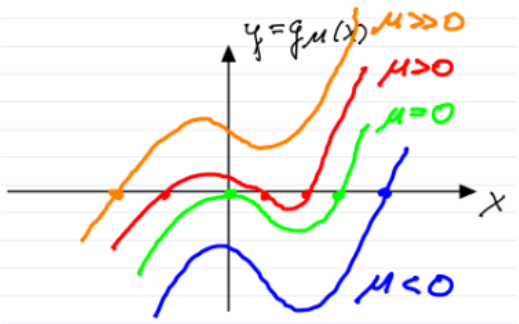


Figure 4.6: The functions g_μ for different values of μ.

as shown in Fig. 4.7.

The fold bifurcation (see caption of Fig. 4.7) is created by quartic (order 4) terms, which become more important away from the origin. The pitchfork bifurcation is already captured by

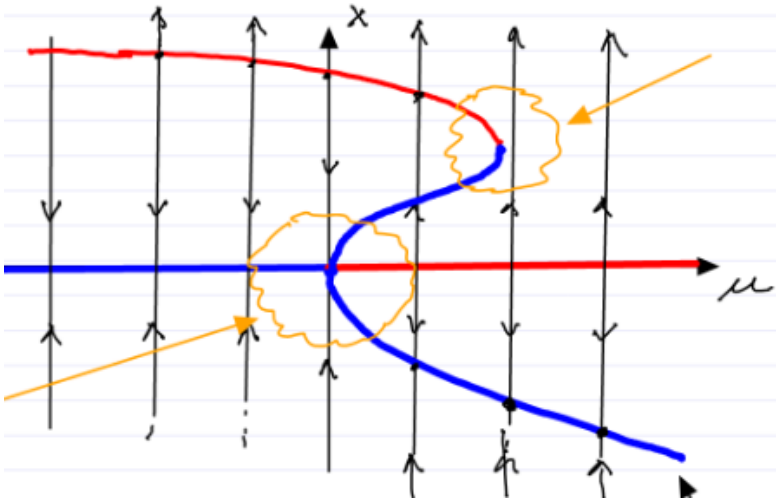


Figure 4.7: Bifurcation diagram for the 1-dimensional dynamical system. Red and blue demarcate if the fixed point at the given (x, μ) pair is stable (blue) or unstable (right). The arrow on the right points towards a *fold bifurcation* and the arrow on the left towards a *pitchfork bifurcation*. The curve is given by implicitly solving $g_\mu(x) = 0$.

the cubic truncation. We would like to know when the truncation captures the full dynamics correctly near the origin. Poincaré showed that, in fact, the truncated system is topologically equivalent to the full system near the origin by using a change of coordinates to remove $\mathcal{O}(4)$ terms.

(i) Let $x = y + h_4(y) = y + ay^4 + \mathcal{O}(y^5)$, which is near the identity near the origin, hence it is also invertible near the origin (by the Implicit Function Theorem). Further, this preserves the ODE up to the $\mathcal{O}(3)$ terms.

(ii) Plug in $x(t)$ and $y(t)$ and take the derivative with respect to time to get

$$\dot{x} = \dot{y}(1 + 4ay^3 + \mathcal{O}(y^4)).$$

(iii) Now use the ODE and find

$$\dot{x} = \mu x - x^3 + x^4 = \mu(y + ay^4) - (y + ay^4)^3 + (y - ay^4)^5 + \dots.$$

(iv) Combine the previous two steps and calculate

$$\dot{y} = \left[1 + 4ay^3 + \mathcal{O}(y^4)\right]^{-1} \left[\mu y + a\mu y^4 - y^3 + y^4 + \mathcal{O}(5)\right].$$

At this point recall the alternating series (it could be verified with Taylor expansion)

$$\frac{1}{1+z} = 1 - z + \mathcal{O}(z^2); \quad 0 \leq |z| \ll 1.$$

We also note that a generalized result holds for an operator $A \in R^{m\times m}$, which could be used in higher dimensional calculations. This is referred to as the Neumann series. Applying this to the left term in the formula for $\dot{y}$ yields

$$\left[1 + 4ay^3 + \mathcal{O}(y^4)\right]^{-1} = 1 - 4ay^3 + \mathcal{O}(y^4).$$

Therefore we find

$$\dot{y} = \mu y - y^3 + y^4 \underbrace{(-4a\mu + a\mu + 1)}_{\text{choose } a \text{ such that this } =0} + \mathcal{O}(y^5).$$

We may now choose the parameter a to make the coefficient of the quartic term 0. The a that fulfills this is $a = \frac{1}{3\mu}$, using this we find

$$\boxed{\dot{y} = \mu y - y^3 + \mathcal{O}(y^5).}$$

This transformation has removed the quartic terms from the equation.

(v) Now we remove the $\mathcal{O}(y^5)$ terms similarly. First set

$$y = \xi + h_5(\xi) = \xi + b\xi^5 + \mathcal{O}(\xi^6)$$

and then continue as before, but with y now playing the role of x and ξ playing the role of y.

(vi) The successive sequence of near identity coordinate changes turns out to converge usually. In general, it depends on the type of problem, sometimes resonant terms are not removeable and must stay as they are crucial to the dynamics. These resonant terms depend only on the linear part of the RHS, for more information see Chapter 3 of [Guckenheimer and Holmes, 1990].

Thus

$$\boxed{\dot{x} = \mu x - x^3}$$

is the *normal form* for the ODE for the study of bifurcations at the origin for $0 \leq \mu \ll 1$. It is topologically equivalent to the full system near $x = 0$ and captures the pitchfork bifurcation.

4.5 Bifurcations

A *bifurcation* is a qualitative change in the dynamical system

$$\boxed{\dot{x} = f(x, \mu); \quad x \in \mathbb{R}^n; \quad \mu \in \mathbb{R}^p.} \tag{4.10}$$

Linear stability analysis led to reducing to the center manifold (depending on parameters). From there we moved to normal forms which enable the analysis of qualitative dynamics under varying parameters.

Definition 4.2 (Local bifurcation)**.** A *local bifurcation* is a qualitative change in the behavior of a system near a non-hyperbolic equilibrium. More precisely, a *bifurcation* occurs in (4.10) at $\mu = \mu_0$ near the fixed point $x = 0$ if there exists no neighborhood of $x = 0$ in which $\dot{x} = f(x, \mu_0)$ is topologically equivalent to all systems $\dot{x} = f(x, \mu)$ for $\|\mu - \mu_0\|$ small enough.

This idea of a bifurcation can be illustrated by a bifurcation surface which separates the space of dynamical systems into components. Within each component the dynamical systems are topologically equivalent, however elements from separate components are not. This is sketched in Fig. 4.8.

Figure 4.8: Illustration of a bifurcation surface (blue). The near side of the surface is one component, and the far side the other. The red path is a family of dynamical systems $\{\dot{x} = f(x, \mu)\}_{\mu \in \mathbb{R}^p}$, going through the bifurcation point μ_0. Around this point, a neighborhood as outlined in the definition is sketched in orange.

We wish to understand what happens in the case that a given family of dynamical systems is nongeneric (atypical). For instance in the case that the family forms a tangency to the bifurcation surface, in which case a bifurcation does not take place, hence the family of dynamical systems is not general enough to capture all possible dynamics.

Example 4.6 (Nongeneric family of dynamical systems). In comparison to the family taken previously consider

$$\dot{x} = -a^2 x - x^3; \quad \mu = -a^2 \leq 0.$$

For every μ which we consider, there is only one fixed point $x = 0$ and it is stable for all values of μ. However, we have unwittingly missed the full picture, as for $\mu > 0$ there are three fixed points, one of which are unstable. This situation is shown in Fig. 4.9.

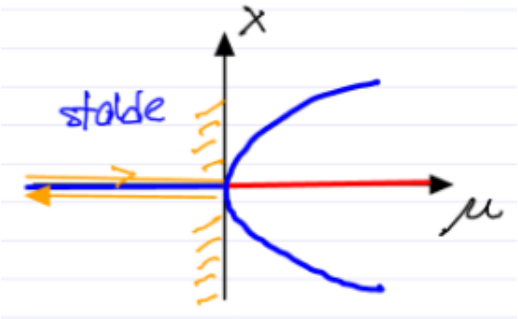


Figure 4.9: A nongeneric family of dynamical systems. We only see what happens to the left of the x-axis, and miss everything to the right, hence our family is tangent to the bifurcation surface.

This idea of nongeneric families motivates our next definition, which is also depicted in Fig. 4.10.

Definition 4.3 (Universal unfolding)**.** A paramaterized family of dynamical systems crossing all nearby topological equivalence classes as the parameters vary is called a *universal unfolding*.

Definition 4.4 (Codimension of a bifurcation)**.** The *codimension of a bifurcation* is the minimum number of parameters required for a universal unfolding. Thus a more degenerate bifurcation requires a larger codimension.

4.6 Codimension 1 bifurcations

We begin by considering different types of center manifolds. We start by classifying these by the number of zero eigenvalues which appear in the linearization.

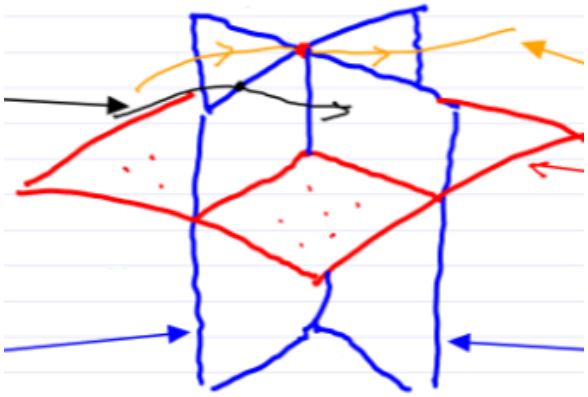


Figure 4.10: An example of universal unfolding (red) for the red bifurcation point which crosses the four topologically equivalent classes (components) created by the two bifurcation surfaces (blue). Furthermore, a nonuniversal unfolding is shown by the yellow 1-dimensional path at the top. Another universal unfolding in for the black bifurcation point, is shown by the 1-dimensional black family.

(i) **Single zero eigenvalue** Our system is as follows

$$\dot{y} = f(y,\mu),\ y \in \mathbb{R}^n,\ \mu \in \mathbb{R}^p; \quad f(0,0) = 0; \quad \dim E^C = 1 \implies \dim W^C_\mu(0) = 1.$$

Thus we have a single zero eigenvalue, an example eigenvalue constellation which fulfills this is given in Fig. 4.11. The universal unfolding (generally parameterized by the normal

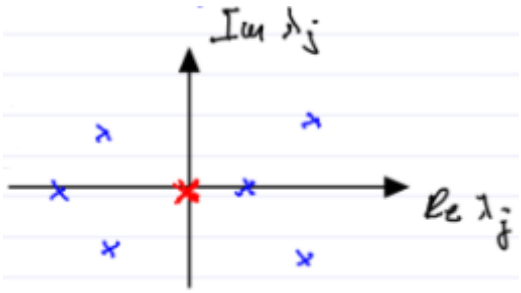


Figure 4.11: An example eigenvalue constellation with a single zero eigenvalue.

form on $W^C_\mu(0)$) is given by

$$\boxed{\dot{x} = \tilde{\mu} \pm x^2; \quad x \in \mathbb{R}; \quad \tilde{\mu} \in \mathbb{R}.}$$

This is a codimension 1 bifurcation, i.e. only 1 parameter is needed for the universal unfolding. Hence we find a bifurcation diagram as shown in Fig. 4.12. This is called a

fold or *saddle node* bifurcation, in this example it is *supercritical* as we show the $-x^2$ case. *Subcriticality* occurs when the diagram is mirrored with respect to the x-axis, i.e. when the fixed points disappear for $\mu > 0$. The sign which is actually used is problem dependent.

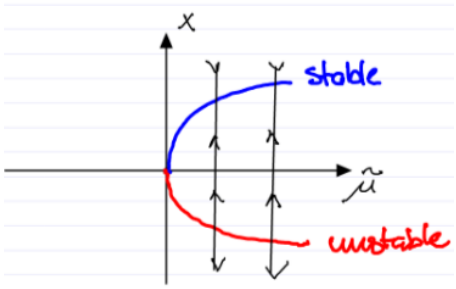


Figure 4.12: A codimension 1 fold bifurcation, which is supercritical.

Remark 4.4. *Hysteresis* is the dependence of the output on past input and is often related to interplay between two such fold bifurcations as shown in Fig. 4.13.

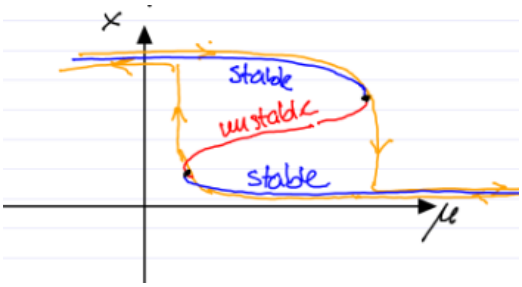


Figure 4.13: An example of hysteresis caused by two interacting fold bifurcations. In this case, the hysteresis is a result of purely deterministic dynamics, no additional memory (or delay) terms were needed.

(ii) **Single zero eigenvalue & origin remains a fixed point** Now we have more conditions and our setup is as follows

$$\dot{y} = f(y,\mu),\ y \in \mathbb{R}^n,\ \mu \in \mathbb{R}^p; \quad f(0,\mu) = 0,\ \forall \mu; \quad \dim E^C = 1 \implies \dim W^C_\mu(0) = 1.$$

Thus we have the additional degeneracy $D_\mu f(0,0) = 0$. We get the universal unfolding

$$\dot{x} = \tilde{\mu}x \pm x^2 = x(\tilde{\mu} \pm x); \quad x \in \mathbb{R}; \quad \tilde{\mu} \in \mathbb{R}.$$

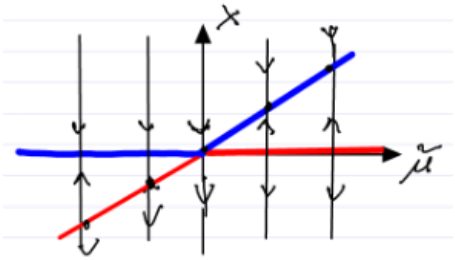


Figure 4.14: A transcritical bifurcation.

This so called *transcritical bifurcation* is depicted in Fig. 4.14. In such bifurcations the observed stable state does not vary until the bifurcation, after which it begins to vary linearly.

(iii) **Single zero eigenvalue, origin remains a fixed point, & RHS is an odd function** We have the same setup as previously, with the additional condition that

$$f(y,0) = -f(-y,0) \implies D_y^2 f(y,0) = -D_y^2 f(-y,0) \implies D_y^2 f(0,0) = 0.$$

This leads to the universal unfolding

$$\dot{x} = \tilde{\mu}x \pm x^3; \quad x \in \mathbb{R}; \quad \mu \in \mathbb{R}.$$

This type of bifurcation is called a *pitchfork bifurcation*. It is *supercritical* for the $+x^3$ case (cf. Fig. 4.15), and subcritical for the $-x^3$ case. The actual sign is problem dependent.

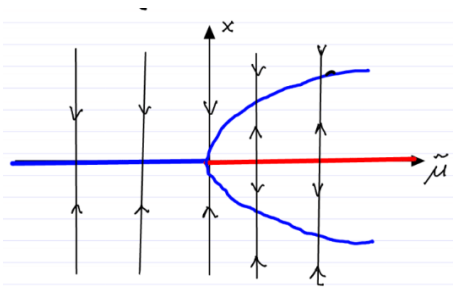


Figure 4.15: A supercritical (i.e with a + sign) pitchfork bifurcation.

(iv) **Purely imaginary pair of eigenvalues** Now we have the following setup, with the eigenvalue constellation shown in Fig. 4.16.

$$f(0,0) = 0; \quad \dim E^C = 2; \quad \lambda_n(0) = \overline{\lambda}_{n-1}(0) = \pm i\omega \neq 0.$$

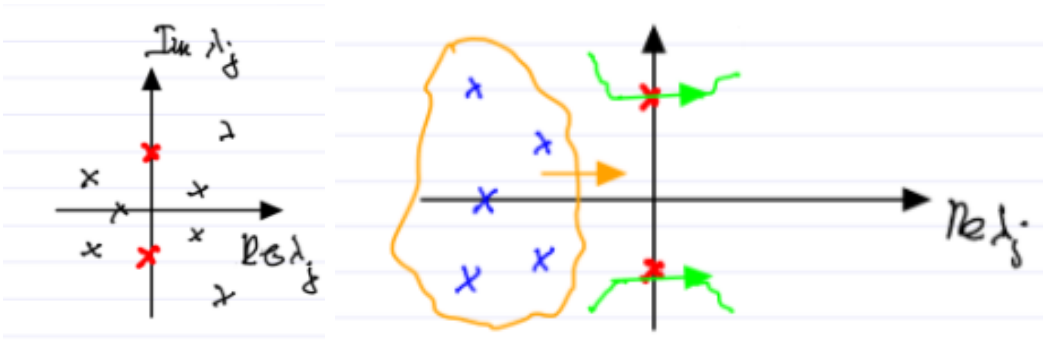


Figure 4.16: Left: Eigenvalue constellation for purely imaginary pair of eigenvalues. Right: Loss of stability in oscillating system as the pair of eigenvalues cross the imaginary axis.

Such constellations are common in oscillatory systems at the loss of stability. This loss of stability is sketched in the right pane of Fig. 4.16. The reduced system dynamics on the 2-dimensional $W^C(0)$ for $\mu = 0$ is

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \tilde{f}(u,v) \\ \tilde{g}(u,v) \end{pmatrix}. \tag{4.11}$$

The coordinates (u, v) correspond to the basis given by $[\mathrm{Re}(e_n), \mathrm{Im}(e_n)]$. The normal form on $W^C_\mu(0)$ is

$$\begin{aligned} \dot{r} &= r\left(d(\mu) + a(\mu)r^2\right) + \mathcal{O}(r^5); & d(\mu) &= \mathrm{Re}(\lambda_n(\mu)) \\ \dot{\theta} &= \omega(\mu) + e(\mu)r^2 + \mathcal{O}(r^4); & \omega(\mu) &= \mathrm{Im}(\lambda_n(\mu)). \end{aligned}$$

The functions a and e depend on the given $\tilde{f}$ and $\tilde{g}$ (thereby also on f). We then define

$$\begin{aligned} d_0 &= \frac{d}{d\mu}\mathrm{Re}\left[\lambda_n(\mu)\right]\big|_{\mu=0}; \quad \omega_0 = \mathrm{Im}(\lambda_n(0)) \\ a_0 = a(0) &= \frac{1}{16}\left[\tilde{f}_{uuu} + \tilde{f}_{uvv} + \tilde{g}_{uuv} + \tilde{g}_{vvv}\right]\Big|_{(u,v)=0} \\ &\quad + \frac{1}{16\omega}\left[\tilde{f}_{uv}(\tilde{f}_{uu} + \tilde{f}_{vv}) - \tilde{g}_{uv}(\tilde{g}_{uu} + \tilde{g}_{vv}) - \tilde{f}_{uu}\tilde{g}_{uu} + \tilde{f}_{vv}\tilde{g}_{vv}\right]\Big|_{(u,v)=0}. \end{aligned}$$

Note that we must start from the standard form (4.11).

Building on this last case, we explore the extended center manifold in the next theorem.

Theorem 4.5 (Hopf-Bogdanov)**.** *Assume the following*

(i) The eigenvalues λ_n and $\overline{\lambda}_n$ do cross the imaginary axis for $\mu \neq 0$, i.e. $d_0 \neq 0$.

(ii) The leading order nonlinearity in $\dot{r}$ is nonzero, i.e. $a_0 \neq 0$. Then there exists a unique extended center manifold $W_\mu^C(0)$ for $0 \leq \mu \ll 1$, on which the dynamical system is locally topologically equivalent to the universal unfolding

$$\boxed{\dot{r} = r\left(d_0\mu + a_0 r^2\right); \quad \dot{\theta} = \omega_0 + e_0 r^2 + b_0\mu,}$$

for $b_0, e_0 \in \mathbb{R}$.

From here we explore two cases for the signs of d_0 and a_0.

