Lecture Notes For: Dynamical Systems

Ali Fele Paranj
alifele@student.ubc.ca

May 6, 2024

Contents

1 Random Tips and Tricks

In this section I will cover the basics of calculus which I have used throughout the text.

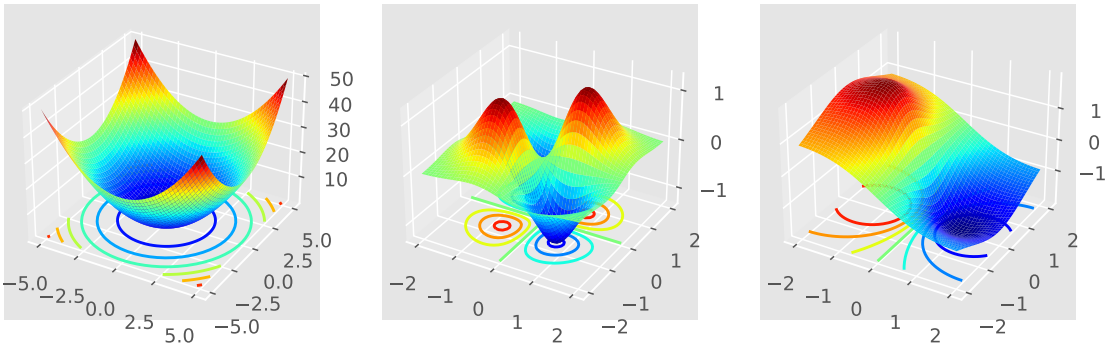
1.1 Level Curves

Here, I will focus my arguments to 2D scalar function and vector fields, as it is easier to imagine and also plot. However, it can easily been generalized to higher dimensions.

Let $F : \mathbb{R}^2 \to \mathbb{R}$ be a function. This function is called a scalar function, as it assigns an scalar value to each point in the $\mathbb{R}^2$ plane. This function can be visualized using it graph which is

$$\text{Graph}(F) = \{(x, y, z) \in \mathbb{R}^3 : z = F(x, y)\},$$

which is basically a surface in 3D. Consider the following plots which represent the graph of different scalar function.



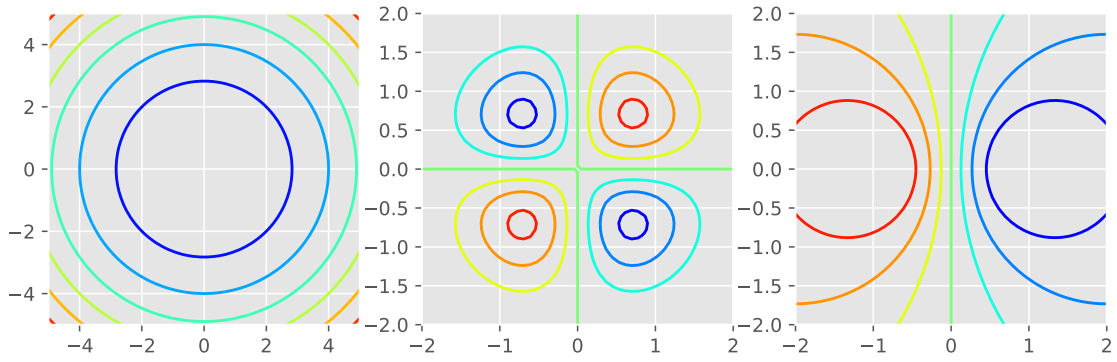
As we can see in the figure above, graph of a scalar function is not always very informative, as certain portions of the function might not be visible due to the projection of the 3d plot. Another idea is to use level curves of the function to represent it. Level curves of a scalar function is defined as

$$LC_c(F) = \{(x, y) \in \mathbb{R}^2 : F(x, y) = c,\ c \in \mathbb{R}\}.$$

Or as an alternative definition, a level curve of F is a path $\gamma : \mathbb{R} \to \mathbb{R}^2$ that satisfies

$$F(\gamma(t)) = c.$$

The following figure represents the level curves of the functions represented in the figure above.



The time derivative of the function $F \circ \gamma$ is zero. The time derivative of $F \circ \gamma$ is the directional derivative of F in the direction of $\gamma'(t)$ evaluated at $\gamma(t)$. Because

$$\frac{d}{dt}F(\gamma(t)) = \frac{\partial F}{\partial x}\Big|_{\gamma(t)}\gamma_1'(t) + \frac{\partial F}{\partial y}\Big|_{\gamma(t)}\gamma_2'(t) = \nabla F\Big|_{\gamma(t)} \cdot \gamma'(t) = D_{\gamma'(t)}F\Big|_{\gamma(t)}.$$

Thus a level curve of $F(x, y)$ passing through $p = (p_1, p_2) \in \mathbb{R}^2$, is a curve $\gamma(t) = (\gamma_1(t), \gamma_2(t))$ that is the solution of the following initial value problem

$$F_x(\gamma(t))\gamma_1'(t) + F_y(\gamma(t))\gamma_2'(t) = 0, \qquad \gamma(0) = p.$$

Let $x = x(t) = \gamma_1(t)$ and $y = y(t) = \gamma_2(t)$. Then the equation above can be written as

$$F_x(x, y)x' + F_y(x, y)y' = 0, \qquad x(0) = p_1,\ y(0) = p_2.$$

This differential equation determines the level curve corresponding to $F(x, y) = F(p_1, p_2)$. We can write y in terms of x as

$$y = f(x) = (\gamma_2 \circ \gamma_1^{-1})(x)$$

Thus the time derivative of y can be written as $y' = \frac{dy}{dx}x'$. Assuming $x' \neq 0$ we can simplify the differential equation above as

$$\frac{df}{dx} = \frac{-F_x(x, y)}{F_y(x, y)}, \qquad f(p_0) = p_1.$$

Solving this initial value problem will determine the desired level curve.

Example 1.1

We want to find the level curve of $F(x, y) = x^2 + y^2$ which passes through

$(1,1) \in \mathbb{R}^2$. To do this, we need to solve the following initial value problem

$$\frac{dy}{dx} = -\frac{2x}{2y}, \qquad y(1) = 1.$$

By the method of separation of variables we will get

$$y^2 + x^2 = 1,$$

which is a circle passing through the origin and radius 1.

Note that we could do the whole business with following the concept of implicit differentiation (see a Calculus text book like Stewart).

Example 1.2

In this example we are calculating the level-curves of the Lotka-Voltera model given as (nondimensionalized)

$$\dot{x} = x(1-y), \qquad \dot{y} = \delta y(x-1).$$

To find the level curves, assume there exists a function $H : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\frac{d}{dt} H(X(t)) = 0,$$

where $X(t) = (x(t), y(t))^T \in \mathbb{R}^2$ such that $\dot{X} = F(X)$, where $F : \mathbb{R}^2 \to \mathbb{R}^2$ is the RHS of the Lotka-Voltera system. In other words, $X(t)$ is a solution of the model that satisfies the differential equation. By chain rule we can write

$$H_x(x(t), y(t))\dot{x}(t) + H_y(x(t), y(t))\dot{y}(t) = 0. \qquad (\text{☺})$$

The equation (☺) is the central equation and is very important. First, assume that $H_y(x(t), y(t)) \neq 0$ for all $t \in \mathbb{R}$. Then we can use the implicit function theorem and write $y = y(x)$. Thus $\dot{y}(t) = y'\dot{x}(t)$. And assuming $\dot{x}(t) \neq 0$ for all $t \in \mathbb{R}$ we can write (☺) as

$$y' = -\frac{H_x(x(t), y(t))}{H_y(x(t), y(t))}.$$

On the other hand, we know that $\dot{x}(t) = f_1(x, y)$ and $\dot{y}(t) = f_2(x, y)$, where f_1 and f_2 are the component functions of the vector field F. Thus (☺) we lead to

$$-\frac{H_x(x(t), y(t))}{H_y(x(t), y(t))} = \frac{f_2(x, y)}{f_1(x, y)}.$$

Thus we can write

$$\frac{dy}{dx} = \frac{f_2(x,y)}{f_1(x,y)}.$$

Using the specific vector field for the Lotka-Voltera model we can write

$$\frac{dy}{dx} = \frac{\delta y(x-1)}{x(1-y)}.$$

Now we can solve this ODE by the method of the separation of variables.

$$\frac{1-y}{y}dy = \delta\frac{x-1}{x}dx.$$

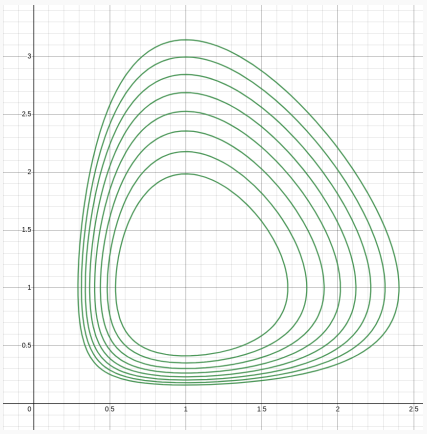
By integrating the two sides of the equation we will have

$$\ln(y) - y = \delta(x - \ln(x)) + c.$$

Thus the level curves of the function

$$H(x,y) = \ln(y) - y - \delta(x - \ln(x))$$

will yield the trajectories of the Lotka-voltera system in the (x,y) plane. The following figure represents these curves for $\delta = 1.9$.



1.2 Complexifying and Realifying A Vector Field

Sometimes, it is very beneficial to express a particular vector field in complex variables, as it can possibly simplify the expression a lot. For instance, consider the

following vector field.

$$\dot{x}_1 = \mu x_1 - \omega x_2 + b(x_1^3 + x_1 x_2^2),$$
$$\dot{x}_2 = \omega x_1 + \mu x_2 + b(x_2^3 + x_1^2 x_2).$$

This vector field looks very messy in the rectangular coordinates. However, if we assume x_1 and x_2 are the real and imaginary parts of a complex variable z, i.e. $z = x_1 + ix_2$, then we can simplify the vector field a lot. After doing the transformation, we will have only one expression for the ODE system, as a complex variable is naturally 2D. To do the change of variable, we have several options. The very first elementary option is to follow our nose, and use the linear algebra constructs. Note that since x_1 is the real part of z and x_2 is the imaginary part, then the following identities hold

$$x_1 = \frac{z + \bar{z}}{2}, \quad x_2 = \frac{z - \bar{z}}{2i}.$$

Thus in vector notation we have

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \end{pmatrix}.$$

Now we can substitute the equations in the expression for the ODE system to find the expression in terms of the new variable z. The second option, is to be a little bit careful and try to compute $\dot{z} = \dot{x}_1 + i\dot{x}_2$, and try to group the terms in a neat way to produce $z = x_1 + ix_2$ terms to be substituted with z, while trying to make terms like $\mu + i\omega$ to replace it with λ.

After following either of ways, we will have

$$\dot{z} = \lambda z + bz|z|^2,$$

where $\lambda = \mu + i\omega$. This is a huge step forward. The expression in terms of z is much more neat and we can easily see what is going on with the dynamics. It is super easy to determine the contribution of the higher order terms in the dynamics, which was quite hard to accomplish when we were working with the rectangular coordinates.

