Ordinary Differential Equations.

Valeriy Slastikov

Spring 2008

1 Existence and uniqueness

Notation: $\mathbf{x} = (x_1, ..., x_n)$ is a vector in $\mathbf{R}^n$.

In this section we are going to study the solutions of the following **Initial Value Problem (IVP)**:

Given an open set $D \subset \mathbf{R} \times \mathbf{R}^n$, a continuous function $\mathbf{f} : D \to \mathbf{R}^n$, and a point $(t_0, \mathbf{x}_0) \in D$, find a solution of the following equation

$$\mathbf{x}'(t) = \mathbf{f}(t, \mathbf{x}(t)) \tag{1}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \tag{2}$$

The solution of (1), (2) is defined as

Definition 1 *Let $I \subset \mathbf{R}$ be an open interval and $t_0 \in I$. We call a function $\mathbf{x} : I \to \mathbf{R}^n$ a solution of IVP (1), (2) if*

1. $\mathbf{x}(t_0) = \mathbf{x}_0$*;*
2. $(t, \mathbf{x}(t)) \in D$ *for all* $t \in I$*;*
3. $\mathbf{x}'(t) = \mathbf{f}(t, \mathbf{x}(t))$ *for all* $t \in I$*.*

There are several natural questions:

- Does any solution of IVP exist?
- If a solution exists, is it unique?
- Is it possible to extend a solution to a bigger interval $J \supset I$?

In what follows below we will show that under some additional assumptions on $\mathbf{f}(t, \mathbf{x})$ the solution of IVP exists and is unique.

Let us reformulate problem (1), (2) in the integral terms.

Lemma 1 *Let us assume that $I \subset \mathbf{R}$ be an open interval with $t_0 \in I$, $D \subset \mathbf{R} \times \mathbf{R}^n$ be an open set, $\mathbf{f} : D \to \mathbf{R}^n$ be a continuous function and $\mathbf{x} : I \to \mathbf{R}^n$ be continuous* *function that satisfies* $(t, \mathbf{x}(t)) \in D$) *for all* $t \in I$. *Then* $\mathbf{x}(t)$ *is a solution of IVP (1), (2) if and only if for all* $t \in I$

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s)) ds. \tag{3}$$

Proof 1 *If* $\mathbf{x}(t)$ *is a solution of IVP (1), (2), then it is an antiderivative of* $\mathbf{f}(t, \mathbf{x}(t))$ *on* I *and therefore by Fundamental Theorem of Calculus*

$$\mathbf{x}(t) - \mathbf{x}(t_0) = \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s)) ds.$$

Since $\mathbf{x}(t_0) = \mathbf{x}_0$ *we obtain (3).*

Now, let us assume (3) is satisfied. Then obviously we have $\mathbf{x}(t_0) = \mathbf{x}_0$*; by Fundamental Theorem of Calculus we know that* $\mathbf{x}(t)$ *is differentiable and* $\mathbf{x}'(t) = \mathbf{f}(t, \mathbf{x}(t))$ *for all* $t \in I$. *The lemma is proved.*

Now we are ready to prove local existence and uniqueness theorem for IVP. In order to do this we will assume that $D_{\mathbf{x}}\mathbf{f}$ is continuous function on D ($D_{\mathbf{x}}\mathbf{f} = \{\frac{\partial f_i}{\partial x_j}\}_{i,j=1}^{n}$).

Theorem 1 *Let* $D \subset \mathbf{R} \times \mathbf{R}^n$ *be an open set,* $\mathbf{f} : D \to \mathbf{R}^n$ *and* $D_{\mathbf{x}}\mathbf{f}$ *be continuous functions on* D *and* $(t_0, \mathbf{x}_0) \in D$. *Let* $P \subset D$ *be a closed rectangle defined as*

$$P = \{(t, \mathbf{x}) : |t - t_0| \le a, |x - x_0| \le b\},$$

where $a, b > 0$. *Let us define* $A = \min\{a, b/M\}$ *with* $M = \max_{(t,\mathbf{x}) \in P} |\mathbf{f}(t, \mathbf{x})|$, $K = \max_{(t,\mathbf{x}) \in P} |D_{\mathbf{x}}\mathbf{f}(t, \mathbf{x})|$. *Then there exists a function* $\mathbf{x}(t)$ *defined on the interval* $J = [t_0 - A, t_0 + A]$ *that satisfies integral equation (3) for all* $t \in J$.

Proof 2 *We are going to construct a sequence of functions* $\{\mathbf{x}_m(t)\}_{m=0}^{\infty}$ *with* $\mathbf{x}_m : J \to \mathbf{R}^n$ *as*

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

$$\mathbf{x}_m(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_{m-1}(s)) ds, \quad m > 0.$$

We will show that $\mathbf{x}_m(t) \to \mathbf{x}(t)$ *in some sense and* $\mathbf{x}(t)$ *satisfies (3).*

Step 1. *Here we show that a sequence* $\{\mathbf{x}_m\}$ *is well defined, meaning that* $(t, \mathbf{x}_m(t)) \in P$ *for all* $t \in J$. *Let's check it first for* $m = 0$*: since* $t \in J$ *we have* $|t - t_0| \le A \le a$, *we also have* $|\mathbf{x}_0(t) - \mathbf{x}_0| = 0 < b$. *Therefore* $(t, \mathbf{x}_0(t)) \in P$.

Now let's check for any m *using induction: assume* $(t, \mathbf{x}_{m-1}(t)) \in P$ *and show this for* $\mathbf{x}_m$. *Since* $t \in J$ *we have* $|t - t_0| \le A \le a$, *we also have*

$$|\mathbf{x}_m(t) - \mathbf{x}_0| = \left| \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_{m-1}(s)) ds \right|.$$

Since $(t, \mathbf{x}_{m-1}(t)) \in P$ *we know that* $|\mathbf{f}(t, \mathbf{x}_{m-1}(t))| \le M$ *and therefore*

$$\left| \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_{m-1}(s)) ds \right| \le M|t - t_0| \le MA \le b.$$

This implies $(t, \mathbf{x}_m(t)) \in P$.

Step 2. *Now we want to show that* $\mathbf{x}_m(t)$ *converges. We first show that*

$$|\mathbf{x}_m(t) - \mathbf{x}_{m-1}(t)| \le M|t-t_0|^m \frac{K^{m-1}}{m!}. \tag{4}$$

The proof of this fact is again by induction. For $m = 1$ *we obviously have*

$$|\mathbf{x}_1(t) - \mathbf{x}_0(t)| = \left| \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_0) ds \right| \le M|t-t_0|.$$

Let's assume the inequality is true for $m-1$ *and show it for* m*:*

$$|\mathbf{x}_m(t) - \mathbf{x}_{m-1}(t)| \le \left| \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_{m-1}(s)) - \mathbf{f}(s, \mathbf{x}_m(s)) ds \right|.$$

For any $(t,y),(t,z) \in P$ *we have*

$$|\mathbf{f}(t,y) - \mathbf{f}(t,z)| = \left| \int_0^1 \frac{\partial \mathbf{f}}{\partial s}(t, y + s(z-y)) ds \right| \le K|y-z|.$$

Using this inequality we obtain:

$$|\mathbf{x}_m(t) - \mathbf{x}_{m-1}(t)| \le \left| \int_{t_0}^{t} K|\mathbf{x}_{m-1}(s) - \mathbf{x}_{m-2}(s)| ds \right| \le$$

$$\le KM \frac{K^{m-2}}{(m-1)!} \left| \int_{t_0}^{t} |s-t_0|^{m-1} ds \right| = M|t-t_0|^m \frac{K^{m-1}}{m!}.$$

This completes the induction.

Now we define the function

$$\mathbf{x}(t) = \mathbf{x}_0(t) + \sum_{k=0}^{\infty} (\mathbf{x}_{k+1}(t) - \mathbf{x}_k(t)).$$

Since the series converges uniformly on J *(using the previous inequality (4)) the function* $\mathbf{x}(t)$ *is well defined. Moreover we know that*

$$\mathbf{x}_m(t) = \mathbf{x}_0(t) + \sum_{k=0}^{m-1} (\mathbf{x}_{k+1}(t) - \mathbf{x}_k(t))$$

and therefore using uniform convergence of the series we obtain $\mathbf{x}_m(t) \to \mathbf{x}(t)$ *uniformly on* J.

Step 3. *Here we show that* $\mathbf{x}(t)$ *is a solution of (3) on* J*. First of all, it is easy to show that* $(t, \mathbf{x}(t)) \in P$*. Since* $t \in J$ *and* $|\mathbf{x}(t) - \mathbf{x}_0| \le b$ *for* $t \in J$ *and* $\mathbf{x}_m(t) \to \mathbf{x}(t)$ *we see that* $|\mathbf{x}(t) - \mathbf{x}_0| \le b$*, therefore* $(t, \mathbf{x}(t)) \in P$.

Now, since $\mathbf{x}_m(t) \to \mathbf{x}(t)$ *uniformly on* J *and* f *is continuous on* P*, we know that* $\mathbf{f}(t, \mathbf{x}_m(t)) \to \mathbf{f}(t, \mathbf{x}(t))$ *uniformly on* J*. Therefore passing to the limit as* $m \to \infty$ *in*

$$\mathbf{x}_m(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}_{m-1}(s)) ds,$$

we obtain

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{x}(s)) ds.$$

The theorem is proved.

So far we showed that there exists at least one solution of IVP (1), (2) near point t_0. We would like to have only one solution close to t_0. Let's show that this is the case.

Lemma 2 *If we are in the assumptions of the previous theorem, there exists only one solution of the integral problem (3) on the interval* $L = [t_0 - B, t_0 + B]$, *where* $B = \min\{A, \frac{\alpha}{K}\}$, *for some* $0 < \alpha < 1$.

Proof 3 *Since* $B \leq A$ *we know that there exists at least one solution on* L. *Suppose there are two solutions on* L, *call them* $\mathbf{x}(t)$ *and* $\tilde{\mathbf{x}}(t)$. *Taking a difference between these solutions we obtain*

$$|\mathbf{x}(t) - \tilde{\mathbf{x}}(t)| = \left| \int_{t_0}^{t} (\mathbf{f}(s, \mathbf{x}(s)) - \mathbf{f}(s, \tilde{\mathbf{x}}(s)))ds \right| \leq$$

$$\leq K \left| \int_{t_0}^{t} |\mathbf{x}(s) - \tilde{\mathbf{x}}(s)|ds \right| \leq BK \max_{s\in L} ||\mathbf{x}(s) - \tilde{\mathbf{x}}(s)|.$$

Taking maximum over all $t \in L$ *from both parts and recalling that* $BK \leq \alpha < 1$ *we obtain*

$$\max_{t\in L} ||\mathbf{x}(t) - \tilde{\mathbf{x}}(t)| < \max_{s\in L} ||\mathbf{x}(s) - \tilde{\mathbf{x}}(s)|.$$

This is a contradiction, therefore there is only one solution on L.

In the proof of this theorem we used the continuity of the derivative $D_x\mathbf{f}$. This condition is very important for uniqueness. It is possible to show existence of the solution using just continuity of f but for uniqueness you need additional assumptions, like, for instance, continuity of $D_x\mathbf{f}$.

Example of non-unique solution if $D_x\mathbf{f}$ is not continuous.

Let's consider the following IVP:

$$x' = \sqrt{x}, \quad x(1) = 0.$$

By assumptions of uniqueness lemma, in order to have unique solution of this IVP we need $f(t, x)$ to be continuous and differentiable around point (t_0, x_0). In our case $(t_0, x_0) = (1, 0)$, $f(t, x) = \sqrt{x}$ and we see that $\frac{df}{dx}$ is singular at $x = 0$. Let's see if this is important.

Obviously $x(t) = 0$ is a solution of this IVP, separating variables and solving this equation we may find another solution $\tilde{x}(t) = \frac{(t-1)^2}{4}$. Therefore for any interval L with $1 \in L$, we have at least 2 solutions of this IVP. Actually, any function

$$\bar{x}(t) = \begin{cases} 0, & t \leq C \\ \frac{(t-C)^2}{4}, & t > C \end{cases}$$

with constant $C \geq 1$ is a solution of this problem.

This example shows that continuity of $D_x\mathbf{f}$ is an important assumption for uniqueness of the solution.

2 First order differential equations

In this section we will study ODEs where $x(t)$ is a scalar function, not a vector. These are the simplest possible equations we may consider and yet we will see that analytical solutions are not always possible. We will start with some types of first order ODEs that we may solve analytically and then we will proceed to approximate and numerical techniques.

2.1 Linear first order ODEs.

Definition 2 *The differential equation is called linear if it can be represented in the form $x' = p(t)x + q(t)$ for some continuous functions $p(t)$ and $q(t)$.*

We consider the following IVP

$$x' = p(t)x + q(t), \quad x(t_0) = x_0 \tag{5}$$

We assume that $p(t)$ and $q(t)$ are continuous functions on some interval (a, b). In order to solve this problem we will use a method of integrating factor.

Let's call $P(t) = \int_{t_0}^{t} p(s)ds$ and multiply both parts of ODE in (5) by $e^{-P(t)}$. We obtain

$$e^{-P(t)}(x' - p(t)x) = e^{-P(t)}q(t).$$

We notice that

$$e^{-P(t)}(x' - p(t)x) = \left(e^{-P(t)}x\right)',$$

therefore we now obtain

$$\left(e^{-P(t)}x\right)' = e^{-P(t)}q(t).$$

Integrating both parts from t_0 to t and recalling that $P(t_0) = 0$ we have

$$e^{-P(t)}x - x(t_0) = \int_{t_0}^{t} e^{-P(s)}q(s)ds,$$

and therefore

$$x(t) = x_0 e^{P(t)} + e^{P(t)} \int_{t_0}^{t} e^{-P(s)}q(s)ds \tag{6}$$

Obviously this solution is defined on the whole interval (a, b) and it is unique on this interval. We just proved the following result.

Theorem 2 *Let $p(t)$ and $q(t)$ be continuous functions on the interval (a, b) and $t_0 \in (a, b)$. Then there exists unique solution of IVP (5) defined on the whole interval (a, b). Moreover it can be found by formula (6).*

Example. Solve the following IVP

$$x' = x + e^t, \quad x(0) = 0$$

In order to solve this problem we may do all the steps we did before or just use (6):

$$p(t) = 1, \quad q(t) = e^t, \quad P(t) = \int_0^t 1ds = t$$

$$x(t) = e^t \int_0^t e^{-s}e^s ds = te^t.$$

2.2 Separable equations

Definition 3 *The differential equation is called separable if it can be represented in the following form*

$$x' = h(t)g(x)$$

for some functions $h(t)$ depending **only** *on t and $g(x)$ depending* **only** *on x.*

If functions $h(t)$, $g(x)$ and $\frac{dg}{dx}(x)$ are continuous on (a, b) then by existence and uniqueness theorem we may show that there exists unique solution of the following IVP

$$x' = h(t)g(x), \quad x(t_0) = x_0 \tag{7}$$

on some interval I, containing t_0.

It is quite easy to solve this IVP: by putting all the functions depending on x to the rleft and the functions depending on t to the right, we obtain

$$\frac{dx}{g(x)} = h(t)dt$$

Now integrating left part from x_0 to x and right part from t_0 to t we obtain

$$\int_{x_0}^{x} \frac{dy}{g(y)} = \int_{t_0}^{t} h(s)ds \tag{8}$$

This is implicit form of the solution, where x depends on t through the above formula.

Exapmle. Let's consider the following IVP

$$x' = xt, \quad x(0) = 1.$$

separating variables we obtain $\int_1^x \frac{dy}{y} = \int_0^t sds$, this provides

$$\ln(x) = \frac{t^2}{2}$$

hence $x(t) = e^{t^2/2}$.

It is not always possible to integrate both parts of (8), consider, for instance, the following IVP

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

This means that we should either leave the solution in the integral form or to find an approximate solution.

It is not always easy to see if the equation is separable. We formulate the important result about separable equations

Theorem 3 *The differential equation $x' = f(t, x)$ is separable if and only if*

$$f(t, x)\frac{\partial^2 f(t, x)}{\partial x \partial t} = \frac{\partial f(t, x)}{\partial t}\frac{\partial f(t, x)}{\partial x}$$

Proof 4 *Proof is left as an exercise.*

2.3 Homogeneous equations.

Definition 4 *A differential equation is called homogeneous if it can be written in the form* $x' = f(\frac{x}{t})$

It is quite easy to reduce a homogeneous ODE to a separable ODE using the following change of variables: $z = \frac{x}{t}$. Then $x = tz$ and

$$x' = tz' + z = f(z)$$

and we obtain a simple equation for z:

$$z' = \frac{f(z) - z}{t}$$

Solving this one by separation of variables we obtain a solution to homogeneous equation.

Example Find the family of solutions for the following equation

$$x' = \frac{x^2 + tx}{t^2}$$

We notice that this equation is homogeneous

$$x' = \left(\frac{x}{t}\right)^2 + \frac{x}{t}$$

Using a change of variables $z = \frac{x}{t}$ we obtain the equation for z

$$tz' = z^2$$

Separating variables and integrating we have

$$-\frac{1}{z} = \ln|t| + C_1$$

Hence $z = Ce^{-1/t}$ and $x = Cte^{-1/t}$.

2.4 Exact equations.

Definition 5 *A differential equation that can be written in the form $d\phi(t,x) = 0$ for some continuous and differentiable function $\phi(t,x)$ is called exact. The one parameter family of solutions of this equation is $\phi(t,x) = C$ (the constant C is arbitrary and for IVP is determined from initial conditions). Other form of exact equation is*

$$\frac{\partial \phi}{\partial t}dt + \frac{\partial \phi}{\partial x}dx = 0 \tag{9}$$

or

$$\frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x}x' = 0$$

We see that our usual form of ODE $x' = f(t,x)$ can be quite easily put in the following form:

$$M(t,x)dt + N(t,x)dx = 0 \tag{10}$$

by rewriting $f(t,x) = -\frac{M(t,x)}{N(t,x)}$ (note that this representation is obviously not unique).