(i) First, when each are strictly positive $d_0, a_0 > 0$. This yields a subcritical pitchfork bifurcation in r and a continued rotation in θ. The bifurcation of r and the dynamics of u and v for $\mu < 0$ are illustrated in Fig. 4.17, for $|\mu|$ and $|r|$ small enough. There is

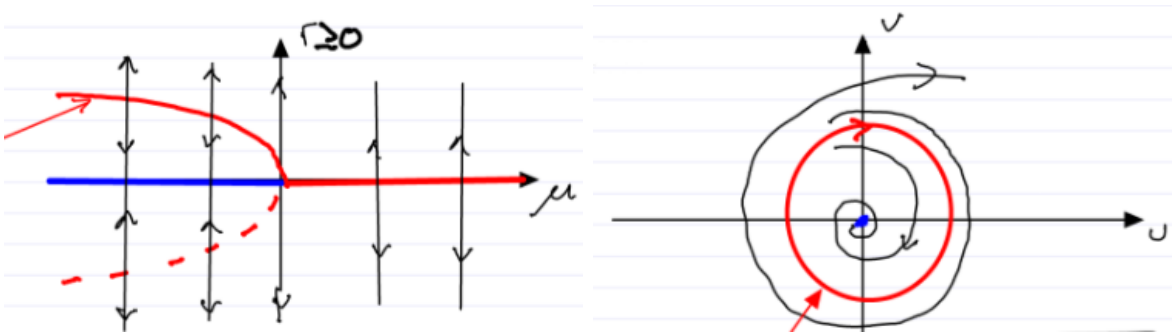


Figure 4.17: Left: The subcritical pitchfork bifurcation of r. The red arrow points towards the $r = \sqrt{\frac{-d_0\mu}{a_0}}$, the symmetric negative part of this curve is dashed, as negative values of r are not relevant for this problem (the radius cannot be negative). Right: The effect of $\mu < 0$ on the system dynamics leading to an unstable limit cycle (red) in the $u - v$ plane.

an increased sensitivity to perturbation as μ increases (the domain of attraction shrinks). This is shown in Fig. 4.18. Such behavior can be observed in the transition to turbulence in pipe flows. In that context, the equation is referred to as the Stuart-Landau equation.

(ii) Next, the two coefficients have opposite signs $d_0 > 0$ and $a_0 < 0$. This leads to the supercritical Hopf bifurcation and a stable limit cycle as shown in Fig. 4.19

There is a sudden development of stable periodic oscillation as μ is increased. Often the parameter μ is a steady speed in the system which is changed gradually. Examples of this are

- increasing wind speed leading to bridge oscillations and collapse,

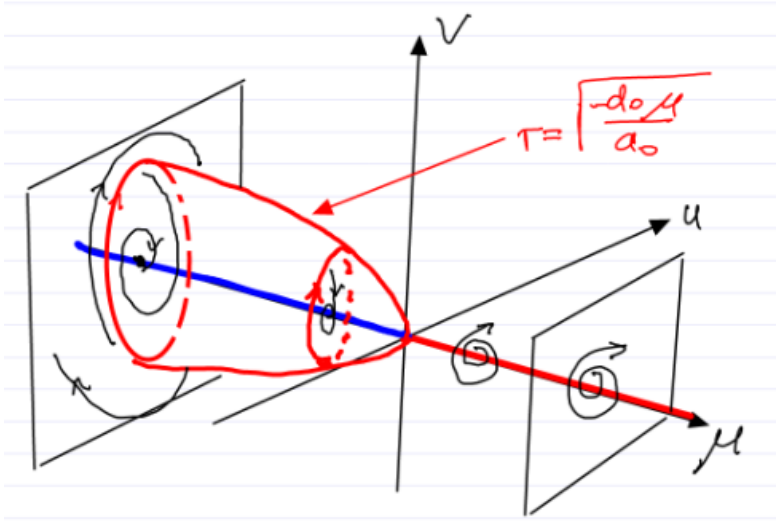


Figure 4.18: The full bifurcation diagram for the dynamical system in case (i).

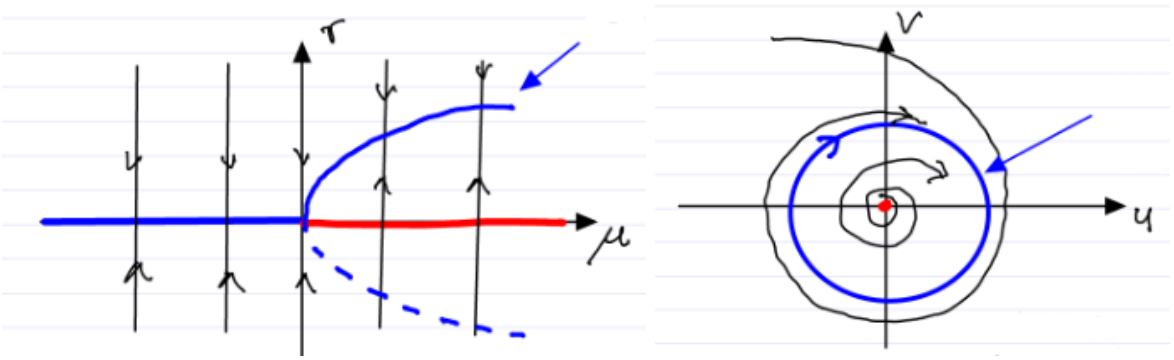


Figure 4.19: Left: The bifurcation in r, the blue arrow designates the curve given by $r = \sqrt{\frac{-d_0\mu}{a_0}}$, again the negative part is not relevant as the radius is always non-negative. Right: The stable limit cycle that arises for this system.

- increasing flight speed leading to wing flutter,
- increasing driving speed leading to death wobble on motorcycles.

Chapter 5

Nonlinear dynamical systems on the plane

For this chapter the general setup will be

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad f \in \mathcal{C}^1.$$

5.1 One degree of freedom conservative mechanical systems

Write $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, with x_1 denoting the position and $x_2 = \dot{x}_1$ denoting the speed. The total mechanical energy is given by $E(x) = \frac{1}{2}mx_2^2 + V(x_1)$, with the mass m and the potential function V. Newton's law then gives the equations of motion

$$m\ddot{x}_1 = -\frac{dV(x_1)}{dx_1} \implies \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\frac{1}{m}\frac{dV(x_1)}{dx_1}. \end{cases}$$

All of the forces derive from a time-independent potential. The energy is a first integral (conserved quantity of such systems), i.e.

$$\frac{d}{dt}E(x(t)) = \frac{\partial E}{\partial x_1}\dot{x}_1 + \frac{\partial E}{\partial x_2}\dot{x}_2 = 0.$$

Therefore on any given trajectory $x(t)$ the energy $E(x(t)) = E_0$ is constant. Hence we can derive an explicit equation for the velocity along a trajectory which suppresses the time dependence

$$x_2 = \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))}.$$

This leads to multiple consequences for these systems

(i) Trajectories form symmetric pairs (w.r.t. the x_1-axis) of the same energy.

(ii) There is a clockwise orientation for trajectories due to

$$\dot{x}_1 = x_2 \implies \begin{cases} x_2 > 0 \implies x_1 \text{ increases} \\ x_2 < 0 \implies x_1 \text{ decreases.} \end{cases}$$

This is shown graphically in Fig. 5.1.

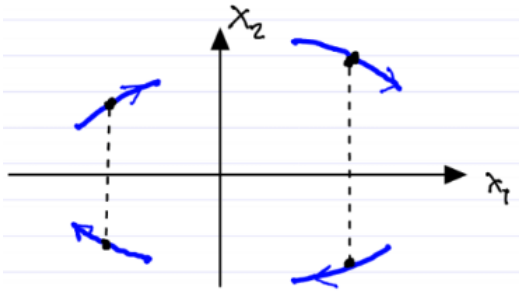


Figure 5.1: The clockwise orientation of trajectories in 1 degree of freedom conservative mechanical systems.

(iii) Local of the potential give rise to a center fixed point surrounded by closed orbits. The minimum x^* is a fixed point if $x_2^* = 0$ (which occurs if $V(x_1^*) = E_0$) as $\dot{x}_2 \propto V'(x_1^*) = 0$. The trajectories of the closed orbits can be found using $x_2 = \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))}$. Such a fixed point at a local minimum of the potential surrounded by closed orbits is shown in Fig. 5.2.

(iv) At local maxima of the potential saddle type fixed points arise. The maximum x^* is a fixed point if $x_2^* = 0$ (which occurs if $V(x_1^*) = E_0$) as $V'(x_2^*) = 0$. The velocity can be approximated as follows

$$\begin{aligned} x_2 &= \pm\sqrt{\frac{2}{m}(E_0 - V(x_1))} \\ &= \pm\sqrt{\frac{2}{m}}\sqrt{\underbrace{(E_0 - V(x_1^*))}_{=0} - \underbrace{V'(x_1^*)}_{=0}(x_1 - x_1^*) - \frac{1}{2}\underbrace{V''(x_1^*)}_{<0}(x_1 - x_1^*)^2 + \mathcal{O}((x_1 - x_1^*)^3)} \\ &= \pm\sqrt{\frac{2}{m}}\sqrt{-\frac{V''(x_1^*)}{2}}\,|x_1 - x_1^*| + \text{higher order terms.} \end{aligned}$$

The behavior is sketched in Fig. 5.2

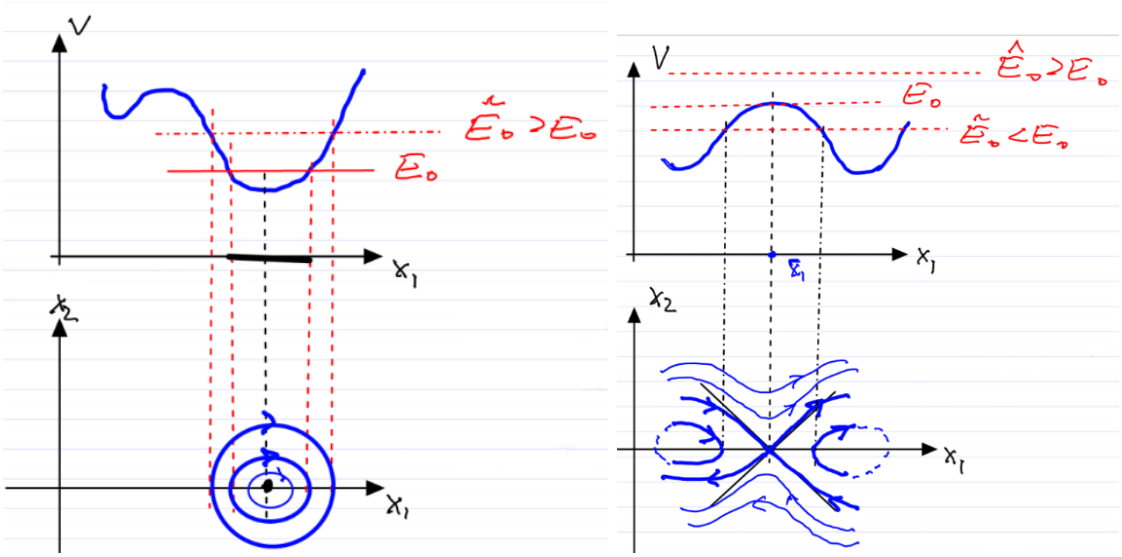


Figure 5.2: Left: Closed orbits around the fixed point at the local minimum of the potential function. Right: The saddle type fixed point at the local maxima of the potential. The clockwise orientations comes from consequence (ii).

(v) A local minimum has local maxima to the left and right, unless V is monotonously nondecreasing to infinity. Then we have two cases to differentiate. One where the local maxima are at the same level, and the other where one local maximum has a smaller potential than the other (or in one direction V monotonously increases towards infinity). *Homoclinic orbits* are special periodic orbits that are formed on a level set that contains a single saddle point. *Heteroclinic orbits* are formed when the level set contains two saddle points, i.e. when the two maxima have the same potential. These two cases are illustrated in Fig. 5.3.

These consequences are illuminated in a basic example.

Example 5.1 (The Duffing oscillator). Let the equation of motion be given by

$$\ddot{x} - x + x^3 = 0; \quad x_1 = x; \quad x_2 = \dot{x}.$$

Therefore we get the first order ODE

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_1 - x_1^3 = -V'(x_1) \end{cases} \implies E(x) = \frac{1}{2}x_2^2 + V(x_1) = \frac{1}{2}x_2^2 + \left(-\frac{1}{2}x_1^2 + \frac{1}{4}x_1^4\right).$$

The full phase portrait of two homoclinic orbits is shown in Fig. 5.4. We find two stable fixed points at the potential minima and they are separated by a potential maximum in the middle.

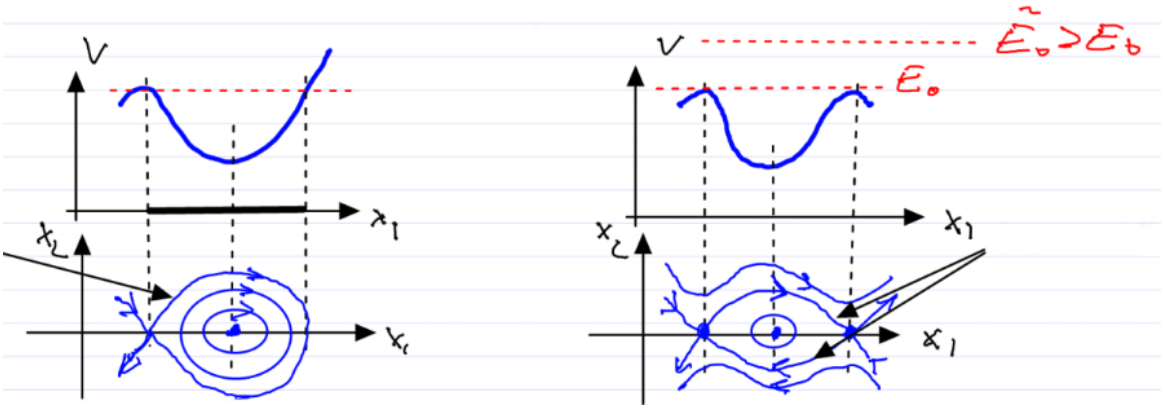


Figure 5.3: Left: One maximum has a lower potential than the other (or V is nondecreasing towards infinity on the right). The black arrow denotes a homoclinic orbit. Right: All maxima have the same potential. The black arrows denote heteroclinic orbits.

Periodic orbits surround the stable fixed points. These orbits are separated from the rest of the phase space by the two homoclinic orbits of the saddle. Outside, we have periodic orbits that correspond to energy level-sets higher than the potential of the saddle.

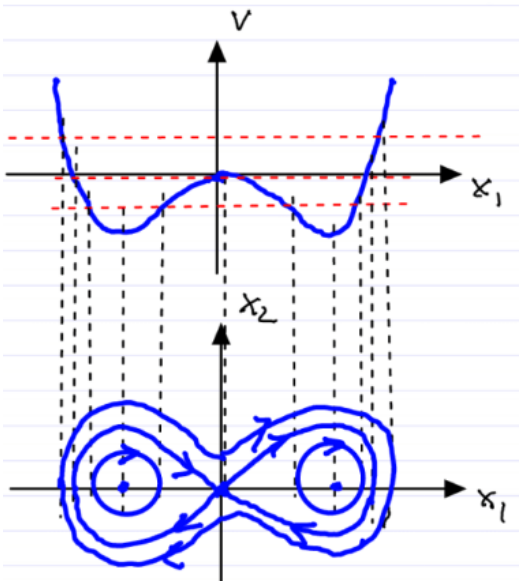


Figure 5.4: Phase portrait of the Duffing oscillator using the potential V.

5.2 Global behavior in two dimensional autonomous dynamical systems

Consider the following dynamical system

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; \quad f \in \mathcal{C}^1.$$

Further assume that solutions exist for all times, therefore there exists a unique solution $x(t; x_0)$ for all $t \in \mathbb{R}$ such that $x(0; x_0) = x_0$. Next, a few definitions are introduced enabling a deeper exploration of such systems.

Definition 5.1. A point $p \in \mathbb{R}^2$ is an *ω-limit point* of x_0 if there exists a monotone increasing unbounded sequence of times $\{t_i\}_{i=1}^{\infty}$ with $t_1 \geq 0$ such that

$$\lim_{i \to \infty} x(t_i; x_0) = p.$$

Definition 5.2. A point $q \in \mathbb{R}^2$ is an *α-limit point* of x_0 if it is an ω-limit point in backward time.

Definition 5.3. The *ω-limit set* of x_0, denoted $\omega(x_0)$ is the set of all ω-limit points of x_0. The *α-limit set* of x_0, denoted $\alpha(x_0)$ is the set of all α-limit points of x_0.

Remark 5.1. Note that $\omega(x_0)$ and $\alpha(x_0)$ is the same for all x_0 along a given trajectory, thus limit sets can be associated with full trajectories.

Example 5.2 (Examples of limit sets). Below (Fig. 5.5) three different limit sets are depicted. For the leftmost dynamical system we have $\omega(p) = \alpha(p) = \{p\}$ and

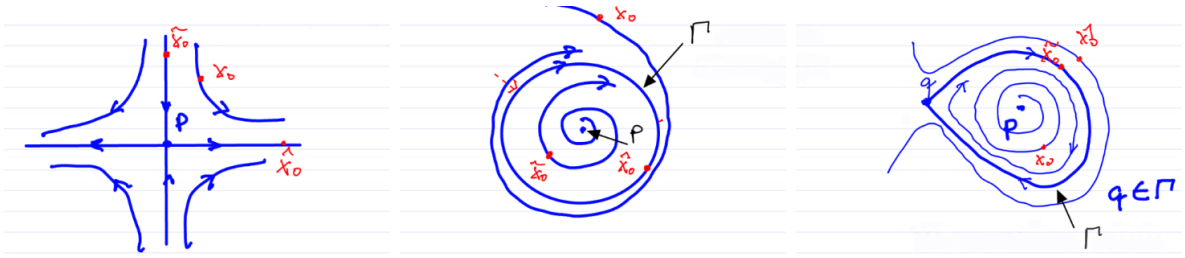


Figure 5.5: Three example systems for which we examine the respective limit sets. The arrows labelled Γ denote a stable limit cycle (middle) and a homoclinic orbit (right).

$$\begin{aligned} \omega(x_0) &= \emptyset, & \omega(\tilde{x}_0) &= \{p\}, & \omega(\hat{x}_0) &= \emptyset, \\ \alpha(x_0) &= \emptyset, & \alpha(\tilde{x}_0) &= \emptyset, & \alpha(\hat{x}_0 &= \{p\}. \end{aligned}$$

For the middle dynamical system we have

$$\begin{aligned} \omega(x_0) &= \Gamma, & \omega(\tilde{x}_0) &= \Gamma, & \omega(\hat{x}_0) &= \Gamma, \\ \alpha(x_0) &= \text{unclear}, & \alpha(\tilde{x}_0) &= \{p\}, & \alpha(\hat{x}_0) &= \Gamma. \end{aligned}$$

For the rightmost dynamical system we have

$$\begin{aligned} \omega(x_0) &= \Gamma, & \omega(\tilde{x}_0) &= \{q\}, & \omega(\hat{x}_0) &= \text{unclear}, \\ \alpha(x_0) &= \{p\}, & \alpha(\tilde{x}_0) &= \{q\}, & \alpha(\hat{x}_0) &= \text{unclear}. \end{aligned}$$

We now introduce two theorems which give us more insight into these limit sets.

Theorem 5.2. *If $x(t; x_0)$ is bounded, then $\omega(x_0)$ and $\alpha(x_0)$ are nonempty, closed, and connected.*

Theorem 5.3 (Poincaré-Bendixon)**.** *If $x(t; x_0)$ is bounded, then $\omega(x_0)$ and $\alpha(x_0)$ must be one of the following*

(i) A connected set of fixed points,

(ii) A limit cycle,

(iii) A set of fixed points and their homo-/heteroclinic orbits (illustrated in Fig. 5.6).

There is no complex (chaotic) liminting behavior in 2-dimensional autonomous nonlinear systems.

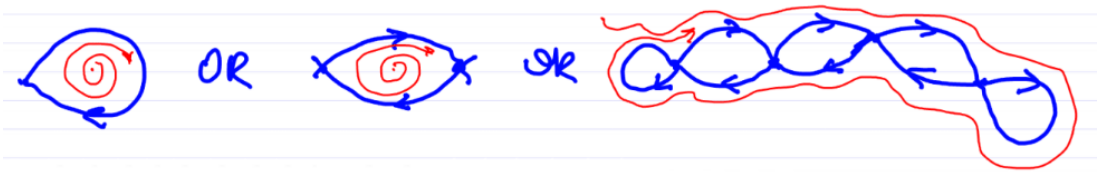


Figure 5.6: Examples of a set of fixed points connected by homo-/heteroclinic orbits.

Remark 5.4. The homo-/heteroclinic orbits are generally not robust, thus we should not expect to see them in a typical dynamical system. We can imagine that under small perturbations the continuous line of the orbit breaks up into separate unstable and stable manifolds as shown in Fig. 5.7. However, these orbits are robust within the class of conservative systems, and conservative pertubations do not lead to global bifurcations as seen in Fig. 5.8.