Now, in order to get the expression back in the rectangular coordinates, we can substitute $z = x + iy$ in the expression for $\dot{z}$ and try to group the terms and separate the real and imaginary parts to determine the expressions for x_1 as well as for x_2. Equivalently, we can use the following coordinate change

$$\begin{pmatrix} z \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

1.3 What Happens When Turning a Matrix to a Jordan Normal Form

Linear algebra is amazing as it handles many messy processes and computations in a very neat and clean way. However, it is always a great practice to ask this

question that "What I would have done if I didn't know linear algebra?". Here we will explain the answer to this question in turning a matrix to Jordan normal form.

For any real matrix, the complex eigenvalues always comes in pairs. I.e. if a real matrix happen to have a complex eigenvalue $\lambda_1 = \mu + i\omega$, then there is certainly another eigenvalue $\lambda_2 = \mu - i\omega$. Also, the eigenvectors associated with any pair of complex conjugate eigenvalues are always complex conjugate themselves. I.e. assume that q is the eigenvector associated with the eigenvalue λ_1, then $\bar{q}$ is the eigenvector associated with the eigenvector λ_2. The following shows the process of turning a 2×2 matrix into its Jordan normal form to observe that what is really happening under the hood.

Consider the following system of ODE

$$\dot{u} = Au, \qquad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

where A has two complex eigenvalues $\lambda_1 = \mu + i\omega$ and $\lambda_2 = \mu - i\omega$, and associated with them two complex conjugate eigenvectors $q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + i \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and $\bar{q} = \begin{pmatrix} \bar{q_1} \\ \bar{q_2} \end{pmatrix}$, where $q_1, q_2 \in \mathbb{C}$. Now there are two distinct paths that can take us to the Jordan normal form of the matrix. One is to turn the matrix into it Real normal form first and then turn it into the Jordan normal form, and then other one is to directly transform the matrix to its Jordan normal form. For the first path, in order to transfer the matrix A to its real normal form, we do the following coordinate change

$$u = \begin{pmatrix} \mathrm{Re}(q_1) & -\mathrm{Im}(q_1) \\ \mathrm{Re}(q_2) & -\mathrm{Im}(q_2) \end{pmatrix} v = \begin{pmatrix} x_1 & -y_1 \\ x_2 & -y_2 \end{pmatrix} v = P_1 v.$$

where we have assumed that λ_1 is the eigenvalue with positive imaginary part. Then the ODE system will be

$$\dot{v} = P_1^{-1} A P_1 v = Rv.$$

After this transformation, the matrix R will be like below

$$R = \begin{pmatrix} \mu & -\omega \\ \omega & \mu \end{pmatrix}.$$

Now turning this matrix to the Jordan normal form (a diagonal matrix with the complex eigenvalues sitting in the main diagonal) has a clear explanation that we described in the previous section as Complexifying a vector field. By evaluating the matrix R it turns out that it has two eigenvectors as $r_1 = \frac{1}{2} \begin{pmatrix} 1 \\ -i \end{pmatrix}$ and $r_2 = \frac{1}{2} \begin{pmatrix} 1 \\ i \end{pmatrix}$.

Thus we need to do the following change of variable to turn the matrix into its normal form

$$v = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\xi = P_2\xi.$$

In the components we will have

$$v_1 = \frac{1}{2}(\xi_1 + \xi_2), \qquad v_2 = \frac{1}{2i}(\xi_1 - \xi_2).$$

And then the ODE system will be

$$\dot{\xi} = P_2^{-1}RP_2\ \xi = P_2^{-1}P_1^{-1}AP_1P_2\ \xi = J\xi,$$

where the matrix J will be

$$J = \begin{pmatrix} \mu + i\omega & 0 \\ 0 & \mu - i\omega \end{pmatrix}.$$

Thus it turns out that in fact ξ_1 is the complex conjugate of ξ_2, and this step in changing the real normal form to Jordan normal form is the same as simply letting

$$\xi = v_1 + iv_2.$$

I.e. letting the components of the vector v be the real and imaginary part of the complex variable ξ.

However, it is also possible to get here directly from A, by simply doing the following coordinate change

$$u = P\hat{z} = \begin{pmatrix} q_1 & \bar{q_1} \\ q_2 & \bar{q_2} \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

and we can predict that z_2 is in fact the complex conjugate of z_1. By letting $z_1 = \bar{z_1}$ we can write

$$u = \begin{pmatrix} q_1 & \bar{q_1} \\ q_2 & \bar{q_2} \end{pmatrix}\begin{pmatrix} z_1 \\ \bar{z_1} \end{pmatrix}$$

which is simply listing the eigenvectors in the column of a vector when we do the general diagonalization. In this direct transformation we in fact set

$$u_1 = q_1z_1 + \overline{q_1z_1}, \qquad u_2 = q_2z_1 + \overline{q_2z_1},$$

or in a vector notation (let $z = z_1 = z_2$) we will have

$$u = qz + \overline{qz}.$$

Now, the only question that remains is the is there a way to express z_1 neatly in terms of u_1 and u_2, as we did in transforming the real normal form to the Jordan

normal form? The answer is yes, but it required a little bit more careful treatment. We know that q is the eigenvector of the matrix A with eigenvalue λ. We define the eigenvector of the adjoint matrix A^* with eigenvalue $\overline{\lambda}$ to be p. I.e.

$$A^* p = \overline{\lambda} p.$$

Then $\langle p, \bar{q} \rangle = 0$. That is because

$$\lambda \langle p, \bar{q} \rangle = \langle \overline{\lambda} p, \bar{q} \rangle = \langle A^* p, \bar{q} \rangle = \langle p, A\bar{q} \rangle = \langle p, \overline{\lambda q} \rangle = \overline{\lambda} \langle p, \bar{q} \rangle,$$

and since $\lambda - \overline{\lambda} = 2i\omega \neq 0$, then $\langle p, \bar{q} \rangle = 0$. Also we can show (I don't know how yet) that

$$\langle p, q \rangle \neq 0,$$

and we can normalize p in a way that $\langle p, q \rangle = 1$. With this setting, and observing that $u = zq + \overline{zq}$ it is very easy to observe that

$$z = \langle p, u \rangle.$$

It is now clear that in the special case where we were transforming the real normal form to Jordan normal form the vector p was actually

$$p = \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

2 Basic Definitions and Concepts

We will devote this section to basic definitions of dynamical systems and the core ideas. For this reason, this chapter, will be served as a main references for the definitions used anywhere in this note. Because of this nature, there might be a little consistency in the material in this section and the concepts will be scattered all over the place!

Definition 2.1. Stable set of an invariant set

Let S be in invariant set of a dynamical system, then it stable set $W^s(S)$, is the set of all states, whose orbits approach S forward in time. In other words

$$W^s(S) = \{x \in X : \ d(\varphi^t x, S) \to 0, \ as \ t \to +\infty\},$$

in which X is the state space, $d : X \times X \to \mathbb{R}$ is a metric function.

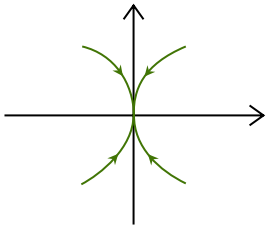
For instance, consider the following system

$$\dot{X} = \begin{pmatrix} -\lambda_1 & 0 \\ 0 & -\lambda_2 \end{pmatrix} X,$$

where $X \in \mathbb{R}^2,\ 0 < \lambda_1 < \lambda_2$. The solution of this system is

$$X(t) = e^{At}X_0 = \begin{pmatrix} e^{-\lambda_1 t} & 0 \\ 0 & -\lambda_2 t \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}.$$

For this system, the invariant sets are the origin $S_1 = (0,0)$, $S_2 = \text{Span}\{(0,1)^T\}$, and $S_3 = \text{Span}\{(1,0)^T\}$. The following figure shows a basic phase portrait of this system.



Thus the stable sets of the invariant sets of this system will be

$$W^s(S_1) = \mathbb{R}^2, \qquad W^s(S_2) = \mathbb{R}^2, \qquad W^s(S_3) = \mathbb{R}^2.$$

With a similar logic, we can define the **unstable set** $W^u(S)$ of in invariant set, which is the set of all states whose orbits approach S *backward* in time.

Corollary: 2.1

The stable, and unstable sets of an invariant set, are invariant sets themselves.

Proof. Here we proof the statement for the unstable set only, however, the proof logic for both of them is similar. We proceed with the proof by contradiction. Assume that $W^s(S)$ is not a invariant set. This means that

$$\exists x_0 \in W^s(S),\ \exists t^* \in \mathbb{R},\ \text{s.t.}\ \ \varphi^{t^*}x_0 = z_0 \notin W^s(S).$$

So $\lim_{t\to\infty} d(\varphi^t z_0, S) \neq 0$. On the other hand, because of $\varphi^{s+t}x_0 = \varphi^s(\varphi^t x_0)$ we have

$$\varphi^{t^*}x_0 = z_0 \ \Leftrightarrow\ \varphi^{t-t^*}(\varphi^{t^*}x_0) = \varphi^{t-t^*}z_0 \ \Leftrightarrow\ \varphi^t x = \varphi^{t-t^*}z_0,$$

which implies that the distance between $\varphi^{t-t^*}z_0$ and the set S goes to zero, which contradicts our assumption. So we conclude that the stable set of an invariant set, is an invariant set itself.

□

Definition 2.2. Stable, Unstable and Center Subspace

Consider the dynamical system

$$\dot{x} = f(x) \approx [\mathcal{D}f](x^*)x,$$

in which $x \in X$, $f : X \to X$. Also $[\mathcal{D}f]$ represents Jacobian matrix evaluated at $x^* \in X$ which is an equilibrium point, i.e. $f(x^*) = 0$. The eigenspace (and possibly generalized eigenspace) $\text{Span}\{v_j^{[1]}, \cdots, v_j^{[m]}\}$ associated with the eigenvalue λ_j is

- stable subspace, if $\text{Re}(\lambda_j) < 0$,
- unstable subspace, if $\text{Re}(\lambda_j) > 0$,
- center subspace, if $\text{Re}(\lambda_j) = 0$.

For instance, consider the linear dynamical system

$$\dot{X} = AX,$$

where $X \in \mathbb{R}^6$, and the matrix A is given as

$$A = \begin{pmatrix} \lambda_1 & -\omega_1 & 0 & 0 & 0 & 0 \\ \omega_1 & \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda_2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & 0 & 0 & \omega_2 & 0 \end{pmatrix},$$

where $\lambda_j > 0$ for $j = 1, 2, 3$. Also $\omega_1 > 0, \omega_2 > 0$. Then the stable, unstable, and center subspaces associated with the origin will be like the following figure.