If we can find a *potential function* $\phi(t,x)$ such that

$$M(t,x) = \frac{\partial \phi}{\partial t}, \quad N(t,x) = \frac{\partial \phi}{\partial x}$$

then by the formula for full differential of $\phi(t,x)$ equation (10) becomes exact

$$d\phi(t,x) = \frac{\partial \phi}{\partial t}dt + \frac{\partial \phi}{\partial x}dx = 0.$$

Since there are a lot of ways to put ODE in the form (10), in general, it is not easy to recognize exact equation. However we may prove the following result.

Theorem 4 *A differential equation*

$$M(t,x)dt + N(t,x)dx = 0$$

is exact if and only if

$$\frac{\partial M(t,x)}{\partial x} = \frac{\partial N(t,x)}{\partial t} \tag{11}$$

and the solution of this equation is $\phi(t,x) = C$, where $M(t,x) = \frac{\partial \phi}{\partial t}$ and $N(t,x) = \frac{\partial \phi}{\partial x}$.

Proof 5 *Let's assume that equation $M(t,x)dt + N(t,x)dx = 0$ is exact, then by definition there exists a function $\phi(t,x)$ such that $M(t,x) = \frac{\partial \phi}{\partial t}$ and $N(t,x) = \frac{\partial \phi}{\partial x}$. Therefore*

$$\frac{\partial M(t,x)}{\partial x} = \frac{\partial}{\partial x}\frac{\partial \phi}{\partial t} = \frac{\partial}{\partial t}\frac{\partial \phi}{\partial x} = \frac{\partial N(t,x)}{\partial t}$$

and this shows (11).

Let's prove converse now. Assume that (11) holds, in order to prove that (10) is exact we have to find a function $\phi(t,x)$ such that $M(t,x) = \frac{\partial \phi}{\partial t}$ and $N(t,x) = \frac{\partial \phi}{\partial x}$. We are going to construct this function as follows.

$$0 = \frac{\partial N(t,x)}{\partial t} - \frac{\partial M(t,x)}{\partial x} = \frac{\partial N(t,x)}{\partial t} - \frac{\partial}{\partial t}\int_t \frac{\partial M(s,x)}{\partial x}ds.$$

Now we get

$$\frac{\partial}{\partial t}\left(N(t,x)-\int_t \frac{\partial M(s,x)}{\partial x}ds\right)=0$$

and this implies that expression in brackets is independent of variable t*. Therefore we may find some function* $G(x)$ *such that*

$$G(x)=N(t,x)-\int_t \frac{\partial M(s,x)}{\partial x}ds$$

Finding antiderivative $h(x)$ *of* $G(x)$ *(*$h'(x) = G(x)$*) and changing the order in integration-differentiation we may show that*

$$N(t,x)=\frac{\partial}{\partial x}\int_t M(s,x)ds+h'(x)=\frac{\partial}{\partial x}\left(\int_t M(s,x)ds+h(x)\right).$$

Since we want to find $\phi(t,x)$ *such that* $N(t,x)=\frac{\partial\phi}{\partial x}$*, we define*

$$\phi(t,x)=\int_t M(s,x)ds+h(x).$$

It is easy to check that with this definition $M(t,x)=\frac{\partial\phi}{\partial t}$ *and* $N(t,x)=\frac{\partial\phi}{\partial x}$*, therefore our equation becomes*

$$\frac{\partial\phi}{\partial t}dt+\frac{\partial\phi}{\partial x}dx=0$$

and it is exact. Theorem is proved.

This theorem not just tells us when equation is exact, but also how to construct the solution of exact equation. Let's put it in a step by step algorithm.

Construction of solution for exact equation $M(t,x)dt+N(t,x)dx=0$**.**

1. Check that equation is exact, i.e. verify that $\frac{\partial M(t,x)}{\partial x}=\frac{\partial N(t,x)}{\partial t}$.
2. Since $M(t,x)=\frac{\partial\phi}{\partial t}$, define $\phi(t,x)=\int_t M(s,x)ds+h(x)$. Here $h(x)$ is unknown function and we determine it in the next steps.
3. Using the above function $\phi(t,x)$, calculate $\frac{\partial\phi}{\partial x}$.
4. Since we know that $N(t,x)=\frac{\partial\phi}{\partial x}$, equate $N(t,x)$ and the expression from Step 3. From this you will be able to find a simple differential equation for $h(x)$.
5. Solve the differential equation for $h(x)$ and obtain $h(x)$ (additive constant does not play any role here).
6. Substitute the expression for $h(x)$ to the expression for $\phi(t,x)$ from Step 2 and obtain $\phi(t,x)$.
7. The solution of exact equation $M(t,x)dt+N(t,x)dx=0$ is $\phi(t,x)=const$.

Example. Solve the following IVP

$$(2t+3x)dt+(3t+2x)dx=0,\quad x(1)=1$$

The equation is already in the form $M(t,x)dt+N(t,x)dx=0$, with $M(t,x)=2t+3x$ and $N(t,x)=3t+2x$. Let's check if it is exact: $\frac{\partial M(t,x)}{\partial x}=3=\frac{\partial N(t,x)}{\partial t}$, therefore this equation is exact and we may solve it using our algorithm.

- We define $\phi(t,x) = \int_t (2s + 3x)ds + h(x) = t^2 + 3xt + h(x)$.
- We find $\frac{\partial \phi}{\partial x} = 3t + h'(x)$.
- Since $N(t,x) = \frac{\partial \phi}{\partial x}$, we have $3t + 2x = 3t + h'(x)$ and we obtain equation for $h(x)$:
$$h'(x) = 2x$$
- Solving equation for $h(x)$ we obtain $h(x) = x^2$.
- Plugging $h(x)$ to $\phi(t,x)$ we have $\phi(t,x) = t^2 + 3tx + x^2$.
- The solution of the equation is $\phi(t,x) = t^2 + 3tx + x^2 = C$, since we want to solve IVP, we have to determine C using initial conditions: $C = 1^2 + 31 \cdot 1 + 1^2 = 5$.

The solution of the IVP is $t^2 + 3tx + x^2 = 5$. It is possible to find out x as a function of t from this expression, but it is not necessary.

2.5 Integrating Factors

Sometimes the equation $M(t,x)dt + N(t,x)dx = 0$ may be not exact, but if we multiply both parts of it by some factor $\mu(t,x)$ we may obtain exact equation

$$\mu(t,x)M(t,x)dt + \mu(t,x)N(t,x)dx = 0 \tag{12}$$

Let's consider a simple example

Example. Find a solution of $(x^2+1)dt + txdx = 0$. Let's check if equation is exact: $\frac{\partial M(t,x)}{\partial x} = 2x$, $\frac{\partial N(t,x)}{\partial t} = x$ and we see that they are different, so equation is not exact. Let's try to correct it by multiplying both parts of the equation by t. Our new equation is
$$(x^2+1)tdt + t^2xdx = 0$$
It is obviously equivalent to the original one (i.e. it has the same solutions). Let's check if it is exact: $\frac{\partial M(t,x)}{\partial x} = 2xt = \frac{\partial N(t,x)}{\partial t}$. So this one is exact and we can solve it, but since it has the same solutions as the previous one, we thus solve our original equation. So, let's solve it using our algorithm:

- $\phi(t,x) = \int_t s(x^2+1)ds + h(x) = \frac{t^2}{2}(x^2+1) + h(x)$;
- $\frac{\partial \phi}{\partial x} = xt^2 + h'(x)$;
- $N(t,x) = t^2x = xt^2 + h'(x)$;
- $h'(x) = 0$ implies $h(x) = const$;
- $\phi(t,x) = \frac{t^2}{2}(x^2+1)$ (constant is not important here);
- $\frac{t^2}{2}(x^2+1) = C$ is a solution of ODE.

We see that this simple trick of multiplying both parts of ODE by t allows us to solve the equation that originally was not exact.

Let's try to understand what we can do to find the function $\mu(t,x)$. We call

$$m(t,x) = \mu(t,x)M(t,x), \quad n(t,x) = \mu(t,x)N(t,x)$$

and want equation (12) to be exact. Therefore, by the previous result, we must have

$$\frac{\partial m(t,x)}{\partial x} = \frac{\partial n(t,x)}{\partial t}$$

and this implies

$$\frac{\partial \mu}{\partial x}M + \mu\frac{\partial M}{\partial x} = \frac{\partial \mu}{\partial t}N + \mu\frac{\partial N}{\partial t}$$

In general it is very difficult to solve this partial differential equation, however we may find some very specific cases:

1. Let's rewrite the above expression in the form

$$\frac{\partial \mu}{\partial x}\frac{M}{N} - \frac{\partial \mu}{\partial t} = \frac{\mu}{N}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$$

and assume that $Q(t) = \frac{1}{N}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$ is a function of t only. Then we may look for μ as a function of one variable t and obtain the following ODE

$$\mu'(t) = -Q(t)\mu$$

and this gives $\mu(t) = e^{-\int_t Q(s)ds}$. Therefore we found an integrating factor μ.

2. By the same token if we rewrite it as

$$\frac{\partial \mu}{\partial x} - \frac{\partial \mu}{\partial t}\frac{N}{M} = \frac{\mu}{M}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$$

and assume that $P(x) = \frac{1}{M}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$ is a function of x only, we may look for μ as a function of x variable and obtain

$$\mu'(x) = P(x)\mu$$

Therefore integrating factor in this case is $\mu(x) = e^{\int_x P(s)ds}$.

There are other cases when you may find the integrating factor, but the idea is always the same: you want μ to be a function of one variable α (for instance α may be yt, y/t, t/y); you express partial derivatives in the left hand side in terms of the derivative with respect to this variable α; then combine terms in such a way that $\mu'(\alpha)$ is in the left side and the rest is in the right hand side; assume that expression in the right hand side depends only on α and call the coefficient in front of μ as $Q(\alpha)$; solve ODE and find μ. This is your integrating factor.

Obviously this method will work just for particular equations, because in 1, for example, $\frac{1}{N}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$ should be a function of t only, and if this is not satisfied then the procedure fails. Therefore when solving equation

$$M(t,x)dt + N(t,x)dx = 0$$

that is not exact, one has to check if $\frac{1}{N}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$ is a function of t only (or if $\frac{1}{M}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right)$ is a function of x only) and if this is true, then obtain the integrating factor as in 1 (or in 2).

Example Let's solve $xdt + (2t - xe^x)dx = 0$. We check if equation is exact

$$\frac{\partial N}{\partial t} = 2, \quad \frac{\partial M}{\partial x} = 1$$

So this equation is not exact. Let's check if we can find an integrating factor

$$\frac{1}{M}\left(\frac{\partial N}{\partial t} - \frac{\partial M}{\partial x}\right) = \frac{1}{x}$$

Therefore by 2 $\mu(x)$ is a function of x only and it is $\mu(x) = e^{\int_x \frac{1}{s} ds} = x$. We multiply equation by x and see that it becomes exact. Now using the procedure for finding solutions for exact equations we may obtain the solution

$$\phi(t, x) = x^2 t + e^x(2 + 2x - x^2) = const$$

2.6 Bernoulli equation

Bernoulli equation has the following form

$$x' = p(t)x + q(t)x^n, \quad n \neq 0, 1 \tag{13}$$

As you may see it is nonlinear equation, however we may find a substitution that makes it linear. To solve this equation we substitute

$$y = x^{1-n}$$

Then we have $y' = (1-n)x^{-n}x'$, recalling the equation for x' we have

$$y' = (1-n)x^{-n}(p(t)x + q(t)x^n) = (1-n)p(t)y + (1-n)q(t)$$

This is a linear equation and we already know how to solve it.

Example. Solve the following IVP

$$y' = y + e^t y^2, \quad y(0) = 1$$

We recognize Bernoulli equation and make a substitution $z = y^{-1}$. Therefore

$$z' = -z - e^t, \quad z(0) = 1$$

The solution of this linear equation is

$$z(t) = e^{-t} + e^{-t} \int_0^t e^{2s} ds = \frac{1}{2}(e^t + e^{-t}).$$

Therefore $y(t) = \frac{2}{e^t + e^{-t}}$.

2.7 Autonomous equations

In this section we will study the important class of first order equations in which the independent variable t does not appear in the right hand side

$$\frac{dx}{dt} = f(x) \tag{14}$$

Such equations are called autonomous. As it is easy to see, we can separate variables and find the solution by integrating both parts

$$t - t_0 = \int_{x_0}^{x} \frac{dx}{f(x)},$$

but that's not what we want to do here. We would like to study a qualitative behavior of (14).

Definition 6 *A constant solution of (14) is called an equilibrium (or a critical point). It's easy to see that equilibria are roots of $f(x)$.*

Example. Let's find equilibria of

$$x' = x^2 - 1.$$

We solve $x^2 - 1 = 0$ and obtain $x = 1$ and $x = -1$ these are critical points of the above ODE.

It is easy to show that the following theorem is true.

Theorem 5 *If $f(x)$ is continuous and has continuous derivative then there exists unique solution of*

$$x' = f(x), \quad x(t_0) = x_0 \tag{15}$$

defined on some interval (a, b) ($t_0 \in (a, b)$).

This theorem tells us that there is only one solution of autonomous equation with some initial conditions. We are interested in finding the behavior of this solution for arbitrary initial conditions.

Theorem 6 *If $f(x)$ is continuously differentiable then all non-equilibrium solutions of (14) are strictly monotone. In particular, the solution of IVP (15) is an equilibrium if $f(x_0) = 0$, strictly increasing if $f(x_0) > 0$ and strictly decreasing if $f(x_0) < 0$.*

Proof 6 *Let $x(t)$ be a non-equilibrium solution on the maximal interval (a, b). Obviously the solution $x(t)$ can not be equal to any equilibrium solution x_e at any point. Since if at point t_* $x(t_*) = x_e$ by taking t_* as initial point and x_e as initial condition we obtain tow solutions to (15) and that is impossible. Therefore we see that $f(x(t))$ does not change sign on (a, b) and it's either always positive or always negative. This implies that $x'(t) = f(x(t)) > 0$ on (a, b) and $x(t)$ is strictly increasing function or $x'(t) = f(x(t)) < 0$ on (a, b) and $x(t)$ is strictly decreasing function .*

Example. We consider the following equation

$$x' = x^2 - 1.$$

There are tow critical points $x = -1$ and $x = 1$. We may split the whole interval in five subintervals $\mathbf{R} = (-\infty, -1) \cup -1 \cup (-1, 1) \cup 1 \cup (1, \infty)$. If our initial condition $x_0 \in (-\infty, -1) \cup (1, \infty)$ we obviously have $f(x_0) > 0$ and therefore $x(t)$ is increasing. If $x_0 = -1$ or $x_0 = 1$ then $x(t) = x_0$ all the time. If $x_0 \in (-1, 1)$ we have $f(x_0) < 0$ and therefore $x(t)$ is decreasing. It i not difficult to verify by solving this equation explicitly.

We can refine the above result

Theorem 7 *Assume $f(x)$ is continuously differentiable for all x and consider IVP*

$$x' = f(x), \quad x(t_0) = x_0.$$

Let $x(t)$ be a non-equilibrium solution (i.e. $f(x_0) \neq 0$) and (a, b) be it's maximal interval of existence. If $f(x_0) > 0$ then $x(t)$ is strictly increasing and one of the following alternatives hold:

1. *if there is no equilibrium greater than x_0 then $\lim_{t \to b} x(t) = +\infty$;*

2. *if there is equilibrium greater than x_0 then $b = +\infty$ and $\lim_{t\to\infty} x(t) = x_e$, where x_e is the smallest of such equilibriums.*

If $f(x_0) < 0$ then $x(t)$ is strictly decreasing and one of the following alternatives hold

1. *if there is no equilibrium smaller than x_0 then $\lim_{t\to b} x(t) = -\infty$;*
2. *if there is equilibrium smaller than x_0 then $b = +\infty$ and $\lim_{t\to\infty} x(t) = x_e$, where x_e is the largest of such equilibriums.*

In cases 1 b may be finite or infinite and the solution "blows up".

Using this theorem we may determine the asymptotic dynamics of autonomous ODE just by finding the roots of $f(x)$ and determining the sign of $f(x)$ between the roots.

2.8 Phase line portraits

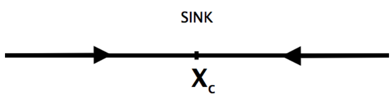
In order to understand the stability of critical points of the autonomous equation

$$x' = f(x)$$

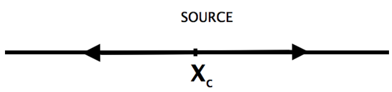
we can draw a diagram called *phase line portrait* . It contains all critical points $\{x_i\}_{i=1}^n$ and directions of the arrows corresponding to the monotonicity properties of the trajectories in the intervals between critical points, i.e. if $x' = f(x) > 0$ on some interval (x_i, x_{i+1}) then we know that trajectory $x(t)$ is increasing and the arrow is pointing in the right direction, on the other hand if $x' = f(x) < 0$ then solution $x(t)$ is decreasing and the arrow is pointing to the left.

There are three types of critical points:

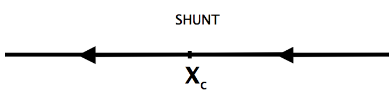
- If arrows point to the critical point from both sides of critical point then it is called *sink* (stable)

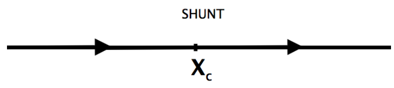


- If arrows point out of critical point from both sides of critical point then it is called *source* (unstable)



- If directions arrows are the same from both sides of critical point then it is called *shunt* (unstable)



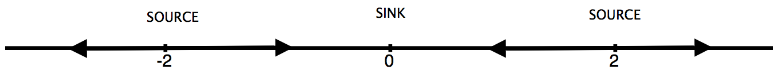


Let's consider a simple example

Example We study the following equation

$$x' = x^3 - 4x$$

It is easy to find critical points by solving $x^3 - 4x = 0$, that gives us $x = 0$, $x = 2$, $x = -2$. Now we want to find the sign of $f(x)$ on the intervals $(-\infty, -2)$, $(-2, 0)$, $(0, 2)$, and $(2, \infty)$. It is clear that $f(x) < 0$ on $(-\infty, -2)$ and $(0, 2)$; $f(x) > 0$ on $(-2, 0)$ and $(2, \infty)$. It follows that phase line portrait looks like



2.9 Bifurcations

Consider the following autonomous ODE

$$x' = f(x, \mu), \tag{16}$$

where $x \in \mathbf{R}$, $\mu \in \mathbf{R}$ - bifurcation parameter and $f : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$. We would like to understand how the behavior of the system described by this ODE depends on parameter μ.