Figure 5.7: The loss of the homoclinic orbit under a small perturbation. Before the perturbation $\Gamma = W^S(p) = W^U(p)$ (p is the hyperbolic fixed point), meanwhile after the perturbation $W^S(p) \neq W^U(p)$. The unstable manifold is given in red, and the stable manifold in blue.

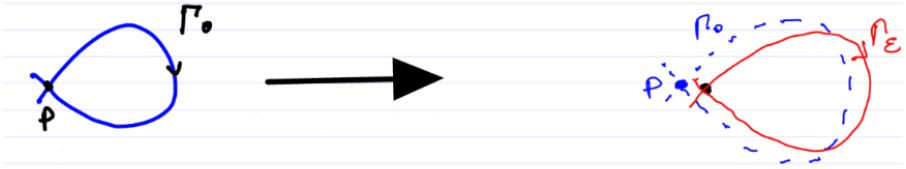


Figure 5.8: Small conservative perturbation does not lead to a global bifurcation in conservative systems. The energy level containing Γ_0 is robust, and perturbs to a new Γ_ε.

Remark 5.5. A consequence of the Poincaré-Bendixon Theorem is that a forward-invariant, bounded open set without fixed points must contain a limit cycle. This is due to the elimination of possibilities (i) and (iii) in the theorem.

Proposition 5.6 (Bendixson criterion)**.** *Consider the following*

$$\dot{x} = f(x); \quad x \in \mathbb{R}^2; f \in \mathcal{C}^1; \quad U \text{ simply connected.}$$

And further assume

$$x \in U \implies div\, f(x) \neq 0.$$

This in turn implies $\nabla \cdot f(x)$ has constant sign on U. Then there cannot exist a limit cycle in U.

Proof. Assume, towards a contradiction, that there exists a closed orbit Γ of period τ. Calculate

$$f^\perp(x) = \begin{pmatrix} f_2(x) \\ -f_1(x) \end{pmatrix} \implies \oint_\Gamma f^\perp \cdot \underbrace{dr}_{=\dot{r}dt} = \int_0^\tau f^\perp(r(t)) \cdot f(r(t))dt = 0.$$

On the other hand we have by Green's Theorem

$$\oint_\Gamma f^\perp \cdot dr = \int\int_{\Gamma^\circ} \left(\nabla \times f^\perp\right)_z dA = -\int\int_{\Gamma^\circ} \underbrace{\nabla \cdot f}_{\neq 0} dA \neq 0.$$

Here, Γ° denotes the interior of the area enclosed by Γ. These two equations stand in contradiction to each other, therefore no limit cycle can exist in U. □

Remark 5.7. Another consequence of the Poincaré-Bendixson Theorem is that purely damped or purely forced perturbations of a conservative system cannot have limit cycles.

$$f(x) = \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -V'(x_1) + g(x_1, x_2) \end{cases}; \quad \nabla \cdot f = \nabla \cdot \begin{pmatrix} 0 \\ g(x) \end{pmatrix} = \frac{\partial g}{\partial x_2}.$$

If we assume pure damping or forcing, the divergence of f cannot be 0 on any open set. If it is positive, we have pure forcing, alternatively if it is negative we have pure damping. In either case no limit cycle can exist.

Chapter 6

Time-dependent dynamical systems

Our setup for this chapter will be of nonautonomous n-dimensional systems, i.e. the following

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n; \quad t \in \mathbb{R}.$$

Note that by adding a dimension to our phase space, we can make the system autonomous

$$X = \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^{n+1}; \quad F(X) = \begin{pmatrix} f \\ 1 \end{pmatrix} \implies \dot{X} = F(X).$$

6.1 Nonautonomous linear systems

For this section, consider

$$\dot{x} = A(t)x + b(t); \quad x \in \mathbb{R}^n; \quad A(t) \in \mathbb{R}^{n\times n}; \quad b(t) \in \mathbb{R}^n. \tag{6.1}$$

We separate (6.1) into two parts. The homogeneous part is

$$\dot{x} = A(t)x. \tag{6.2}$$

For the homogeneous part we have the solution $x(t) = \Phi_{t_0}^t x_0$, for the fundamental matrix solution Φ with the property $\Phi_{t_0}^{t_0} = I \in \mathbb{R}^{n\times n}$. The general solution of (6.1) is the sum of the general solution of (6.2) and a particular solution of (6.1). We can find the particular solution by, for example, the variation of constants formula. We refer to [Arnold, 1992] for the derivation.

$$\dot{x} = b(t).$$

Thus we find the general solution to (6.1) to be

$$\boxed{x(t) = \Phi_{t_0}^t x_0 + \Phi_{t_0}^t \int_{t_0}^t \left(\Phi_{t_0}^s\right)^{-1} b(s) ds.}$$

Note that due to the group property of the flow map, the solution can be equivalently written as

$$x(t) = \Phi_{t_0}^t x_0 + \int_{t_0}^t \Phi_s^t b(s) ds.$$

From here we can see that it is enough to understand the homogeneous part. The stability of $x = 0$ in (6.2) is equivalent to the stability of $x_p(t) = \Phi_{t_0}^t \int_{t_0}^t \left(\Phi_{t_0}^s\right)^{-1} b(s) ds$ in (6.1). Unfortunately the general solution to (6.2) is unknown.

6.2 Time-periodic nonlinear systems

We now specify our system to be T-periodic, thus the evolution rule defined by $\dot{x} = A(t)x$ repeats itself with period T, i.e.

$$A(t + T) = A(t).$$

This implies

$$\Phi_{t_0}^{t_0+T} = \Phi_{t_0+T}^{t_0+2T} = \Phi_{t_0+2T}^{t_0+3T} = \ldots =: P_{t_0},$$

The operator P_{t_0} is called the *Poincaré map* (or time T map) based at t_0. Thus we have

$$x_0(t_0 + nT) = \underbrace{P_{t_0} \ldots P_{t_0}}_{n \text{ times}} = P_{t_0}^n x_0.$$

The n-th iterate of the Poincaré map is denoted by $P_{t_0}^n$.

Time being a periodic variable allows us to study the geometry in the extended phase space $\mathbb{R}^n \times S^1$ as depicted in Fig. 6.1. Importantly, the stability of $x = 0$ can be studied as the stability of the $x = 0$ fixed point of the map P_{t_0}. Label the eigenvalues of P_{t_0} $\rho_1, \ldots, \rho_n \in \mathbb{C}$ these are called *Floquet multipliers*. Since this is now a discrete dynamical system, we know how to assess the stability type of its fixed point. We have two cases

(i) If there exists an eigenvalue with $|\rho_i| > 1$ for P_{t_0} or there exists a repeated eigenvalue with $|\rho_j| = 1$ for which the algebraic multiplicity is greater than the geometric multiplicity ($a > g$), then $x = 0$ is unstable for (6.2);

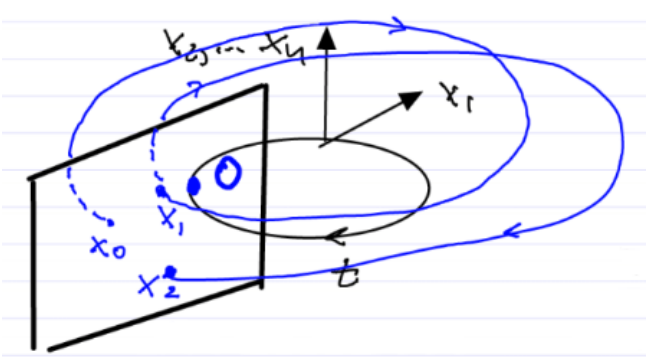


Figure 6.1: The map P records snapshots of the trajectory $x(t)$ taken at times nT. These snapshots are depicted along a single plane in the extended phase space. Due to periodicity, the extended phase space is made up as the product of the plane and the circle S^1. The snapshots shown are $Px_0 = x_1$ and $P^2x_0 = x_2$.

(ii) Otherwise, if each eigenvalue of P_{t_0} has magnitude strictly less than 1, i.e. $|\rho_j| < 1$, then $x = 0$ is asymptotically stable for (6.2).

Floquet theory tells us that the following holds in general

$$\boxed{\Phi_{t_0}^t = B(t)e^{\Lambda(t-t_0)}; \quad B(t+T) = B(t) \in \mathbb{R}^{n\times n}; \quad \Lambda \in \mathbb{R}^{n\times n} \text{ is constant.}}$$

Definition 6.1 (Floquet exponents)**.** The eigenvalues of Λ are $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and they are called *Floquet exponents*. The previously defined Floquet multipliers are related to λ_j as

$$\boxed{\rho_j = e^{\lambda_j T} = e^{(\alpha_j + i\beta_j)T}.}$$

Hence, λ_j is well-defined only up to addition of $i2k\pi/T$ for $k \in \mathbb{Z}$. Therefore we have

$$|\rho_j| \underset{>}{\overset{\leq}{}} 1 \Leftrightarrow \mathrm{Re}\lambda_j \underset{>}{\overset{\leq}{}} 0.$$

This is generally only numerically computable.

We now explore a sufficient criterion for instability of $x = 0$. For this, recall Liouville's Theorem (Abel's Theorem)

$$\boxed{\det(\Phi_{t_0}^t) = \det(\Phi_{t_0}^t)e^{\int_{t_0}^t \mathrm{Tr}[A(s)]ds}.}$$

This holds true for any dynamical system $\dot{x} = A(t)x$ with a periodic time-dependence.

Proposition 6.1 (Sufficient criterion for instability)**.** *If we have* $\int_0^T Tr[A(s)]ds > 0$ *then* $x = 0$ *is unstable for* $\dot{x} = A(t)x$ *with* $A(t) = A(t+T)$

Proof. Apply Liouville's Theorem to our T-periodic system to find

$$\prod_{j=1}^{n} \rho_j = e^{\sum_{j=1}^{n} \lambda_j T} = \det(P_{t_0}) = \det(\Phi_{t_0}^{t_0+T}) = e^{\int_{t_0}^{t_0+T} \mathrm{Tr}[A(s)]ds}.$$

Now we can compare the arguments of the exponential functions to derive

$$\boxed{\sum_{j=1}^{n} \lambda_j = \frac{1}{T} \int_0^T \mathrm{Tr}\left[A(s)\right] ds.}$$

Therefore our claim holds due to the case differentiation for stability above. Thus we can verify the stability without the numerical solution of (6.2). □

Remark 6.2. The change of variables $x = P(t)y$, for some matrix $P(t)$ transforms the system (6.2) into the form $\dot{y} = By$, an autonomous linear ODE.

Carrying out the change of coorinates, we obtain that

$$\dot{y} = P^{-1}(A(t)P - \dot{P})y.$$

But

$$\frac{d}{dt}\left(P(t)e^{Bt}\right) = APe^{Bt} \implies \dot{P} + PB = AP.$$

Together this implies $\dot{y} = By$.

6.3 Averaging

We will now attempt to study the mean evolution of oscillatory systems

$$\dot{x} = f(x,t); \quad x \in \mathbb{R}^n; \quad f(x,t+T) = f(x,t).$$

In general the averaged system

$$\dot{y} = \frac{1}{T} \int_0^T f(y,t)$$

does not capture the overall (mean) behavior of the system. This can be demonstrated in an example.

Figure 6.2: A swing (pendulum) with varying length and a mass m.

Example 6.1 (Averaging for a swing). Mechanically, we examine the stability of a pendulum with a periodically varying length as shown in Fig. 6.2. The first order model for this system is

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\omega_0^2(1 + a(t))x_1 \end{cases}; \quad a(t) = a(t+T), \quad 0 < |a| \ll 1.$$

Averaging these equations yields

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -\omega_0^2(1 + \overline{a})x_1. \end{cases}$$

Where $\overline{a}$ is the average of $a(t)$, a constant. The phase portrait of this is known, as it is just a standard pendulum, i.e. $x = 0$ is stable. This means that the small forcing, when averaged, shouldn't move us from the stable position at $x = 0$. But we know that by periodically swinging, we don't just stay stuck at the equilibrium point of the swing, but instead are able to reach larger amplitude oscillations. Therefore, averaging clearly does not work in this situation.

However, there are situations in which averaging does work.

Example 6.2 (Averaging in slowly varying periodic systems). Slowly varying periodic systems are given by

$$\dot{x} = \varepsilon f(x, t, \varepsilon); \quad x \in \mathbb{R}^n; \quad 0 \leq \varepsilon \ll 1; \quad f(x, t, \varepsilon) = f(x, t + T, \varepsilon).$$

Such systems are also called *adiabatic* systems. The averaged system in this case is

$$\dot{y} = \varepsilon \frac{1}{T} \int_0^T f(y, t, 0) = \varepsilon \overline{f_0}(y).$$

Here only the leading order terms with respect to ε in the integrand were taken, as $\epsilon \ll 0$. Now we transform our original equation into the averaged equation via a near-identity, periodic change of variables

$$x = y + \varepsilon w(y, t); \quad w(y, t) = w(y, t + T).$$

Plugging this transformation into our differential equation, we find

$$\begin{aligned}\dot{x} &= [I + \varepsilon D_y w]\,\dot{y} + \varepsilon \frac{\partial w}{\partial t} \\ \dot{y} &= \underbrace{[I + \varepsilon D_y w]^{-1}}_{I - \varepsilon D_y w + \mathcal{O}(\varepsilon^2)} \left[\varepsilon \underbrace{f(y + \varepsilon w(y,t), t, \varepsilon)}_{=\dot{x}} - \varepsilon \frac{\partial w}{\partial t} \right] \\ &= \varepsilon \left[f(y,t,0) - \frac{\partial w}{\partial t} \right]_t + \mathcal{O}(\varepsilon)^2.\end{aligned}$$

In the last equation we used the Taylor expansion. We cannot set $w(y,t) = \int_0^t f(y,s,0)ds$ as that would make ω aperiodic. Instead we separate f into two parts, the average and the deviation, $f(y,t,0) = \overline{f}_0(y) + \tilde{f}(y,t)$. Note that the deviation has mean 0, therefore we set $w(y,t) = \int_0^t \tilde{f}(y,s)ds$ which is periodic. Therefore we find $\dot{y} = \varepsilon \overline{f}_0(y) + \mathcal{O}(\varepsilon^2)$, noting here that the order 2 term is still T periodic.

This has given us that the original system is a T-periodic, $\mathcal{O}(\varepsilon^2)$ perturbation of the averaged system in the y coordinates.

Averaging Principle Solutions of the original equation starting $\mathcal{O}(\varepsilon)$ close to those of the averaged system, stay $\mathcal{O}(\varepsilon)$ close for times of $\mathcal{O}\left(\frac{1}{\varepsilon}\right)$ (very long).

The averaging principle is depicted in Fig. 6.3.

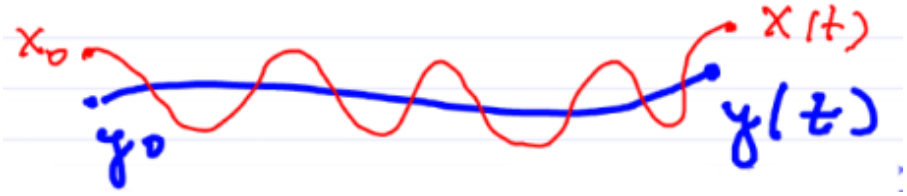


Figure 6.3: For the starting point x_0 which is $\mathcal{O}(\varepsilon)$ close to y_0, the trajectories remain order ε close, i.e. $x(t)$ is $\mathcal{O}(\varepsilon)$ close to $y(t)$, for $t \ll \frac{K}{\varepsilon}$.

This has implications for the Poincaré map of the original equation

$$P^\varepsilon_{t_0} : x_0 \mapsto x(t_0 + T, x_0; \epsilon); \quad P^\varepsilon_{t_0} = P^0_{t_0} + \mathcal{O}(\epsilon) \implies \left(P^\varepsilon_{t_0}\right)^n = \left(P^0_{t_0}\right)^n + \mathcal{O}(\varepsilon).$$

This only holds if $n \leq \frac{1}{\varepsilon}$. The time T map of the averaged system was denoted as $P^0_{t_0}$.

Example 6.3 (Averaged system with a saddle-type fixed point). Consider a system whose Poincaré map P_0 has a hyperbolic fixed point p_0. By the persistence of hyperbolic fixed points, the perturbed Poincaré map P_ε also has a hyperbolic fixed point p_ε which is $\mathcal{O}(\varepsilon)$ close to p_0. Furthermore, the $\mathcal{C}^1$ manifolds $W^U_{\text{loc}}(p_\varepsilon)$ and $W^S_{\text{loc}}(p_\varepsilon)$ are $\mathcal{O}(\varepsilon)$ close to their unperturbed counterparts. The unperturbed system and geometry of the perturbed system are shown in Fig. 6.4.

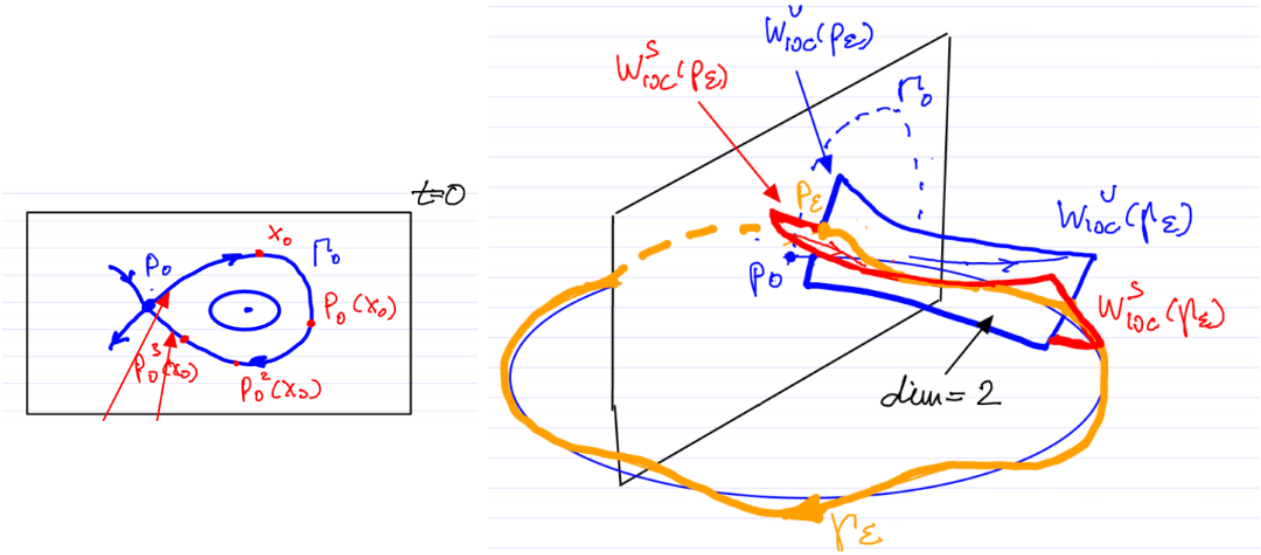


Figure 6.4: Left: The averaged system with the saddle type fixed point p_0. The upper red arrow designates the local unstable manifold, and the lower arrow the local stable manifold. Right: The same in the extended phase space, time is represented by S_1 and runs clockwise. The hyperbolic limit cycle Γ_ε is given by the yellow trajectory. Furthermore, the non-averaged local stable and unstable manifolds are drawn in red and blue respectively.

Remark 6.3. We can show that fixed points persist for P_ε as follows. First we take the flow map for the system $F^t_{t_0}(y_0;\varepsilon)$, then we use the Taylor expansion to find

$$\begin{aligned} F^{t_0+T}_{t_0}(y_0;\varepsilon) &= F^{t_0+T}_{t_0}(y_0;0) + \varepsilon\frac{\partial}{\partial\varepsilon}F^{t_0+T}_{t_0}(y_0;0) + \mathcal{O}(\varepsilon^2) \\ &= y_0 + \varepsilon\left.\frac{\partial y(t_0+T;t_0;\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=0} + \mathcal{O}(\varepsilon^2) \\ &= y_0 + \varepsilon\int_{t_0}^{t_0+T}\overline{f}(y_0)dt + \mathcal{O}(\varepsilon^2). \end{aligned}$$

In the last equality we used that

$$\left.\frac{\partial y(t_0+T;t_0,\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = \left.\frac{\partial}{\partial \varepsilon}\int_{t_0}^{t_0+T} \dot{y}(s;t_0,\varepsilon)ds\right|_{\varepsilon=0} = \int_{t_0}^{t_0+T} \left.\frac{\partial}{\partial \varepsilon}\left(\varepsilon \overline{f}(y(s)) + \mathcal{O}(\varepsilon^2)\right)\right|_{\varepsilon=0} ds$$
$$= \int_{t_0}^{t_0+T} \overline{f}(y_0)ds.$$

From this we have

$$P_\varepsilon(y_0) - y_0 \Leftrightarrow y_0 + \varepsilon T\overline{f}(y_0) + \mathcal{O}(\varepsilon^2) - y_0 = 0 \Leftrightarrow T\overline{f}(y_0) + \mathcal{O}(\varepsilon) = 0.$$

And now we can use the Implicit Function Theorem to conclude.