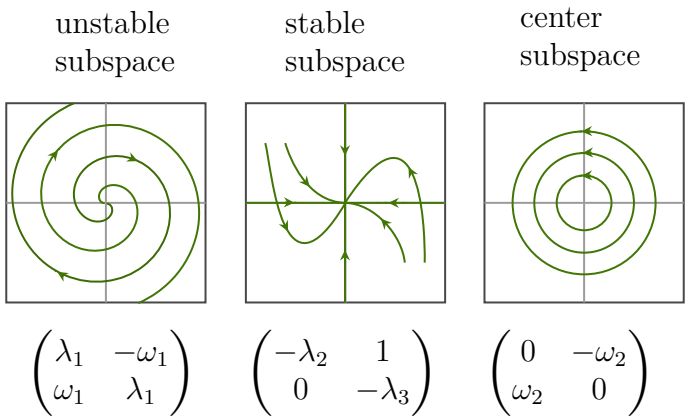
3 Systems of Linear Differential Equations

In dealing with the systems of linear differential equations we encounter

$$\dot{x} = A(t)x, \tag{♫}$$

where $x = x(t) \in \mathbb{R}^n$. It turns out that this system of linear differential equations has n linearly independent solutions

$$\{x_1(t), x_2(t), \ldots, x_n(t)\}.$$



It is important to note that the span of all of these solutions is a n-dimensional subspace of the all continuous function from $\mathbb{R}^n$ to $\mathbb{R}^n$. Thus any particular solution of the ODE can be expressed as the linear combination of the solutions stated above, i.e.

$$x(t) = \sum_{i=1}^{n} c_i x_i(t). \tag{1}$$

To show all of these ideas in a neat matrix form, we construct the matrix $\Psi(t)$ called a fundamental matrix (note that this is not unique) as

$$\Psi(t) = [x_1(t), x_2(t), \ldots, x_n(t)],$$

that is a matrix whose jth column is $x_j(t)$. With this notation in hand we can express any particular solutions of the ODE system as

$$x(t) = \Psi(t)c. \tag{2}$$

Note that so far nothing serious is happening. What happened is that we just defined the matrix $\Psi(t)$ just to be able to write the equation (1) in a more fancy matrix notation (2).

So far, the functions $\{x_1(t), x_2(t), \ldots, x_n(t)\}$ where acting like the basis of the space they span. However, we can find another basis that is more interesting. Consider the functions $\{X_1(t), X_2(t), \ldots, X_n(t)\}$ where satisfy the (♫) as well as

$$X_1(t_0) = \hat{e}_1,\ X_2(t_0) = \hat{e}_2,\ \ldots\ ,\ X_n(t_0) = \hat{e}_n,$$

where $\hat{e}_j$ is the standard basis of $\mathbb{R}^n$. Note that we can calculate these functions easily since by equation (2) we can write

$$X_j(t) = \Psi(t)c_j \implies c_j = \Psi^{-1}(t)X_j(t) \implies c_j = \Psi^{-1}(t_0)\hat{e}_j.$$

Thus any of $X_j(t)$ can be written as

$$X_j(t) = \Psi(t)\Psi^{-1}(t_0)\hat{e}_j$$

So we can now construct a new fundamental matrix $M(t, t_0)$ as

$$M(t, t_0) = [X_1(t), X_2(t), X_3(t), \ldots, X_n(t)].$$

This fundamental matrix is useful when dealing with initial values problems, i..e

$$\dot{x} = A(t)x, \qquad x(t_0) = x_0. \tag{♪}$$

In this case the solution to initial values problem can be written as

$$x(t) = \varphi(t, t_0, x_0) = M(t, t_0)x_0.$$

That is simply because, since $M(t, t_0)$ is a fundamental matrix, then any solution can be written as $x(t) = M(t, t_0)c$, where $c \in \mathbb{R}^n$. Let $t = t_0$, then we will have $x_0 = x(t_0) = M(t_0, t_0)c = Ic = c$, thus $c = x_0$.

Apart from being useful in solving initial values problems, $M(t, t_0)$ has the following properties as well. In fact, $M(t, t_0)$ is an example of a *flow operator*. Also, derivative of $M(t, t_0)$ is given by

$$\frac{d}{dt}M(t, t_0) = A(t)M(t, t_0).$$

That is because substituting $x(t) = M(t, t_0)x_0$ in (♪) will yield in

$$\frac{d}{dt}M(t, t_0)x_0 = A(t)M(t, t_0)x_0,$$
$$(\frac{d}{dt}M(t, t_0) - A(t)M(t, t_0))x_0 = 0.$$

Since this is true for all x_0, then

$$\frac{d}{dt}M(t, t_0) = A(t)M(t, t_0).$$

□

4 Poincare Map

When dealing with dynamical systems that has periodic orbits, one useful machinery to study the stability of those orbits is to use Poincare map. We will explore the idea in the following example

Question 1. Consider the dynamical system described by the following ODE system.

$$\dot{x}_1 = x_1 - \omega x_2 - x_1^3 - x_1 x_2^2,$$
$$\dot{x}_2 = \omega x_1 + x_2 - x_1^2 x_2 - x_2^3,$$

in which $\omega > 0$. Analyze the behaviour of this system near the equilibrium points and the periodic orbits.

Answer.

The study the behaviour of the system near the equilibrium points, we first need to find them. So we solve the following equation

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

By analyzing the ODE system we can infer that $X = (0,0)^T$ a solution of the equation above. So the point $X = (0,0)^T$ is an equilibrium point. To study the stability of this equilibrium point, we need to linearize the system near the equilibrium point.

$$F_X(X)|_{X=0} = \begin{pmatrix} (f_1)_{x_1} & (f_1)_{x_2} \\ (f_2)_{x_1} & (f_2)_{x_2} \end{pmatrix}_{X=0} = \begin{pmatrix} 1 & -\omega \\ \omega & 1 \end{pmatrix}$$

Since the trance and determinant are both positive, then the origin is unstable.

However, to analyze the stability of the periodic orbits, we need to find them in the first place. To find the periodic orbit in this example, it is more convenient to convert this ODE system to the polar coordinate in which we have

$$x_1 = r\cos\theta, \qquad x_2 = r\sin\theta,$$

in which $r > 0$ and $\theta = \mathbb{S}^1$, in which $\mathbb{S}^1$ is the unit circle. Substituting in the ODE system and utilizing the chain rule, then we can write

$$\begin{aligned} \dot{r}\cos\theta - r\dot{\theta}\sin\theta &= r\cos\theta - r\omega\sin\theta - r^3\cos^3\theta - r^3\cos\theta\sin^2\theta, \\ \dot{r}\sin\theta + r\dot{\theta}\cos\theta &= r\omega\cos\theta + r\sin\theta - r^3\cos^2\theta\sin\theta - r^3\sin^3\theta. \end{aligned}$$

Multiplying the first equation in $\sin\theta$ and the second one in $\cos\theta$ and then subtracting them, And then multiplying the second equation in $\cos\theta$ and the first one in $\sin\theta$ and then adding the equations will yield in

$$\begin{aligned} \dot{r} &= r - r^3, \\ \dot{\theta} &= \omega. \end{aligned}$$

It is clear that

$$p^0(t) = \begin{pmatrix} r(t) \\ \theta(t) \end{pmatrix} = \begin{pmatrix} 1 \\ \omega t \pmod{2}\pi \end{pmatrix},$$

a solution of the this ODE system (because, simply, it satisfies the differential equation). Also, this is a periodic orbit with period $T = 2\pi/\omega$, since

$$p^0(t + 2\pi/\omega) = \begin{pmatrix} 1 \\ (\omega t + 2\pi) \pmod{2}\pi \end{pmatrix} = \begin{pmatrix} 1 \\ \omega t \end{pmatrix}.$$

We can also write this periodic orbit in the original x_1 and x_2 coordinates

$$p_X^0(t) = \begin{pmatrix} r(t)\cos\big(\theta(t)\big) \\ t(t)\sin\big(\theta(t)\big) \end{pmatrix} = \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix}.$$

Note that we write $\sin(\theta)$ instead of $\sin\big(\theta \pmod{2}\pi\big)$, because they are the same and the modulu operator is kind of defined in the definition of sin and cos functions.

Now we can evaluate the stability of this periodic orbit in two ways: 1. staying in the polar coordinate, 2. translating stuff to the original rectangular coordinate.

Poincare map in the polar coordinate.

Drawing the phase space of this system in the polar coordinates makes to build intuition what is happening here.

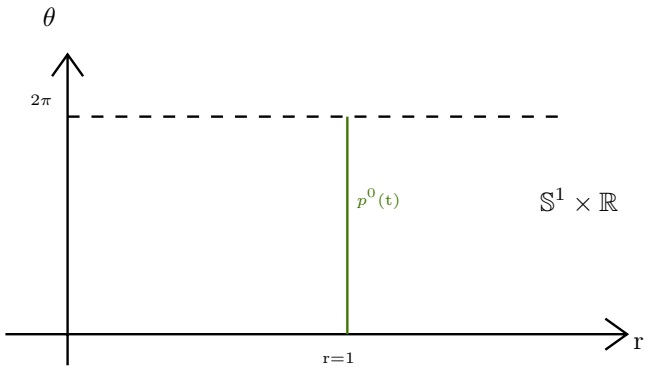


Figure 4.1: State space of the system in polar coordinates.

Now we can choose $\Sigma \subset \mathbb{R} \times \mathbb{S}^1$ on which $F(r, \theta) \neq 0$. We choose

$$\Sigma = \{(r, \theta) : \theta = 0 \pmod{2}\}.$$

Obviously, in this subset of state space we have

$$F(r, \theta) = \begin{pmatrix} r - r^3 \\ \omega \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \forall (r, \theta) \in \Sigma.$$

Let $p_0^0 = p^0(t) = (1, 0)^T \in \Sigma$. Now let's define a coordinate $\xi \in \mathbb{R}$ for Σ, by

$$r = 1 + \xi.$$

With this definition $\xi = 0$ corresponds to the point $p_0^0 \in \Sigma$. Fix an arbitrary value of ξ_0. Now we are interested to know starting from $(1 + \xi_0, 0)$ as initial value of the dynamical system, when and where we will return to Σ. So we need to solve

$$\begin{aligned} \dot{r} &= r - r^3, \qquad & r(0) &= 1 + \xi_0, \\ \dot{\theta} &= \omega, \qquad & \theta(0) &= 0. \end{aligned}$$

We can use elementary methods to solve this initial value problem. So we will get

$$r(t, \xi_0) = \frac{e^t}{\sqrt{(1+\xi)^{-2} - 1 + e^{2t}}},$$
$$\theta(t) = \omega t \pmod{2}\pi.$$

From the definition of Σ, it is clear that the time of first return is when $\theta(T_0) = 2\pi$, so $T_0 = 2\pi/\omega$. Thus the value r in the first return will be $r_1 = r(T_0, \xi_0)$, and the value of ξ in the first return will be $\xi_1 = r_1 - 1$. So, we basically got

$$\xi_1 = r(2\pi/\omega, \xi_0).$$

Since the return time is always $T = 2\pi/\omega$, then we can conclude

$$\xi_{k+1} = r(2\pi/\omega, \xi_k).$$

This is the Poincare map $P : \mathbb{R} \to \mathbb{R}$ and we have

$$P(\xi) = \frac{e^{2\pi/\omega}}{\sqrt{(1+\xi)^{-2} - 1 + e^{4\pi/\omega}}} - 1.$$

We can analyze the fixed point of this map using linearization. First observe that $\xi = 0$ is a fixed point since $P(0) = 0$. With linearization argument at $\xi = 0$ we have

$$P_\xi(\xi)|_{\xi=0} = e^{-4\pi/\omega} < 1.$$

Thus the origin is a stable equilibrium point. This is also clear form the cobweb plot of this map

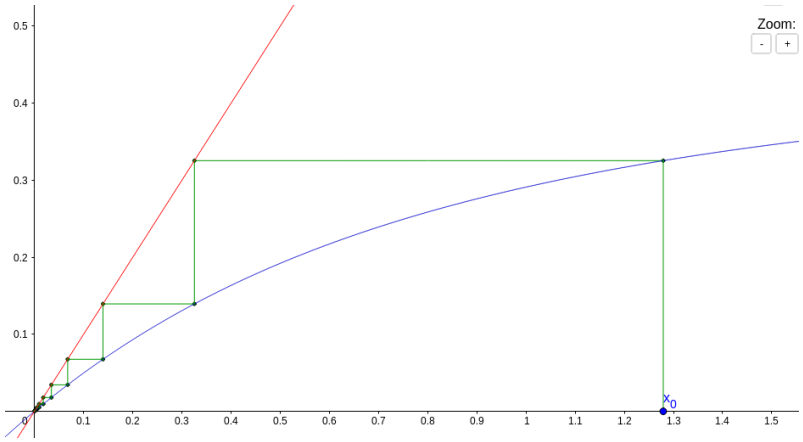


Figure 4.2: Cobweb plot of the poincare map.