Let $x_c(\mu)$ be a family of critical points of (16). It is clear that $f(x_c(\mu), \mu) = 0$. It is clear that the phase line portrait of (16) depends on parameter μ. Consider the following simple example

Example

$$x' = \mu - x^2$$

- For $\mu < 0$ there are no critical points
- For $\mu = 0$ there is one critical point that is a shunt
- For $\mu > 0$ there are two critical points: one is a source and another one is a sink.

We see that at $\mu = 0$ there is a qualitative change in phase line portrait (number of critical points is different for $\mu < 0$ and $\mu \geq 0$). This is called a bifurcation and $\mu = 0$ is a bifurcation point.

Two phase line portraits are qualitatively different if the number of critical points or the sequence of arrows in the intervals between critical points are different.

There is another way to find a bifurcation point. It is clear that critical points of

$$x' = f(x, \mu)$$

satisfy $f(x, \mu) = 0$. On the other hand it is clear that qualitative change in phase line portrait for each fixed μ occurs when $\frac{\partial f(x,\mu)}{\partial x} = 0$. Therefore if we want to find bifurcation points we need to solve the following system

$$\begin{cases} f(x, \mu) = 0 \\ \frac{\partial f(x,\mu)}{\partial x} = 0 \end{cases}$$

Now we want to consider the main types of bifurcations for one dimensional autonomous equations.

2.9.1 The saddle node bifurcation

We consider the following equation

$$x' = \mu - x^2, \text{ where } \mu \in R \text{ is a bifurcation parameter}$$

We first find the critical points depending on parameter μ by solving

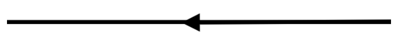
$$\mu - x^2 = 0$$

It is clear that real solutions of this equation exist only for $\mu > 0$ and they are

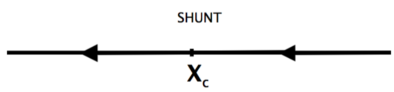
$$x = \pm\sqrt{\mu}.$$

Therefore the phase line portraits depending on parameter μ look like

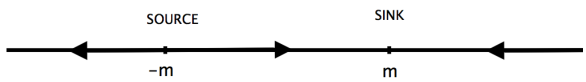
- $\mu < 0$ – no critical points



- $\mu = 0$ – one critical point $x = 0$



- $\mu > 0$ – two critical points $x = \pm\sqrt{\mu}$



Therefore the change in phase line portrait occurs when $\mu = 0$ and it is a bifurcation point. We can also solve the system

$$\begin{cases} \mu - x^2 = 0 \\ 2x = 0 \end{cases}$$

to obtain $\mu = 0$ is a bifurcation point.

Now we want to draw a *bifurcation diagram* that contains the dependence of critical points of the equation on bifurcation parameter μ.

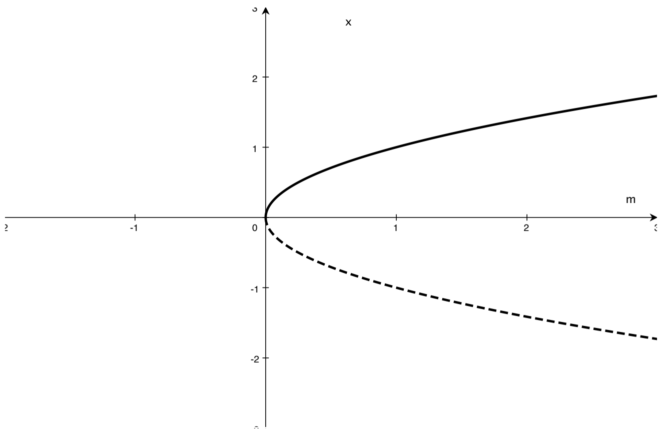


Figure 1: Saddle node bifurcation. Solid line is stable branch, dashed line is unstable branch.

Since for $\mu < 0$ there are no critical points the graph does not have any branches there. For $\mu = 0$ there is only one unstable critical point $x = 0$ and we see it on the graph. For $\mu > 0$ there are two branches of critical points $x = \sqrt{\mu}$ is stable sink and $x = -\sqrt{\mu}$ is unstable source.

2.9.2 The pitchfork bifurcation

We consider the following equation

$$x' = x(\mu - x^2), \text{ where } \mu \in R \text{ is a bifurcation parameter}$$

We first find the critical points depending on parameter μ by solving

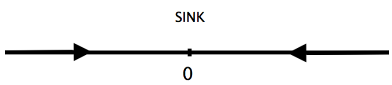
$$x(\mu - x^2) = 0$$

It is clear that real solutions of this equation exist only for $\mu > 0$ and they are

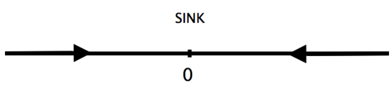
$$x = 0, \; x = \pm\sqrt{\mu}.$$

Therefore the phase line portraits depending on parameter μ look like

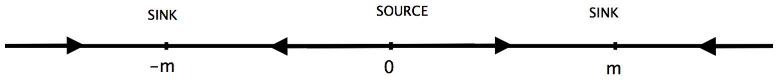
- $\mu < 0$ – one critical point $x = 0$



- $\mu = 0$ – one critical point $x = 0$



- $\mu > 0$ – three critical points $x = 0$, $x = \pm\sqrt{\mu}$



Therefore the change in phase line portrait occurs when $\mu = 0$ and it is a bifurcation point. We can also solve the system

$$\begin{cases} x(\mu - x^2) = 0 \\ \mu - 3x^2 = 0 \end{cases}$$

to obtain $\mu = 0$ is a bifurcation point.

Now we want to draw a *bifurcation diagram* that contains the dependence of critical points of the equation on bifurcation parameter μ.

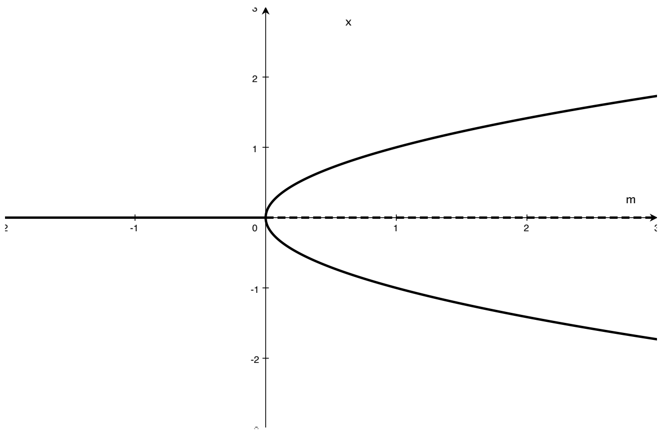


Figure 2: Pitchfork bifurcation. Solid line is stable branch, dashed line is unstable branch.

For $\mu \le 0$ there is only one stable branch $x = 0$. For $\mu > 0$ there are three branches: two stable $x = \pm\sqrt{\mu}$ and one unstable $x = 0$.

2.9.3 The transcritical bifurcation

We consider the following equation

$$x' = \mu x - x^2, \text{ where } \mu \in R \text{ is a bifurcation parameter}$$

We first find the critical points depending on parameter μ by solving

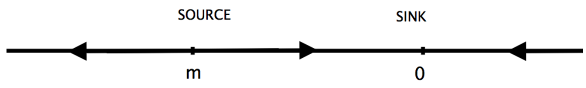
$$\mu x - x^2 = 0$$

It is clear that real solutions of this equation exist only for $\mu > 0$ and they are

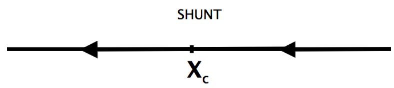
$$x = 0, \ x = \mu.$$

Therefore the phase line portraits depending on parameter μ look like

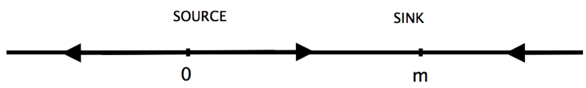
- $\mu < 0$ – two critical points $x = 0$, $x = \mu$



- $\mu = 0$ – one critical point $x = 0$



- $\mu > 0$ – two critical points $x = 0$, $x = \mu$



Therefore the change in phase line portrait occurs when $\mu = 0$ and it is a bifurcation point. We can also solve the system

$$\begin{cases} \mu x - x^2 = 0 \\ \mu - 2x = 0 \end{cases}$$

to obtain $\mu = 0$ is a bifurcation point.

Now we want to draw a *bifurcation diagram* that contains the dependence of critical points of the equation on bifurcation parameter μ.

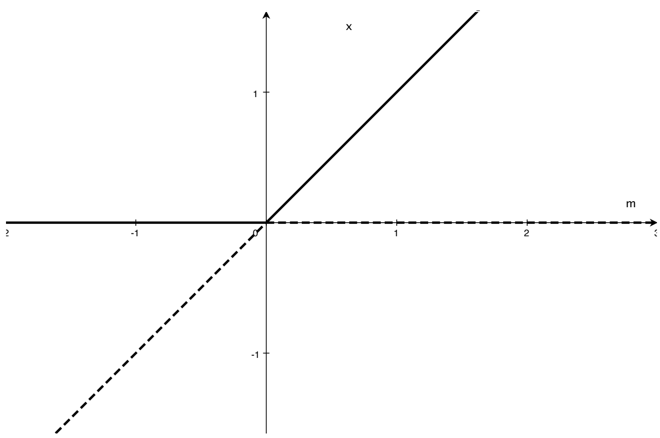


Figure 3: Transcritical bifurcation. Solid line is stable branch, dashed line is unstable branch.

3 Linear Systems of ODEs

In this section we are going to study the following problem

$$\mathbf{x}' = A(t)\mathbf{x} + \mathbf{b}(t), \tag{17}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0 \tag{18}$$

where $A(t)$ is an $n \times n$ matrix, $\mathbf{b}(t)$ and $\mathbf{x}(t)$ are vectorial functions in $\mathbf{R}^n$.

Using the results from section 1 we may prove the following theorem.

Theorem 8 *If $A(t)$ and $\mathbf{b}(t)$ are continuous functions on the interval (a,b) then there exists unique solution to the problem (17), (18) defined on the whole interval (a,b).*

We want to understand how to solve (17), (18), using the analogy with linear first order ODEs we first investigate a homogeneous problem

$$\mathbf{x}' = A(t)\mathbf{x}. \tag{19}$$

Definition 7 *The set of m n-dimensional vectors $\{\mathbf{x}_i\}$ is called linearly independent if*

$$\sum_{i=1}^{m} \alpha_i \mathbf{x}_i = 0$$

implies that all $\alpha_i = 0$.

Lemma 3 *The set of all solutions of (19) is an n-dimensional vector space.*

Proof 7 Step 1. *We check that the set of solutions of (19) is a vector space. Let $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ are to solutions of (19). Then it is easy to see that $\alpha\mathbf{x}_1 + \beta x_2$ is also a solution of (19). Therefore a set of solutions is a vector space.*

Step 2. *We want to show that there are n linearly independent solutions, i.e. the vector space of solutions of (19) is at least n-dimensional. We choose $t_0 \in (a,b)$ and for $i = 1, ..., n$ let $\mathbf{x}_i(t)$ be a solution of IVP (19)with initial condition $\mathbf{x}_i(t_0) = \mathbf{e}_i$, where $\mathbf{e}_i^T = (0, ..., 0, 1, 0, ..., 0)$ is the i-th basis vector of $\mathbf{R}^n$. The functions $\mathbf{x}_i(t)$ are linearly independent on (a,b), i.e. no linear combination of these functions $\sum_{i=1}^{n} \alpha_i \mathbf{x}_i(t)$ can vanish identically on (a,b) unless all coefficients $\{\alpha_i\}$ are zero. Let's prove it: suppose there exists a point $t_1 \in (a,b)$ such that $\mathbf{x}(t_1) = \sum_{i=1}^{n} \alpha_i \mathbf{x}_i(t_1) = 0$, then by uniqueness of the solution $\mathbf{x}(t) = 0$ for all $t \in (a,b)$ and therefore for t_0. But at point t_0 we have $0 = \mathbf{x}(t_0) = (\alpha_1, ..., \alpha_n)$ and we get a contradiction. Therefore $\mathbf{x}_i(t)$ are linearly independent for all $t \in (a,b)$.*

Step 3. *Now we want to show that a dimension of our vector space is exactly n. To do this we show that any solution of (19) can be represented as a linear combination of solutions $\mathbf{x}_i(t)$. Let $\mathbf{y}(t)$ be any solution of (19) then $\mathbf{y}(t_0) = \beta = (\beta_1, ..., \beta_n)$ for some vector $\beta \in \mathbf{R}^n$. But then $\mathbf{x}_(t) = \sum_{i=1}^{n} \beta_i \mathbf{x}_i(t)$ satisfies the same initial condition as $\mathbf{y}(t)$ and therefore by uniqueness of solution it coincides with $\mathbf{y}(t)$ on the whole interval (a,b). Thus we proved that $\mathbf{x}_1(t), ..., \mathbf{x}_n(t)$ form a basis of the solution vector space and therefore it is n-dimensional.*

Definition 8 *A set of n linearly independent solutions of (19) is called a fundamental set of solutions.*

Definition 9 *An $n \times n$ matrix $X(t)$ in which the columns are formed from a fundamental set of solutions is called a fundamental matrix.*

Note that a fundamental matrix $X(t)$ satisfies

$$X' = A(t)X \tag{20}$$

Let us summarize the properties of fundamental matrices in the following theorem. Proof of this theorem is not difficult and is left as an exercise.

Theorem 9 *Let $X(t)$ be an $n \times n$ matrix satisfying (23) and denote the n columns of $X(t)$ by $\mathbf{x}_1(t), ..., \mathbf{x}_n(t)$. Then*

1. *The following four conditions are equivalent*
   (a) $\mathbf{x}_1(t), ..., \mathbf{x}_n(t)$ *are linearly dependent functions;*
   (b) $det X(t) = 0$ *for all* $t \in (a, b)$*;*
   (c) $det X(t) = 0$ *at some* $t_1 \in (a, b)$*;*
   (d) $\mathbf{x}_1(t_1), ..., \mathbf{x}_n(t_1)$ *are linearly dependent vectors.*
2. *If $X(t)$ is a fundamental matrix then any solution of (19) has the form $\mathbf{x}(t) = X(t)\mathbf{c}$ for some constant vector $\mathbf{c} \in \mathbf{R}^n$.*
3. *If $X(t)$ is a fundamental matrix and C is invertible constant $n \times n$ matrix then $Y(t) = X(t)C$ is a fundamental matrix. Conversely, if $X(t)$ and $Y(t)$ are fundamental matrices then $Y(t) = X(t)C$ for some invertible constant matrix C.*

Now we want to relate our knowledge of homogeneous problem (19) to a non-homogeneous problem (17).

Lemma 4 *Let $\mathbf{x}_p(t)$ be a solution of a non-homogeneous problem (17), and $\mathbf{x}_h(t)$ be a solution of a homogeneous problem (19). Then $\mathbf{x}(t) = \mathbf{x}_p(t) + x_h(t)$ is a solution of a non-homogeneous problem (17). Moreover, any solution $\mathbf{x}(t)$ of (17) can be represented as $\mathbf{x}(t) = \mathbf{x}_p(t) + x_h(t)$ for some choice of $\mathbf{x}_h(t)$.*

The proof of this lemma is obvious and is left to a reader.

This lemma tells us that if we know all solutions of a homogeneous problem (19) (or a fundamental matrix $X(t)$) and we know just **one** solution of a non-homogeneous problem (17). Then we can we can find any solution of (17), i.e. solve the non-homogeneous problem (17), (18) with **any** initial condition.

In fact, in order to obtain any solution of a non-homogeneous problem (17), (18) we have to know just a fundamental matrix $X(t)$. Let's show this:

We are looking for a solution of $\mathbf{x}' = A(t)\mathbf{x} + \mathbf{b}(t)$ with initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$. Let $X(t)$ be a fundamental matrix for the homogeneous equation $\mathbf{x}' = A\mathbf{x}$. Then acting on a non-homogeneous equation by inverse of a fundamental matrix from the left we obtain

$$X^{-1}(t)\mathbf{x}' = X^{-1}(t)A\mathbf{x} + X^{-1}(t)\mathbf{b}(t) \tag{21}$$

Let's see what we have here: $\left(X^{-1}(t)\mathbf{x}\right)' = \left(X^{-1}(t)\right)'\mathbf{x} + X^{-1}(t)\mathbf{x}'$. It is not difficult to show that

$$\left(X^{-1}(t)\right)' = -X^{-1}(t)X'(t)X^{-1}(t)$$

(Use the product rule: $\left(X(t)X^{-1}(t)\right)' = X'(t)X^{-1}(t) + X(t)\left(X^{-1}(t)\right)'$ and the fact that $X(t)X^{-1}(t) = I$.)

Therefore, using the expression for $\left(X^{-1}(t)\right)'$ and the fact that $X'(t) = A(t)X(t)$ we have

$$\left(X^{-1}(t)\mathbf{x}\right)' = -X^{-1}(t)X'(t)X^{-1}(t)\mathbf{x} + X^{-1}(t)\mathbf{x}' = -X^{-1}(t)A(t)\mathbf{x} + X^{-1}(t)\mathbf{x}'$$

and our equation (21) is

$$\left(X^{-1}(t)\mathbf{x}\right)' = X^{-1}(t)\mathbf{b}(t).$$

Integrating both parts from t_0 to t we obtain

$$X^{-1}(t)\mathbf{x}(t) - X^{-1}(t_0)\mathbf{x}(t_0) = \int_{t_0}^{t} X^{-1}(s)\mathbf{b}(s)ds.$$

This gives us the solution of (17), (18):

$$\mathbf{x}(t) = X(t)X^{-1}(t_0)\mathbf{x}_0 + X(t)\int_{t_0}^{t} X^{-1}(s)\mathbf{b}(s)ds \tag{22}$$

Compare this reasoning with what we did for linear first order ODEs.

So the only information we need to know about linear system of ODEs in order to get the solution is the fundamental matrix. In general, it can be quite difficult task to find it, but there are some special cases when it is possible. Below we consider the problem when matrix A is constant.