Example 6.4 (Averaged system with a stable focus). Consider the averaged system with a stable focus around a fixed point p_0. The transition to the perturbed system is shown in Fig. 6.5.

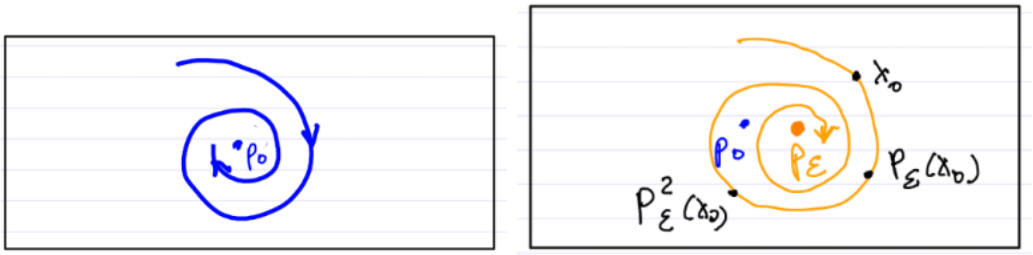


Figure 6.5: The stable focus around p_0 in the averaged system is preserved in the non-averaged system.

Example 6.5 (Averaged system with a limit cycle). Consider the averaged system with a limit cycle Γ_0. The non-averaged system also has a stablelimit cycle Γ_ε with $\dim W^S(\Gamma_\varepsilon) = \mathbb{R}^{n+1}$ in extended phase space. Thus in the full system we have an attracting 2-dimensional invariant torus. The averaged and non-averaged system, along with this torus are shown in Fig. 6.6.

Example 6.6 (Weakly nonlinear oscillations). An important application of averaging is for weakly nonlinear oscillations. This is covered in-depth in [Guckenheimer and Holmes, 1990]. The system is given by

$$\ddot{x} + \omega_0^2 x = \varepsilon f(x,\dot{x},t) = \varepsilon f(x,\dot{x},t+T); \quad T = 2\pi\omega.$$

Averaging applies after a change of coordinates moving with the solutions of the $\varepsilon = 0$ limit. First transform the system to be a first-order ODE with $X = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$.

$$\dot{X} = AX + \varepsilon F(X,t); \quad A = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix}; \quad F(X,t) = \begin{pmatrix} 0 \\ f(X_1,X_2,t) \end{pmatrix}. \tag{6.3}$$

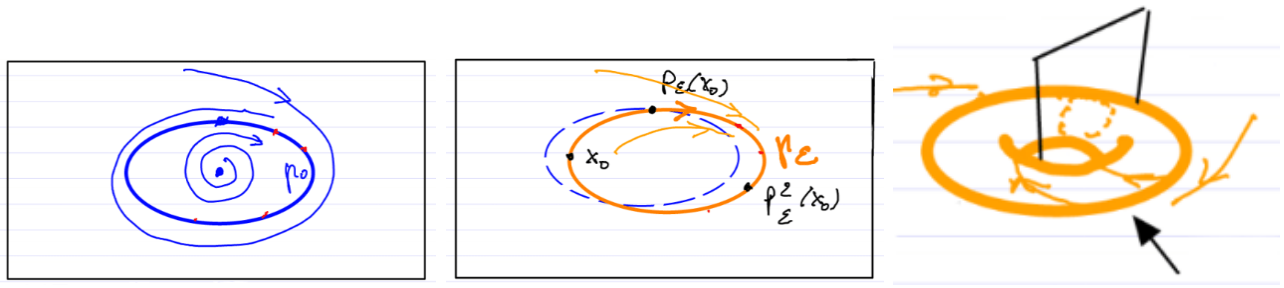


Figure 6.6: Left: The averaged system with the stable limit cycle Γ_0. Middle: The non-averaged system with the limit cycle Γ_ε. Right: The attracting 2-dimensional invariant torus (black arrow), trajectories can be seen as the thin yellow lines.

We cannot directly apply averaging to (6.3). Instead we introduce a coordinate change along the trajectories of the system. Let $\Phi(t;\omega_0)$ be the fundamental matrix of solutions. In fact, this is explicitly computable for the harmic oscillator as

$$\Phi(t;\omega_0) = \begin{pmatrix} \cos(\omega_0 t) & -\sin(\omega_0 t) \\ -\omega_0 \sin(\omega_0 t) & -\omega_0 \cos(\omega_0 t) \end{pmatrix}.$$

It can be checked that $\dot{\Phi} = A\Phi$. Thus we change coordinates with $X(t) = \Phi(t,\omega_0)Y(t)$. As Φ is invertible, we have $Y(t) = \Phi^{-1}(t,\omega_0)X(t)$ with the inverse

$$\Phi^{-1}(t,\omega_0) = \begin{pmatrix} \cos(\omega_0 t) & -\frac{1}{\omega_0}\sin(\omega_0 t) \\ -\sin(\omega_0 t) & -\frac{1}{\omega_0}\cos(\omega_0 t) \end{pmatrix}.$$

Now we substitute this transformation into the differential equation (6.3) and use the fact that $\dot{\Phi} = A\Phi$ to find

$$\dot{\Phi}Y + \Phi\dot{Y} = A\Phi Y + \varepsilon F(\Phi^{-1}Y, t) \implies \dot{Y} = \varepsilon\Phi^{-1}F(\Phi Y, t).$$

This appears to be in the required form for averaging, except the right hand side is not generally periodic, because Φ has frequency ω_0, while the driving has frequency ω. It is only periodic if $\omega = k\omega_0$ for $k \in \mathbb{Z}$, i.e. resonance of order k. Often it is important to analyze near-resonant oscillations, i.e. $\omega \approx k\omega_0$. Thus our expectation is to find an almost sinusoidal response of frequency $\frac{\omega}{k}$. Therefore we instead use the transformation $\Phi(t;\frac{\omega}{k})$ and find

$$\dot{Y} = \Phi^{-1}\left[A\Phi - \dot{\Phi}\right]Y + \varepsilon\Phi^{-1}F(\Phi Y, t).$$

Now write $Y = \begin{pmatrix} u \\ v \end{pmatrix}$, and we find

$$\begin{cases} \dot{u} = -\dfrac{k}{\omega}\left[\left(\dfrac{\omega^2 - k^2\omega_0^2}{k^2}\right)x + \varepsilon f(x,\dot{x},t)\right]\sin\left(\dfrac{\omega t}{k}\right) \\ \dot{v} = -\dfrac{k}{\omega}\left[\left(\dfrac{\omega^2 - k^2\omega_0^2}{k^2}\right)x + \varepsilon f(x,\dot{x},t)\right]\cos\left(\dfrac{\omega t}{k}\right) \end{cases} \tag{6.5}$$

Now we have that if $\omega^2 - k^2\omega_0^2 = \mathcal{O}(\varepsilon)$, averaging applies to (6.5).

Next we consider a specific example, the Duffing oscillator

$$f(x,\dot{x},t) = \beta\cos(\omega t) - \delta\dot{x} - \alpha x^3,$$

with $\omega_0^2 - \omega^2 = \varepsilon\Omega$, i.e. a nearly order one resonance. The transformed system then takes the form

$$\begin{cases} \dot{u} = \dfrac{\varepsilon}{\omega}[& \Omega\left(u\cos(\omega t) - v\sin(\omega t)\right) - \omega\delta\left(u\sin(\omega t) + v\cos(\omega t)\right) \\ & +\alpha\left(u\cos(\omega t) - v\sin(\omega t)\right)^3 - \beta\cos(\omega t)]\sin(\omega t) \\ \dot{v} = \dfrac{\varepsilon}{\omega}[& \Omega\left(u\cos(\omega t) - v\sin(\omega t)\right) - \omega\delta\left(u\sin(\omega t) + v\cos(\omega t)\right) \\ & +\alpha\left(u\cos(\omega t) - v\sin(\omega t)\right)^3 - \beta\cos(\omega t)]\cos(\omega t). \end{cases}$$

Averaging over one period $T = \frac{2\pi}{\omega}$ we find

$$\frac{1}{T}\int_0^T \sin(\omega t)\cos(\omega t)dt = \frac{1}{2T}\int_0^T \sin(2\omega t)dt = \left[-\frac{\cos(2\omega t)}{4\omega T}\right]_0^T = 0,$$

and

$$\frac{1}{T}\int_0^T \sin^2(\omega t)dt = \frac{1}{T}\int_0^T \frac{1-\cos(2\omega t)}{2}dt = \frac{1}{T}\left[\frac{1}{2} - \frac{\sin(2\omega t)}{2\omega}\right]_0^T = \frac{1}{2}.$$

Using trigonometric identities then leads to the equation

$$\begin{cases} \dot{u} = \dfrac{\varepsilon}{2\omega}\left[-\omega\delta u - \Omega v - \dfrac{3\alpha}{4}(u^2+v^2)v\right] \\ \dot{v} = \dfrac{\varepsilon}{2\omega}\left[\Omega u - \omega\delta v + \dfrac{3\alpha}{4}(u^2+v^2)u - \beta\right]. \end{cases}$$

This can be further simplified by using the polar coordinates (r,ϕ), i.e.

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} r\cos(\phi) \\ r\sin(\phi) \end{pmatrix} \implies \begin{cases} \dot{r} = \dfrac{\varepsilon}{2\omega}\left[-\omega\delta r - \beta\sin(\phi)\right] \\ r\dot{\phi} = \dfrac{\varepsilon}{2\omega}\left[\Omega r + \dfrac{3\alpha}{4}r^3 - \beta\cos(\phi)\right]. \end{cases} \tag{6.6}$$

Next, recall that $x(t) = u(t)\cos(\omega t) - v(t)\sin(\omega t) = r(t)\cos(\omega t - \phi(t))$, i.e. we have slowly varying amplitude $r(t)$ and phase $\phi(t)$. The hyperbolic fixed points of (6.6) correspond to steady, almost sinusoidal response of the original system. We obtain the *forced response curves* by fixing α, β, and δ and plotting the fixed point $(\overline{r}, \overline{\phi})$ of (6.6) as a function of the frequency parameter Ω (or $\frac{\omega}{\omega_0}$) as in Fig. 6.7.

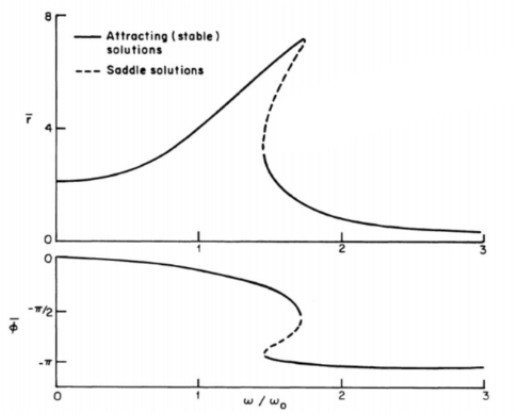


Figure 6.7: The forced response curved for the Duffing equation.

6.4 The Harmonic Balance Method

Averaging and other perturbation techniques are often restrictive due to their small-parameter assumptions and result in cumbersome expressions limiting their application in high-dimensional systems. To avoid these constraints, the *Harmonic Balance* method computes the Fourier series approximation of the periodic response and is formally applicable without the previous restrictions. Consider the setup

$$\dot{x} = f(x,t) = f(x,t+T); \quad x \in \mathbb{R}^n; \quad f \in \mathcal{C}^r_x,\ r \geq 1,\ f \in \mathcal{C}^0_t. \tag{6.7}$$

Now we seek a T-periodic response via the Fourier series.

Remark 6.4. It is not immediately clear that a periodic response exists in a periodically forced multi-degree of freedom mechanical system. See for instance [Breunung and Haller, 2019] for more detail.

Assume (6.7) exhibits a T-periodic response, say $x_p(t) = x_p(t+T)$. Then it can be expressed as a convergent Fourier series

$$x_p(t) = \sum_{k\in\mathbb{Z}} \hat{x}_k e^{ik\omega t} = x_0 + \sum_{j\in\mathbb{N}} c_j \cos(j\omega t) + s_j \sin(j\omega t).$$

Substituting this into (6.7) and using that f is T-periodic (in both variables) and therefore admits a convergent Fourier series we find

$$\begin{aligned}\frac{d}{dt}x_p(t) &= f(x_p(t),t)\\ \frac{d}{dt}\sum_{k\in\mathbb{Z}}\hat{x}_k e^{ik\omega t} &= \sum_{k\in\mathbb{Z}}\hat{f}_k(x_p(t),t)e^{ik\omega t}\\ \sum_{k\in\mathbb{Z}}ik\omega\hat{x}_k e^{ik\omega t} &= \sum_{k\in\mathbb{Z}}\hat{f}_k(x_p(t),t)e^{ik\omega t}.\end{aligned}$$

Now we can compare coefficients at different harmonics to find

$$ik\omega\hat{x}_k = \hat{f}_k(x_p(t),t) = \hat{f}_k\left(\sum_{j\in\mathbb{Z}}\hat{x}_j e^{ij\omega t},t\right) \quad \forall k.$$

These are algebraic equations which give a solution for each coefficient. In general they are coupled, with a noteworthy exception if f is linear. However, there are infinitely many equations to solve, so in practice we only use a finite truncation of the Fourier series to approximate the response.

$$x_H = \sum_{k=-H}^{H}\hat{x}_k e^{ik\omega t}.$$

Now we instead solve $\dot{x}_H = f_H(x_H(t),t)$, leading to only $n(2H+1)$ equations. Recall a few basic results from Fourier Analysis.

Theorem 6.5. *For any $F \in \mathcal{C}^0$ which is T-periodic, the truncation $F_H(t) = \sum_{k=-H}^{H}\hat{F}_k e^{ik\omega t}$ converges uniformly to $F(t)$ as $H \to \infty$.*

Theorem 6.6 (Decay of Fourier coefficients)**.** *For $F \in \mathcal{C}^r$ which is T-periodic, the Fourier coefficients $\hat{F}_k$ decay at least as fast as $\frac{1}{|k|^{r+1}}$. For an analytic function $F \in \mathcal{C}^a$, the coefficients decay faster than any power of k, i. e. exponentially.*

Remark 6.7. Note that the decay estimate is asymptotic, hence it is generally unclear exactly where to truncate. However, the method seems to work well for many practical applications and produces fast results.

We will now apply this method in a few examples.

Example 6.7 (Linear spring-mass-damper oscillator). Consider the following damped linear spring

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x = F(t) = F_0\cos(\omega t).$$

As only a single harmonic is sufficient to describe F, the harmonic coefficients are decoupled. By taking the Fourier series of x and $F = \frac{1}{2}\left(F_0 e^{i\omega t} + F_0 e^{-i\omega t}\right)$ we compare coefficients

$$\begin{aligned} k = 1: &\quad -\omega^2 \hat{x}_1 e^{i\omega t} + i2\zeta\omega_0\omega\hat{x}_1 e^{i\omega t} + \omega_0^2 \hat{x}_1 e^{i\omega t} = F_0 e^{i\omega t}; \\ k = -1: &\quad -\omega^2 \hat{x}_{-1} e^{-i\omega t} - i2\zeta\omega_0\omega\hat{x}_{-1} e^{-i\omega t} + \omega_0^2 \hat{x}_{-1} e^{i\omega t} = F_0 e^{-i\omega t}. \end{aligned}$$

We therefore calculate

$$\hat{x}_1 = \frac{1}{2}\left(\frac{F_0}{\omega_0^2 - \omega^2 + i2\zeta\omega_0\omega}\right); \quad \hat{x}_{-1} = \frac{1}{2}\left(\frac{F_0}{\omega_0^2 - \omega^2 - i2\zeta\omega_0\omega}\right).$$

For all other values of k $\hat{x}_k$ must equal zero. Now we can calculate the famous formula from linear vibration text books for the magnitude of the response

$$|x| = \frac{F_0}{\omega_0^2\sqrt{\left(1 - \left(\frac{\omega}{\omega_0}\right)^2\right)^2 + \left(2\zeta\frac{\omega}{\omega_0}\right)^2}}.$$

Example 6.8 (Duffing Oscillator). Recall the Duffing oscillator

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2 x + \alpha x^3 = F_0\cos(\omega t).$$

We consider the Ansatz

$$q_1(t) = c_1\cos(\omega t) + s_1\sin(\omega t) \implies \ddot{q}_1 + 2\zeta\omega_0\dot{q}_1 + \omega_0^2 q_1 + \alpha q_1^3 = F_0\cos(\omega t).$$

Now calculate q_1^3 using multi-angle trigonometric identities

$$\begin{aligned} q_1^3 &= \left(c_1\cos(\omega t) + s_1\sin(\omega t)\right)^3 \\ &= c_1^3\cos^3(\omega t) + 3c_1^2 s_1\cos^2(\omega t)\sin(\omega t) + 3c_1 s_1^2\cos(\omega t)\sin^2(\omega t) + s_1^3\sin^3(\omega t) \\ &= \frac{1}{4}\left(3c_1^3 + 3c_1 s_1^2\right)\cos(\omega t) + \frac{1}{4}\left(c_1^3 - 3c_1 s_1^2\right)\cos(3\omega t) + \frac{1}{4}\left(3c_1^2 s_1 + 3s_1^3\right)\sin(\omega t) \\ &\quad + \frac{1}{4}\left(3c_1^2 s_1 - s_1^3\right)\sin(3\omega t). \end{aligned}$$

Now we neglect the harmonics higher than 1 and balance the remaining

$$\begin{aligned} 0 = &\left[\left(\omega_0^2 - \omega^2\right)c_1 + 2\zeta\omega_0\omega s_1 + \frac{3}{4}\alpha\left(c_1^3 + c_1 s_1^2\right) - F_0\right]\cos(\omega t) \\ &+ \left[\left(\omega_0^2 - \omega^2\right)s_1 - 2\zeta\omega_0\omega c_1 + \frac{3}{4}\alpha\left(s_1^3 + c_1^2 s_1\right)\right]\sin(\omega t) + \text{higher order harmonics.} \end{aligned}$$

Therefore we have two algebraic equations in c_1 and s_1. Using the polar coordinates $c_1 = A\cos(\phi)$ and $s_1 = A\sin(\phi)$ we find

$$\left(\omega_0^2 - \omega^2\right) A\cos(\phi) + 2\zeta\omega_0\omega A\sin(\phi) + \frac{3}{4}\alpha\left(A^3\cos^3(\phi) + A^3\cos(\phi)\sin^2(\phi)\right) = F_0$$
$$\left(\omega_0^2 - \omega^2\right) A\sin(\phi) + 2\zeta\omega_0\omega A\cos(\phi) + \frac{3}{4}\alpha\left(A^3\sin^3(\phi) + A^3\sin(\phi)\cos^2(\phi)\right) = 0.$$

Adding $\cos(\phi)$ times the first equation to $\sin(\phi)$ times the second we obtain

$$\left(\omega_0^2 - \omega^2\right) A + \frac{3}{4}\alpha A^3 = F_0\cos(\phi).$$

Similarly, adding $\sin(\phi)$ times the first equation to $\cos(\phi)$ times the second we obtain

$$2\zeta\omega_0\omega A = F_0\sin(\phi).$$

We now take the sum of the squares of these two new equations, which yields

$$A^2\left[\left(\omega_0^2 - \omega^2 + \frac{3}{4}\alpha A^2\right)^2 + 4\zeta^2\omega_0^2\omega^2\right] = F_0^2.$$

This equation now allows us plot the amplitude A versus the frequency ω as in Fig. 6.8.

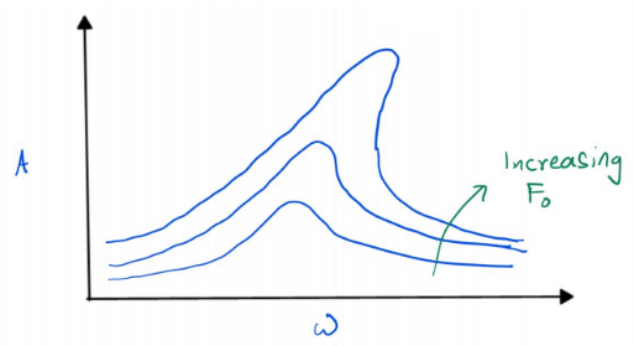


Figure 6.8: Forced response curves derived via harmonic balance for the Duffing oscillator.