5 Hamiltonian Systems and Lyapunov Function

5.1 Hamiltonian Systems

Assuming extra structures in a vector field leads to systems that are easier to analyze. One of such systems is the Hamiltonian systems. A Hamiltonian system is

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

where $x \in \mathbb{R}^{2s}$, and $x = (q, p)$ where $q = (q_1, \cdots, q_s) \in \mathbb{R}^s$, $p = (p_1, \cdots, p_s) \in \mathbb{R}^s$. Also, the vector field $f : \mathbb{R}^{2s} \to \mathbb{R}^{2s}]$ has a special property which is

$$f(x) = f(q,p) = (H_p(q,p), -H_q(q,p)) = (\frac{\partial}{\partial p_1}, \cdots, \frac{\partial}{\partial p_s}, -\frac{\partial}{\partial q_1}, \cdots, -\frac{\partial}{\partial q_s})H(q,p).$$

where $H : \mathbb{R}^{2s} \to R$ is the Hamiltonian function. In other words we have

$$\dot{q} = H_p(q,p), \qquad \dot{p} = -H_q(q,p).$$

Corollary: 5.1

If $x(t)$ is a solution of a Hamiltonian system, then $\frac{d}{dt}H(x(t)) = 0$, hence $H(x(t)) = c$ is a constant. Thus, all solutions remain on the level sets of the Hamiltonian function, which is known as conservation of energy.

Proof. We need to calculate $\frac{d}{dt}H(x(t))$ via the chain rule.

$$\frac{d}{dt}H(x(t)) = \sum_{i=1}^{2s} \frac{\partial H}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{s} \frac{\partial H}{\partial q_i}\frac{dq_i}{dt} + \sum_{i=1}^{s} \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} = \sum_{i=1}^{s} -\frac{dp_i}{dt}\frac{dq_i}{dt} + \sum_{i=1}^{s} \frac{dq_i}{dt}\frac{dp_i}{dt} = 0$$

□

This is a very useful property of Hamiltonian systems. That is because we can easily draw the phase portrait as the set of all level sets of the Hamiltonian functions. Following two examples will help to illustrate this point.

Example 5.1

Consider the following non-Linear oscillator:

$$\ddot{u} - u + u^3 = 0.$$

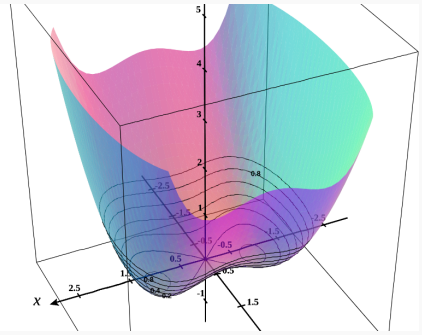
To analyze this system, first we need to write it in the form of a system of

first order ODEs. Let $q = u, p = \dot{u}$. Then we can write the system as

$$\dot{q} = v, \qquad \dot{p} = q - q^3.$$

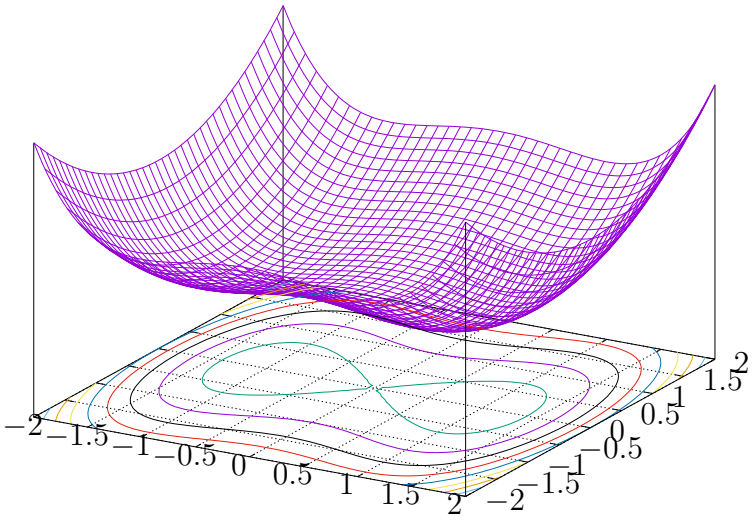
We can analyze this system using the level curves of the Hamiltonian function. We can find the Hamiltonian as

$$H(q,p) = \frac{1}{2}p^2 + \frac{1}{2}q^2 - \frac{1}{4}q^4.$$



The figure above shows the graph of this function. The orbits in the phase portrait is simply the level curves of this function.

The level curves of the dynamical system analyzed above worth more analysis. The following is the phase portrait of the dynamics



As we can see in the figure above, the equilibria $p^\circ = (0, 0)$ is sort of special. By linearization argument at the origin, we see that the Jacobian matrix has two eigenvalues, one of which is positive and the other one is negative. Thus we conclude that the equilibria is unstable and is in fact a saddle point. However, we can see that some orbits emerge from it and comeback to it again! We call such orbits homoclinic orbits.

Definition 5.1.

Homoclinic orbits are the orbits that emerge from one equilibria point and return to it again. In other words, the points of homoclinic orbit approach the equilibrium point from which the orbit emerged, as $t \to \pm\infty$.
More formally, consider the continuous dynamical system

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

and assume there is an equilibrium point at x_0. A solution $x(t)$ is a homoclinic orbit if

$$x(t) \to x_0, \quad as \quad t \to \pm\infty.$$

Homoclinic orbits are in fact the intersection of stable and unstable manifolds (see Def 2.1) of an equilibrium point.

5.2 Lyapunov Function

Lyapunov functions can be thought of generalization of Hamiltonian systems. Lyapunov function are tools that enable us to determine the invariant sets of the dynamics as well as the stability of the equilibrium points. The concept behind using Lyapunov functions is that for every scalar function $F : X \to \mathbb{R}$, there is a corresponding natural vector field. This vector field is the gradient of the scalar function. The gradient determines the iso-curves of the scalar function, i.e. the curves on which the value of the scalar function is constant. For a scalar field chosen carefully, these level curves are also closed, thus we will have the notions like the interior of the level curves, etc. These regions (i.e. interior of the level curves) along with the gradient vector field can help in determining the behaviour of a dynamical system.

The argument here focuses on $\mathbb{R}^2$, as it is easier to visualize. However, the general idea can easily be extended to higher dimensions. Let a dynamical system be defined as

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

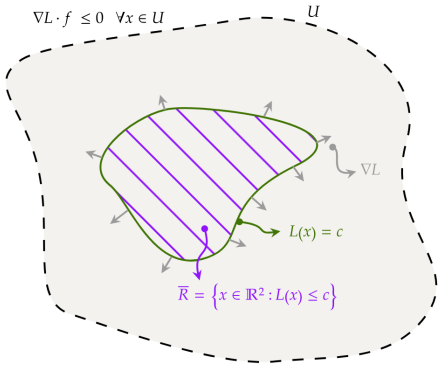
where $x \in \mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}^2$. Also, let $L : \mathbb{R}^2 \to \mathbb{R}$ be a scalar function. The gradient of this scalar is a vector field ∇L. If for a region $U \subseteq \mathbb{R}^2$ we have

$$\nabla L\Big|_x \cdot f(x) \leq 0 \quad \forall\, x \in U,$$

And the compact (closed and bounded) set

$$\overline{R} = \{x \in \mathbb{R}^2 : \ L(x) \leq c, \ c \in \mathbb{R}\}$$

be contained in U, then $\overline{R}$ is a *positively invariant* set. Considering the figure below, this fact can be justified intuitively.



Here are some intuitive explanations of this fact

- $\nabla L \cdot f$ evaluated at point $p \in \mathbb{R}^2$ is in fact the directional derivative of the scalar field at point p along the direction determined by f. Thus if $\nabla L \cdot f \leq 0$ in some region, then the value of L along any path whose tangent is determined by $f(x)$ (which is in fact the solution of the ODE) will decrease or remain constant. Thus any such path whose one of its points is in $\overline{R}$ will remain in $\overline{R}$.

- $\nabla L \cdot f$ determines the projection of f on ∇L at each point $x \in \mathbb{R}^2$. Since $\nabla L \cdot \gamma(t) = 0$ for any level curve $\gamma(t)$ for which $L(\gamma(t)) = c$, then $\nabla L \cdot f \leq 0$ means that any path whose tangent is determined by $f(x)$ (i.e. solutions of the ODE) won't leave the region $\overline{R}$.

- We can treat this fact more formally which is easier to follow. Let $x(t)$ be the solution of the ODE provided above, thus $\dot{x}(t) = f(x(t))$. Then by the chain rule we can write

$$\frac{d}{dt} L(x(t)) = \nabla L \Big|_{x(t)} \cdot f(x(t)).$$

  Then $\nabla L \cdot f \leq 0$ in a region U means that the value of $L(x(t))$ can not increase. Thus if a compact region $\overline{R} = \{x \in \mathbb{R}^2 : L(x) \leq c, \ c \in \mathbb{R}\}$ is contained at U, then the solution $x(t)$ can not leave this region since the the value of $L(x(t))$ can not increase for $t \to \mathbb{R}$.

Such a smartly chosen scalar function is called a Lyapunov function. Other than determining positively invariant sets, Lyapunov function can also help to determine the stability of the equilibrium points of the dynamics.