3.1 Linear ODEs with constant coefficients.

In this section we are going to study the problem (17) with the constant matrix A. We saw before that in order to solve this problem it is enough to find a fundamental matrix $X(t)$, satisfying $X' = AX$. That's our main goal here.

Definition 10 *We define a norm of $n \times n$ matrix A as*

$$\|A\| = \sup_{\mathbf{x}\in\mathbf{R}^n,\ x\neq 0} \frac{|A\mathbf{x}|}{|\mathbf{x}|},$$

where $|\mathbf{x}| = \sqrt{\sum_{i=1}^{n} x_i^2}$ is a Euclidean norm of a vector $\mathbf{x} = (x_1, ..., x_n)$.

The same definition works for complex matrices if you change $\mathbf{R}^n$ to $\mathbf{C}^n$.

We know that for scalar case $x' = ax$ there is a solution $x(t) = e^{at}$. So we hope that something like this should happen in the general case. But what is e^A if A is an $n \times n$ matrix?

Definition 11 *If A is an $n \times n$ matrix (maybe complex), we define the exponent of A as*

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

It is clear that this exponent is well defined, i.e. the series converges. Since

$$\left\| \sum_{k=0}^{n} \frac{A^k}{k!} - \sum_{k=0}^{m} \frac{A^k}{k!} \right\| \leq \sum_{k=m+1}^{n} \frac{\|A\|^k}{k!} \to 0$$

as $n, m \to \infty$.

Lemma 5 *Let A, B be an $n \times n$ matrices.*

- *if $AB = BA$ then $e^{A+B} = e^A e^B$;*
- $\left(e^{tA}\right)' = Ae^{tA}$.

Proof 8 *It is easy to see that*

$$e^{A+B} = \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = \sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{1}{k!} \left(\begin{array}{c} k \\ m \end{array} \right) A^m B^{k-m} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{A^k}{k!} \frac{B^l}{l!} = e^A e^B$$

Now let's show the second identity. We first verify the formula for $t = 0$. Since $e^0 = I$ we have to show $\lim_{t \to 0} \frac{1}{t} \left(e^{tA} - I\right) = A$

$$\left\| \frac{1}{t} \left(e^{tA} - I\right) - A \right\| = \left\| \sum_{k=1}^{\infty} \frac{t^{k-1} A^k}{k!} - A \right\| \leq |t| \|A\|^2 \sum_{k=0}^{\infty} \frac{|t|^{k-1} A^k}{(k+2)!} \leq |t| \|A\|^2 e^{|t| \|A\|}$$

As $t \to 0$ we obtain the result. Now for any $t_1 \in \mathbf{R}$ we have

$$e^{tA} = e^{(t-t_1)A} e^{t_1 A}$$

and $\frac{de^{tA}}{dt}(t_1) = \frac{de^{tA}}{d(t-t_1)}(t_1) = \frac{de^{(t-t_1)A}}{d(t-t_1)}(t_1) e^{t_1 A} = Ae^{t_1 A}$

From this simple lemma we immediately see that $X(t) = e^{tA}$ is a fundamental matrix for the system $\mathbf{x}' = A\mathbf{x}$ since

- It satisfies $X' = AX$.
- It is invertible with $X^{-1}(t) = e^{-tA}$.

Therefore it seems that we solved the problem (17) for the case when A is a constant matrix. It remains to understand how we actually compute e^{tA} since infinite summation is quite difficult to implement.

3.2 Computation of a fundamental matrix e^{tA}.

We begin with a simple observation: if $A = PBP^{-1}$ for some invertible matrix P then $e^A = Pe^B P^{-1}$ (matrices A and B are called similar). Indeed $A^2 = PBP^{-1}PBP^{-1} =$

PB^2P^{-1}, using induction it is quite easy to show that $A^k = PB^kP^{-1}$ for all k and therefore

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} PB^kP^{-1} = Pe^BP^{-1}$$

This simple observation gives us extremely useful result when matrix B is diagonal.

Lemma 6 *Let $A = PDP^{-1}$ with D being diagonal matrix*

$$D = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ .... & & & \\ 0 & ... & 0 & \lambda_n \end{pmatrix}$$

Then $e^{tA} = Pe^{tD}P^{-1}$, where

$$e^{tD} = \begin{pmatrix} e^{t\lambda_1} & 0 & ... & 0 \\ 0 & e^{t\lambda_2} & ... & 0 \\ .... & & & \\ 0 & ... & 0 & e^{t\lambda_n} \end{pmatrix} \tag{23}$$

Proof 9 *To prove this result we just have to compute e^{tD}. By definition*

$$e^{tD} = \sum_{k=0}^{\infty} \frac{t^kD^k}{k!} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{t^k\lambda_1^k}{k!} & 0 & ... & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^k\lambda_2^k}{k!} & ... & 0 \\ .... & & & \\ 0 & ... & 0 & \sum_{k=0}^{\infty} \frac{t^k\lambda_n^k}{k!} \end{pmatrix} =$$

$$= \begin{pmatrix} e^{t\lambda_1} & 0 & ... & 0 \\ 0 & e^{t\lambda_2} & ... & 0 \\ .... & & & \\ 0 & ... & 0 & e^{t\lambda_n} \end{pmatrix}$$

This gives us the result.

We just showed that in case when matrix A can be diagonalized, if λ_i is the i-th eigenvalue and $\mathbf{v}_i$ is the corresponding eigenvector, satisfying $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, then

$$e^{tA} = Pe^{tD}P^{-1},$$

where

$$P = (\mathbf{v}_1|\mathbf{v}_2|...|\mathbf{v}_n), \; e^{tD} = \begin{pmatrix} e^{t\lambda_1} & 0 & ... & 0 \\ 0 & e^{t\lambda_2} & ... & 0 \\ .... & & & \\ 0 & ... & 0 & e^{t\lambda_n} \end{pmatrix}$$

Moreover we know that $e^{tA}P$ is also fundamental (see theorem in section 3). Therefore the fundamental matrix in the case of diagonalizable matrix A is

$$X(t) = Pe^{tD}$$

Unfortunately not every matrix is diagonalizable, it happens if the eigenvalue of A has a multiplicity $k > 1$ but there are only $l < k$ linearly independent eigenvectors corresponding to this eigenvalue. In this case we have to work a bit more.

Definition 12 *A* **Jordan block** *is a square matrix (maybe complex) in which all diagonal entries are equal, all entries immediately above the diagonal are one and all other entries are zero:*

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

Definition 13 *A square matrix J is in a* **Jordan form** *if it has a block decomposition where all diagonal blocks are Jordan blocks and all other entries are zero*

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & 0 \\ 0 & .... & & \\ & & & J_m \end{pmatrix}$$

Here J_i are Jordan blocks of size $n_i \times n_i$ ($\sum_{i=1}^{m} n_i = n$).

Theorem 10 *Every $n \times n$ matrix A is similar to some matrix J in a Jordan form. In particular, let λ_1, ..., λ_j be the eigenvalues of A corresponding to the distinct linearly independent eigenvectors $\mathbf{v}_1$, ..., $\mathbf{v}_j$. Denote the multiplicity of the eigenvalue λ_i by k_i for $i = 1, 2, ..., j$ (i.e. $\sum_{i=1}^{j} k_i = n$). Let $\mathbf{w}_{i1}, ..., \mathbf{w}_{ik_i}$ be the generalized eigenvectors for λ_i with $\mathbf{w}_{i1} = \mathbf{v}_i$, i.e*

$$(A - \lambda_i)\mathbf{w}_{i1} = 0,$$

$$(A - \lambda_i)\mathbf{w}_{i2} = \mathbf{w}_{i1},$$

$$...$$

$$(A - \lambda_i)\mathbf{w}_{ik_i} = \mathbf{w}_{i(k_i-1)}$$

Then $A = PJP^{-1}$, where

$$P = (\mathbf{w}_{11} | \ ... \ | \ \mathbf{w}_{1k_1} \ | \ ... \ | \ \mathbf{w}_{j1} \ | \ ... \ | \ \mathbf{w}_{jk_j})$$

and

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & 0 \\ 0 & .... & & \\ & & & J_j \end{pmatrix}$$

with J_i being a $k_i \times k_i$ Jordan block

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

Moreover J is uniquely determined by A except maybe for reordering the Jordan blocks.

The proof of Jordan theorem is usually given in Linear Algebra course and therefore will not be given here. What we need is the implications of Jordan decomposition theorem.

We want to calculate $X(t) = Pe^{tJ}$. Since we already know how to calculate P from Jordan decomposition theorem, it is left to calculate e^{tJ}. It is not difficult to see that

$$e^{tJ} = \begin{pmatrix} e^{tJ_1} & & \\ & e^{tJ_2} & 0 \\ 0 & .... & \\ & & e^{J_j} \end{pmatrix}$$

So what we have to find is e^{tJ_i}: Let represent $J_i = \lambda_i I + U_i$, where

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

It is obvious that $e^{tJ_i} = e^{t\lambda_i I} e^{tU_i} = e^{t\lambda_i} e^{tU_i}$, moreover

$$e^{tU_i} = \sum_{k=0}^{\infty} \frac{t^k U_i^k}{k!}$$

It is not difficult to calculate U_i^k (it is $k_i \times k_i$ matrix):

$$U_i^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 0 & 1 & ... & 0 \\ .... & & & & & \\ 0 & ... & 0 & 0 & 0 & 1 \\ 0 & ... & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 & 0 \end{pmatrix}, \quad U_i^{k_i-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & ... & 1 \\ 0 & 0 & 0 & 0 & ... & 0 \\ .... & & & & & \\ 0 & ... & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 & 0 \\ 0 & ... & 0 & 0 & 0 & 0 \end{pmatrix}, \quad U_i^p = 0, \; p \geq k_i$$

Therefore we obtain

$$e^{tU_i} = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{6} & ... & \frac{t^{k_i-1}}{(k_i-1)!} \\ 0 & 1 & t & \frac{t^2}{2} & ... & \frac{t^{k_i-2}}{(k_i-2)!} \\ .... & & & & & \\ 0 & ... & 0 & 1 & t & \frac{t^2}{2} \\ 0 & ... & 0 & 0 & 1 & t \\ 0 & ... & 0 & 0 & 0 & 1 \end{pmatrix}$$

This gives us the final expression for e^{tJ_i}:

$$e^{tJ_i} = e^{t\lambda_i} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{6} & ... & \frac{t^{k_i-1}}{(k_i-1)!} \\ 0 & 1 & t & \frac{t^2}{2} & ... & \frac{t^{k_i-2}}{(k_i-2)!} \\ .... & & & & & \\ 0 & ... & 0 & 1 & t & \frac{t^2}{2} \\ 0 & ... & 0 & 0 & 1 & t \\ 0 & ... & 0 & 0 & 0 & 1 \end{pmatrix}$$

Using this expression we obtain a fundamental matrix $X(t)$. This general procedure allows you to solve any linear system of ODEs with constant coefficients. Sometimes the eigenvalues are complex and therefore a fundamental matrix is complex. In this case you have to take the linear combination of two complex solutions corresponding to conjugate complex eigenvalues in order to get two linearly independent real solutions. The basic idea of this rearrangement will be seen in the next section, where we consider 2D linear systems.

4 Classification of 2D linear systems with constant coefficients and phase planes.

In this section we are going to consider the critical points of homogeneous 2D linear systems with constant coefficients. In the spirit it is similar to investigating the stability of equilibriums for autonomous equations. For instance, let's consider the following simple first order ODE

$$x' = px$$

and ask the question: is equilibrium point 0 is stable? For non-zero p the answer obviously depend on the sign of p. If $p > 0$ then we know that 0 is an unstable source, if $p < 0$ then 0 is a stable sink. (If $p = 0$ then the solutions are constants and you never reach 0 unless you start from 0.) So, for non-zero p there are only two possible behavior of this system near point 0: converging to 0 or diverging from 0 with exponential speed. Based on this knowledge we can also analyze autonomous systems (see the corresponding section). For liner systems the behavior near $\mathbf{0}$ is much more complex and we want to understand it in order to apply this knowledge later to autonomous systems of equations.

We consider the following system

$$\mathbf{x}' = A\mathbf{x}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{24}$$

and want to investigate the behavior of the solutions near the equilibrium point $\mathbf{0}$ (note that here $\mathbf{0} = (0, 0)^T$).

4.1 $\det A \neq 0$

In this section we assume that $\det A \neq 0$. In order to understand the behavior of the system (24) we solve it explicitly and look at the behavior of trajectories depending on different initial conditions. In order to solve this system we need to find a fundamental matrix $X(t)$ and to do this we have to compute the eigenvalues and eigenvectors of A. We know that eigenvalues of A satisfy

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

or

$$\lambda^2 - \lambda \mathrm{tr} A + \det A = 0 \tag{25}$$

We call equation (25) characteristic equation of the system. Depending on the roots of characteristic equation (or eigenvalues of A) there solutions of (24) exhibit different behavior. There are several different cases.

Case 1: two different real roots.

When roots λ_1, λ_2 of the characteristic equation are different there are two linearly independent eigenvectors $\mathbf{v}$ and $\mathbf{w}$, corresponding to these roots, i.e. $A\mathbf{v} = \lambda_1\mathbf{v}$, $A\mathbf{w} = \lambda_2\mathbf{w}$. The fundamental matrix of the system (24) is

$$X(t) = (e^{\lambda_1 t}\mathbf{v} \mid e^{\lambda_2 t}\mathbf{w})$$

and the general solution can be found as $\mathbf{x}(t) = X(t)\mathbf{c}$ for some constant vector $\mathbf{c}$. Let's write down the components of $\mathbf{x}(t)$:

$$x_1(t) = c_1 v_1 e^{\lambda_1 t} + c_2 w_1 e^{\lambda_2 t}$$

$$x_2(t) = c_1 v_2 e^{\lambda_1 t} + c_2 w_2 e^{\lambda_2 t}$$

In order to see the behavior of the solutions near the origin we put the components of solutions on the phase plane (the axis of the plane are x_1 and x_2). What kind of behavior we observe depending on eigenvalues and eigenvectors?

1. $\lambda_1, \lambda_2 < 0$.

   In this case we see that as $t \to \infty$ $x_1(t) \to 0$ and $x_2(t) \to 0$ independently of $\mathbf{c}$ or eigenvectors $\mathbf{v}$ and $\mathbf{w}$. We can divide $x_2(t)$ by $x_1(t)$ to obtain

   $$\frac{x_2}{x_1} = \frac{c_1 v_2 e^{\lambda_1 t} + c_2 w_2 e^{\lambda_2 t}}{c_1 v_1 e^{\lambda_1 t} + c_2 w_1 e^{\lambda_2 t}}$$

   Without loss of generality we take $\lambda_2 < \lambda_1$ (i.e. $|\lambda_2| > |\lambda_1|$ since they both are negative).

   Suppose our we start from the points corresponding to $c_1 = 0$, then $\frac{x_2}{x_1} = \frac{w_2}{w_1}$ and we put this trajectory on the phase plane (it is a straight line). The direction of the flow along this trajectory as we increase time t will be towards the origin, since as $t \to \infty$ both $x_1(t)$ and $x_2(t)$ converge to 0. By the same token, if we start from the points corresponding to $c_2 = 0$, then $\frac{x_2}{x_1} = \frac{v_2}{v_1}$. If both components of $\mathbf{c}$ are non-zero, after simple calculations we obtain

   $$\frac{x_2}{x_1} = \frac{c_1 v_2 + c_2 w_2 e^{(\lambda_2 - \lambda_1)t}}{c_1 v_1 + c_2 w_1 e^{(\lambda_2 - \lambda_1)t}}$$

   Since $\lambda_2 < \lambda_1$, as $t \to \infty$ we obtain $\frac{x_2}{x_1} = \frac{v_2}{v_1}$. Combining this information with what we already know about the behavior of trajectories we have the following picture

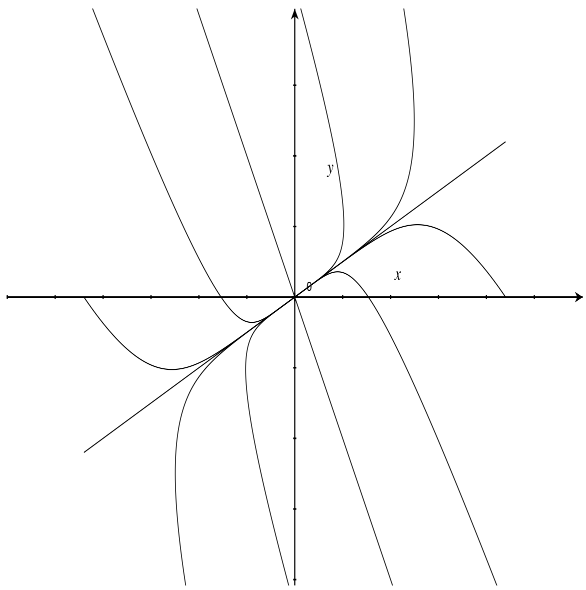


Figure 4: *Stable node or sink.*

2. $\lambda_1, \lambda_2 > 0$.