Unfortunately, in contrast to averaging, harmonic balance by itself does not return any stability information. We can, however, compute the Floquet multipliers of the periodic orbit by directly solving the equation of variations.

Example 6.9 (Unforced and undamped Duffing oscillator). For unforced systems harmonic balance can also be used to obtain internally parameterized periodic orbits. However, the expected period is a priori unknown. Consider the following

$$\ddot{x} + \omega_0^2 x + \alpha x^3 = 0.$$

Since the system in autonomous, the phase of the response is irrelevant. Thus the simplest Ansatz for a single harmonic response is

$$x_1(t) = A\cos(\omega t),$$

for an unknown ω. Applying the procedure as before and neglecting the higher harmonics we find

$$\omega^2 = \omega_0 + \frac{3}{4}\alpha A^2.$$

For unforced, conservative, systems the frequency-amplitude relation gives the so-called *backbone curve*, illustrated in Fig. 6.9.

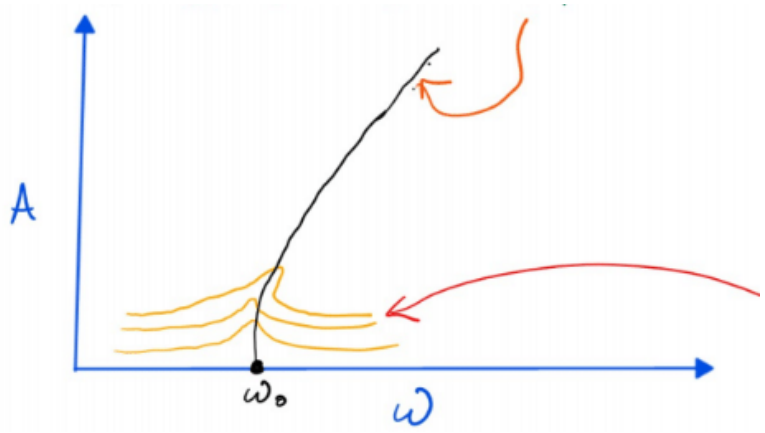


Figure 6.9: The backbone curve (denoted by the orange arrow) for the unforced and undamped Duffing oscillator. Under small forcing and damping, the backbone curve has been observed to connect the forced response curves' peaks (red arrow).

Part II

Nonlinear Dynamics and Chaos 2

Chapter 7

Indroduction to chaotic dynamics

We begin with a few motivating examples.

Example 7.1 (Periodically forced slender beam). A beam is hanging on the inside of a rectangular frame, attached to the upper edge. Two permanent magnets are attached to the lower edge. Furthermore, there is a $T = 2\pi$-periodic forcing in the horizontal direction to the frame. The deflection of the beam is measured by the variable x. The setup is illustrated in Fig. 7.1. This leads to the equation of motion

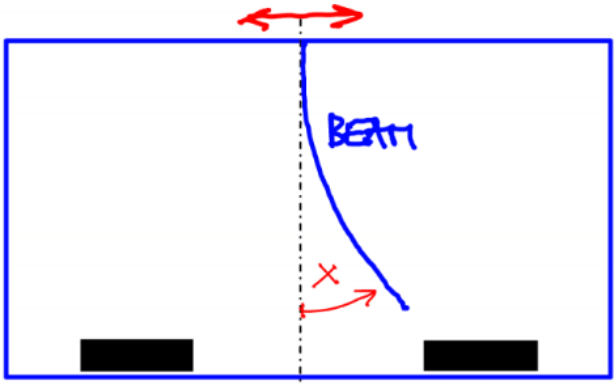


Figure 7.1: Depiction of the periodically forced slender beam. The frame is given by the blue rectangle. The two permanent magnets are represented by the black boxes.

$$\ddot{x} + \dot{x} - x + x^3 = \varepsilon \cos(t); \quad 0 \leq \varepsilon \ll 1.$$

Therefore we have a perturbed Duffing oscillator. We transform the system into a first order

ODE

$$\begin{cases} \dot{x} = y \\ \dot{y} = -y + x - x^3 + \varepsilon \cos(t). \end{cases}$$

For $\varepsilon = 0$ driving, two homoclinic orbits arise for the three fixed points. However, for nonzero driving, we get seemingly chaotic behavior. Both of these regimes are depicted in Fig. 7.2.

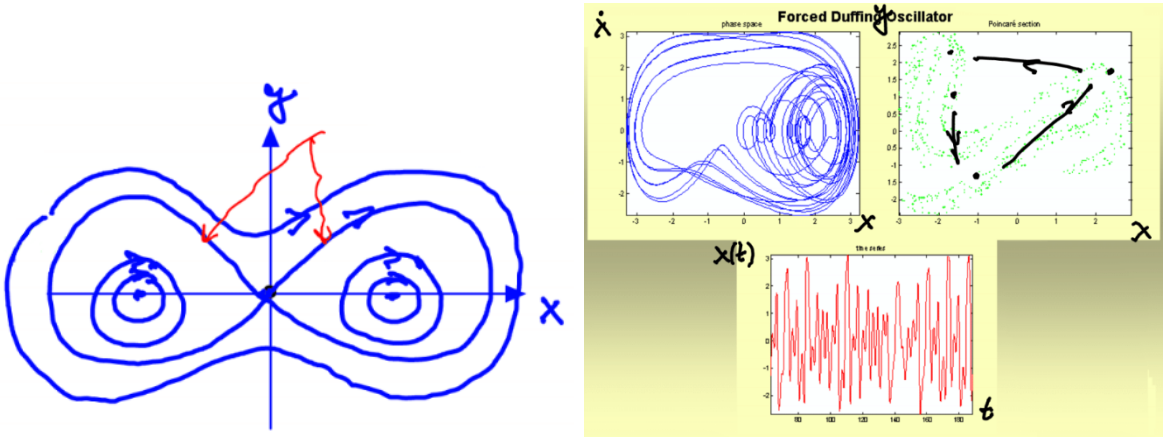


Figure 7.2: Left: Duffing oscillator with $\varepsilon = 0$ driving, the red arrows designate the homoclinic orbits. Right: (Clockwise from top left) The phase space of the driven Duffing oscillator; The Poincaré section of the driven Duffing oscillator, with the Poincaré map illustrated by the black arrows; The value of $x(t)$ over time, with no apparent pattern.

Example 7.2 (2-dimensional Rayleigh-Bénard convection). A fluid is held between two plates. The upper plate is cold and the lower plate is hot. This causes convections to start as hotter fluid rises and colder fluid falls. So called convection cells then form for low Rayleigh numbers. This process is illustrated in Fig. 7.3.

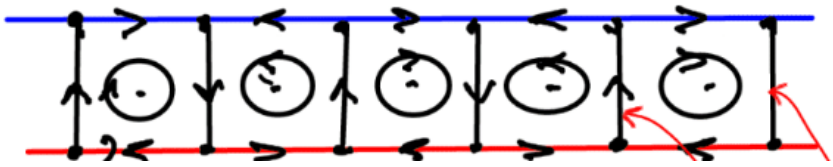


Figure 7.3: Convection cells forming between two plates of different temperature. The cold plate (blue) is vertically above the hot plate (red). Each black dot represents a fixed point of the system, with heteroclinic orbits connecting them (red arrows). This is in an unperturbed setting ($\varepsilon = 0$).

If the Rayleigh number R_a exceeds a critical value $R_{a_{\text{crit}}}$, we have a time periodic perturbation to the velocity field. The fluid trajectories have the following equations of motion

$$\begin{cases} \dot{x} = u(x,y) + \varepsilon u_1(x,y,t) \\ \dot{y} = v(x,y) + \varepsilon v_1(x,y,t) \end{cases}; \quad \varepsilon > 0.$$

Here, the functions u_1 and v_1 are T-periodic. The chaotic nature of this system is shown in Fig. 7.4.

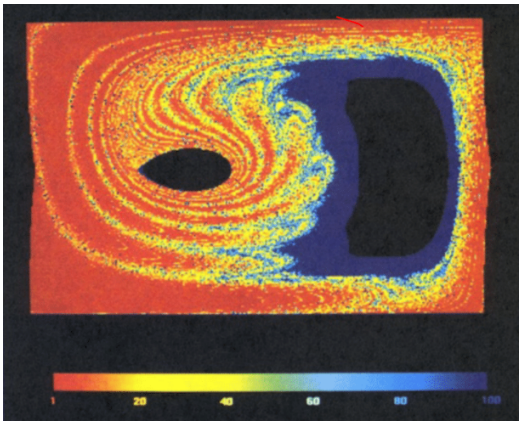


Figure 7.4: Escape time from one convection cell. The number of iterations of the Poincaré map needed for escape is given by the color scale. Here $0 < \varepsilon \leq 1$.

Moving forward we will have a common mathematical setting

$$\dot{x} = f(x) + \varepsilon g(x,t); \quad x \in \mathbb{R}^2, \quad g(x,t) = g(x,t+T); \quad 0 \leq \varepsilon \ll 1; \quad f,t \in \mathcal{C}^1. \tag{7.1}$$

We have a small T-periodic perturbation of a planar ODE which can be studied via a Poincaré map $P_\varepsilon^{t_0} : x_0 \mapsto x(t_0+T; t_0, x_0)$. Assume that for $\varepsilon = 0$ the system (7.1) has a saddle type fixed point with a homoclinic (or heteroclinic) orbit $x^0(t-t_0)$. Such a system is depicted in Fig. 7.5.

Remark 7.1. The fixed point of $P_0^{t_0}$ given by p_0 is hyperbolic

$$P_0^{t_0}(p_0) = p_0; \quad DP_0^{t_0}(p_0) \text{ has eigenvalues } \lambda_1, \lambda_2 : \ |\lambda_1| < 1, \ |\lambda_2| > 1.$$

More generally we define a hyperbolic fixed point for a map.

Definition 7.1. For a map $F : \mathbb{R}^n \to \mathbb{R}^n$ and a dynamical system with $x_{k+1} = F(x_k)$, the fixed point p_0 (i.e. $F(p_0) = p_0$) is *hyperbolic* if the linearization's $DF(p_0) \in \mathbb{R}^{n\times n}$ eigenvalues $\lambda_1, \ldots, \lambda_n$ never have unitary length: $|\lambda_i| \neq 1$ for $i = 1, \ldots, n$.

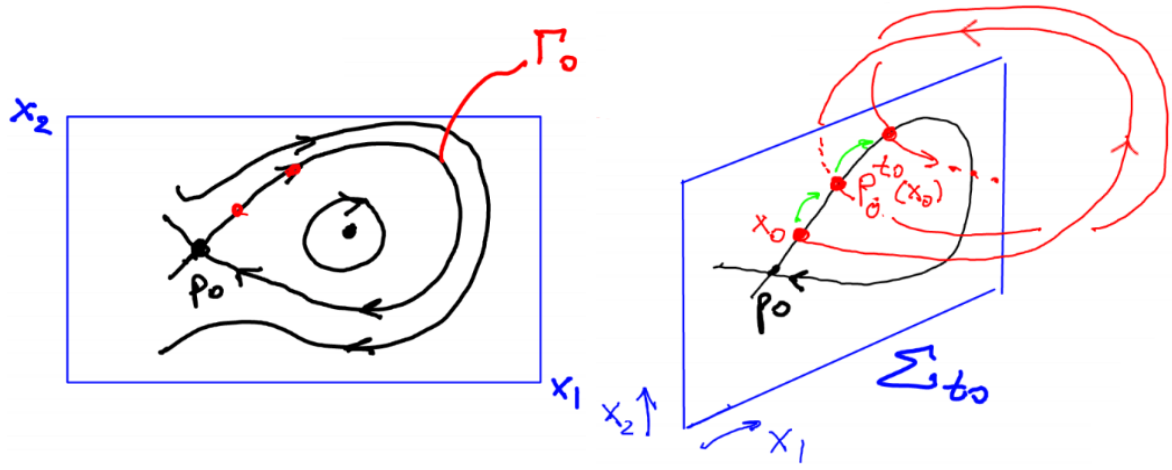


Figure 7.5: Left: An example of a system following the assumptions. The homoclinic orbit Γ_0 is equal to the local unstable and local stable manifolds. Right: The Poincaré map for $\varepsilon = 0$, with time running counterclockwise about the x_1 axis with period T.

For each eigenvalue λ_i of the linearization the corresponding eigenvector is given by s_i. Assume that $DF(p_0)$ is semisimple, i.e. all of the eigenvectors are linearly independent. The linearized dynamics at p_0 for $y = x - p_0$ small is

$$y_{k+1} = DF(p_0)y_k \implies y_k = \lambda_1^k c_1 s_1 + \ldots + \lambda_n^k c_n s_n.$$

If all of the eigenvalues have less than unitary magnitude, $|\lambda_i| < 1$ for $i = 1, \ldots, n$, then p_0 is asymptotically stable. Otherwise if there exists an eigenvalue with modulus strictly larger than 1, p_0 is unstable. The relationship of the nonlinear and linearized dynamics are shown in Fig. 7.6.

The fixed point p_0 along with the stable and unstable manifolds ($W^S(p_0)$ and $W^U(p_0)$) persist smoothly under small smooth perturbations to F near p_0.

7.1 Consequences of hyperbolicity

Hyperbolicity has a few important consequences in our setting. Foremost, the perturbed hyperbolic fixed point $p_\varepsilon^{t_0}$ translates to a hyperbolic periodic orbit for the ODE. Furthermore, the solutions of the ODE are smooth in ε, thus the solutions within $W^U(p_\varepsilon^{t_0})$ and $W^S(p_\varepsilon^{t_0})$ remain $\mathcal{O}(\varepsilon)$ close to Γ_0, i.e.

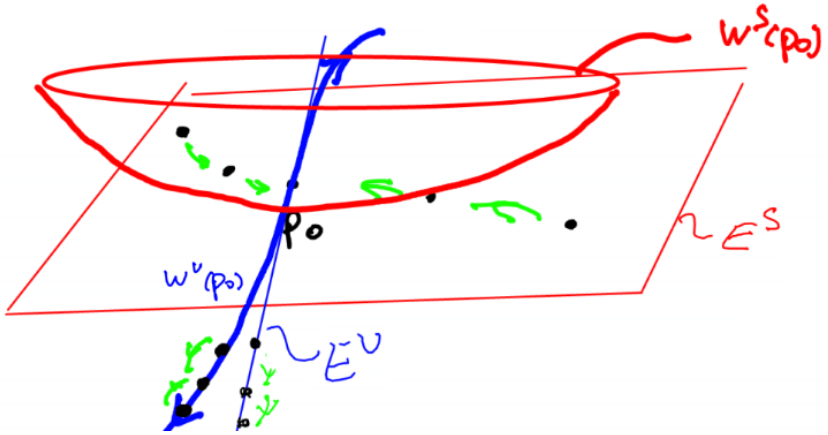


Figure 7.6: The stable ($W^S(p_0)$) and unstable ($W^U(p_0)$) manifolds drawn in relation to the stable (E^S) and unstable (E^U) subspaces. The subspaces are given as the span of the stable (resp. unstable) eigenvectors. The green arrows signify steps of the linearized (resp. original) system.

$$x_\varepsilon^S(t;t_0) = x^0(t-t_0) + \varepsilon a^S(t) + \mathcal{O}(\varepsilon^2); \quad t \in [t_0, \infty)$$
$$x_\varepsilon^U(t;t_0) = x^0(t-t_0) + \varepsilon a^U(t) + \mathcal{O}(\varepsilon^2); \quad t \in (-\infty, t_0].$$

Now we examine what the global shape of these manifolds are for $\varepsilon > 0$ and if they interact. To do this we will follow an idea from Poincaré, Arnold, and Melnikov. To this end we define the perpendicular vector to f

$$f^\perp(x^0(0)) = \begin{pmatrix} -f_2(x^0(0)) \\ f_1(x^0(0)) \end{pmatrix}.$$

We would like to use this to measure the distance between $x_\varepsilon^U(t_0;t_0)$ and $x_\varepsilon^S(t_0;t_0)$. The outlook for this is shown in Fig. 7.7. Thus we have a signed distance function

$$d_\varepsilon(t_0) = \frac{\left\langle f^\perp(x^0(0)), x_\varepsilon^U(t_0;t_0) - x_\varepsilon^S(t_0;t_0)\right\rangle}{|f^\perp(x^0(0))|} = \varepsilon \frac{\left\langle f^\perp(x^0(0)), a^U(t_0) - a^S(t_0)\right\rangle}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon^2).$$

The numerator in the second equation is called the *Melnikov function*

$$M(t_0) = \left\langle f^\perp(x^0(0)), a^U(t_0) - a^S(t_0)\right\rangle.$$

In fact Melnikov proved an identity for this function.

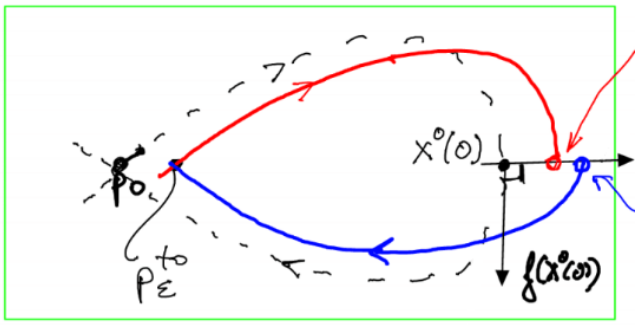


Figure 7.7: The Poincaré-Arnold-Melnikov idea. The dotted black line signifies Γ, the red dot $x_\varepsilon^U(t_0;t_0)$, the blue dot $x_\varepsilon^S(t_0;t_0)$, and the arrow pointing to the right $f^\perp$. The plane is Σ_{t_0}.

Theorem 7.2 (Melnikov)**.**

$$M(t_0) = \int_{-\infty}^{\infty} \langle f^\perp(x^0(t-t_0)), g(x^0(t-t_0),t)\rangle dt.$$

For the proof, see [Guckenheimer and Holmes, 1990].

Remark 7.3. The integral in Melnikov's theorem converges as $|g(x^0(t-t_0),t)|$ is globally bounded as $t \to \pm\infty$ and $|f^\perp| = |f|$ and

$$\lim_{t\to\pm\infty} \left|f(x^0(t-t_0))\right| = 0,$$

exponentially as p_0 is a hyperbolic fixed point.

Remark 7.4. In order to evaluate $M(t_0)$ we do not need to solve the perturbed ODE $\dot{x} = f(x) + \varepsilon g(x,t)$, instead we only need $x^0(t-t_0)$.

Now observe that

$$d_\varepsilon(t_0) = 0 \iff \varepsilon\frac{M(t_0)}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon^2) = 0 \overset{\varepsilon\neq 0}{\iff} \underbrace{\frac{M(t_0)}{|f^\perp(x^0(0))|} + \mathcal{O}(\varepsilon) = 0}_{F(t_0,\varepsilon)=0}.$$

Now we would like to know when we can find a solution $t_0(\varepsilon)$ such that $d_\varepsilon(t_0(\varepsilon)) = 0$ for $\varepsilon > 0$. To do this we use the Implicit Function Theorem. First assume that $F(\bar{t}_0, 0) = 0$, i.e. $M(\bar{t}_0) = 0$, and $\frac{\partial F}{\partial t_0}(\bar{t}_0), 0) \neq 0$, i.e. $\frac{\partial}{\partial t_0}M(\bar{t}_0) \neq 0$. If this condition is fulfilled the root is called *transverse*. Then there exists a unique $t_0(\varepsilon) = \bar{t}_0 + \mathcal{O}(\varepsilon)$ which solves $F(t_0(\varepsilon),\varepsilon) = 0$ for $\varepsilon \neq 0$ small enough. Also $t_0(\varepsilon)$ is smooth if $F(t_0,\cdot)$ is smooth.