Proposition: 5.1

Consider the dynamical system described by

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

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$. Let $L : \mathbb{R}^n \to \mathbb{R}$ be a Lyapunov function such that $\nabla L \cdot f \leq 0$ in a region $U \subseteq \mathbb{R}^n$. Then

- If $\overline{R} = \{x \in \mathbb{R}^n : \ L(x) \leq c, \ c \in \mathbb{R}\}$ is compact and $\overline{R} \subseteq U$ then

$$\boxed{\overline{R} \text{ is positively invariant}}$$

- If p^0 is an equilibrium point, and $L(p^0)$ is an *isolated local minimum* of L, then

$$\boxed{p^0 \text{ is Lyapunov stable.}}$$

- If **in addition to the condition above**, we have $\nabla F \cdot f < 0$ in $U \backslash \{p^0\}$, then

$$\boxed{p^0 \text{ is stable.}}$$

In the following example we explore this useful tool.

Example 5.2

In this example we want to analyze the damped planar pendulum system described by

$$\ddot{\theta} + \delta\dot{\theta} + \sin\theta = 0, \quad \theta \in \mathbb{S}^1.$$

To analyze this system, we first need to write it in the form of a system of ODEs. Let $x_1 = \theta$ and $x_2 = \dot{\theta}$, where $x = (x_1, x_2) \in \mathbb{S}^1 \times \mathbb{R}$. Then we can write

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\delta x_2 - \sin x_1, \end{aligned}$$

The equilibria points are

$$p^0_{[1]} = (0 \ (\text{mod } 2\pi), 0), \quad p^0_{[2]} = (\pi \ (\text{mod } 2\pi), 0).$$

Now by the linearization argument we can determine the stability of these equilibria. The Jacobian matrix is

$$[Df] = \begin{pmatrix} 0 & 1 \\ -\cos x_1 & -\delta \end{pmatrix}.$$

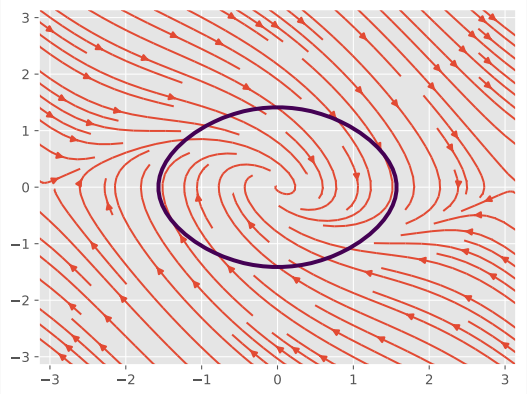
Evaluating this matrix at the equilibria points we will get

$$[Df](p^0_{[1]}) = \begin{pmatrix} 0 & 1 \\ -1 & -\delta \end{pmatrix}, \quad [Df](p^0_{[2]}) = \begin{pmatrix} 0 & 1 \\ 1 & -\delta \end{pmatrix}$$

For $p^0_{[1]}$ we have $\Delta > 0$ and $\sigma < 0$, thus $\lambda^1_1 \leq \lambda^1_2 < 0$ which implies that $p^0_{[1]}$ is stable equilibrium and a hyperbolic sink. However, for $p^0_{[2]}$ we have $\Delta < 0$ and $\sigma < 0$ which implies $\lambda^2_1 < 0 < \lambda^2_2$, implying the equilibrium is a hyperbolic saddle.
Let $L = \frac{1}{2}x_2^2 - \cos(x_1)$ be a Lyapunov function. Note that this is in fact the Hamiltonian of a simple harmonic oscillator (whose ODE is $\ddot{\theta} + \sin\theta = 0$). This is a Lyapunov function since

$$\nabla L \cdot f = (\sin(x_1), x_2) \cdot (x_2, -\delta x_2 - \sin(x_1)) = -\delta x_2^2 \leq 0.$$

Let $\overline{R} = \{x \in \mathbb{R}^2 : L(x) \leq 0\}$. The following figure shows the boundary of the set $\overline{R}$. As we can visually see, all of the arrows determined by the vector field f are pointing towards the interior of this region, which is also reflected by $\nabla L \cdot f \leq 0$ inside any open ball containing $\overline{R}$. Thus we can conclude that $\overline{R}$ is **positively invariant**.



Furthermore, since $p^0_{[1]}$ is an isolated local minimum for $L(x)$ (check $\nabla L\Big|_{x=(0,0)} = 0$), thus we can conclude that $p^0_{[1]}$ is Lyapunov stable.

6 Bifurcation Theory

Here in this section we will explore the ideas of the bifurcation theory in the dynamical systems. This is a very important topic since it can give explanation for some sudden change that we see in the nature (for example the type I and type II phase transition in the thermodynamics). We will study them with the following classification.

6.0.1 A Quick Review

Consider the following family of dynamical systems parameterized by one parameter α.

$$\dot{x} = f(x, \alpha),$$

where $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and $x, \alpha \in \mathbb{R}$. And assume $f(x_0, \alpha_0) = 0$, thus for $\alpha = \alpha_0$ the point x_0 is an equilibrium point. Now we can do the linearization at the equilibrium point and determine the type of stability. Thus after linearization we will have

$$\dot{x} = f_x(x_0, \alpha_0)(x - x_0).$$

By letting $\xi = x - x_0$ we can now write

$$\dot{\xi} = f_x(0, \alpha_0)\xi.$$

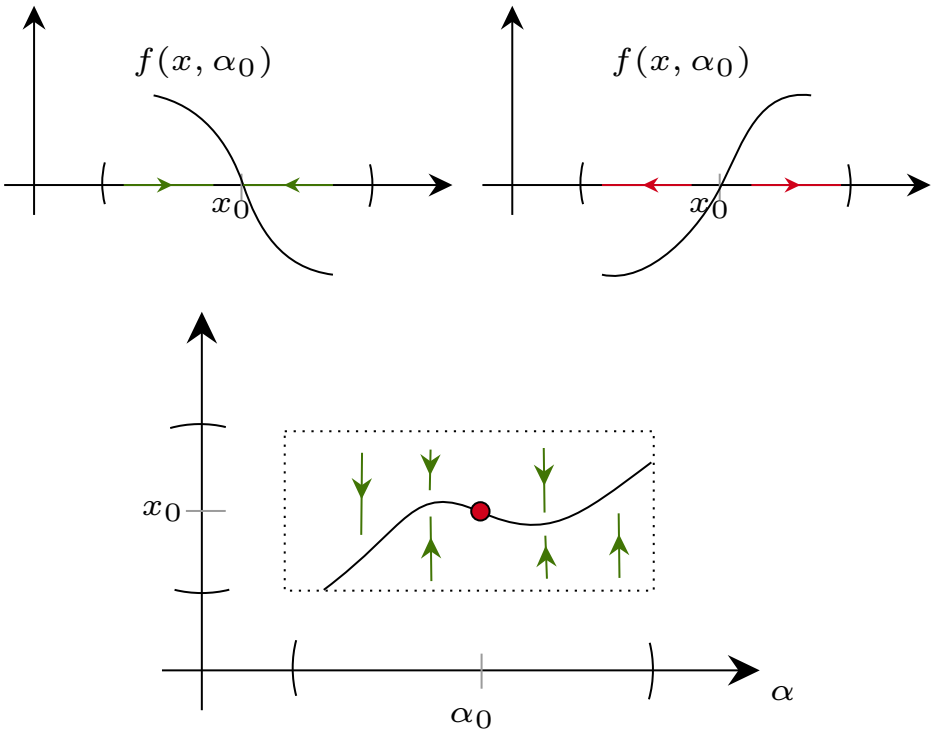
This is a very simple ODE whose solution is the exponential function. Thus we will have

$$\begin{aligned} f_x(x_0, \alpha_0) < 0 &\implies x_0 \text{ is stable.} \\ f_x(x_0, \alpha_0) > 0 &\implies x_0 \text{ is unstable.} \\ f_x(x_0, \alpha_0) = 0 &\implies \text{more terms from the Taylor's series is needed.} \end{aligned}$$

Another way of seeing this is the following figure. Note that we can still determine the type of stability by looking at the graph, despite the fact that in the classification above we have said that in that case more terms from the Taylor's series is needed. The point is that when we are looking at the graph, we automatically infer the higher order terms of the Taylor's series as well and that is why we arrive at the right conclusion often.

It turns out that because of the smooth dependence of the function f on both parameters, by sufficiently small enough perturbation of the parameter α from α_0 the stability type of the equilibrium point will remain the same up to topological equivalence. For example in the figure below, the type of equilibrium point has remained the same (in this case stable equilibrium) with sufficiently small perturbation of parameter α.

However, when $f_x(x_0, \alpha_0) = 0$, then it is possible that any perturbation of parameter α leads to the emergence or disappearance of a new type of equilibrium point, in which we say we have a *bifurcation*.



6.1 Intuitive Explanation of Bifurcation

In this section we will cover the intuitive explanation behind the notion of bifurcation and different types, and the next section we will cover the theory behind it and discuss the topological normal forms.

6.1.1 Fold (Saddle-Note) Bifurcation

The fold bifurcation is one of the simple type bifurcations where a saddle and a node emerges from blue sky as we vary the parameter α. Following the line of reasoning from the section above, assume $f(x, \alpha_0)$ looks like the following in the neighborhood of the point x_0.

Then as we can clearly see, any small perturbation of α will lead ot emergence or disappearance of new equilibria out of blue sky. We can not intuitively say that the requirements for such a bifurcation is the followings

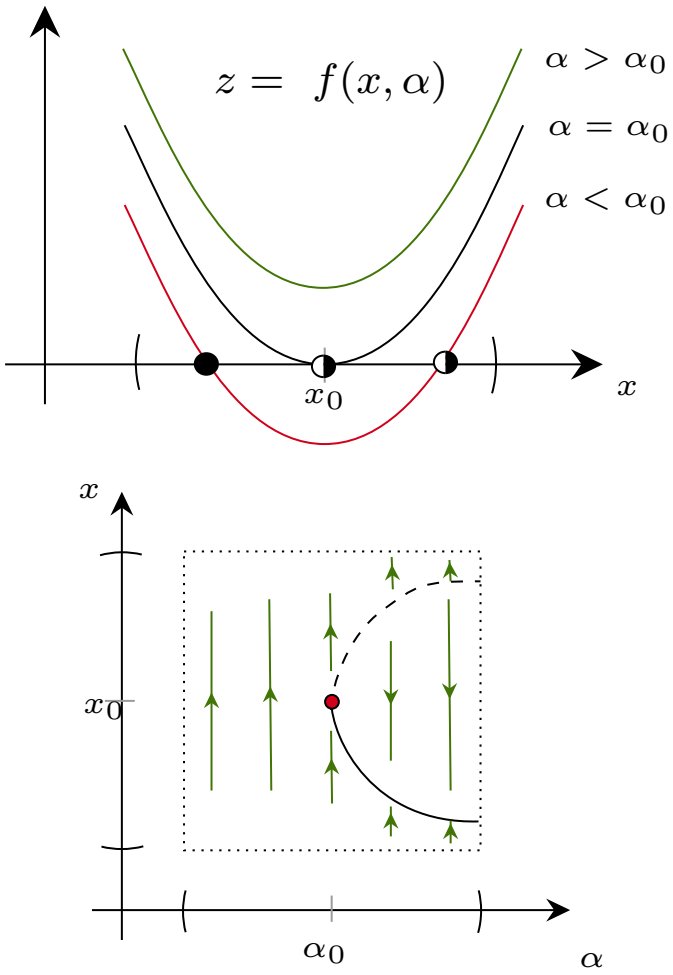
(i) $f(x_0, \alpha_0) = 0$,

(ii) $f_x(x_0, \alpha_0) = 0$,

(iii) $f_\alpha(x_0, \alpha_0) \neq 0$,

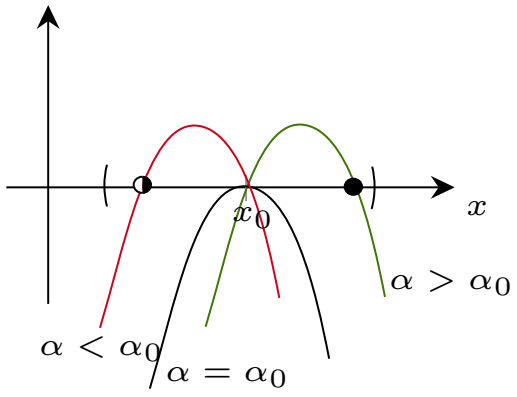
(iv) $f_{xx}(x_0, \alpha_0) =\neq 0$.