   Without loss of generality we take $\lambda_2 > \lambda_1$. In this case we see that as $t \to \infty$ $x_1(t)$ and $x_2(t)$ diverge independently of $\mathbf{c}$ or eigenvectors $\mathbf{v}$ and $\mathbf{w}$. Also if we reverse the flow, i.e. take $t \to -\infty$, both $x_1(t)$ and $x_2(t)$ converge to 0. We can do the same trick as before to obtain two trajectories $\frac{x_2}{x_1} = \frac{w_2}{w_1}$ (if $c_1 = 0$) and

$\frac{x_2}{x_1} = \frac{v_2}{v_1}$ (if $c_2 = 0$). The flow along these trajectories goes out of the origin. Therefore even if we choose initial condition very close to the origin we are going to diverge to infinity as $t \to \infty$. If both components of $\mathbf{c}$ are non-zero, after simple calculations we obtain

$$\frac{x_2}{x_1} = \frac{c_1 v_2 + c_2 w_2 e^{(\lambda_2 - \lambda_1)t}}{c_1 v_1 + c_2 w_1 e^{(\lambda_2 - \lambda_1)t}}$$

As $t \to -\infty$ we see that $\frac{x_2}{x_1} \to \frac{v_2}{v_1}$. Therefore the phase plane diagram looks like this

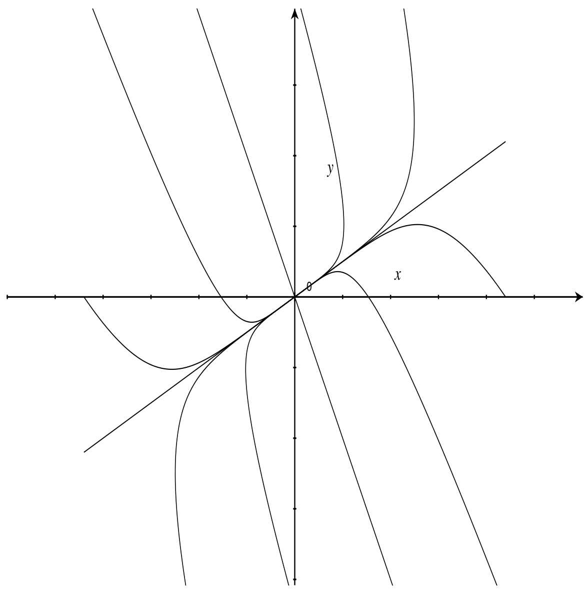


Figure 5: *Unstable node or source.*

3. λ_1 and λ_2 have opposite signs.

   Without loss of generality we take $\lambda_1 < 0 < \lambda_2$. We have two straight line trajectories $\frac{x_2}{x_1} = \frac{w_2}{w_1}$ with flow direction outside of the origin ($\lambda_2 > 0$), corresponding to $c_1 = 0$, and $\frac{x_2}{x_1} = \frac{v_2}{v_1}$ with flow direction towards the origin ($\lambda_1 < 0$), corresponding to $c_2 = 0$. The rest of trajectories is quite easy to obtain by calculations similar to the previous cases. The phase plane diagram is

Case 2: two complex roots.

When roots λ_1, λ_2 of the characteristic equation are complex, they obviously are related as $\lambda_2 = \bar{\lambda}_1$. There are two linearly independent complex eigenvectors $\mathbf{u}$ and $\bar{\mathbf{u}}$, corresponding to λ_1 and $\bar{\lambda}_1$. We don't want to deal with complex solutions and therefore we have to find two linearly independent **real** solutions of our system. Let $\lambda_1 = \alpha + i\beta$ and $\mathbf{u} = \mathbf{v} + i\mathbf{w}$, the complex linearly independent solutions are

$$\mathbf{x}_1(t) = \mathbf{u}e^{\lambda_1 t}, \quad \mathbf{x}_2(t) = \bar{\mathbf{u}}e^{\bar{\lambda}_1 t}$$

We may rewrite them as

$$\mathbf{x}_1(t) = e^{\alpha t}\left[\mathbf{v}\cos(\beta t) - \mathbf{w}\sin(\beta t) + i(\mathbf{v}\sin(\beta t) + \mathbf{w}\cos(\beta t))\right]$$

$$\mathbf{x}_1(t) = e^{\alpha t}\left[\mathbf{v}\cos(\beta t) - \mathbf{w}\sin(\beta t) - i(\mathbf{v}\sin(\beta t) + \mathbf{w}\cos(\beta t))\right]$$

Since any linear combination of solutions is itself a solution we may find two linearly independent real solutions as

$$\mathbf{y_1} = \frac{\mathbf{x}_1 + \mathbf{x}_2}{2}, \quad \mathbf{y_2} = \frac{\mathbf{x}_1 - \mathbf{x}_2}{2i}$$

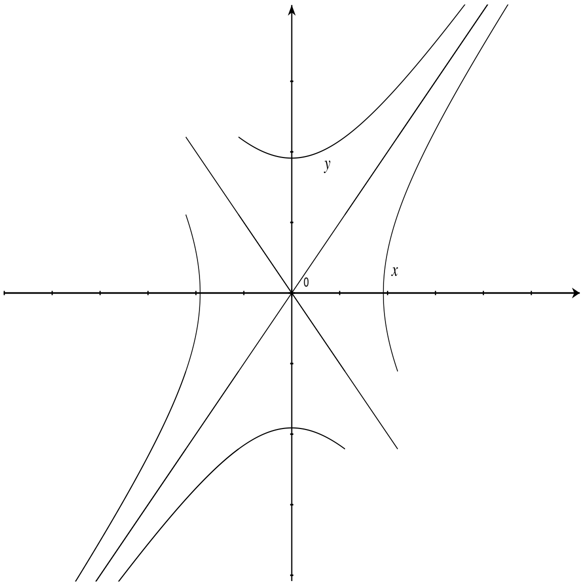


Figure 6: *Saddle point.*

The general solution is $\mathbf{x}(t) = c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t)$ and has the following components

$$x_1(t) = e^{\alpha t}\left[c_1(v_1\cos(\beta t) - w_1\sin(\beta t)) + c_2(v_1\sin(\beta t) + w_1\cos(\beta t))\right]$$

$$x_2(t) = e^{\alpha t}\left[c_1(v_2\cos(\beta t) - w_2\sin(\beta t)) + c_2(v_2\sin(\beta t) + w_2\cos(\beta t))\right]$$

In order to see the behavior of the solutions near the origin we put the components of solutions on the phase plane.

1. $\alpha = 0$.

   In this case it is possible to show that the trajectories will be closed curves (circles or ellipses) around origin.

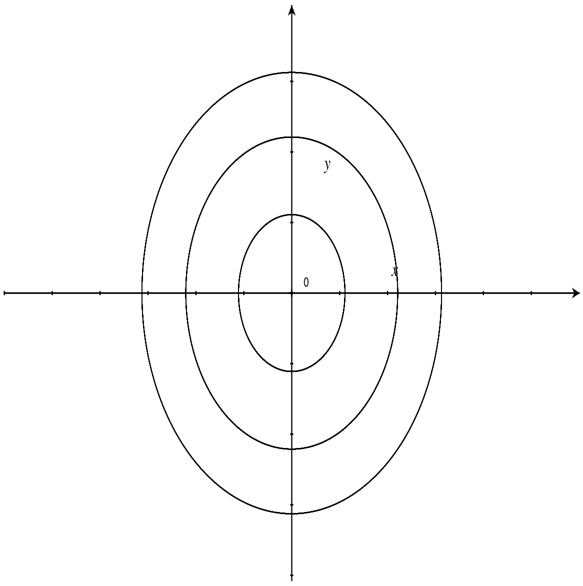


Figure 7: *Stable center.*

2. $\alpha < 0$.

   In this case we observe that as $t \to \infty$ both components $x_1(t)$ and $\mathbf{x}_2(t)$ tend to 0. The trajectory will be a spiral with a flow going towards the origin.

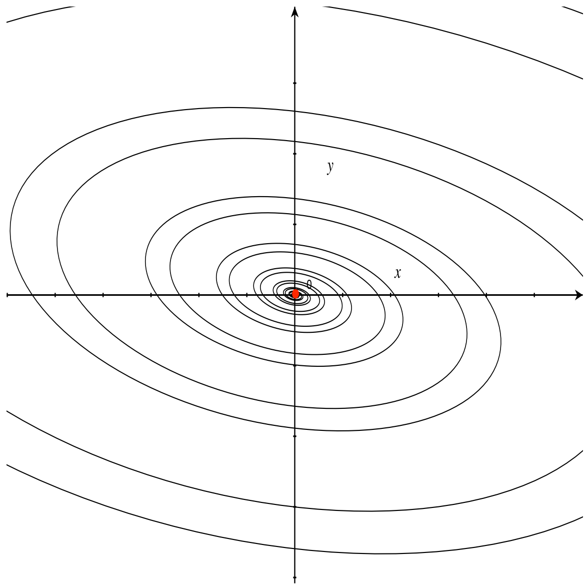


Figure 8: *Stable spiral.*

3. $\alpha > 0$.

   In this case we observe that as $t \to \infty$ both components $x_1(t)$ and $\mathbf{x}_2(t)$ diverge to infinity. The trajectory will be a spiral with a flow going outside of the origin.

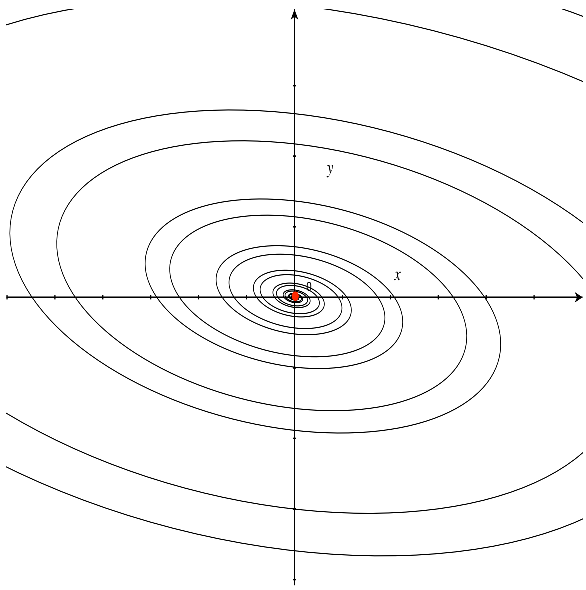


Figure 9: *Unstable spiral.*

Case 2: two real equal roots.

If eigenvalues of A are equal then there are two possibilities:

a) Matrix $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$. This is the case when our 2D system reduces to the same first order ODEs

$$x_1' = \lambda x_1, \quad x_2' = \lambda x_2$$

It is quite easy to see that $x_1 = c_1 e^{\lambda t}$ and $x_2 = c_2 e^{\lambda t}$. This gives us a simple relation $\frac{x_2}{x_1} = \frac{c_2}{c_1}$ and the phase plane diagram is

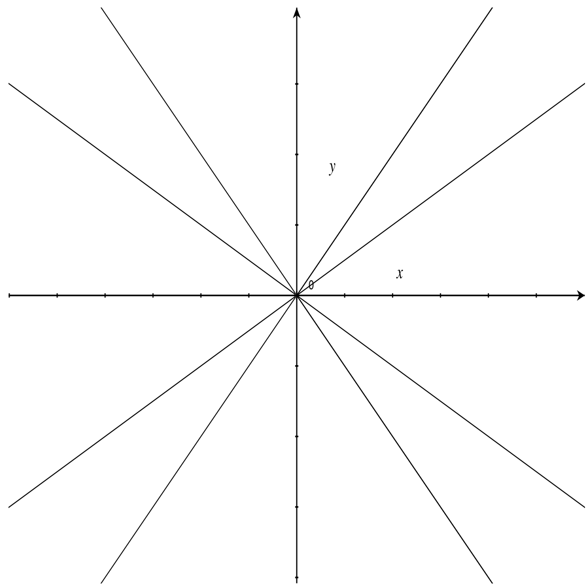


Figure 10: *Stable star.*

b) Some non-diagonal components of A are non-zero.

In this case we know that there is one eigenvector $\mathbf{v}$, satisfying $A\mathbf{v} = \lambda v$. This suggests that one solution is

$$\mathbf{x}_1(t) = \mathbf{v} e^{\lambda t}.$$

In order to find the second solution we can either guess it

$$\mathbf{x}_2(t) = (\mathbf{v}t + \mathbf{w})e^{\lambda t},$$

where $(A - \lambda I)\mathbf{w} = \mathbf{v}$ or use Jordan form decomposition. Let's use Jordan form. We know that $A = PJP^{-1}$, where

$$P = (\mathbf{v}\ |\mathbf{w}),\ J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

By the above results for linear systems $e^{tA} = Pe^{tJ}P^{-1}$ and we choose our fundamental matrix $X(t)$ to be

$$X(t) = Pe^{tJ}.$$

It's clear that

$$e^{tJ} = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

and therefore

$$X(t) = e^{\lambda t}(\mathbf{v}\ |\mathbf{w}) \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = e^{\lambda t}(\mathbf{v}\ |\mathbf{v}t + \mathbf{w}).$$

We see that two linearly independent solutions are

$$\mathbf{x}_1(t) = \mathbf{v} e^{\lambda t},$$

$$\mathbf{x}_2(t) = (\mathbf{v}t + \mathbf{w})e^{\lambda t},$$

and the general solution is

$$\mathbf{x}(t) = c_1\mathbf{v}e^{\lambda t} + c_2(\mathbf{v}t + \mathbf{w})e^{\lambda t}.$$

Obviously

$$x_1(t) = c_1 v_1 e^{\lambda t} + c_2(v_1 t + w_1)e^{\lambda t},$$
$$x_2(t) = c_1 v_2 e^{\lambda t} + c_2(v_2 t + w_2)e^{\lambda t}.$$

Obviously

$$\frac{x_2}{x_1} = \frac{c_1 v_2 + c_2(v_2 t + w_2)}{c_1 v_1 + c_2(v_1 t + w_1)},$$

and if $c_2 = 0$ we have $\frac{x_2}{x_1} = \frac{v_2}{v_1}$; if $c_1 = 0$ we have $\frac{x_2}{x_1} \to \frac{v_2}{v_1}$ as $t \to \infty$ (positive or negative).

1. If $\lambda > 0$ we see that as $t \to \infty$ both $x_1(t)$ and $x_2(t)$ diverge independently of $\mathbf{c}$, $\mathbf{v}$ and $\mathbf{w}$. The phase plane portrait is

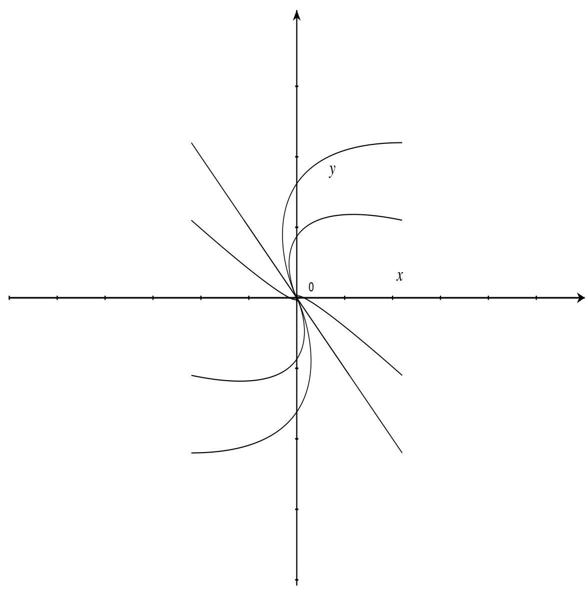


Figure 11: *Unstable improper node.*

2. If $\lambda < 0$ we see that as $t \to \infty$ both $x_1(t) \to 0$ and $x_2(t) \to 0$ independently of $\mathbf{c}$, $\mathbf{v}$ and $\mathbf{w}$. The phase plane portrait is

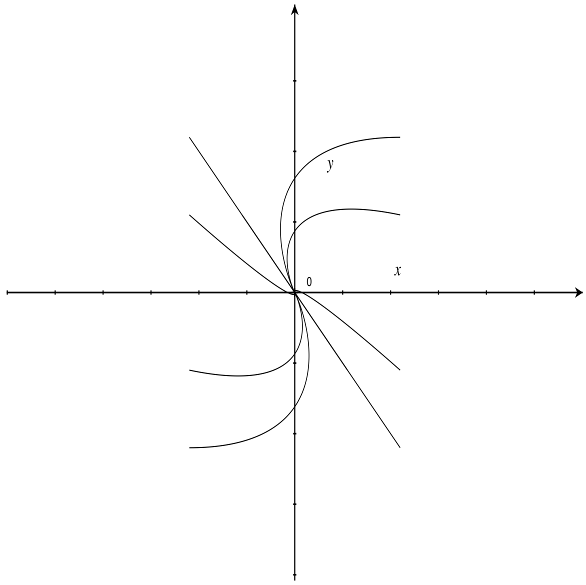


Figure 12: *Stable improper node.*

5 Linear high order ODEs

5.1 Equations with constant coefficients

General form of linear high order ODE with constant coefficients is

$$a_n x^{(n)} + a_{n-1} x^{(n-1)} + ... + a_1 x' + a_0 x = f(t). \tag{26}$$

In order to study IVP we have to supply initial conditions

$$x^{(n-1)}(t_0) = x_{n-1}, \ \dots, x(t_0) = x_0. \tag{27}$$

In this section we consider second order equations, all results obtained for $n = 2$ can be generalized for higher order equations.

Theorem 11 *The second order IVP ($a \neq 0$)*

$$ax'' + bx' + cx = f(t), \tag{28}$$

$$x(t_0) = x_0, \ x'(t_0) = x_1. \tag{29}$$

has the unique solution.

If we want to find a solution of (28) we first have to solve a homogeneous equation ($f(t) = 0$) and then find a particular solution of (28). Obviously, there are two linearly independent solutions of homogeneous ODE

$$ax'' + bx' + cx = 0,$$

call them $x_{1,h}(t)$ and $x_{2,h}(t)$. The general solution is a simple linear combination of them

$$x_h(t) = c_1 x_{1,h}(t) + c_2 x_{2,h}(t).$$

It is not difficult to show (see section for linear systems) that any solution of (28) maybe represented as

$$x(t) = x_h(t) + x_p(t),$$

where $x_p(t)$ is some particular solution of (28). In order to solve IVP (28), (29) we just have to find constants c_1 and c_2 using initial conditions (29).

We first solve a homogeneous equation $ax'' + bx' + cx = 0$. Let's try to look for a solution in the form $x(t) = e^{\lambda t}$. We substitute this into equation and obtain

$$a\lambda^2 + b\lambda + c = 0.$$

Solving this algebraic equation we obtain two roots λ_1 and λ_2. There are three possibilities:

1. λ_1, λ_2 are real and not equal.

   In this case we obtain two linearly independent solutions $x_{1,h}(t) = e^{\lambda_1 t}$, $x_{2,h}(t) = e^{\lambda_2 t}$ and the general solution is $x_h(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$.

2. $\lambda_1 = \lambda_2 = \lambda$.

   In this case one solution is $x_{1,h}(t) = e^{\lambda t}$ and the second solution is $x_{2,h}(t) = e^{\lambda t} t$ (it's quite easy to check this by plugging it into homogeneous ODE). The general solution is $x_h(t) = c_1 e^{\lambda t} + c_2 e^{\lambda t} t$.