A transverse zero for $M(t_0)$ implies that the Melnikov distance $d_\varepsilon(t_0(\varepsilon) = 0$, in turn implying that the stable and unstable manifolds intersect, $W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0}) \neq \emptyset$. Therefore we have an element $q \in W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0})$, for this q we also have that for every $n \in \mathbb{Z}$

$$P_\varepsilon^n(q) \in W^S(p_\varepsilon^{t_0}) \cap W^U(p_\varepsilon^{t_0}).$$

Therefore we have that for each iterate of the Poincaré map there is a unique point in both the stable and unstable manifolds, so these must intersect each other infinitely many times. This behavior is shown in Fig. 7.8. In the homoclinic tangle, we can see the accumulation of

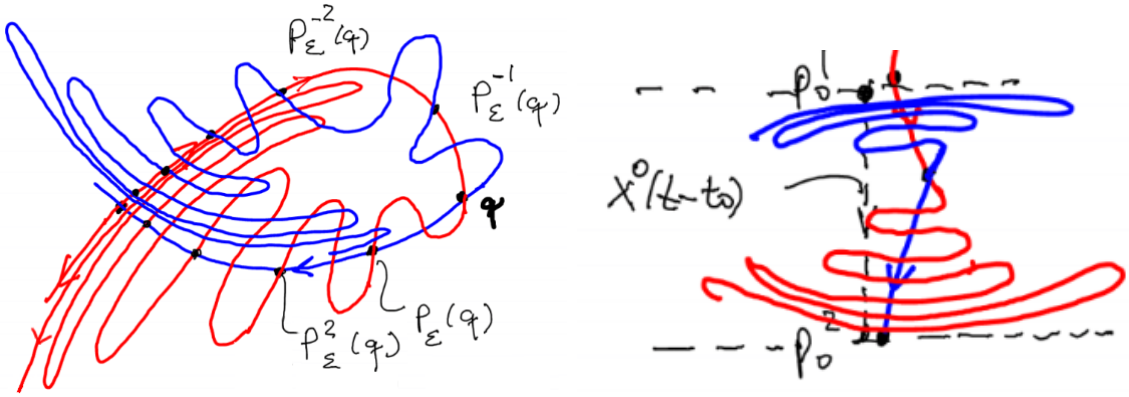


Figure 7.8: The infinite intersections of the stable and unstable manifolds for a transverse zero. Left: The homoclinic case, this is aptly called a *homoclinic tangle*. Right: The heteroclinic case.

trajectories that are caused by the accumulation of the iterates of the Poincaré map. For more see the Λ-*lemma* [Palis, 1967, Palis, 1969].

Remark 7.5. It is impossible for stable and unstable manifolds to intersect themselves. This is usually argued by stating that q cannot have two distinctive preimages under a diffeomorphism. However, this is not sufficient as a self intersection does not imply that two distinct preimages exist, for instance a looping manifold structure and suitable Poincaré map as in Fig. 7.9. A priori, this is possible to occur. In this case the existence of one loop actually implies the existence of infinitely many converging to p_ε. Thereby there must exist a loop in every arbitrarily small neighborhood of p_ε, this then contradicts the Hartman-Grobman Theorem as our system must be topologically equivalent to the linear saddle in a small enough neighborhood.

Remark 7.6. For every pair of intersections $P_\varepsilon^k(q)$ and $P_\varepsilon^{k+1}(q)$, there exists at least another intersection between the stable and unstable manifolds. This is due to the fact that P_ε is orientation preserving (i.e. $DP_\varepsilon(q)$ is an orientation preserving linear map). The implication here is illustrated in Fig. 7.10.

Figure 7.9: An example for a self-intersecting manifold without q having multiple preimages.

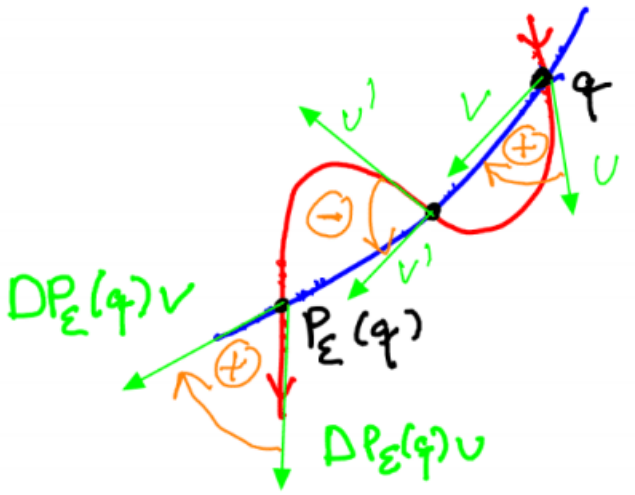


Figure 7.10: Another intersection of the stable and unstable manifolds exists between two sequential iterates of the Poincaré map due to the preservation of orientation under this map.

The orientation preserving nature of P_ε can be seen by using Liouville's Theorem

$$\det(DF_{t_0}^t(p)) = \exp\left(\int_{t_0}^t \operatorname{div}(f(x(s;t_0,p),s))ds\right),$$

for the flow map $F_{t_0}^t : x_0 \mapsto x(t;t_0,x_0)$ of the dynamical system $\dot{x} = f(x,t)$. In our case we find

$$\det(DP_\varepsilon(q)) = \exp\left(\int_{t_0}^t \operatorname{div}\,(f+\varepsilon g)|_{x_\varepsilon(t);\ x_\varepsilon(t_0)=q})ds\right) > 0$$

Therefore $DP_\varepsilon(q)$ is orientation preserving, in fact Poincaré maps in general are orientation preserving.

Example 7.3 (The forced-damped Duffing equation). Recall the dynamical system

$$\begin{cases} \dot{x} = y \\ \dot{y} = x - x^3 + \varepsilon(\gamma\cos(\omega t) - \delta y) \end{cases}; \quad |\varepsilon| \ll 1.$$

Here the forcing amplitude is given by γ and the linear damping coefficient by δ. We separate the right hand side into two parts

$$f(x,y) = \begin{pmatrix} y \\ x - x^3 \end{pmatrix}, \quad g(x,y) = \begin{pmatrix} 0 \\ \gamma\cos(\omega t) - \delta y \end{pmatrix}.$$

A Hamiltonian for this system is given by

$$H = \frac{1}{2}y^2 - \frac{1}{2}x^2 + \frac{1}{4}x^4 = E_0 = \text{const.}$$

Further the phase portrait for $\varepsilon = 0$ is known and shown in Fig. 7.11. We have that locally

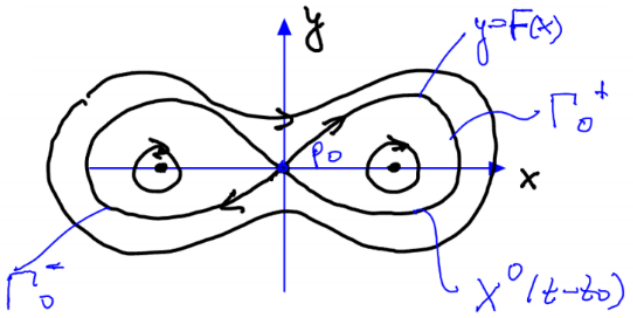


Figure 7.11: The phase portrait for the forced-damped Duffing oscillator for $\varepsilon = 0$.

$\dot{x} = F(x)$, thus we have a separable equation. Colloquially this can be seen as follows (the implication is not rigorous and should not be interpreted as usual)

$$\frac{dx}{dt} = F(x) \implies \int_0^x \frac{d\tilde{x}}{F(\tilde{x})} = \int_0^t d\tilde{t}.$$

Hence, on Γ_0^+, we have

$$x_0(t) = \begin{pmatrix} \sqrt{2}\text{sech}(t) \\ -\sqrt{2}\text{sech}(t)\tanh(t) \end{pmatrix}.$$

Now for $\varepsilon > 0$ we calculate the Melnikov function

$$\begin{aligned} M^+(t_0) &= \int_{-\infty}^{\infty} \langle f^\perp(x^0(t-t_0), g(x^0(t-t_0), t)\rangle dt \\ &= -\frac{4\delta}{3} + \sqrt{2}\gamma\pi\omega\text{sech}\left(\frac{\pi\omega}{2}\right)\sin(\omega t_0). \end{aligned}$$

Next, we ask if this is equal to zero and look for a transverse zero. To this end we define $R^0(\omega) = \frac{4\cosh\left(\frac{\pi\omega}{2}\right)}{3\sqrt{2}\pi\omega}$. Thus the zeros of the Melnikov function are given by

$$R^0(\omega) = \frac{\gamma}{\delta}\sin(\omega t_0).$$

I wasn't sure how he got the graph and what the marks meant, so I thought it wouldn't be clear to students, so I tried to do this part myself. I get the inverse result of him. A solution is then a transverse zero if the partial derivative with respect to t_0 is nontrivial

$$\sqrt{2}\gamma\pi\omega^2\text{sech}\left(\frac{\pi\omega}{2}\right)\cos(\omega t_0) \neq 0$$

The sech function is nonzero, thus we only need to know when $\cos(\omega t_0) = 0$, which occurs for $\omega t_0 = (2k+1)\pi$, or as a function of t_0 when $\frac{\omega}{\pi} = \frac{2k+1}{t_0}$. Thus at the odd integers scaled by $\frac{1}{t_0}$, we have nontransverse zeros, and all other roots of the Melnikov function are transverse. This is depicted in Fig. 7.12. We need to go over this part I think.

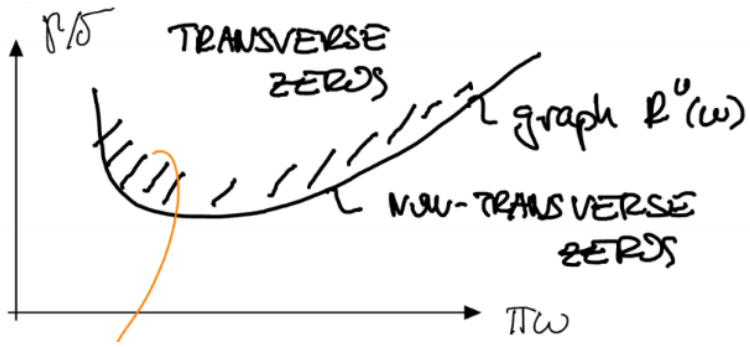


Figure 7.12: The yellow line showing that a homoclinic tangle exists for the forced-damped Duffing oscillator.

7.2 Dynamics near the homoclinic tangle & Smale's horseshoe map

We now continue and examine the dynamics near the homoclinic tangle. Consider the dynamical system

$$\dot{x} = f(x) + \varepsilon g(x,t); \quad x \in \mathbb{R}^2; \quad f, g \in \mathcal{C}^1; \quad g(x,t) = g(x,t+T).$$

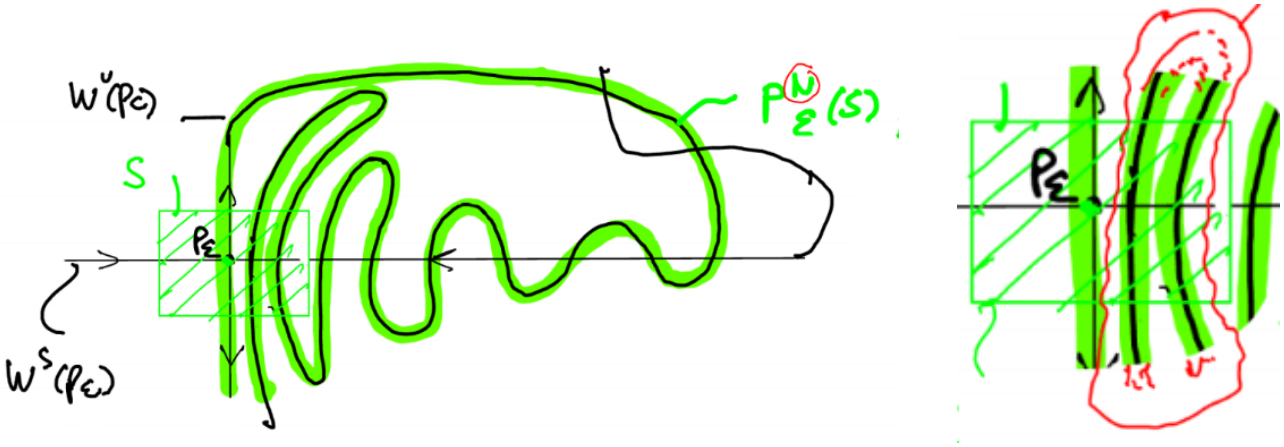


Figure 7.13: Smale's construction is found through an appropriate change of coordinates. Left: The full construction, the N circled in red is large enough such that $P_\varepsilon^N(S) \cap S = \emptyset$, where S is the green rectangle around p_ε. Right: Focus on the set S, the red lines designate a horseshoe-like structure.

After an appropriate change of coordinates we find the geometry called *Smale's construction*, which is as shown as in Fig. 7.13.

A model for the right panel of Fig. 7.13 is given by *Smale's Horseshoe map*. This is defined by first letting the set S be equal to $[0,1]^2 \in \mathbb{R}^2$, next the map itself is given by $f : S \to \mathbb{R}^2$. The map f will not be defined explicitly, and instead geometrically. The unit square is divided into three horizontal strips, the upper one being labelled H_2 and the lower H_1. The square is then stretched vertically, and then smoothly bent to a horseshoe such that the previously upper horizontal strip now forms a vertical strip on the right side of the box V_2. Similarly, the lower horizontal strip now forms a vertical strip on the left V_1. Now the unit square can also be divided into three vertical strips, on the left V_1, an area in the middle, and on the right V_2. What was previously part of the middle horizontal strip now forms the connecting curve between V_1 and V_2. Furthermore, these vertical strips are straight, without any curvature on their boundary. The horseshoe map is depicted in Fig. 7.14. This map models the Poincaré map as seen before.

Using this construction, we would like to know if any initial conditions stay in S for long times (after many iterations of the Horseshoe map), despite the overall instability coming from P_ε. To further study this we introduce a couple definitions.

Definition 7.2. The set of all points which stay in S under k backwards iterations is given by V^k.

Definition 7.3. For a map $f : X \to Y$, we denote the *preimage* of a set $U \subset Y$ (or point

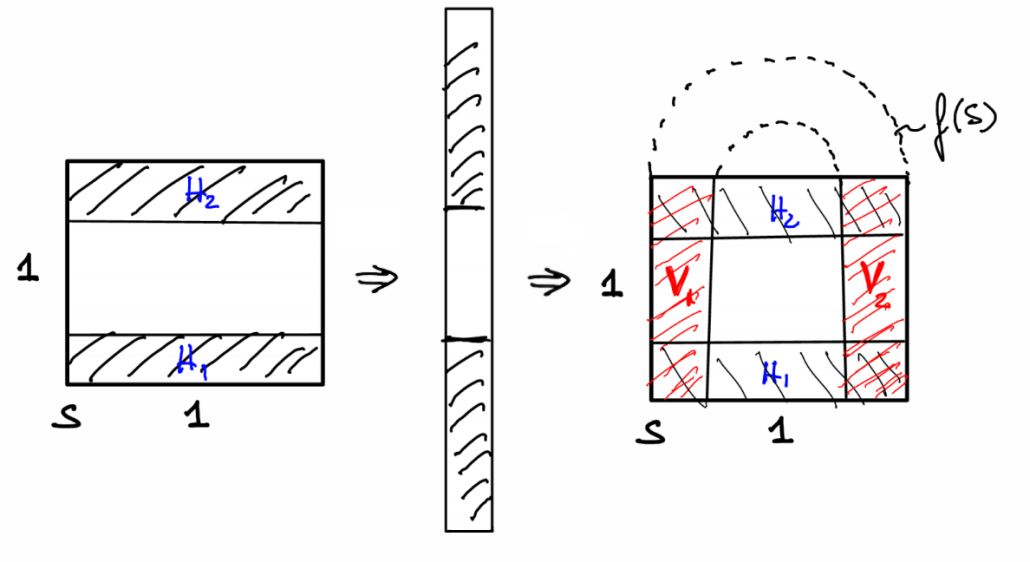


Figure 7.14: The two geometric steps of the Smale map, first a vertical stretch and then a fold into the horseshoe.

$p \in Y$) as

$$\boxed{f^{-1}(U) = \{x \in X : \ f(x) \in U\}\,.}$$

Under 1 backward iteration the set V^1 is given by the closed and nonempty set

$$V^1 = f\left(S \cap f^{-1}(S)\right) = f(S) \cap S = V_1 \cup V_2.$$

For 2 backward iterations we have

$$V^2 = f^2\left(S \cap f^{-1}(S) \cap f^{-2}(S)\right) = f^2(S) \cap f(S) \cap S = V_{11} \cup V_{12} \cup V_{22} \cup V_{21} = \bigcup_{i_k \in \{1,2\}} V_{i_1 i_2}.$$

The sets V_{ij} are defined in Fig. 7.15, the indices for each component are given by their orbit, those which after one iteration are in V_i ($i \in \{1,2\}$) have the first index i, those which after two iterations are in V_j have the second index j. Also the set V^2 is closed, nonempty, contained in V^1, and has 2^2 components.

This continues for any number of backward iterates k

$$V^k = f^k\left(S \cap f^{-1}(S) \cap \ldots \cap f^k(S)\right) = f^k(S) \cap \ldots \cap f^1(S) \cap S = \bigcup_{i_j \in \{1,2\}} V_{i_1 \ldots i_k}.$$

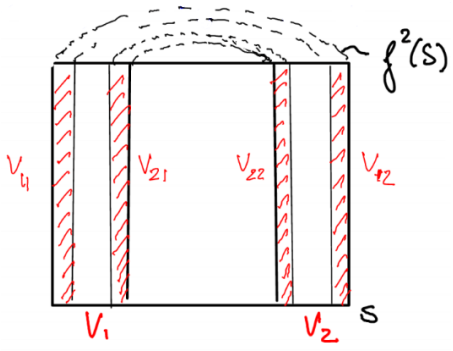


Figure 7.15: The four components of V^2 with their index labels.

This set has 2^k disjoint components (vertical strips) and is closed, nonempty, and contained in V^{k-1}. Therefore we have an infinite, nested sequence of nonempty and closed sets, therefore by Cantor's theorem we have

$$\boxed{V^\infty = \bigcap_{i=1}^{\infty} V^i \neq \emptyset.}$$

Remark 7.7. Cantor's theorem comes from the famous Cantor Dust which arises by first removing the open middle third of the unit interval in $\mathbb{R}$, and then iteratively removing the open middle third of each remaining interval. The Cantor Dust is a nonempty (in fact uncountably infinite) and disconnected set of points. For more on Cantor's theorem and Cantor's Dust, see any book on measure theory.

Definition 7.4. A set U is called *perfect* if it is closed and has no isolated points, i.e. given any point $x \in U$ and any distance ε we can find a distinct element $y \in U$ such that $\|x - y\| < \varepsilon$.

Remark 7.8. This definition uses that the topology on U was induced by a metric, which may not be the case in general, but is not relevant to the content here.

In our case V^∞ is in fact a Cantor set, i.e. it is a closed, nonempty set which is totally disconnected (the largest connected component is a line), and perfect. this set isn't totally disconnected, totally disconnected means that every connected subset is a singleton, here every connected set is a line. The projection of V^∞ onto the unit interval in $\mathbb{R}$ is totally disconnected. We would need to define the quotient topology $S/\sim$ for $[x] = \left\{ x + \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \ \lambda \in [0,1] \right\}$ in order for each line to be a single element. I think this statement should apply to Λ, we should make that change and move this and the definition of perfect to after the definition of Λ.

Definition 7.5. Analogous to before we define the set of all points which stay in S under k forwards iterations is given by H^k. For each k the set H^k is a nonempty set of 2^k closed, disjoint, horizontal strips, forming a nested sequence.

By similar argumentation as before we have

$$\boxed{H^\infty = \bigcap_{k=1}^{\infty} H^k \neq \emptyset.}$$

Definition 7.6. The set of all points which stay in S for all forward and backward iterates of the Smale Horseshoe map is given by

$$\boxed{\Lambda = H^\infty \cap V^\infty.}$$

The geometry of this set is shown in Fig. 7.16.

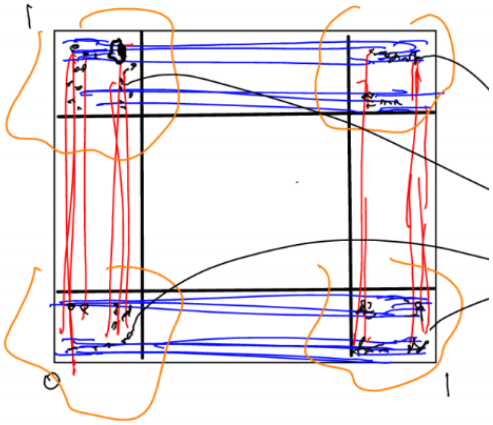


Figure 7.16: The geometry of the set Λ, with the four regions containing points circled in orange. The red vertical lines represent the set V^∞ and the blue horizontal lines the set H^∞.

Remark 7.9. The set Λ is a Cantor set of points.

Returning to our case study of a Poincaré map with a homoclinic tangle, the points in Λ keep coming back forever to the unstable periodic orbit as shown in Fig. 7.17.