It is also informative to look at the branching diagram on the $(\alpha - x)$ plane.

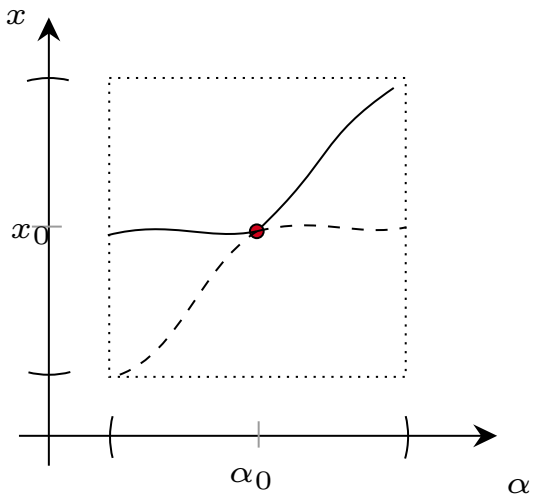


6.1.2 Transcritical Bifurcation

Another possible way that the function $f(x, \alpha)$ might change in the neighborhood of x_0 is the following diagram.



We can observe that the $f(x, \alpha)$ has a form like $f(x, \alpha) = \alpha(x - x_0) - (x - x_0)^2$. We can see that two equilibria corresponding to $\alpha > \alpha_0$ turns into a saddle type equilibria when $\alpha = \alpha_0$ and then turn into two equilibria points for $\alpha < \alpha_0$. The following branching diagram summarizes this fact.



Again, we can intuitively say that the function $f(x, \alpha)$ should have the following conditions in order to exhibit a Transcritical bifurcation.

(i) $f(x_0, \alpha_0) = 0$,

(ii) $f_x(x_0, \alpha_0) = 0$,

(iii) $f_{xx}(x_0, \alpha_0) \neq 0$,

(iv) $f_{\alpha x}(x_0, \alpha_0) \neq 0$.

Note that $f_{\alpha x} = \frac{d}{d\alpha}\frac{df}{dx}$, which simply is the rate of change of the $x-$slope of $f(x, \alpha)$ with change in α.

Theoretical Treatment of the Transcritical Bifurcation

Consider the following family of vector fields

$$\dot{x} = f(x, \alpha),$$

where $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^1$. Furthermore suppose that the vector field is constrained in such a way that

$$f(0, \alpha) = 0,$$

for all α in some interval containing the special value α_0. Also, assume that $(0, \alpha_0)$ is a non-hyperbolic equilibrium i.e.

$$f_x(0, \alpha_0) = 0.$$

Thus there is a chance that we observe a bifurcation as we change the value of α through α_0. Now we can expand the vector field around the point $(0, \alpha_0)$.

$$\begin{aligned} f(x, \alpha) &= \underbrace{f(0, \alpha)}_{=0} + f_x(0, \alpha)x + 1/2 f_x(0, \alpha)x^2 + \mathcal{O}(|x|^3) \\ &= x(f_x(0, \alpha) + f_{xx}(0, \alpha)x/2 + \mathcal{O}(|x|^2)) = x\hat{f}(x, \alpha). \end{aligned}$$

Now in this representation of the vector field, it is clear that due to the constraint that $f(0, \alpha) = 0$, $x = p_1^0(\alpha) \equiv 0$ is a branch of equilibria. The other branch can be computed by setting $\hat{f}(x, \alpha) = 0$. Before doing that, we need to expand the $\hat{f}(x, \alpha)$ in x as well. Thus

$$\hat{f}(x, \alpha) = \underbrace{f_x(0, \alpha_0)}_{=0} + \underbrace{f_{x\alpha}(0, \alpha_0)}_{a}(\alpha - \alpha_0) + \underbrace{1/2 f_{xx}(0, \alpha_0)}_{b} x + \mathcal{O}(|\alpha - \alpha_0|^2 + |\alpha - \alpha_0||x| + |x|^2)$$

In short, we can write the vector field as

$$f(x, \alpha) = x(a(\alpha - \alpha_0) + bx + \mathcal{O}(|\alpha - \alpha_0|^2 + |\alpha - \alpha_0||x| + |x|^2))$$

We require to more conditions in order to have the transcritical bifurcation, i.e. $a \neq 0$ and $b \neq 0$. So, now we can easily solve $\hat{f}(x, \alpha) = 0$ locally to get the second branch of equilibrium locally. We can do it fancy and use the fact that $\hat{f}_x(x, \alpha)b =\neq 0$ and also $\hat{f}(0, \alpha_0) = 0$, to obtain a locally smooth function $x = g(\alpha)$ on which $\hat{f}(g(\alpha), \alpha) = 0$, and then using the Taylor expansion of $g(\alpha) = c_0 + c_1\alpha + c_2\alpha^2 + \cdots$, and pluging it in in the equation $\hat{f}(g(\alpha), \alpha) = 0$ we can get the coefficients. However, a much more simpler way is to simply ignore the higher order terms in $\hat{f}(x, \alpha)$ and get x in terms of α which will be $x = \frac{-a}{b}(\alpha - \alpha_0)$. This defines the second branch of equilibrium. To determine the stability of the equilibria branches, we need to determine the the sign of the vector field. The following figure shows this procedure for $a > 0$ and $b < 0$ case

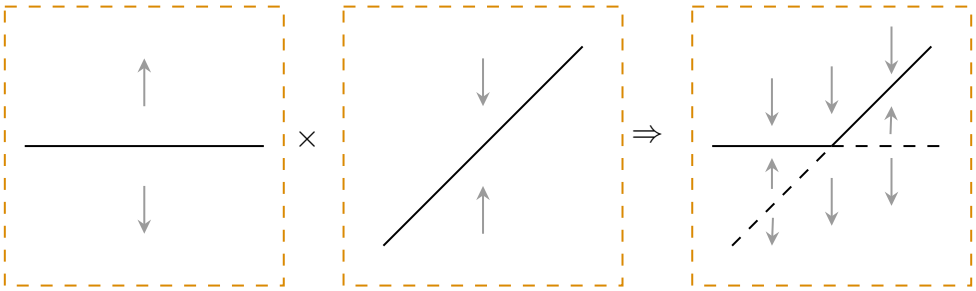
General Transcritical Bifurcation

Consider the following family of vector fields

$$\dot{x} = f(x, \alpha)$$

where $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^1$. Furthermore, assume that the vector field is constrained to be zero on the smooth curve $p_1^0(\alpha)$ for all α in some interval containing the special value α_0 and $p_1^0(\alpha_0) = 0$. I.e.

$$f(p_1^0(\alpha), \alpha) = 0, \quad \text{for all } \alpha \text{ in some interval containing } \alpha_0.$$



In addition, assume that the point $(0, \alpha_0)$ is a non-hyperbolic equilibrium point, i.e.

$$f_x(0, \alpha_0) = 0.$$

In this case, there is a chance that we observe a bifurcation, as it is possible that the linearization change sign as we pass α through α_0. To analyze the system further, we do the following change of variable

$$x = p_1^0(\alpha) + u.$$

By substituting in the ODE for x and expanding it we will get

$$\dot{x} = f(p_1^0(\alpha)+u, \alpha) = \underbrace{f(p_1^0(\alpha), \alpha)}_{=0} + \underbrace{f_x(p_1^0(\alpha), \alpha)}_{\lambda(\alpha)} u + \underbrace{\underbrace{1/2 f_{xx}(p_1^0(\alpha), \alpha)}_{\underbrace{1/2 f_{xx}(p_1^0(\alpha_0), \alpha_0)}_{b} + \mathcal{O}(|\alpha-\alpha_0|)}}_{} u^2 + \mathcal{O}(|u|^3).$$

And by expanding $\lambda(\alpha)$ we will get

$$\lambda(\alpha) = \underbrace{\lambda(\alpha_0)}_{=0} + \lambda'(\alpha_0)(\alpha - \alpha_0) + \mathcal{O}(|\alpha - \alpha_0|^2).$$

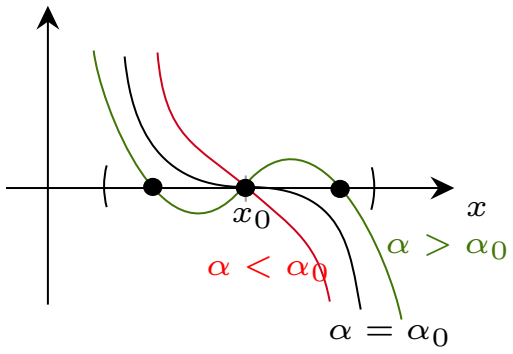
Furthermore, to have the transcritical bifurcation, we require $a = \lambda'(\alpha_0) \neq 0$, and by the chain rule, we can get a much more simpler expression for $\lambda'(\alpha_0)$

$$a = \lambda'(\alpha_0) = \frac{d}{d\alpha} f_x(p_1^0(\alpha), \alpha)\Big|_{\alpha=\alpha_0} = \underbrace{f_{xx}(0, \alpha_0)}_{2b} (p_1^0)'(\alpha_0) + f_{x\alpha}(0, \alpha_0) \neq 0.$$

Note that $(p_1^0)'(\alpha_0)$ is know to us because $p_1^0(\alpha_0)$ is already given (this is the smooth curve on which the vector field is constrained to be zero). Also, $a \neq 0$ means that the value of $\lambda(\alpha)$ crosses $\lambda(\alpha_0) = 0$ transversally.