3. $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$ are complex conjugate roots.
   In this case there are two linearly independent solutions $x_{1,h}(t) = e^{(\alpha+i\beta)t}$ and $x_{2,h}(t) = e^{(\alpha-i\beta)t}$. These are complex solutions and in order to obtain real ones we have to take suitable linear combinations $\frac{x_{1,h}(t)+x_{2,h}(t)}{2} = e^{\alpha t}\cos\beta t$ and $\frac{x_{1,h}(t)-x_{2,h}(t)}{2i} = e^{\alpha t}\sin\beta t$. The general solution is $x_h(t) = c_1 e^{\alpha t}\cos\beta t + c_2 e^{\alpha t}\sin\beta t$.

In order to find particular solution one may employ the following method. The characteristic equation of (28) is

$$a\lambda^2 + b\lambda + c = 0.$$

It has two roots λ_1 and λ_2 and we may rewrite the differential operator as

$$ax'' + bx' + cx = aD^2x + bDx + cx = a(D-\lambda_1)(D-\lambda_2)x_p = f(t).$$

Here D is a differential operator, i.e. $Dx = x'$. We define $y = (D-\lambda_2)x_p = x_p' - \lambda_2 x_p$, then equation for y becomes

$$a(D-\lambda_1)y = f(t) \text{ or } ay' - ay = f(t).$$

This is a first order equation that can be easily solved to obtain y. Now we may solve another first order equation

$$x_p' - \lambda x_p = y(t)$$

to obtain x_p. This procedure yields

$$x_p(t) = \frac{1}{a}e^{\lambda_2 t}\int_t e^{(\lambda_1-\lambda_2)r}\int_r e^{-\lambda_1 s}f(s)dsdr$$

The algorithm for finding a solution of (28), (29) is:

1. Find the general solution of the homogeneous equation x_h (see above).
2. Find particular solution x_p by the above procedure.
3. The general solution is $x(t) = x_h(t) + x_p(t)$.
4. Use initial conditions (29) to find the constants c_1 and c_2 in $x_h(t)$.

This method can be easily generalized to a higher order equations.

Example. Find the solution of the following second order ODE

$$x'' + 5x' + 6x = 0.$$

The characteristic equation $\lambda^2 + 5\lambda + 6 = 0$ has two roots $\lambda_1 = -2$ and $\lambda_2 = -3$. Therefore the general solution is

$$x(t) = c_1 e^{-2t} + c_2 e^{-3t}$$

Example. Find the solution of the following second order ODE

$$x'' + 2x' + x = 0.$$

The characteristic equation $\lambda^2 + 2\lambda + 1 = 0$ has only one root $\lambda = -1$ and therefore one solution is $x_1(t) = e^{-t}$. The second solution $x_2(t)$ can be found as $x_2(t) = te^{-t}$. Therefore the general solution is

$$x(t) = c_1 e^{-t} + c_2 te^{-t}$$

Example. Find the solution of the following second order ODE

$$x'' + x = 0.$$

The characteristic equation $\lambda^2 + 1 = 0$ has two complex conjugate roots $\lambda_1 = i$ and $\lambda_2 = -i$. One solution is $x_1(t) = e^{it}$ and the second one is $x_2(t) = e^{-it}$. Therefore the general solution is

$$x(t) = c_1 e^{it} + c_2 e^{-it}.$$

However we don't want our solution to be complex and since $x_1(t)$ and $x_2(t)$ are complex conjugate solutions we may take their linear combinations $y_1(t) = \frac{x_1(t)+x_2(t)}{2} = \cos(t)$ and $y_2(t) = \frac{x_1(t)-x_2(t)}{2i} = \sin(t)$. Since equation is linear these are also solutions of our ODE and the general real solution is

$$x(t) = c_1 \cos(t) + c_2 \sin(t).$$

Example. Find the solution of the following IVP

$$x'' - x = t, \quad x(0) = 1, \; x'(0) = 0$$

We first solve a homogeneous equation $x'' - x = 0$. The characteristic equation is $\lambda^2 - 1 = 0$ has roots $\lambda_1 = 1$ and $\lambda_2 = -1$. Therefore a general solution of a homogeneous equation is $x_h(t) = c_1 e^t + c_2 e^{-t}$. Now let's find a particular solution x_p. It is easy to see that $x_p(t) = -t$ since $-t'' + t = t$. The general solution is

$$x(t) = x_h(t) + x_p(t) = c_1 e^t + c_2 e^{-t} - t.$$

Now we have to find c_1 and c_2 such that $x(t)$ solve given IVP. We know that $x(0) = 1$ and $x'(0) = 0$ hence we have

$$c_1 + c_2 = 1, \quad c_1 - c_2 = 1.$$

Solving this system we get $x(t) = e^t - t$.

5.2 Equations with variable coefficients.

In this section we will study the method of order reduction for homogeneous second order equations

$$x'' + p(t)x' + q(t)x = 0. \tag{30}$$

Definition 14 *Let $X(t)$ be a fundamental matrix associated with the equation (30)*

$$X(t) = \begin{pmatrix} x_1(t) & x_2(t) \\ x_1'(t) & x_2'(t) \end{pmatrix}$$

(here $x_1(t)$ and $x_2(t)$ are two linearly independent solutions of (30)). Wronskian of (30) is a function $W(t) = det X(t)$.

Lemma 7 *Wronskian $W(t)$ of the equation (30) satisfies the following ODE*

$$W'(t) = -p(t)W(t)$$

and therefore $W(t) = Ce^{-\int_t p(s)ds}$.

Proof 10 *By definition* $W(t) = x_1x_2' - x_2x_1'$ *and therefore*

$$W'(t) = x_1x_2'' - x_1'x_2' - x_2'x_1' - x_2x_1'' = -x_1(p(t)x_2' + q(t)x_2) + x_2(p(t)x_1' + q(t)x_1) =$$

$$= -p(t)(x_1x_2' - x_2x_1') = -p(t)W(t).$$

Method of order reduction.

Suppose we know one solution $u(t)$ of (30) and want to find the second solution $v(t)$. To do this we apply the following procedure. We know that Wronskian is defined as $W(t) = u(t)v'(t) - v(t)u'(t)$ and by previous lemma we know that $W(t) = Ce^{-\int_t p(s)ds}$. Therefore we have

$$u(t)v'(t) - v(t)u'(t) = e^{-\int_t p(s)ds}$$

(constant C may be absorbed into $v(t)$) and obtain the following equation for $v(t)$

$$v' - \frac{u'}{u}v = \frac{W(t)}{u}.$$

Since $u(t)$ and $W(t)$ are known functions we solve first order ODE for v and get the desired second solution.

Example. Find the second and the general solutions of

$$t^2x'' + 3tx' - 3x = 0$$

if first solution is given as $x(t) = t$.

We first put equation in a canonical form

$$x'' + \frac{3}{t}x' - \frac{3}{t^2}x = 0.$$

We find that Wronskian is $W(t) = e^{-\int_t \frac{3}{s}ds} = \frac{1}{t^3}$, on the other hand

$$W(t) = tx_2' - x_2.$$

Therefore we obtain a simple first order ODE

$$tx_2' - x_2 = \frac{1}{t^3}.$$

Solving this equation we obtain $x_2 = -\frac{1}{4t^3}$. The general solution is

$$x(t) = c_1t + c_2\frac{1}{t^3}$$

6 Power series solutions

We want to find a solution of the following IVP

$$a_n(t)x^{(n)} + a_{n-1}(t)x^{(n-1)} + ... + a_1(t)x' + a_0(t)x = f(t) \tag{31}$$

$$x(t_0) = x_0, ..., x^{(n-1)}(t_0) = x_{n-1} \tag{32}$$

In general this is quite difficult task, however, if the coefficients $a_k(t)$ and $f(t)$ are analytic then we may look for a power series solution

$$x(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k \tag{33}$$

Definition 15 *A function $f(t)$ is said to be analytic at point $t = t_0$ if there exists converging power series expansion in some neighborhood $\mathcal{N}$ of t_0, such that*

$$f(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k$$

for all $t \in \mathcal{N}$.

Taylor's theorem tells us that if a function $f(t)$ is infinitely differentiable in $\mathcal{N}$ then $f(t)$ is analytic at t_0 and

$$f(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(t_0)}{k!}(t - t_0)^k$$

The method of power series solutions says that we can find a solution of (31) by plugging in (33) into (31) and solving for coefficients c_k. There are several possibilities here that we must explore

Definition 16 *A point t_0 is called* **an ordinary point** *of differential operator*

$$L_n x = a_n(t)x^{(n)} + a_{n-1}(t)x^{(n-1)} + ... + a_1(t)x' + a_0(t)x$$

if all functions $p_k(t) = \frac{a_k(t)}{a_n(t)}$ are analytic at $t = t_0$; otherwise t_0 is called **a singular point**.

Definition 17 *A singular point t_0 of differential operator*

$$L_n x = a_n(t)x^{(n)} + a_{n-1}(t)x^{(n-1)} + ... + a_1(t)x' + a_0(t)x$$

is called **a regular singular point** *if all functions $q_k(t) = \frac{a_k(t)}{a_n(t)}(t - t_0)^{(n-k)}$ are analytic, otherwise t_0 is* **irregular singular point**.

We will consider two separate cases: series solutions at ordinary points and regular singular points.

6.1 Series solutions at ordinary points

The result for this case is based on the following theorem.

Theorem 12 *If $p_k(t)$ are analytic functions at $t = t_0$ then IVP*

$$x^{(n)} + p_{n-1}(t)x^{(n-1)} + ... + p_0(t)x = 0 \tag{34}$$

$$x(t_0) = x_0, ..., x^{(n-1)} = x_{n-1} \tag{35}$$

has an analytic solution at $t = t_0$ such that

$$x(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k.$$

The coefficients c_k are found by method of undetermined coefficients.

Let's give an example of how this theorem works.

Example. Find a solution of the following IVP

$$x'' + tx' + x = 0, \quad x(0) = 1,\ x'(0) = 0$$

By theorem 12 we may look for a solution in the form

$$x(t) = \sum_{k=0}^{\infty} c_k t^k.$$

We plug this series in our ODE and obtain

$$\sum_{k=0}^{\infty} k(k-1)c_k t^{k-2} + t\sum_{k=0}^{\infty} kc_k t^{k-1} + \sum_{k=0}^{\infty} c_k t^k = 0$$

We have to combine the coefficients in front of t^k. By simple calculations we obtain

$$\sum_{k=0}^{\infty}[(k+2)(k+1)c_{k+2} + c_k(k+1)]t^k = 0.$$

By linear independence of functions t^k we see that in order for the above relation to be satisfied all coefficients in front of t^k must be zeros, i.e.

$$c_{k+2}(k+2)(k+1) + c_k(k+1) = 0$$

for all k. From initial conditions we also obtain $c_0 = 1$ and $c_1 = 0$, therefore

$$c_{k+2} = -\frac{c_k}{k+2},\ c_0 = 1,\ c_1 = 0.$$

From this we see that all odd coefficients are zero and even coefficients satisfy

$$c_{2n} = \frac{(-1)^n}{2^n n!}.$$

The solution is

$$x(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n n!} t^{2n} = e^{-\frac{t^2}{2}}.$$

Based on this example we may introduce the algorithm for finding a solution at ordinary point.

Algorithm. In order to find a solution of IVP

$$x^{(n)} + p_{n-1}(t)x^{(n-1)} + ... + p_0(t)x = f(t) \tag{36}$$

$$x(t_0) = x_0, ..., x^{(n-1)} = x_{n-1} \tag{37}$$

near ordinary point t_0 we do the following steps

1. We look for a solution as
$$x(t) = \sum_{k=0}^{\infty} c_k (t - t_0)^k$$
2. First n coefficients c_k for $k = 0, .., n-1$ are found as
$$c_k = \frac{x_k}{k!}$$
3. Calculate the first n derivatives of $x(t)$ and plug them into ODE.
4. Expand all $a_k(t)$ and $f(t)$ in Taylor series around t_0 (if they are not polynomials already) and plug these expansions into ODE.
5. Combine the terms to form the equation of the form
$$\sum_{k=0}^{\infty} b_k (t - t_0)^k = 0,$$
where b_k is a function of some set of c_j.
6. Using linear independence of $(t - t_0)^k$ set $b_k = 0$ and find all c_k.

This method may also work for some regular singular points, however it's **not guaranteed**. In order to find a solution near regular singular points we have to use method of Frobenius and this will be explained in the next section.

6.2 Series solutions at regular singular points and Frobenius method

We will concentrate on second order ODEs and study the following equation

$$x'' + b(t)x' + c(t)x = 0. \tag{38}$$

Assume that t_0 is a regular singular point for this equation, then the functions

$$p(t) = (t - t_0)b(t), \quad q(t) = (t - t_0)^2 c(t)$$

are analytic at point t_0. If we multiply (38) by $(t - t_0)^2$ we obtain

$$(t - t_0)^2 x'' + (t - t_0)p(t)x' + q(t)x = 0. \tag{39}$$

This is the canonical form of the second order ODE with regular singularity at point t_0. We are going to look for a solution of (39) in the following form

$$x(t) = (t - t_0)^r S(t) = (t - t_0)^r \sum_{k=0}^{\infty} c_k (t - t_0)^k. \tag{40}$$

Here r may be a complex number and $S(t)$ is an analytic function at point t_0. Before formulating theorems let's try to see what this form of the solution suggests. We plug (40) into (39) to obtain

$$x' = r(t-t_0)^{r-1}S + (t-t_0)^r S',$$

$$x'' = r(r-1)(t-t_0)^{r-2}S + 2r(t-t_0)^{r-1}S' + (t-t_0)^r S'',$$

$$(t-t_0)^2[r(r-1)(t-t_0)^{r-2}S + 2r(t-t_0)^{r-1}S' + (t-t_0)^r S'']+$$

$$+(t-t_0)p(t)[r(t-t_0)^{r-1}S + (t-t_0)^r S'] + q(t)(t-t_0)^r S(t) = 0.$$

Combining terms we have

$$(t-t_0)^r\left[(t-t_0)^2 S'' + (2r+p(t))(t-t_0)S' + (r(r-1) + rp(t) + q(t))S\right] = 0 \quad (41)$$

If $t \neq t_0$ we have

$$(t-t_0)^2 S'' + [2r+p(t)](t-t_0)S' + [r(r-1) + rp(t) + q(t)]S = 0. \quad (42)$$

Since all functions in the above expression are analytic at t_0 we may take a limit as $t \to t_0$ to obtain

$$[r(r-1) + rp(t_0) + q(t_0)]S(t_0) = 0.$$

Therefore either $S(t_0) = 0$ or

$$r(r-1) + rp(t_0) + q(t_0) = 0. \quad (43)$$

Equation (43) is called **indicial equation** of ODE (39). Solving this equation we may find the powers r_1 and r_2 that **may** be relevant for solving (39). Based on above calculations we may formulate the following theorem

Theorem 13 *Let $p(t)$ and $q(t)$ be analytic functions at $t = t_0$ with radii of convergence R, and let r_1, r_2 be the roots of the indicial equation*

$$r(r-1) + rp(t_0) + q(t_0) = 0.$$

Then

1. *if r_1, r_2 are real and $r = \max\{r_1, r_2\}$ there exists a solution of (39) in the form*

$$x(t) = (t-t_0)^r \sum_{k=0}^{\infty} c_k (t-t_0)^k;$$

2. *if r_1, r_2 are real and $r_1 - r_2$ is not an integer then there are exactly two solutions of (39) in the form*

$$x_1(t) = (t-t_0)^{r_1} \sum_{k=0}^{\infty} a_k (t-t_0)^k,$$

$$x_2(t) = (t-t_0)^{r_2} \sum_{k=0}^{\infty} b_k (t-t_0)^k;$$

3. *if r_1, r_2 are complex conjugate roots then there are exactly two solutions of (39) in the form*

$$x_1(t) = (t-t_0)^{r_1} \sum_{k=0}^{\infty} a_k (t-t_0)^k,$$

$$x_2(t) = (t-t_0)^{r_2} \sum_{k=0}^{\infty} b_k (t-t_0)^k;$$

Let's solve some examples.

Example. Using theorem above find a solution of Bessel's equation of order $\frac{1}{2}$

$$t^2x'' + tx' + (t^2 - \frac{1}{4})x = 0.$$

We see that $t = 0$ is a regular singular point and $p(t) = 1$, $q(t) = (t^2 - \frac{1}{4})$ are both analytic functions at $t = 0$. Therefore we are in the conditions of the theorem 13. First we solve indicial equation

$$r(r-1) + r - \frac{1}{4} = 0$$

to obtain $r_1 = \frac{1}{2}$, $r_2 = -\frac{1}{2}$. Since $r_1 - r_2 = 1$ is an integer, by theorem 13 we may hope to find only one solution corresponding to $r = \frac{1}{2}$. The form of the solution is

$$x(t) = \sqrt{t}S(t) = \sqrt{t}\sum_{k=0}^{\infty} a_k t^k = \sum_{k=0}^{\infty} a_k t^{k+1/2}.$$

Plugging this into equation (or using equation (42) for $S(t)$) we obtain

$$\sum_{k=0}^{\infty} a_k k(k+1)t^k + \sum_{k=0}^{\infty} a_k t^{k+2} = 0,$$

and after rearrangement we have

$$2a_1 t + \sum_{k=0}^{\infty}[a_{k+2}(k+2)(k+3) + a_k]t^{k+2} = 0.$$

By linear independence of t^k we obtain

$$a_1 = 0, \quad a_{k+2} = -\frac{a_k}{(k+2)(k+3)}.$$

It's not difficult to show that $a_{2n+1} = 0$, $a_{2n} = \frac{(-1)^n a_0}{(2n+1)!}$, and therefore

$$x(t) = a_0\sqrt{t}\sum_{n=0}^{\infty}(-1)^n \frac{t^{2n}}{(2n+1)!} = a_0\frac{\sin t}{\sqrt{t}}.$$

We can not find the second solution in the form $x(t) = \frac{1}{\sqrt{t}}\sum_{k=0}^{\infty} b_k t^k$ using just the above theorem 13 (actually, in this particular case, we can, but it's pure luck).