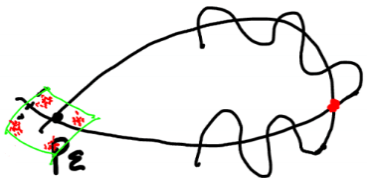


Figure 7.17: The homoclinic tangle with the set S outlined in green and the elements of Λ being denoted by the red dots within S.

7.3 Dynamics of Smale's Horseshoe map on the invariant set Λ

We label any point $p \in \Lambda$ with an infinitely long binary sequence, which reflects the itinerary of p in Λ under iterations of f. For a given p, if we have

$$\ldots,\ f^{-2}(p) \in V_{s_{-3}},\ f^{-1}(p) \in V_{s_{-2}},\ p \in V_{s_{-1}},\ p \in H_{s_0},\ f(p) \in H_{s_1},\ f^2(p) \in H_{s_2},\ \ldots$$

for $s_i \in \{1, 2\}$ then we would have the itinerary

$$\boxed{s = \ldots s_{-3}s_{-2}s_{-1}\bullet s_0 s_1 s_2 \ldots .}$$

This is a unique symbolic encoding for <u>all</u> points in Λ. Therefore Λ is diffeomorphic to a space of doubly infinite sequences of two symbols. Such a space is just the space of outcomes for an infinite coin-tossing experiment. This equivalence of Λ to the doubly infinite sequence of two symbols leads us to *symbolic dynamics*: the analogue of the horseshoe map on the symbol space Σ.

In order to study these symbolic dynamics we need to know what the symbolic encoding of $f(p)$ is. For the analogue $\ldots s_{-2}s_{-1}\bullet s_0 s_1 s_2 \ldots = s \in \Sigma$ of $p \in \Lambda$, we can read off that $f(p) \in H_{s_1}$. Continuing in this fashion, we can see that $f(f(p)) = f^2(p) \in H_{s_2}$ and so on, therefore $f(p) \in H^\infty$. By recalling that $f(H_{s_0}) = V_{s_0}$ we can see that $f(p) \in V_{s_0}$. Applying the definition of f we find that $f^{-1}(f(p)) = p \in V_{s_{-1}}$. Therefore $f(p) \in V^\infty$ and thereby $f(p) \in \Lambda$.

From our examination of the forward iterates of $f(p)$ we know that the symbolic encoding of $f(p)$ to the right of the dot is

$$\bullet s_1 s_2 s_3 \ldots .$$

Similarly, from our examination of the backwards iterates of $f(p)$ we know that the symbolic encoded of $f(p)$ to the left of the dot is

$$\ldots s_{-2}s_{-1}s_0\bullet .$$

Putting these together we get the symbolic encoding for $f(p)$ is

$$\ldots s_{-2}s_{-1}s_0{\bullet}s_1s_2s_3\ldots \in \Sigma.$$

So the encoding of $f(p)$ in the symbolic space is just a shift to the left on the encoding of p.

Definition 7.7. We define the following three maps:

(i) The *symbolic encoding map*

$$\boxed{\phi : \Lambda \to \Sigma; \quad p \mapsto s = \phi(p);}$$

This map is a homeomorphism (i.e. it is continuous and has a continuous inverse).

(ii) The *Smale Horseshoe map restricted to* Λ

$$\boxed{\tilde{f} : \Lambda \to \Lambda; \quad p \mapsto f(p);}$$

(iii) The *Bernoulli shift map*

$$\boxed{\sigma : \Sigma \to \Sigma; \quad \ldots s_{-2}s_{-1}{\bullet}s_0s_1s_2\ldots \mapsto \ldots s_{-2}s_{-1}s_0{\bullet}s_1s_2s_3\ldots.}$$

We have that the following diagram commutes

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ {\scriptstyle\phi}\uparrow & & \uparrow{\scriptstyle\phi} \\ \Lambda & \xrightarrow{\tilde{f}} & \Lambda \end{array}.$$

Therefore $\tilde{f} = f|_\Lambda$ is topologically conjugate to a Bernoulli shift on Σ

$$\tilde{f} = \phi^{-1} \circ \sigma \circ \phi.$$

Hence, orbits of f are taken to those of σ, with their orientation preserved.

Remark 7.10. To denote an infintely repeating string, the bar notation will be used i.e.

$$\overline{1} = 111\ldots, \quad \overline{123} = 123123\ldots.$$

It is possible to classify all of the orbits of σ in Σ, the classes and their cardinality are as follows:

(i) **Fixed points** There are exactly two fixed points of σ, namely $\overline{1}.\overline{1}$ and $\overline{2}.\overline{2}$;

(ii) **Period-two orbits** A single such orbit exists, that of $\overline{12}.\overline{12}$, as $\sigma^2(\overline{12}.\overline{12}) = \sigma(\overline{21}.\overline{21} = \overline{12}.\overline{12}$;

(iii) **Period-three orbits** There exists two such orbits, namely $\overline{112}.\overline{112} \to \overline{121}.\overline{121} \to \overline{211}.\overline{211}$ and $\overline{221}.\overline{221} \to \overline{212}.\overline{212} \to \overline{122}.\overline{122}$.

(iv) **Period-k orbits** For every natural number k, there exists

$$\boxed{N(k) = \frac{1}{k}\left(2^k - \sum_{i|k} iN(i)\right)}$$

orbits of period exactly k, the notation $i|k$ signifies integers i which are divisors of k.

There exists a countable infinity (i.e. the elements can be organized uniquely into an infinite countable sequence) of periodic orbits of arbitrarily high period. The countable sequence ordering can be done by listing each orbit in the row corresponding to its minimal period. The specifig ordering within rows is based on the magnitude of the k-long string representing the orbit, considered as a decimal number, i.e. $\overline{112} \prec \overline{122}$, for each orbit the representative element is the one with the least magnitude as considered above. Thus $\overline{112}.\overline{112}$ would go in the fourth row as it has the minimal period 3 (not third row, as the fixed points have period length 0 and occupy the first row), and in the leftmost position as $\overline{112} \prec \overline{122}$. Each row has a finite number of elements, so ordering each row is not a problem. Then the complete order is done by starting with 0 for the first element in the first row (the fixed point $\overline{1}.\overline{1}$), then reading from left to right before continuing to the next row. This list is well defined and countably infinite, containing all of the orbits of σ.

Before further study, it is important to introducte a notion of distance on the space Σ.

Definition 7.8. The distance on the space Σ between two elements s and s' is given by

$$d(s, s') = \sum_{i=-\infty}^{\infty} \frac{|s_i - s'_i|}{2^{|i|}}.$$

Therefore two symbols are close if they agree on large central blocks.

Remark 7.11. The function d satisfies all the axioms of a distance and tuple (Σ, d) forms a complete metric space.

Proposition 7.12 (The periodic orbits are dense)**.** *The set of all periodic orbits is dense in* Σ.

Proof. this is not included in his notes, he just says 'Show it', but it is quite simple so I am adding here for now. Given an arbitrary element $s \in \Sigma$ and $k \in \mathbb{N}$, we would like to find a 2^{-k} periodic approximation of s. To do this define the two strings

$$S_- = s_{-(k+1)} \dots s_{-1}, \quad S_+ = s_0 \dots s_{k+1}.$$

For any two strings (of infinite length) L and R the element

$$s' = LS_-.S_+R$$

is 2^{-k} close to s. This is because they can only disagree at indices with modulus strictly greater than $k+1$, therefore

$$d(s', s) \leq \sum_{j=-(k+2)}^{-\infty} \frac{1}{2^{-|j|}} + \sum_{j=k+2}^{\infty} \frac{1}{2^{-|j|}} = 2 \cdot 2^{-(k+1)} = 2^{-k}.$$

Now it is clear that the periodic element

$$s^* = \overline{S_+S_-}.\overline{S_+S_-} = L'S_-.S_+R'$$

is a 2^{-k} approximation for s. As s and k were arbitrary, this completes the proof. □

Consider a k-periodic point $s = \overline{a}.\overline{a}$, where a is a k-long binary sequence. Let $s' = \overline{a}a'.\overline{a}$, clearly $s \neq s'$, therefore $d(s, s') \neq 0$, yet when we take the limit we find

$$\lim_{n \to \pm\infty} d(\sigma^n(s), \sigma^n(s')) = 0.$$

Thus the underlying periodic orbit is unstable. Furthermore, s has a countable infinity of homoclinic orbits. As the periodic orbit was arbitrary, these statements hold for all periodic orbits, hence all periodic orbits are unstable and have a countable infinity of homoclinic orbits.

Remark 7.13. Also note that any two periodic orbits (of arbitrary period) are connected by a countable infinity of heteroclinic orbits.

Together, there exists a countable infinity of periodic and asymptotically periodic orbits, a sketch of this behavior is shown in Fig. 7.18.

In fact there also exists a countable infinity of non-periodic (not even asymptotically periodic) orbits in Σ. To show this, we will show that Σ is uncountable, i.e. that it cannot be arranged into a well-defined, countable sequence.

Proposition 7.14. *The set Σ is uncountable.*

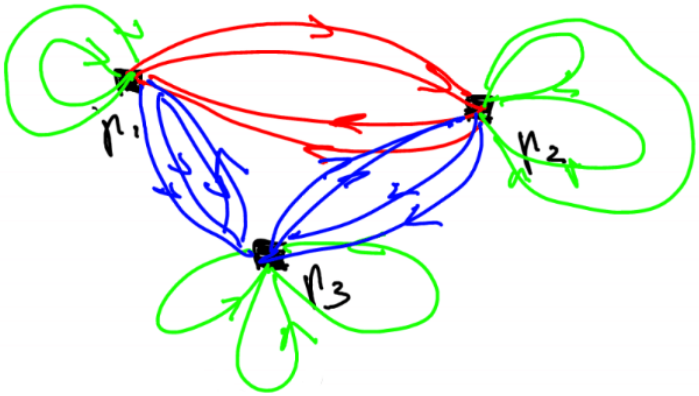


Figure 7.18: A depiction of the (infinitely many) homo- and heteroclinic orbits which connect periodic orbits of σ.

Proof. It is enough to show that the set $\tilde{\Sigma}$ defined by

$$\tilde{\Sigma} = \{.s_1 s_2 s_3 \ldots : \ s_i \in \{0, 1\}\}$$

is uncountable. This suffices as we can injectively embed $\tilde{\Sigma}$ into Σ via the map $.s_1 s_2 \ldots \mapsto \overline{1}.(s_1 + 1)(s_2 + 1)\ldots$.

Every binary expansion for all real numbers in the interval $[0, 1]$ is contained within $\tilde{\Sigma}$, therefore the cardinality of $\tilde{\Sigma}$ is at least as large as that of $[0, 1]$ which is uncountable. In fact, the binary expansion is a bijection, and these two sets have the exact same cardinality. □

Definition 7.9. The binary expansion of a number $x \in [0, 1]$ is the sequence $(b_i)_{i=1}^{\infty}$ of 1s and 0s such that

$$x = \sum_{i=1}^{\infty} \frac{b_i}{2^i}.$$

To calculate this sequence define the map $g(x) = 2x \mod 1$. Then generate the sequence $(x_i)_{i=1}^{\infty} = (x, g(x), g^2(x), \ldots)$. Finally, set each binary coefficient as

$$b_i = \begin{cases} 0 & 2x_i < 1 \\ 1 & 2x_i \geq 1 \end{cases}.$$

Example 7.4 (Binary expansion). Let $x = \frac{1}{5}$, we find the iterates of g

$$\begin{aligned}
x &= \frac{1}{5} && & 2 \cdot \frac{1}{5} < 1 &\implies x_1 = 0 \\
g\left(\frac{1}{5}\right) &= \frac{2}{5} && & 2 \cdot \frac{2}{5} < 1 &\implies x_2 = 0 \\
g^2\left(\frac{1}{5}\right) &= g\left(\frac{2}{5}\right) = \frac{4}{5} && & 2 \cdot \frac{4}{5} > 1 &\implies x_3 = 1 \\
g^3\left(\frac{1}{5}\right) &= g\left(\frac{4}{5}\right) = \frac{8}{5} \mod 1 = \frac{3}{5} && & 2 \cdot \frac{3}{5} > 1 &\implies x_4 = 1 \\
g^4\left(\frac{1}{5}\right) &= g\left(\frac{3}{5}\right) = \frac{6}{5} \mod 1 = \frac{1}{5} && & 2 \cdot \frac{1}{5} < 1 &\implies x_5 = 0.
\end{aligned}$$

Since $g^4\left(\frac{1}{5}\right) = \frac{1}{5}$, we have reached a repeating loop (since our seed value was $\frac{1}{5}$), therefore we have that

$$\frac{1}{5} = 0 \cdot \frac{1}{2} + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{16} + \ldots = .\overline{0011}.$$

Remark 7.15. Recall that for a metric space (X, d) when we write $\mathcal{B}(x, r)$ we refer to the open ball of radius r around x, i.e.

$$\mathcal{B}(x, r) = \{y \in X : \ d(x, y) < r\}.$$

Proposition 7.16. *The map σ has a dense orbit in Σ, i.e. an orbit which approaches any point in Σ arbitrarily closely.*

This is listed as HW, but when I took the course the HW used this proposition, we didn't have to prove it. So I am giving a small proof here... Just finished the proof, bit longer than I expected.

Proof. Without loss of generality, assume that $\delta < 1$. Next, note that the maximal distance between two elements occurs when they disagree at every entry, so they have the distance

$$\sum_{i=-\infty}^{\infty} \frac{1}{2^{|i|}} = 2.$$

Therefore, the entire space is included in the ball $\mathcal{B}\left(\overline{1}.\overline{1}, 2\right)$. Now to show that for any given distance δ, we want to find an element s^* such that its orbit gets δ close to every element of Σ. Let

$$k = \min\left\{j \in \mathbb{N} : \ 2^{-j} \leq \delta\right\},$$

i.e. the first index in the binary expansion of δ which is not 0. Now we would like to find a finite $2^{-(k+1)}$-covering of Σ, i.e. a set of elements s_i^k such that $\bigcup_i \mathcal{B}\left(s_i^k, 2^{-(k+1)}\right)$. We start out with the covering $\bigcup_{j=1}^{\infty} \mathcal{B}\left(q_j, 2^{-(k+1)}\right)$, with the q_j being an enumeration of the periodic orbits. As we previously noted, Σ is bounded as it can be contained in a finite ball, furthermore Σ is also closed (the intersection of two closed sets is always closed due to the de Morgan identity), thus Σ is compact and a finite subcovering exists. Hence we can deduce that only finitely many q_i are needed, call these $\left(s_i^k\right)_{i=1}^N$, and let the string representation of each s_i^k be given by $\overline{S_i^k}.\overline{S_i^k}$.

If we can construct s_k^* such that its orbit passes through each of the $\mathcal{B}\left(s_i^k, 2^{-(k+1)}\right)$, then we know that the orbit of s_k^* gets within distance 2^{-k} of every element in Σ, by the triangle inequality.

For any given s_i^k the element

$$L\underbrace{S_i^k \dots S_i^k}_{(k+3) \text{ times}}.\underbrace{S_i^k \dots S_i^k}_{(k+4) \text{ times}} R$$

is within distance $2^{-(k+1)}$ of s_i^k, for L and R arbirary. This is because the string S_i^k is of at least length 1, so the two strings can disagree only at elements with index modulus strictly greater than $k+2$ (we need the one more copy of S_i^k on the right side as the 0th index is to the right of the .), therefore the distance between them is upper bounded by

$$\sum_{j=-(k+3)}^{-\infty} \frac{1}{2^{-|j|}} + \sum_{j=k+3}^{\infty} \frac{1}{2^{-|j|}} = 2 \cdot 2^{-(k+2)} = 2^{-(k+1)}.$$

Therefore, using $2k+7$ copies of S_i^k we are able to approximate s_i^k within distance $2^{-(k+1)}$. We now construct the element

$$s_k^* = L.\underbrace{S_1^k \dots S_1^k}_{2k+7 \text{ times}} \dots \underbrace{S_N^k \dots S_N^k}_{2k+7 \text{ times}} R$$

for L and R arbitray. We now can see that for every $i \in \{1 \dots N\}$, the element

$$L'\underbrace{S_i^k \dots S_i^k}_{(k+3) \text{ times}}.\underbrace{S_i^k \dots S_i^k}_{(k+4) \text{ times}} R',$$

is contained within the orbit of s_k^* for L' and R' chosen accordingly. Therefore we have now constructed the the desired s_k^*. Note here that this construction is well defined as each S_i^k is of finite length. This construction can be done for each k: start from the infinite cover using periodic orbits, then construct a finite subcover using the compactness of Σ, finally construct the element s_k^* by concatenating the $2^{-(k+1)}$ approximations of the elements which consitute the finite subcover.

To construct s^*, concatenate each of the strictly defined blocks (the block the the right of . and left R) of s_k^* for each $k \in \mathbb{N}$. For any n, we know that we can get within distance 2^{-n} distance of each element in Σ, therefore we have constructed a dense orbit. This construction is well definied as the strictly defined block of any given s_k^* is always of finite length. Thus we have constructed the element, whose orbit is dense in Σ. □

Proposition 7.17. *The map σ is topologically transitive, i.e. for any two open sets A and B in Σ, there exists an $N \in \mathbb{N}$ such that*

$$\sigma^N(A) \cap B \neq \emptyset.$$

This is called the mixing property*.*

Proof. This is a consequence of the existence of a dense orbit. □

Proposition 7.18 (Sensitive dependece on initial conditions)**.** *Regardless of how close two distinct initial conditions s and s' are chosen, they will not remain close. There exists a $\Delta > 0$ such that for every $s \in \Sigma$ and any $\delta > 0$, there exists an $s' \in \mathcal{B}(s, \delta)$ such that for an N large enough the following holds*

$$d(\sigma^N(s), \sigma^N(s')) > \Delta.$$

This is called the uniform instability of all orbits*. This fact is depicted in Fig. 7.19.*

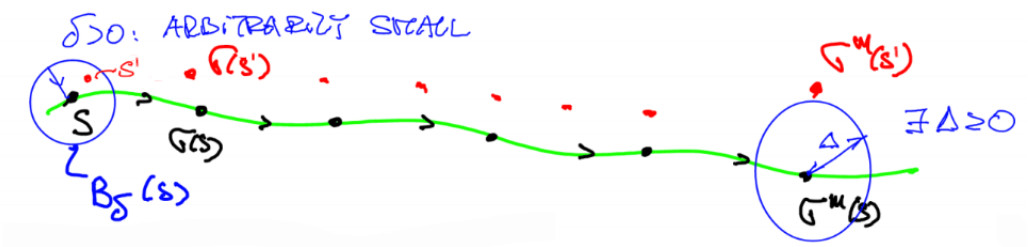


Figure 7.19: The uniform instability of orbits for the map σ.

Remark 7.19. This by itself implies no complexity, the linear saddle

$$\begin{cases} \dot{x} = \lambda x \\ \dot{y} = -\mu y, \end{cases}$$

also has this property.

Remark 7.20. In the case of this map, every $s' \in \mathcal{B}(s, \delta)$ has an N large enough such that the inequality holds. This is my addition.

These observations lead us to the following definition of a chaotic map.

Definition 7.10 (Chaotic map)**.** Let $M : C \to C$ be a $\mathcal{C}^0$ map on a compact metric space C. We call M *chaotic* if

(i) M is topologially transitive (the mixing property);

(ii) M has sensitive dependence on initial conditions (a lack of reproduceability);

(iii) M has a dense set of periodic orbits in C.

Therefore σ is a chaotic map.

Example 7.5 (Chaos in a 1-dimensional model). Consider the logistic map for resource-limited discrete population growth dynamics

$$f : x \mapsto ax(1 - x); a \in \mathbb{R}^+.$$

We will now continue with $a = 4$. This is not an ODE, and we assume there are so small perturbations of known structures. The cobweb diagram is illustrated in Fig. 7.20. We can

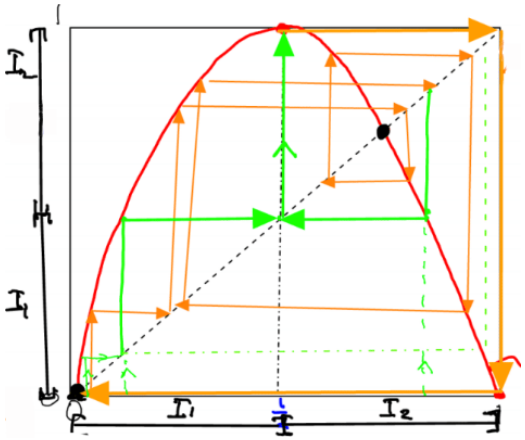


Figure 7.20: The cobweb diagram for the logisitic map with $a = 4$. The sets I_1 and I_2 are defined as $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$.

see that for our choice of a, f maps $[0, 1]$ to $[0, 1]$ and is not invertible. In fact each preimage of $\frac{1}{2}$ reaches 0 in a finite number of iterations. Furthermore, the preimages of $\frac{1}{2}$ converge to either $x = 0$ or $\frac{3}{4}$, of which there are a countable infinity of each. This implies that there is

a countable infinity of homo- and heteroclinic orbits for $x = 0$ and $\frac{3}{4}$. From Fig. 7.20, the dynamics appear complicated, with no stable fixed points (can be seen using the linearization).