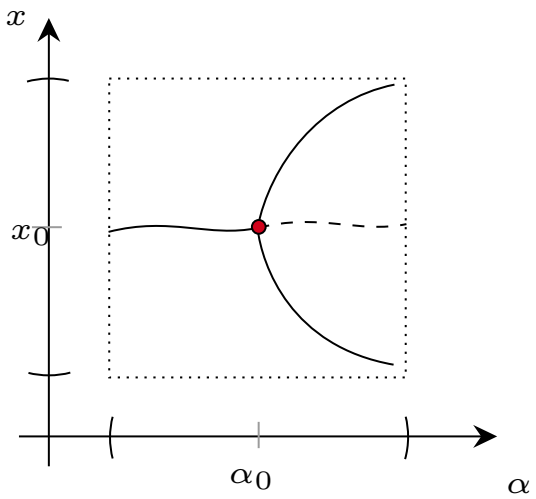
6.1.3 Pitchfork Bifurcation

In the pitchfork bifurcation, a stable equilibrium point turns into two stable equilibrium points separated by an unstable equilibrium point (in the case of the supercritical one). In this bifurcation, the vector field changes as depicted by the following figure.



The dependence of $f(x, \alpha)$ on the variable (x, α) has the form $f(x, \alpha) = \xi_\alpha \xi_x - \xi_x^3$, where $\xi_\alpha = \alpha - \alpha_0$ and $\xi_x = x - x_0$.

Also, the following figure shows that branching diagram for the pitchfork bifurcation.



Again, we can intuitively figure out what conditions the functions $f(x, \alpha)$ should satisfy in order to have this kind of bifurcation.

(i) $f(x - x_0, \alpha)$ be locally odd (in the neighborhood of x_0) with respect to x, i.e. $f(-(x - x_0), \alpha) = -f(x - x_0, \alpha)$ and $f(x_0, \alpha_0) = 0$. The local oddness also leads to

  (a) $f(x_0, \alpha) = 0$,

  (b) $f_{xx}(x, \alpha) = 0$, as well as any other even derivative.

(ii) $f_x(x_0, \alpha_0) = 0$,

(iii) $f_{xxx}(x_0, \alpha_0) \neq 0$,

(iv) $f_{\alpha x}(x_0, \alpha_0) \neq 0$.

More Theoretical Treatment of the Pitchfork Bifurcation

In this section we will derive the results above in a more analytic way. Assume we have the following family dynamical systems

$$\dot{x} = f(x, \alpha)$$

where $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^1$. Also assume that the vector field f is odd with respect to x. I.e.

$$f(-x, \alpha) = -f(x, \alpha).$$

This will immediately lead to

$$f(0, \alpha) = 0, \ f_{xx}(0, \alpha) = 0, \ f_{xxxx}(0, \alpha) = 0, \ \cdots .$$

Also, assume that for the special value of $\alpha = \alpha_0$ we have

$$f_x(0, \alpha_0) = 0.$$

In this case, there is a chance that we have a bifurcation (since changing the value of α through α_0 might change the sign of $f_x(0, \alpha)$ which will lead to a bifurcation.) Now we can expand the vector field around the bifurcation point as follows

$$f(x, \alpha) = \underbrace{f(0, \alpha)}_{=0} + \underbrace{f_x(0, \alpha)}_{=f_x(0,\alpha_0)+f_{x\alpha}(0,\alpha_0)(\alpha-\alpha_0)+\cdots} x + 1/6 \ \underbrace{f_{xxx}(0, \alpha)}_{=f_{xxx}(0,\alpha_0)+\cdots} x^3 + \mathcal{O}(|x^5|)$$

Now let $a = f_{x\alpha}(0, \alpha_0) \neq 0$ and $b = 1/6 f_{xxx}(0, \alpha_0) \neq 0$. So the vector field can be written like

$$f(x, \alpha) = a(\alpha - \alpha_0)x + bx^3 + \cdots = x(a(\alpha - \alpha_0) + bx^2 + \cdots) = x\hat{f}(x, \alpha),$$

in which the higher order terms do not change the topological behaviour. Also, from this expression of the vector field, it is clear that $x = 0$ is a branch of equilibria. We can find the other branch by letting $\hat{f}(x, \alpha) = 0$. We can do so by ignoring all of higher order terms and solve x in terms of α and find the second branch of equilibria. However, we want to be fancy and use implicit function theorem! It is clear from the expression for $\hat{f}(x, \alpha)$ that $\hat{f}_\alpha(0, \alpha_0) = a \neq 0$. Thus we can write $\alpha = g(x)$ where $g(0) = \alpha_0$ and also $\hat{f}(x, g(x)) = 0$ for α belonging to some interval containing the special value α_0. We can do such by the implicit function theorem. The implicit function theorem also states that the function g is smooth. Thus we can write its Taylor expansion as follows

$$g(x) = \underbrace{g(0)}_{=\alpha_0} + \underbrace{g'(0)}_{=c_1} x + \underbrace{g''(0)/2}_{=c_2} x^2 + \mathcal{O}(|x|^3).$$

We can substitute the expression above in $\hat{f}(x, g(x)) = a(g(x) - \alpha_0) + bx^2 + \cdots = 0$, and find the coefficients c_i. By doing so we will find

$$a(c_1 x + c_2 x^2 + \cdots) + bx^2 = 0 \implies c_1 = 0, \quad c_2 = \frac{-b}{a}, \cdots .$$

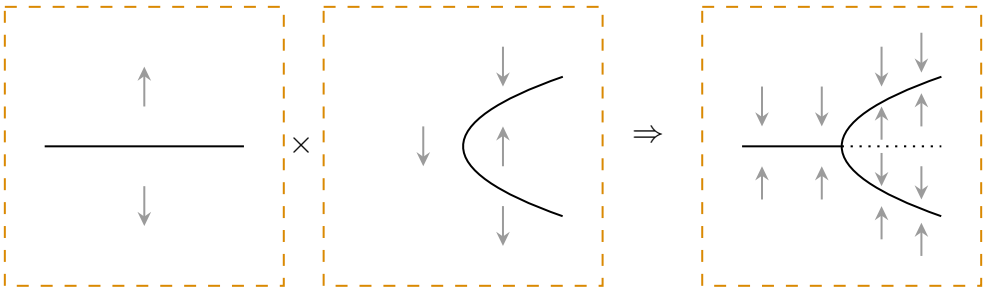
Thus the expression for $g(x)$ will be

$$\alpha = g(x) = \alpha_0 + \frac{-b}{a} x^2.$$

All of this trouble was to be fancy and solve $\hat{f}(x, \alpha) = 0$. The simpler way is to simply ignore the higher order terms in $\hat{f}(x, \alpha)$ and solve the equation $\hat{f}(x, \alpha) = 0$. I.e.

$$\hat{f}(x, \alpha) \approx a(\alpha - \alpha_0) + bx^2 \implies \alpha = \alpha_0 - \frac{b}{a} x^2.$$

This will define the second branch of equilibria. The next step is to determine the stability of these equilibrium points that can be done by determining the sign of the vector field around the equilibrium branches. For instance, fix $b < 0$ and $a > 0$. Then we will have



Note that the vertical axis represents x while the horizontal axis represents α. We can do the same kind of analysis on the sign of the vector field when a and b has other combination of signs.

Generalized Pitchfork Bifurcation

Consider the following family of vector fields

$$\dot{x} = f(x, \alpha),$$

where $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^1$. Also, assume that the vector field is constrained in such a way that on $p_1^0(\alpha)$, where $0 = p_1^0(\alpha_0)$, we have

$$f(p_1^0(\alpha), \alpha) = 0, \quad f_{xx}(p_1^0(\alpha), \alpha) = 0,$$

for all α in some interval containing the special value α_0. Thus, clearly $(p_1^0(\alpha), \alpha)$ is a branch of equilibrium, and in particular, $(0, \alpha_0)$ is an equilibrium point on this branch. Furthermore, assume that

$$f_x(0, \alpha_0) = 0,$$

which implies that $(0, \alpha_0)$ is a non-hyperbolic equilibrium point. Thus there is a chance that we observe a bifurcation as we pass the value of α through α_0. To analyze the system we use the following coordinate change.

$$x = p_1^0(\alpha) + u.$$

Substituting in the ODE for x we will have

$$(p_1^0)'(\alpha)\dot{\alpha} + \dot{u} = f(p_1^0(\alpha) + u, \alpha).$$

Now we can expand the vector field as

$$\dot{u} = \underbrace{f(p_1^0(\alpha), \alpha)}_{=0} + \underbrace{f_x(p_1^0(\alpha), \alpha)}_{\lambda(\alpha)} u + \underbrace{f_{xx}(p_1^0(\alpha), \alpha)}_{=0} u^2/2 + f_{xxx}(p_1^0(\alpha), \alpha)u^3/6 + \mathcal{O}(|u|^4).$$

Thus in short we will have

$$\dot{u} = \lambda(\alpha)u + \frac{1}{6} \underbrace{f_{xxx}(p_1^0(\alpha), \alpha)}_{f_{xxx}(0,\alpha_0)+\mathcal{O}(\alpha-\alpha_0)} u^3 + \mathcal{O}(|u|^4).$$

We require $b = 1/6 f_{xxx}(0, \alpha_0) \neq 0$. Also, we can expand $\lambda(\alpha)$ as

$$\lambda(\alpha) = \underbrace{\lambda(\alpha_0)}_{f_x(0,\alpha_0)=0} + \lambda'(\alpha_0)(\alpha - \alpha_0) + \mathcal{O}(|\alpha - \alpha_0|^2).$$

Further more, we require $a = \lambda'(\alpha_0) \neq 0$. By the chain rule we can write

$$a = \lambda'(\alpha_0) = \frac{d}{d\alpha} f_x(p_1^0(\alpha), \alpha)\Big|_{\alpha=\alpha_0} = \underbrace{f_{xx}(0, \alpha_0)}_{=0}(p_1^0)'(\alpha_0) + f_{x\alpha}(0, \alpha_0).$$

By these conditions ($a \neq 0$ and $b \neq 0$) we will have the simplified vector field as

$$\dot{u} = a(\alpha - \alpha_0)u + bu^3 + \mathcal{O}(|u|^4)$$

Thus as we can see, even in the generalized setting for the Pitchfork bifurcation, by the use of appropriate change of coordinates, we can still write the vector field very similar to the case where the vector field was a symmetric function.