Example. Find a solution of the following equation using Frobenius method

$$2t^2x'' - tx' + x = 0.$$

Obviously $t = 0$ is a regular singular point and $p(t) = -\frac{1}{2}$, $q(t) = \frac{1}{2}$. The indicial equation is

$$r(r-1) - \frac{1}{2}r + \frac{1}{2} = 0$$

and the roots are $r_1 = 1$, $r_2 = \frac{1}{2}$. By theorem 13 we may find two solutions of the ODE

$$x_1 = t\sum_{k=0}^{\infty} a_k t^k,$$

$$x_2 = \sqrt{t}\sum_{k=0}^{\infty} b_k t^k.$$

The first solution is quite easy to see without any calculations: $x_1(t) = a_0 t$. The second solution we may find by plugging $x_2(t)$ into ODE, yielding

$$\sum_{k=0}^{\infty} k(k+\frac{1}{2})b_k t^k = 0.$$

So all $b_k = 0$ except b_0 and $x_2(t) = b_0\sqrt{t}$. The general solution obviously is

$$x(t) = a_0 t + b_0\sqrt{t}.$$

Example. Using theorem above find a solution of the following equation

$$t^2 x'' + tx' + x = 0.$$

We see that $t = 0$ is a regular singular point and $p(t) = 1$, $q(t) = 1$ are both analytic functions at $t = 0$. Therefore we are in the conditions of the theorem 13. We solve indicial equation

$$r(r-1) + r + 1 = 0$$

to obtain two complex conjugate roots $r = i$ and $r = -i$. By theorem 13 we have two solutions of the form

$$x_1(t) = t^i \sum_{k=0}^{\infty} a_k t^k,$$

$$x_2(t) = t^{-i} \sum_{k=0}^{\infty} b_k t^k.$$

These are complex functions and when we plug them into ODE we **may** get complex coefficients a_k and b_k, however since ODE has just real coefficients $x_1(t)$ and $x_2(t)$ are complex conjugate functions and we may always find two real solutions by taking linear combination of x_1 and x_2, like

$$y_1(t) = \frac{x_1(t) + x_2(t)}{2},$$

$$y_2(t) = \frac{x_1(t) - x_2(t)}{2i}.$$

This is a general argument and can be applied for other problems where indicial equation has complex roots. Let's find $x_1(t)$

$$\sum_{k=0}^{\infty} k(k-1)a_k t^k + (2i+1)\sum_{k=0}^{\infty} k a_k t^k = 0$$

Combining the terms we obtain

$$\sum_{k=0}^{\infty} k(k+2i)a_k t^k = 0.$$

From this we see that

$$a_0 \text{ is any },\ a_k = 0 \text{ for all } k \geq 1.$$

So we see that $x_1(t) = t^i$ and obviously $x_2 = t^{-i}$, now we want to get real solutions by noticing that

$$t^i = e^{i\ln t} = \cos(\ln t) + i\sin(\ln t), \quad t^{-i} = \cos(\ln t) - i\sin(\ln t).$$

Therefore two real solutions are $y_1(t) = \cos(\ln t)$ and $y_2(t) = \sin(\ln t)$ and the general solution is

$$x(t) = a\cos(\ln t) + b\sin(\ln t).$$

The first Frobenius theorem does not tell us how to find the second solution of the ODE if the difference between the roots of indicial equation is integer. To find the second solution in this case we need the second Frobenius theorem.

Theorem 14 *Let $p(t)$ and $q(t)$ be analytic functions at $t = t_0$ with radii of convergence R, and let r_1, r_2 be the roots of the indicial equation*

$$r(r-1) + rp(t_0) + q(t_0) = 0$$

such that $r_1 - r_2 \in \mathbf{Z}$ (note that r_1 may coincide with r_2). Then the second solution of (39) is

$$x_2(t) = ax_1(t)\ln|t - t_0| + (t - t_0)^{r_2}\sum_{k=0}^{\infty} a_k(t - t_0)^k.$$

Let's consider the following example.

Example.

Let's find the second solution of Bessel's equation of order $\frac{1}{2}$

$$t^2x'' + tx' + (t^2 - \frac{1}{4})x = 0.$$

By theorem 14 we know that

$$x_2(t) = ax_1(t)\ln t + \frac{1}{\sqrt{t}}\sum_{k=0}^{\infty} a_k t^k = ax_1(t)\ln t + \frac{1}{\sqrt{t}}S(t).$$

We find x_2' and x_2'' to be

$$x_2' = ax_1'\ln t + a\frac{x_1}{t} + \frac{S'}{\sqrt{t}} - \frac{S}{2t^{3/2}}$$

$$x_2'' = ax_1''\ln t + 2a\frac{x_1'}{t} - a\frac{x_1}{t^2} + \frac{S''}{\sqrt{t}} - \frac{S'}{t^{3/2}} + \frac{3S}{4t^{5/2}}$$

We plug this into ODE to obtain

$$at^2\ln t(x_1'' + tx_1' + (t^2 - \frac{1}{4})x_1) + 2atx_1' + \frac{1}{\sqrt{t}}(t^2S'' + t^2S) = 0$$

Since $x_1(t)$ solves original ODE we have

$$(S'' + S) = -2ax_1' = -2a\sum_{n=0}^{\infty}(-1)^n(2n + \frac{1}{2})\frac{t^{2n-1}}{(2n+1)!}.$$

Now we recall that $S(t) = \sum_{k=0}^{\infty} a_k t^k$ and obtain

$$\sum_{k=0}^{\infty} k(k-1)a_k t^{k-2} + \sum_{k=0}^{\infty} a_k t^k = -2a\sum_{n=0}^{\infty}(-1)^n(2n + \frac{1}{2})\frac{t^{2n-1}}{(2n+1)!}.$$

The first term in the LHS corresponds to t^0 and the first term in the RHS corresponds to t^{-1}. Therefore the coefficient in front of t^{-1} must be 0 and hence $a = 0$. But this yields a simple equation for $S(t)$

$$S'' + S = 0,$$

So $S(t) = A\cos t + B\sin t$ and $x_2(t) = \frac{A\cos t + B\sin t}{\sqrt{t}}$ (since $x_1(t) = \frac{\sin t}{\sqrt{t}}$ we may take $x_2(t) = \frac{\cos t}{\sqrt{t}}$).

7 Nonlinear systems

In this section we study nonlinear autonomous 2D systems of ODEs.

$$x' = F(x, y) \tag{44}$$

$$y' = G(x, y).$$

To study initial value problem (IVP) for (44) we have to supply initial conditions

$$x(t_0) = x_0,\ y(t_0) = y_0. \tag{45}$$

In general, it is impossible to solve nonlinear systems of ODEs. We will try to find out some qualitative behavior of the system like we did for first order ODEs. To do this we need several definitions.

Definition 18 *A critical point (equilibrium point) of the system (44) is a point* (x_c, y_c) *such that* $F(x_c, y_c) = G(x_c, y_c) = 0$.

Obviously $x(t) = x_*$, $y(t) = y_*$ is a solution of (44). It is called equilibrium solution.

Example. Let's find equilibrium points of the following system

$$x' = y(x - 2)$$

$$y' = x(y - 1)$$

In order to find critical points we have to solve

$$y(x - 2) = 0 \text{ and } x(y - 1) = 0.$$

This gives us the following critical points $(2, 1)$, $(0, 0)$.

Definition 19 *A critical point* $\mathbf{x}_c = (x_c, y_c)$ *is stable if for every* $\epsilon > 0$ *there exists* $\delta > 0$ *such that if* $|\mathbf{x}_0 - \mathbf{x}_c| < \delta$ *the solution of IVP (44), (45) satisfies* $|\mathbf{x}(t) - \mathbf{x}_c| < \epsilon$.

Definition 20 *A critical point is asymptotically stable if it is stable and if* $|\mathbf{x}_0 - \mathbf{x_c}| < \delta$ *for some delta implies that* $\lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}_c$.

There are four possible types of behavior for solutions (trajectories) of nonlinear system:

1. $\mathbf{x}(t) \to \mathbf{x}_c$ as $t \to \infty$;
2. $\mathbf{x}(t)$ is unbounded;
3. $\mathbf{x}(t)$ is periodic;
4. $\mathbf{x}(t)$ converges to a periodic solution.

7.1 Linearization of nonlinear systems

In order to study qualitative long time behavior of nonlinear systems we are going to linearize the system about critical points. After that we may classify critical points of the linearized system and transfer their stability to nonlinear system (if it is possible). This method will give us asymptotic behavior of the nonlinear system.

Let's briefly describe this method. We study

$$\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}),$$

suppose $\mathbf{x}_c$ is an isolated critical point of $\mathbf{f}(\mathbf{x})$, then if $\mathbf{x}$ is close to $\mathbf{x}_c$ we may approximately get

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x_c}) \approx D\mathbf{f}(\mathbf{x_c})(\mathbf{x} - \mathbf{x}_c).$$

Since $\mathbf{f}(\mathbf{x}) = (F(x,y), G(x,y))$ we know that Jacobi matrix

$$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \frac{\partial F(x,y)}{\partial x} & \frac{\partial F(x,y)}{\partial y} \\ \frac{\partial G(x,y)}{\partial x} & \frac{\partial G(x,y)}{\partial y} \end{pmatrix}.$$

Therefore around critical point instead of original nonlinear system we may study linear system

$$\mathbf{x}'(t) = D\mathbf{f}(\mathbf{x_c})(\mathbf{x} - \mathbf{x}_c),$$

or noticing that $\mathbf{x}_c$ is independent of t we obtain

$$(\mathbf{x}(t) - \mathbf{x}_c)' = D\mathbf{f}(\mathbf{x_c})(\mathbf{x}(t) - \mathbf{x}_c).$$

Changing variables $\mathbf{z}(t) = \mathbf{x}(t) - \mathbf{x}_c$ we get linearization of nonlinear system near critical point $\mathbf{x}_c$:

$$\mathbf{z}'(t) = D\mathbf{f}(\mathbf{x_c})\mathbf{z}$$

Example. Let's consider the following nonlinear system

$$x' = x(y-1)$$

$$y' = y(x+2)$$

There are two critical points here $(0,0)$ and $(-2,1)$, let's linearize around them. It's quite straightforward to calculate $D\mathbf{f}(\mathbf{x})$:

$$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} y-1 & x \\ y & x+2 \end{pmatrix}.$$

At critical point $(0,0)$ we have

$$D\mathbf{f}(\mathbf{0},\mathbf{0}) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

and therefore we suspect that near $(0,0)$ our nonlinear system should behave approximately as the following linear system

$$\mathbf{x}' = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}\mathbf{x}.$$

By the same token we may get

$$D\mathbf{f}(-\mathbf{2},\mathbf{1}) = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$$

and therefore we suspect that near $(-2,1)$ our nonlinear system should behave approximately as the following linear system

$$\mathbf{x}' = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \mathbf{x}.$$

So far we don't know if we may transfer behavior of the linearized system to nonlinear system. In order to make a statement more precise we formulate the linearization theorem.

Theorem 15 *The type and stability of critical points in the nonlinear system and linearized system are the same as long as in the linearized system it's not on the border-line case in the trace-determinant plane, i.e.*

- *Nodes, saddles, spirals are not altered.*
- *Centers, degenerate nodes may be altered.*

Note that critical points that may be altered are exactly those where real part of eigenvalues is 0 or eigenvalues are real and coinciding. It's quite easy to justify since during the transfer you perturb our system a little bit and this may introduce small changes to the eigenvalues. But then real part may become non-zero or in case of equal eigenvalues they may become non-equal, hence the stability will change.

Example. Let's look at the previous example. We found out that critical points are $(0,0)$ and $(-2,1)$. It's quite easy to see that $(0,0)$ is a saddle and $(-2,1)$ is a center in the linearized system. Linearization theorem tells us that $(0,0)$ will be a saddle in the nonlinear system, but the type and stability of the $(-2,1)$ may be different from center. In order to find what is $(-2,1)$ in the nonlinear system we may not use linearization theorem and have to do something else.

7.2 Predator-Prey Model

$$\begin{aligned} x' &= ax - pxy \\ y' &= -by + qxy \end{aligned} \tag{46}$$

Critical points are $(0,0)$ and $(\frac{b}{q}, \frac{a}{p})$. Let's find out their stability. Jacobi matrix $D\mathbf{f}$ is

$$D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} a - py & -px \\ qy & -b + qx \end{pmatrix}.$$

Critical point $(0,0)$ gives us

$$D\mathbf{f}(\mathbf{0}, \mathbf{0}) = \begin{pmatrix} a & 0 \\ 0 & -b \end{pmatrix},$$

hence it is a saddle and therefore it's a saddle in the nonlinear system. Critical point $(\frac{b}{q}, \frac{a}{p})$ gives us

$$D\mathbf{f}(\mathbf{0}, \mathbf{0}) = \begin{pmatrix} 0 & -b\frac{p}{q} \\ a\frac{q}{p} & 0 \end{pmatrix},$$

hence it is a center and we can not transfer this to nonlinear system. Therefore we have to investigate this using different methods. Let's divide second equation by the first one, then

$$\frac{dy}{dx} = \frac{y(-b + qx)}{x(a - py)}.$$

We may separate variables in this equation to obtain

$$\frac{dy}{y}(a - py) = \frac{dx}{x}(-b + qx).$$

Solving this equation we obtain

$$a \ln |y| - py + b \ln |x| - qx = C$$

For each fixed constant C we obtain a trajectory (x, y) on the phase plane. We may introduce a function

$$z(x, y) = a \ln |y| - py + b \ln |x| - qx$$

and say that level curves $z(x, y) = C$ correspond to trajectories of the system on the phase plane. At this point we may already draw the phase plane portrait since we explicitly know it. But it is quite difficult to do without using computers and we want to analyze the system analytically, therefore we investigate the stability of $(\frac{b}{q}, \frac{a}{p})$. In order to do this we have to see how $z(x, y)$ behaves near $(\frac{b}{q}, \frac{a}{p})$. It's not difficult to verify that at $(\frac{b}{q}, \frac{a}{p})$

$$\frac{\partial z}{\partial x} = 0, \ \frac{\partial z}{\partial y} = 0.$$

So it is a critical point of $z(x, y)$. Let's try to see if it's minimum, maximum or a saddle for $z(x, y)$. To find this out we have to take a second derivative of z:

$$D^2 z = \begin{pmatrix} \frac{\partial^2 z}{\partial x^2} & \frac{\partial^2 z}{\partial x \partial y} \\ \frac{\partial^2 z}{\partial x \partial y} & \frac{\partial^2 z}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -\frac{b}{x^2} & 0 \\ 0 & -\frac{a}{y^2} \end{pmatrix}$$

At the critical point $(\frac{b}{q}, \frac{a}{p})$ $D^2 z$ is negative definite and therefore $(\frac{b}{q}, \frac{a}{p})$ is a maximum. Since the level curves of z near maximum are closed curves that shrink as approaching critical point we may deduce that $(\frac{b}{q}, \frac{a}{p})$ is a center.

7.3 Hamiltonian systems

We study the following system

$$\begin{aligned} x' &= F(x, y) \\ y' &= G(x, y) \end{aligned} \tag{47}$$

with $F(x, y) = \frac{\partial H(x,y)}{\partial y}$ and $G(x, y) = -\frac{\partial H(x,y)}{\partial x}$ for some continuously differentiable function $H(x, y)$. Such systems are called Hamiltonian systems and function $H(x, y)$ is called a hamiltonian (or energy of the system).

The main property of Hamiltonian system is conservation of energy.

Theorem 16 *Total energy $H(x, y)$ of Hamiltonian system (47) remains constant along trajectories of (47).*

It's quite easy to see why this is true:

$$\frac{dH}{dt}(x(t), y(t)) = \frac{\partial H(x,y)}{\partial x} x' + \frac{\partial H(x,y)}{\partial y} y' = \frac{\partial H(x,y)}{\partial x} \frac{\partial H(x,y)}{\partial y} - \frac{\partial H(x,y)}{\partial x} \frac{\partial H(x,y)}{\partial y} = 0.$$

This implies $H(x(t), y(t)) = const.$

How do we know if the system is Hamiltonian? It's enough to check if

$$\frac{\partial F(x,y)}{\partial x} = -\frac{\partial G(x,y)}{\partial y}.$$

Example.

$$x' = y$$
$$y' = -\sin x$$

Is this system Hamiltonian? Yes, since

$$\frac{\partial y}{\partial x} = \frac{\partial \sin x}{\partial y} = 0.$$

The hamiltonian is $H(x,y) = \frac{y^2}{2} + (1 - \cos x)$.

The good thing about Hamiltonian systems is that is you know $H(x,y)$, you know the trajectories of the system. This is true since level curves of function $H(x,y)$ contain trajectories of (47). How do we find a Hamiltonian? It's quite easy:

$$H(x,y) = \int_y F(x,s)ds + \phi(x)$$

Here $\phi(x)$ is unknown function that can be found from

$$\frac{\partial H(x,y)}{\partial x} = \frac{\partial}{\partial x}\int_y F(x,s)ds + \phi'(x) = -G(x,y).$$

One of very important properties of Hamiltonian systems is the statement about their critical points. Let (x_c, y_c) be a critical point of Hamiltonian system

$$x' = \frac{\partial H(x,y)}{\partial y}$$

$$y' = -\frac{\partial H(x,y)}{\partial x}$$

Since trajectories correspond to level curves of $H(x,y)$ we may actually show that

1. (x_c, y_c) is a saddle if and only if it is a saddle of $H(x,y)$;
2. (x_c, y_c) is a center if and only if it is a local minimum or maximum of $H(x,y)$;

The proof of this fact is an exercise.

So for hamiltonian systems we don't have sinks, sources or spirals and it's due to the fact that energy is conserved.

7.4 First integrals and conservative systems

We study the following system

$$x' = F(x,y) \tag{48}$$
$$y' = G(x,y)$$

Definition 21 *Let $K(x,y)$ be continuously differentiable function such that*

$$\frac{dK}{dt}(x(t),y(t))=0$$

along trajectories of (48) and K is non-constant on every ball in $\mathbf{R}^n$. Then $K(x,y)$ is called the first integral of the system.

One of examples of first integrals is hamiltonian $H(x,y)$ for Hamiltonian systems.

Definition 22 *The system (48) is called conservative if it has the first integral $K(x,y)$.*

Conservative systems are very similar to Hamiltonian systems since:

1. trajectories of conservative systems lie on the level curves of first integral $K(x,y)$;
2. conservative systems don't have neither attractors, not repellers.