To better understand the dynamics we can encode the itinerary of a point $x \in I = [0, 1]$ by if x is in I_1 or I_2. We get a one-sided itinerary, because f is not invertible. For example if $x \in I_{s_0}$, $f(x) \in I_{s_1}$, $f^2(x) \in I_{s_2}$ then we would encode x as $.s_0 s_1 s_2 \ldots$ for $s_i \in \{1, 2\}$.

Note here that this symbolic encoding is not well defined for every point in I, as all preimages of $x = \frac{1}{2}$ have multiple encodings, since their iterates eventually reach $I_1 \cap I_2 = \frac{1}{2}$. To address this define $\Psi : \Sigma \to I$ for Σ the set of semi-infinite binary symbols, by

$$\Psi : .s_0 s_1 s_2 \mapsto x = I_{s_0} \cap f^{-1}(I_{s_1}) \cap f^{-2}(I_{s_2}) \cap \ldots.$$

We now have to ask if this is well definied, i.e. if it yields a unique x value. Note here that for any subset $J \subset I$ which is closed and connected, we have that $I_{s_i} \cap f^{-1}(J)$ is closed, nonempty, and connected. This is due to the symmetry around $x = \frac{1}{2}$ and is shown visually in Fig. 7.21.

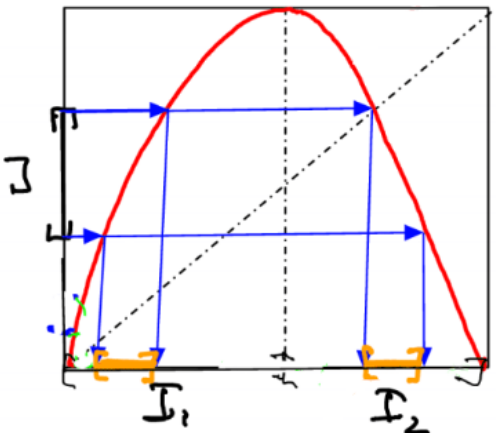


Figure 7.21: The preimage of a subset $J \subset I$, with its closed, nonempty, and connected intersections with I_1 and I_2.

Remark 7.21. For any two sets A and B we have that

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B).$$

Further observe that for $I_{s_1 s_1} = I_{s_0} \cap f^{-1}(I_{s_1})$ is closed, nonempty, and connected for all s_0 and s_1 by the previous note. Continuing this we find that

$$I_{s_0 s_1 s_2} = I_{s_0} \cap f^{-1}(I_{s_1}) \cap f^{-2}(I_{s_2}) = I_{s_0} \cap f^{-1}\left(I_{s_1} \cap f^{-1}(I_{s_2})\right)$$

is closed, nonempty, connected, and contained in $I_{s_0 s_1}$. This line of argument can be continued to show for any $n \in \mathbb{N}$ the set $I_{s_0 \ldots s_n}$ is closed, nonempty, connected, and contained in $I_{s_0 \ldots s_{n-1}}$.

Thus we have a nested sequence of closed, nonempty, and connected sets. By Cantor's theorem a nonempty limit exists, which in this case is also connected. It is also possible to show that $\lim_{n\to\infty} |I_{s_0\dots s_n}| = 0$, where $|\cdot|$ denotes the length of a set. Now we can can conclude that the limit set contains a single point x, therefore as $n \to \infty$ the function $\Psi(s) = x$ is well-defined.

Additionally with the metric

$$d(s, s') = \sum_{i=0}^{\infty} \frac{|s_i - s_i'|}{2^i}$$

the function Ψ is continuous.

By definition, a single iteration of the map f is equivalent to a shift to the left on the corresponding symbols in Σ, i.e.

$$f : x = .s_0 s_1 s_2 \dots \mapsto .s_1 s_2 \dots = f(x).$$

Therefore we have the following commuting diagram

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle \Psi} \\ I & \xrightarrow{f} & I \end{array}.$$

In comparison to the previous example, Ψ is not invertible (although it is well-defined and continuous). The invertibility of Ψ only fails on a countable set of points $\left\{f^{-n}\left(\frac{1}{2}\right)\right\}$. Therefore we have that the logistic map is chaotic on $\tilde{I}$.

The previous example leads us to the definition for maps such as the logistic map

Definition 7.11 (Topological semi-conjugacy)**.** A map f is *topologically semi-conjugate* to σ if there exists a surjection Ψ such that

$$\boxed{f \circ \Psi = \Psi \circ \sigma.}$$

Chapter 8

Generalization of chaotic dynamics

We begin by defining a general symbol space.

Definition 8.1. Denote by Σ^N the symbol space defined as the space of doubly infinite sequences of N symbols (for $N \in \mathbb{N}$), i.e.

$$\boxed{\Sigma^N = \{s = \ldots s_{-2}s_{-1}.s_0s_1s_2\ldots : \ s_i \in \{0, 1, \ldots, N-1\}\}.}$$

We may constrain this symbol space by constraining which sequences are admissible. This can be done by using a transition matrix $A \in \mathbb{R}^{N\times N}$ with entries $A_{ij} \in \{0, 1\}$. Then the constrained symbol space is given by

$$\Sigma_A^N = \left\{s \in \Sigma^N : \ A_{s_i s_{i+1}} \neq 0, \forall i\right\}.$$

In the constrained space, consecutive symbols s_i and s_{i+1} must correspond to a 1 in the transition matrix A. In other words if the symbol $s_i = k$ then the symbol s_{i+1} must be equal to j such that $A_{kj} = 1$.

Example 8.1 (Constrained symbol space). To demonstrate how a transition matrix can constrain the symbol space, lets start by examining Σ_A^2 for

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

Be examining the entries of the transition matrix A we find which strings are admissible

$$\begin{aligned} A_{11} = 1 &\implies \ldots 00 \ldots \quad \text{admissible} \\ A_{12} = 0 &\implies \ldots 01 \ldots \quad \text{inadmissible} \\ A_{21} = 0 &\implies \ldots 10 \ldots \quad \text{inadmissible} \\ A_{22} = 1 &\implies \ldots 00 \ldots \quad \text{admissible.} \end{aligned}$$

Therefore the set Σ_A^2 is comprised only of two elements $\{\overline{0}.\overline{0}, \overline{1}.\overline{1}\}$. As this space is very simple and only contains fixed points of σ, no chaos can occur.

Definition 8.2. The map

$$\sigma : \Sigma_A^N \to \Sigma_A^N; \quad \ldots s_{-1}.s_0 \ldots \mapsto \ldots s_{-1}s_0.s_1s_2 \ldots,$$

is called a *Bernoulli subshift of finite type* (i.e. $N < \infty$) with transition matrix A.

Before more examination of chaos on such constrained spaces, we need to recall a definition from linear algebra.

Definition 8.3. A matrix A is called *irreducible* if there exists a $k \in \mathbb{N}$ such that $\left(A^k\right)_{ij} \neq 0$ for every $i, j \in \{1, \ldots, N\}$.

Theorem 8.1. *Assume that A is irreducible, then*

(i) Σ_A^N is a Cantor set; and

(ii) $\sigma : \Sigma_A^N \to \Sigma_A^N$ is a chaotic map.

Note that in the previous example with $A = I_2$, the transition is not irreducible, which is why chaos is not garunteed.

The role of irreducibility is critical for the theorem to work. Say for instance that the required k such that each entry of A^k is nonzero is $k = 3$. Then we know that $(A^3)_{ij} = \sum_{l,m=1}^{3} a_{il}a_{lm}a_{mj} \neq 0$ for any pair i, j. Therefore there exist l, m such that a_{il}, a_{lm}, and a_{mj} which are not equal to 0. Hence there exists an admissible ultimate transition from state i to state j for all pairs i, j, if the symbol sequence is viewed as an itinerary. Furthermore, different irreducible transition matrices generate different types of chaos.

Example 8.2 (Chaos in the Newton-Raphson iteration). Newton-Raphson iteration is a process to find the roots (zeros) of a nonlinear function f. The iteration is as follows

(i) Start with the seed point x_0;

(ii) Define $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = N(x_0)$;

(iii) Repeat (ii) until $x_{n+1} = N(x_n)$.

In practice this approach is often truncated when $x_{n+1} \approx N(x_n)$. The roots of f coincide with the fixed points of the 1-dimensional dynamical system $x \mapsto N(x)$. If such a fixed point (x^*) is stable and x_0 is in its domain of attraction, then $\lim_{i\to\infty} N^i(x_0) = x^*$. The iteration process is sketched in Fig. 8.1.

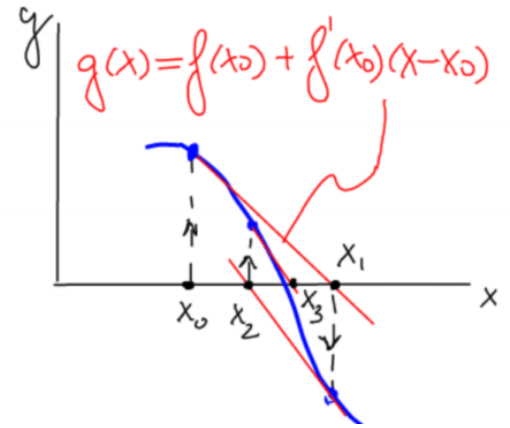


Figure 8.1: An illustration of the iteration of the Newton-Raphson process on the function f (blue). The red designates the function g, when given the point x_0 the next iterate x_1 fufills the property $g(x_1) = 0$.

Note that if $N(x^*) = x^*$ then we have

$$N'(x^*) = 1 - \frac{f'(x^*)^2 - f(x^*)f''(x_0)}{f'(x^*)^2} = 0.$$

The latter equality is due to $f(x^*) = 0$. This implies that $|N'(x^*)| < 1$, i.e. the fixed point is stable, and additionally $N'(x^*) = 0$, i.e. x^* is in fact super stable.

Specifically, once we are near the optimum x^*, we have

$$N(x_n) = \underbrace{N(x^*)}_{=x^*} + \underbrace{N'(x^*)}_{=0}(x_n - x^*) + \frac{1}{2}N''(x^*)(x_n - x^*)^2 + \mathcal{O}\left(|x_n - x^*|^3\right).$$

This implies we have fast quadratic convergence, as

$$|N(x_n) - N(x^*)| \approx \frac{1}{2}|N''(x^*)||x_n - x^*|^2.$$

Despite the fast convergence, erratic iterations may result if the initial point is not close enough to the root [Saari and Urenko, 1984]. In particular, consider a 4th order polynomial

$$f(x) = (x - x_0)(x - x_1)(x - x_2)(x - x_3); \quad x_0 < x_1 < x_2 < x_3.$$

This can be observed to lead to chaotic Newton-Raphson iterations. The proof will be sketched for this chaotic behaviour.

Retain the 4th order polynomial f from above. We then have $N(x_i) = x_i$ for $i = 0, 1, 2, 3$ as these are fixed points of f. Note then that we already know what the shapes of f and f' are, and these are illustrated in Fig. 8.2.

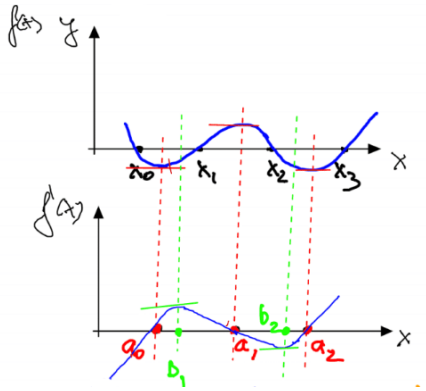


Figure 8.2: The shape of f (above) and f' (below). The points a_i are roots of f' and the points b_i are roots of f''.

As $f'(a_i) = 0$ we also have that $|N(a_i)| = \infty$, also $f''(b_i) = 0$ implies that $N'(b_i) = 0$. Thus via our knowledge of the shape of f, we have been able to derive information about N. If we assume that $b_1 \in (a_0, x_1)$ and $b_2 \in (x_2, a_2)$, then we get the graph of N as shown in Fig. 8.3.

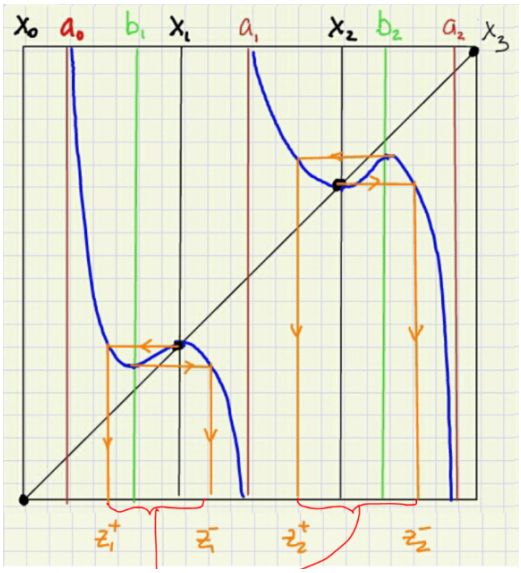


Figure 8.3: The graph of N, the points $z_i^{\pm}$ are lower bounds on the domain of attraction as they surely converge to the fixed points. The are chosen by starting at a local extrema and working backwards.

Next, we seek non-convergent iterations in $[a_0, a_1] \cup [a_1, a_2]$. We begin by finding the preimages of a_0 (called $y_{0,i}$) and a_2 (called $y_{2,i}$) this can be done geometrically by starting from the

point a_k, going down to the diagonal, and then searching horizontally for where the graph of N crosses the level a_k, going down from this point to the x-axis yields a preimage. This is shown in Fig. 8.4. Using these points define the following sets

$$I_1 = [y_{2,1}, z_1^+]; \quad I_2 = [z_1^-, y_{0,1}]; \quad I_3 = [y_{2,2}, z_2^+]; \quad I_4 = [z_2^-, y_{0,2}].$$

These sets are illustrated in Fig. 8.4 as well.

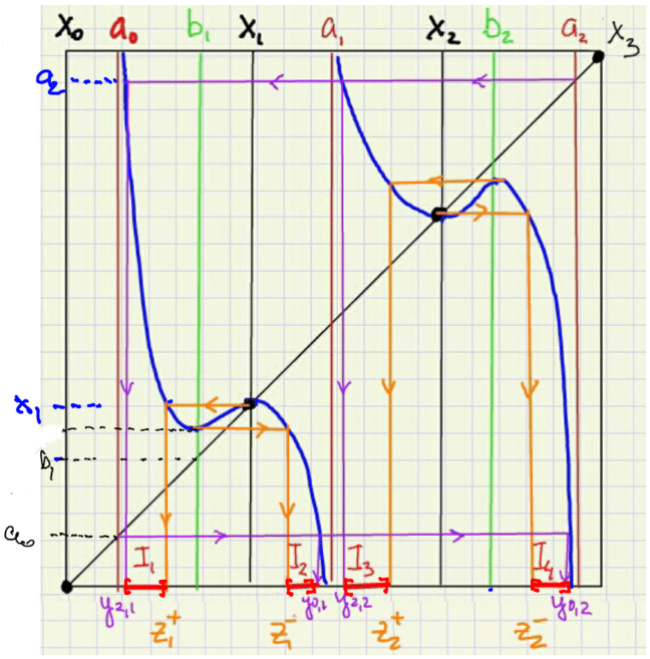


Figure 8.4: The preimages of the values a_0 and a_2, along with the sets I_j for $j = 1, 2, 3, 4$.

Noting that $b_2 > N(b_2)$ and $b_1 < N(b_1)$ observe that the following hold

$$N(I_1) = [x_1, a_2] \supset I_2 \cup I_3 \cup I_4; \quad N(I_3) = [N(b_2), a_2] \supset [b_2, a_2] \supset I_4;$$
$$N(I_2) = [a_0, N(b_1)] \supset [a_0, b_1] \supset I_1; \quad N(I_4) = [a_0, x_2] \supset I_1 \cup I_2 \cup I_3.$$

Therefore we have that the possible transitions from I_i to I_j described by the transition matrix

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

This transition matrix is in fact irreducible with $k = 3$

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{pmatrix}; \quad A^3 = \begin{pmatrix} 2 & 3 & 3 & 3 \\ 2 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 3 & 3 & 3 & 2 \end{pmatrix}.$$

We are not done yet, as we still need to establish a unique encoding. Note that if $a_{ij}^3 \neq 0$ then the set $I_i \cap f^{-1}(I_j) \neq \emptyset$ is closed for $i = 1, 2, 3, 4$. In light of this, with $f = N^3$, define

$$\begin{aligned} I_{s_0 s_1 \ldots s_n} &= I_{s_0} \cap f^{-1}(I_{s_1}) \cap f^{-2}(I_{s_2}) \cap \ldots \cap f^{-n}(I_{s_n}) \\ &= I_{s_0} \cap f^{-1}\left(I_{s_1} \cap f^{-1}\left(I_{s_2} \cap \ldots f^{-1} \underbrace{\left(I_{s_{n-1}} \cap f^{-1}(I_{s_n}) \right)}_{\neq \emptyset, \text{ closed}} \right) \right). \end{aligned}$$

The condition $I_{s_{n-1}} \cap f^{-1}(I_{s_n}) \neq \emptyset$ and is closed is true as each entry $a_{ij}^3 \neq 0$. Therefore we have a nested sequence of nonempty, closed sets and by Cantor's Theorem their intersection yields a nonempty, closed set. By our previous result we have that $\Lambda = \bigcup_{s \in \Sigma_A^N} I_s$ is a Cantor set. In this case $N = 4$ and A is as above. We then have the encoder function $\phi : \Lambda \to \Sigma_A^4$ and the chaotic function $\sigma : \Sigma_A^4 \to \Sigma_A^4$. These yield the commuting diagram

$$\begin{array}{ccc} \Sigma_A^4 & \xrightarrow{\sigma} & \Sigma_A^4 \\ \phi \uparrow & & \phi \uparrow \\ \Lambda & \xrightarrow{N} & \Lambda \end{array}.$$

Thus we have that $N = \phi^{-1} \circ \sigma \circ \phi$, which implies that N is a chaotic map. Hence, the Newton-Raphson iteration has an invariant Cantor set, on which the iteration is equivalent to a Bernoulli subshift of finite type on 4 symbols.

Bibliography

[Arnold, 1992] Arnold, V. (1992). *Ordinary Differential Equations*. Springer.

[Breunung and Haller, 2019] Breunung, T. and Haller, G. (2019). When does a periodic response exist in a periodic forced multi-degree-of-freedom mechanical system? *Nonlinear Dynamics*, 98:1761–1780.

[Chicone, 1999] Chicone, C. (1999). *Ordinary Differential Equations with Applications*. Springer.

[Guckenheimer and Holmes, 1990] Guckenheimer, J. and Holmes, P. (1990). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, New York, 3. print., rev. and corr. edition.

[Haller, 2001] Haller, G. (2001). Distinguished material surfaces and coherent structures in three-dimensional fluid flows. *Physica D: Nonlinear Phenomena*, 149(4):248–277.

[Kutz et al., 2015] Kutz, J. N., Fu, X., and Brunton, S. L. (2015). Multi-resolution dynamic mode decomposition.

[Palis, 1967] Palis, J. (1967). *On Morse-Smale Diffeomorphisms*. PhD thesis, UC Berkely.

[Palis, 1969] Palis, J. (1969). On morse-smale dynamical systems. *Topology*, 8:385–404.

[Saari and Urenko, 1984] Saari, D. G. and Urenko, J. B. (1984). Newton's method, circle maps, and chaotic motion. *The American Mathematical Monthly*, 91(1):3–17.

[Schmid, 2010] Schmid, P. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5 – 28.

[Strogatz, 2000] Strogatz, S. (2000). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Studies in nonlinearity. Westview.

[Sun et al., 2016] Sun, P., Colagrossi, A., Marrone, S., and Zhang, A. (2016). Detection of lagrangian coherent structures in the sph framework. *Computer Methods in Applied Mechanics and Engineering*, 305:849–868.

[Verhulst, 1989] Verhulst, F. (1989). *Nonlinear Differential Equations and Dynamical Systems*. Springer.

[Williams et al., 2015] Williams, M., Kevrekidis, I., and Rowley, C. (2015). A data-driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, pages 1307–1346.