6.1.4 Some General Notes about Pitchfork Bifurcation

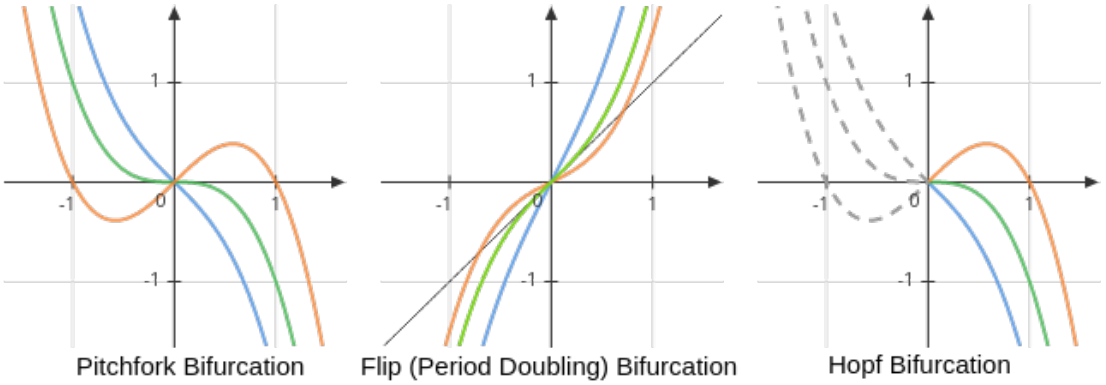
The pitchfork bifurcation is a very important bifurcation as it appears in many systems, and even more important than that, it is the bifurcation lying behind other bifurcations like the Hopf bifurcation, flop (period doubling bifurcation) for maps, etc. In fact, for the Hopf bifurcation we have

$$\begin{cases} \dot{r} = a\alpha r + br^3, \\ \dot{\theta} = w. \end{cases}$$

And for the period doubling bifurcation for maps, the normal form of the second iterate map is

$$y \mapsto y + a\alpha y + by^3.$$

In all of these situations, the curve $a\alpha y + by^3$ is crossing "something" in one way or another. The following figure makes this more clear. Further more, the interesting fact is that we can actually get the normal form of the pitchfork bifurcation by simply integrating (with respect to x) of the normal form of the fold bifurcation.



7 Local Bifurcation Theory for Maps

In this section we will develop our theory to study the bifurcation for the maps.

7.1 Flip or Period Doubling Bifurcation

Consider the following family of maps

$$x \mapsto f(x, \alpha)$$

where $x \in \mathbb{R}^1$, $\alpha \in \mathbb{R}^1$, and $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}$ is a smooth vector family of vector fields. Assume that at (p_0^0, α_0) we have

$$f(p_0^0, \alpha_0) = p_0^0, \qquad f_x(p_0^0, \alpha_0) = -1.$$

This means that the point p_0^0 is an equilibrium point that is non-hyperbolic. Then there is a high chance that some bifurcation is going to happen (since $f_x(p_0^0, \alpha_0) = -1$)). First, notice that there is a smooth branch of equilibria containing p_0^0. To show this, let $G(x, \alpha) = f(x, \alpha) - x$. Then it immediately follows that $G(p_0^0, \alpha_0) = 0$ and $G_x(p_0^0, \alpha_0) = -2 \neq 0$. Thus using the implicit function theorem, we can write $x = p^0(\alpha)$, where $p_0^0 = p^0(\alpha_0)$. Furthermore, we have

$$f(p^0(\alpha), \alpha) = p^0(\alpha)$$

since $p^0(\alpha)$ is actually a branch of equilibria. To determine the behavior of the system near the equilibrium points, we can do the following change of variable

$$x = p^0(\alpha) + u.$$

By substituting the equation above in the recursive form of the map we will get

$$\begin{aligned} p^0(\alpha) + u_{k+1} &= f(p^0(\alpha) + u, \alpha) \\ &= \underbrace{f(p^0(\alpha), \alpha)}_{p^0(\alpha)} + \underbrace{f_x(p^0(\alpha), \alpha)}_{\mu(\alpha)} u_k + \underbrace{1/2 f_{xx}(p^0(\alpha), \alpha)}_{\hat{f}_2(\alpha)} u_k^2 + \underbrace{1/6 f_{xxx}(p^0(\alpha), \alpha)}_{\hat{f}_3(\alpha)} u_k^3 + \mathcal{O}(|u_k|^4). \end{aligned}$$

So we can now write the map as

$$u \mapsto \mu(\alpha) u + \hat{f}_2(\alpha) u^2 + \hat{f}_3(\alpha) u^3 + \mathcal{O}(|u|^4).$$

First, notice that, in order to have the bifurcation, the value of $\mu(\alpha)$ should cross -1 at α_0 transversally (i.e. have values other than -1 at the points arbitrary close to α_0). Thus we need to have $-a = \mu'(\alpha_0) \neq 0$. The next step is a very crucial one. The following example makes this step much more sensible.

Example 7.1

Let $f(x) = x + ax^2 + bx^3 + cx^4$ where $a = -2.6, b = 6.8$, and $c = -10$. Define a new identity non-linear transformation as

$$x = y + dy^2 + gy^3 + hy^4 + jy^5 + ky^6 + ly^7 + my^8.$$

Find the coefficients of the transformation such that $f(y)$ has zero or small coefficients for the terms with lower order.

Answer.

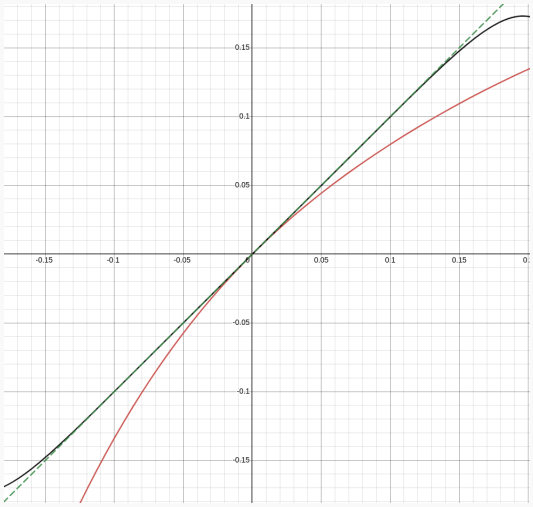
To find the answer, we simply substitute the transformation into $f(x)$ and then find the coefficients such that the terms with lower order (lower powers) be zero or be minimum. We can do so using symbolic tools in python or other programming languages. The final results for the coefficients will be

$$d = 2.6,\ g = 6.72,\ h = 9.48,\ j = 30.8,\ k = 266.1,\ l = 515.2,\ m = 5619.4$$

And the expression for $f(z)$ will be

$$\begin{aligned} f(z) = &- 9.9718 \times 10^{15} z^{32} + 3.6575 \times 10^{15} z^{31} + 1.3862 \times 10^{15} z^{30} \\ &- 2.7004 \times 10^{11} z^{29} - 2.1478 \times 10^{11} z^{28} - 3.1614 \times 10^{10} z^{27} \\ &+ 9.3739 \times 10^{9} z^{26} + 4.1132 \times 10^{9} z^{25} + 4.1196 \times 10^{9} z^{24} \\ &- 3.6990 \times 10^{8} z^{23} - 5.1383 \times 10^{8} z^{22} - 1.0269 \times 10^{8} z^{21} \\ &+ 1.1804 \times 10^{7} z^{20} + 1.1487 \times 10^{7} z^{19} + 2.9077 \times 10^{6} z^{18} \\ &+ 4.9065 \times 10^{5} z^{17} - 4.7806 \times 10^{5} z^{16} - 1.2066 \times 10^{5} z^{15} \\ &- 3.7331 \times 10^{3} z^{14} + 5.7317 \times 10^{3} z^{13} + 1.8873 \times 10^{3} z^{12} \\ &+ 239.716 z^{11} - 46.1615 z^{10} - 66.1136 z^{9} + z. \end{aligned}$$

In the figure below, we can easily see that the $f(y)$ is much more linear than $f(x)$ near the origin.



Back to our business with the map we were analyzing that now look like this:

$$u \mapsto \mu(\alpha)u + \hat{f}_2(\alpha)u^2 + \hat{f}_3(\alpha)u^3 + \mathcal{O}(|u|^4).$$

We do the following transformation

$$u = v + h_2(\alpha)v^2.$$

Now we substitute the expression for u and try to choose $h_2(\alpha)$ in a way that makes the coefficient of v^2 zero or very small. By substituting we will have

$$v_{k+1}+h_2(\alpha)v_{k+1}^2 = \mu(\alpha)(v_k+h_2(\alpha)v_k^2)+\hat{f}_2(\alpha)(v_k+h_2(\alpha)v_k^2)^2+\hat{f}_3(\alpha)(v_k+h_2(\alpha)v_k^2)^3+\cdots .$$

(assume $v_{k+1} = \mu(\alpha)v_k + \cdots$ to the leading order. I don't know the reason yet but I will update the notes as soon as I notice.) Then we can write

$$v_{k+1} = \mu(\alpha)v_k + (\mu(\alpha)h_2(\alpha) + \hat{f}_2(\alpha) - h_2(\alpha)\mu^2(\alpha))v_k^2 + (2\hat{f}_2(\alpha)h_2(\alpha) + \hat{f}_3(\alpha))v_k^3 + \cdots .$$

Thus simply by letting

$$h_2(\alpha) = \frac{\hat{f}_2(\alpha)}{\mu^2(\alpha) - \mu(\alpha)}$$

the coefficient of v_k^2 will be zero and the map will transform into

$$v \mapsto \mu(\alpha)v + g_3(\alpha)v^3 + \mathcal{O}(|v|^4),$$

where $-b = g_3(\alpha_0) = [\hat{f}_2(\alpha_0)]^2 + \hat{f}_3(\alpha_0)$. Now expanding the map around α_0 will result in

$$v \mapsto -(1 + a(\alpha - \alpha_0))v + g_3(\alpha_0)v^3 + \mathcal{O}(|v|^4 + |\alpha - \alpha_0|^2|v| + |\alpha - \alpha_0||v|^3).$$

In short, the vector field will be

$$v \mapsto -v - a\beta v - bv^3 + \cdots ,$$

where we used the transformation $\alpha = \beta + \alpha_0$.

8 Poincare Normal Form and Projection Method

The nonlinear close to identity transformation that we did above, is make the vector field more linear close to the origin and make the lower order coefficients of the expansion around the origin to be as small as possible. Note that all of the nonlinear close to identity transformations are one-to-one when we are sufficiently near the origin. Thus the transformation has inverse, from which we can retrieve the original values.

The nature of the non-linear close to identity transformation is quite messy. Thus we need to develop some theoretical setting in order to handle the mess of it. There are two ways to do that: Poincare-Birkhoff normal form theorem and the projection method. No surprise that we see the use of linear algebra to do some large scale computations pretty much automatically. We start with the Poincare-Birkhoff normal form theorem.

8.1 Poincare-Birkhoff Normal Form Theorem

As we explained above, this method is used to do some large scale calculations pretty much automatically by using the ideas form linear algebra.

8.1.1 Test

Let

$$\dot{z} = \lambda_1 z + \frac{1}{2} g_{20} z^2 + g_{11} z\overline{z} + \frac{1}{2} g_{02} \overline{z}^2 + \frac{1}{6} g_{30} z^3 + \frac{1}{2} g_{21} z^2 \overline{z} + \frac{1}{2} g_{12} z \overline{z}^2 + g_{03} \overline{z}^3 + O(\|z, \overline{z}\|^4)$$

Now do the following near identity change of variable

$$z = \xi + h_1 \xi^2 + h_2 \xi \overline{\xi} + h_3 \overline{\xi}^2 + h_4 \xi^3 + h_5 \xi^2 \overline{\xi} + h_6 \xi \overline{\xi}^2 + h_7 \overline{\xi}^3$$

and choose the coefficients h_i such that there are no second order terms in the expression for $\dot{\xi}$, and also determine the third order coefficients. Note that all of the variables are complex.