Example. Find first integral of the following system if it exists

$$x'=xy^2$$
$$y'=yx^2$$

Dividing one equation by another and separating variables we obtain

$$ydy=xdx$$

Therefore $K(x,y)=y^2-x^2$, let's check it:

$$\frac{dK}{dt}(x(t),y(t))=2xx'-2yy'=2x^2y^2-2x^2y^2=0.$$

So the system is conservative.

7.5 Lyapunov functions

Let's consider the following system

$$x'=y \tag{49}$$
$$y'=-y-\sin x$$

It describes a pendulum with friction. We already know that for pendulum without friction the energy is conserved and hamiltonian is

$$H(x,y)=\frac{y^2}{2}+(1-\cos x).$$

Obviously $H(x,y)>0$ except at points $(2\pi n,0)$. Let's try to see the behavior of time derivative of $H(x,y)$ along trajectories of (49)

$$\frac{dH}{dt}(x(t),y(t))=\frac{\partial H(x,y)}{\partial x}x'+\frac{\partial H(x,y)}{\partial y}y'=-y^2\leq 0.$$

We recall that points $(2\pi n,0)$ were stable spirals for system (49). Is there any relation between behavior of $H(x,y)$ and stability of critical points? The answer is yes. It turns out that $H(x,y)$ in this example is a Lyapunov function for system (49).

Definition 23 *Let (x_c, y_c) be a critical point of the system*

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

$$y' = G(x, y)$$

A continuously differentiable function $V(x, y)$ is called a weak Lyapunov function for this system if

1. $V(x_c, y_c) = 0$;
2. $V(x, y) > 0$ *in an open neighborhood of* (x_c, y_c);
3. $\frac{dV}{dt}(x(t), y(t)) \leq 0$ *for all* (x, y) *in some open neighborhood of* (x_c, y_c);

and it is strong Lyapunov function if in 3) inequality is strict.

Theorem 17 *Let $(x_c, y_c) \in E$ - open set and $F(x_c, y_c) = G(x_c, y_c) = 0$ (i.e. (x_c, y_c) is a critical point of the system*

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

$$y' = G(x, y).$$

Suppose that for (x_c, y_c) there exists a weak (strong) Lyapunov function $V(x, y)$ then critical point (x_c, y_c) is stable (asymptotically stable).

Lyapunov functions are important since they may provide information about stability of critical points with minimum information and where other methods fail.

Example. Let's consider system

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

$$y' = -x - G(x, y)$$

Suppose we know that $F(0, 0) = G(0, 0) = 0$ and that $xF + yG > 0$ in some open neighborhood of $(0, 0)$ (excluding $(0, 0)$). Can we say anything about stability of the origin? It turns out to be that yes, we can. Define $V(x, y) = x^2 + y^2$, obviously

$$\frac{dV}{dt} = -2xF(x, y) - 2yG(x, y) < 0.$$

Therefore $V(x, y)$ is a strong Lyapunov function for the origin and $(0, 0)$ is asymptotically stable.

8 1D Calculus of Variations.

We are going to study the following general problem:

- Minimize functional
$$I(u) = \int_a^b f(x, u(x), u'(x))\, dx$$
subject to boundary conditions $u(a) = \alpha,\ u(b) = \beta$.

In order to correctly set up this problem we have to assume certain properties of $f(x, u, \xi)$ and a function $u(x)$. Since we know how to integrate continuous functions it is reasonable to assume that function $f : [a, b] \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ is continuous. In fact in this section we will need stronger assumption that $f \in C^2([a, b] \times \mathbf{R} \times \mathbf{R})$, meaning that f, f' and f'' are continuous functions.

Concerning function $u : [a, b] \to \mathbf{R}$ we may assume that $u \in C^1([a, b])$ and $u(a) = \alpha$, $u(b) = \beta$. It seems reasonable since minimization problem involves $u'(x)$ (that needs to be continuous since we know how to integrate continuous functions) and boundary conditions. In fact in this section we will use a bit stronger regularity assumption $u \in C^2([a, b])$.

Note: we can change boundary conditions by removing constraints at one or both ends. In general, one can not prescribe u' at the end points.

We are ready to set up the problem with mathematical precision:

Define

$$I(u) \equiv \int_a^b f(x, u(x), u'(x))\, dx,$$

where $f \in C^2([a, b] \times \mathbf{R} \times \mathbf{R})$, $u \in C^1([a, b])$.

Problem:

$$\inf_{u \in X} I(u), \tag{50}$$

where $X = \{u \in C^1([a, b]) : u(a) = \alpha,\ u(b) = \beta\}$.

We want to find a function $u_0(x)$ such that $u_0 \in X$ and u_0 delivers minimum to the functional $I(u)$ (we call such function a *minimizer*). In order to do this we need to understand what it means to minimize the functional $I(u)$.

Definition 24 *Function $u_0 \in X$ is a minimizer (or global minimizer) of the functional $I(u)$ on a set X if*

$$I(u_0) \le I(u) \quad \textit{for all } u \in X.$$

In general it is very difficult to solve (50), however one can try to characterize minimizers. Let's consider a simple example from Calculus.

Example 1 *Let $g \in C^1([a, b])$ and we want to find a minimizer of g. By well known result of Analysis we know that any continuous function on a compact interval (closed and bounded) achieves its minimum (and maximum). Obviously the minimizer is just some point $x_0 \in [a, b]$. In order to find it we have to use the definition, i.e.*

$$g(x_0) \le g(x) \quad \textit{for all } x \in [a, b].$$

Assume that $x_0 \in (a, b)$ then for any $x \in (a, b)$ we have

$$\frac{f(x) - f(x_0)}{x - x_0} \ge 0, \quad \textit{if } x > x_0$$

and

$$\frac{f(x) - f(x_0)}{x - x_0} \le 0, \quad \textit{if } x < x_0.$$

Taking a limit $x \to x_0$ we obtain $f'(x_0) \geq 0$ and $f'(x_0) \leq 0$. Therefore we have $f'(x_0) = 0$. **Note**: we must assume $x_0 \in (a, b)$ in order to obtain $f'(x_0) = 0$, explain it.

In order to find a minimizer we must find all points where $f'(x) = 0$ and compare values at these points with values at the ends of the interval ($f(a)$, $f(b)$).

In order to characterize minimizers of $I(u)$ we proceed in the similar way. We know that if u_0 is a minimizer then $I(u_0) \leq I(u)$ for all $u \in X$. Take any function $h \in C_0^1([a, b])$, obviously $u_0 + th \in X$ and therefore

$$I(u_0) \leq I(u_0 + th) \quad \text{for all } h \in C_0^1([a, b]).$$

We can define a function $F(t) = I(u_0 + th)$ and notice that since u_0 is a minimizer of $I(u)$ then $t = 0$ is a minimizer of $F(t)$ on $\mathbf{R}$. Therefore we must have $F'(0) = 0$ (see example).

Note: we a priori assumed existence of minimizer $u_0 \in X$ for $I(u)$ and just trying to characterize it **if it exists**. *In the example above for 1D function existence was guaranteed by a simple result of Analysis (look it up).*

We characterize minimizers in the following theorem

Theorem 18 *Let $f \in C^2([a, b] \times \mathbf{R} \times \mathbf{R})$ ($f \equiv f(x, u, \xi)$) and*

$$\inf_{u \in X} \left\{ I(u) \equiv \int_a^b f(x, u(x), u'(x))\, dx \right\} = m, \tag{51}$$

where $X = \{u \in C^1([a, b]) \,:\, u(a) = \alpha,\ u(b) = \beta\}$. Then

1. *If (51) admits a minimizer $u_0 \in X \cap C^2([a, b])$ then*
$$\frac{d}{dx}\left(f_\xi(x, u_0(x), u_0'(x))\right) = f_u(x, u_0(x), u_0'(x)) \quad x \in (a, b) \tag{52}$$
$$u_0(a) = \alpha,\ u_0(b) = \beta.$$
2. *If some function $u_* \in X \cap C^2([a, b])$ satisfies (52) and if $(u, \xi) \to f(x, u, \xi)$ is convex (as a function of two variables) for every $x \in [a, b]$ then u_* is a minimizer of (51).*

Proof. 1) As we discussed before, if $u_0 \in X$ is a minimizer of $I(u)$ then $F'(0) = 0$ (see the definition of F above). Let's find $F'(0)$: by definition of derivative

$$F'(0) = \lim_{t \to 0} \frac{F(t) - F(0)}{t} = \lim_{t \to 0} \frac{I(u_0 + th) - I(u_0)}{t}.$$

It's not difficult to see that

$$\frac{I(u_0 + th) - I(u_0)}{t} = \int_a^b \frac{f(x, u_0(x) + th(x), u_0'(x) + th'(x)) - f(x, u_0(x), u_0'(x))}{t}\, dx. \tag{53}$$

Passing to the limit and using standard Analysis results to interchange integral and limit (for instance Dominated Convergence Theorem) we obtain

$$F'(0) = \int_a^b \frac{d}{dt} f(x, u_0(x) + th(x), u_0'(x) + th'(x))\Big|_{t=0}\, dx.$$

Using standard rules of differentiation we have

$$\frac{d}{dt} f(x, u_0(x) + th(x), u_0'(x) + th'(x))\Big|_{t=0}$$
$$= f_u(x, u_0(x), u_0'(x))h(x) + f_\xi(x, u_0(x), u_0'(x))h'(x). \quad (54)$$

Therefore if $u_0 \in X$ is a minimizer of $I(u)$ we must have

$$\int_a^b f_u(x, u_0(x), u_0'(x))h(x) + f_\xi(x, u_0(x), u_0'(x))h'(x)\, dx = 0$$

for all $h \in C_0^1([a, b])$. This is a weak form of Euler-Lagrange equation. It is not clear how to deal with this integral identity and therefore we want to obtain a pointwise identity. We can use integration by parts to obtain

$$\int_a^b f_\xi(x, u_0(x), u_0'(x))h'(x)\, dx$$
$$= -\int_a^b \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right) h(x)\, dx + f_\xi(x, u_0(x), u_0'(x))h(x)\Big|_a^b. \quad (55)$$

Recalling that $h(a) = h(b) = 0$ we have

$$\int_a^b \left[f_u(x, u_0(x), u_0'(x)) - \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right)\right] h(x)\, dx = 0$$

for all $h \in C_0^1([a, b])$. Using Fundamental Lemma of Calculus of Variations (proved after this theorem) we obtain

$$f_u(x, u_0(x), u_0'(x)) - \frac{d}{dx} \left(f_\xi(x, u_0(x), u_0'(x))\right) = 0 \quad x \in (a, b).$$

Conditions $u_0(a) = \alpha$, $u_0(b) = \beta$ follow from the fact $u_0 \in X$.

2) Let $u_* \in X \cap C^2([a, b])$ is a solution of (52). Since $f(x, u, \xi)$ is convex in last two variables we have

$$f(x, u(x), u'(x)) \geq f(x, u_*(x), u_*'(x))$$
$$+ f_u(x, u_*(x), u_*'(x))(u(x) - u_*(x))$$
$$+ f_\xi(x, u_*(x), u_*'(x))(u'(x) - u_*'(x)) \quad (56)$$

for every $u \in X$ (this inequality will be shown later for a convex function of one variable, extend it). We can integrate both parts of this inequality over (a, b) to obtain

$$\int_a^b f(x, u(x), u'(x))\, dx \geq \int_a^b f(x, u_*(x), u_*'(x))\, dx$$
$$+ \int_a^b f_u(x, u_*(x), u_*'(x))(u(x) - u_*(x))\, dx$$
$$+ \int_a^b f_\xi(x, u_*(x), u_*'(x))(u'(x) - u_*'(x))\, dx. \quad (57)$$

Using integration by parts and noting that $u - u_* \in C_0^1([a, b])$ we obtain

$$I(u) \geq I(u_*)$$
$$+ \int_a^b \left[f_u(x, u_*(x), u_*'(x)) - \frac{d}{dx} \left(f_\xi(x, u_*(x), u_*'(x))\right)\right] (u(x) - u_*(x))\, dx. \quad (58)$$

Since u_* satisfies (52) we have $I(u) \leq I(u_*)$ for all $u \in X$.

Theorem is proved.

Remark 1 *Note that we did not prove existence of minimizer in this theorem. In the first statement we showed that if u_0 is a minimizer then it solves Euler-Lagrange equation. In the second statement we showed that if u_* solves Euler-Lagrange equation and* **the function f is convex in last two variables** *then u_* is a minimizer. But we did not prove existence of solution for Euler-Lagrange equation and therefore even the case of convex f we don't know if there is a minimizer.*

Now we prove Fundamental Lemma of Calculus of Variations (FLCV)

Lemma 8 *Let $g \in C([a,b])$ and*

$$\int_a^b g(x)\eta(x)\,dx = 0 \quad \text{for all } \eta \in C_0^\infty([a,b]),$$

then $g(x) = 0$ on $[a,b]$.

Proof. Suppose there exists a point $c \in (a,b)$ such that $g(c) \neq 0$. We may assume without loss of generality that $g(c) > 0$. By continuity of function $g(x)$ there exists an open interval $(s,t) \subset (a,b)$ containing point c such that $g(x) > 0$ on (s,t). Taking $\eta(x)$ to be positive on (s,t) with $\eta(s) = \eta(t) = 0$ we obtain the contradiction with $\int_a^b g(x)\eta(x)\,dx = 0$. Therefore $g(x) = 0$ for all $x \in (a,b)$. By continuity of g we obtain $g(x) = 0$ on $[a,b]$.

Lemma is proved.

For convenience we also prove a simple fact that for $g \in C^1([a,b])$ convexity is equivalent to

$$g(x) \geq g(y) + g'(y)(x-y) \quad \text{for all } x, y \in (a,b).$$

Proof. Let g be convex then by definition

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)g(y)$$

for all $\alpha \in (0,1)$ and x, y in (a,b). After simple rearrangement we have

$$\frac{g(y + \alpha(x-y)) - g(y)}{\alpha} \leq g(x) - g(y).$$

Taking a limit as $\alpha \to 0$ we obtain

$$g'(y)(x-y) \leq g(x) - g(y).$$

We proved it in one direction. Now we assume

$$g(x) \geq g(y) + g'(y)(x-y) \quad \text{for all } x, y \in (a,b).$$

Using this we have

$$g(x) \geq g(\alpha x + (1-\alpha)y) + (1-\alpha)g'(\alpha x + (1-\alpha)y)(x-y)$$

and

$$g(y) \geq g(\alpha x + (1-\alpha)y) - \alpha g'(\alpha x + (1-\alpha)y)(x-y).$$

Multiplying first inequality by α, second inequality by $1-\alpha$ and adding them we obtain

$$\alpha g(x) + (1-\alpha)g(y) \geq g(\alpha x + (1-\alpha)y).$$

Result is proved.

Note that this fact is true also for functions defined on $\mathbf{R}^n$.

Let's consider several examples.

Example 2 (Brachistochrone problem) *In this example we have*

$$f(u,\xi) = \sqrt{\frac{1+\xi^2}{u}},$$

and $X = \{u \in C^1([0,1]) \,:\, u(0) = 0,\, u(1) = \beta,\, u(x) > 0 \text{ for } x \in (0,1)\,\}$. *We would like to find solutions of Euler-Lagrange equation.*

Using theorem 18 we can easily find Euler-Lagrange equations

$$\left(\frac{u'}{\sqrt{u[1+(u')^2]}}\right)' = -\sqrt{\frac{1+(u')^2}{4u^3}}.$$

We can multiply both parts by $\frac{u'}{\sqrt{u[1+(u')^2]}}$ *and obtain*

$$\left(\frac{u'}{\sqrt{u[1+(u')^2]}}\right)' \frac{u'}{\sqrt{u[1+(u')^2]}} = -\frac{u'}{2u^2}.$$

It is clear that this is equivalent to

$$\left(\frac{(u')^2}{u[1+(u')^2]}\right)' = \left(\frac{1}{u}\right)'.$$

Now we obtain

$$\frac{1}{u} = \frac{(u')^2}{u[1+(u')^2]} + C$$

and simplifying this expression we obtain

$$u[1+(u')^2] = 2\mu,$$

where $\mu > 0$ *is some constant that we can find from boundary conditions. Separating variables we obtain*

$$dx = \sqrt{\frac{u}{2\mu - u}}du.$$

Now we integrate both parts

$$x = \int_0^u \sqrt{\frac{y}{2\mu - y}}dy.$$

We can make a natural substitution $y = \mu(1-\cos\theta)$ *to obtain* $x = \mu(\theta - \sin\theta)$. *Therefore solution to this problem can be represented in the following parametric form:*

$$u(\theta) = \mu(1-\cos\theta),\ x(\theta) = \mu(\theta - \sin\theta).$$

It is clear that $u(0) = 0$ *so that first boundary condition is satisfied and we can find* μ *by applying second boundary condition.*

Note that we don't know if this solution is a minimizer of the problem. For this we have to prove either existence of minimizer or convexity of $f(u,\xi)$. *However from physical model we can see that there must be a solution here (therefore minimizer should exist) and since solution to the problem is unique it must be a minimizer (I guess that was the reasoning in 17-th century).*

Example 3 (Minimal surface of revolution) *In this example we have*

$$f(u,\xi) = 2\pi u\sqrt{1+\xi^2},$$

and $X = \{u \in C^1([0,1]) \,:\, u(0) = \alpha,\, u(1) = \beta,\, u(x) > 0\,\}$. *We would like to find solutions of Euler-Lagrange equation.*

Again we can use the theorem 18 to obtain

$$\left(\frac{uu'}{\sqrt{1+(u')^2}}\right)' = \sqrt{1+(u')^2}.$$

Multiplying both parts by $\frac{uu'}{\sqrt{1+(u')^2}}$ *and integrating we obtain*

$$\frac{u^2}{1+(u')^2} = C, \text{ or } (u')^2 = \frac{u^2}{a^2} - 1$$

for some constant $a > 0$. *We can separate variables to obtain the solution here but there is a better way: we search for a solution in the following form*

$$u(x) = a\cosh\frac{f(x)}{a}.$$

Plugging this into equation we obtain $[f'(x)]^2 = 1$ *and therefore either* $f(x) = x + \mu$ *or* $f(x) = -x + \mu$. *Since* $\cosh x$ *is even function and* μ *is any constant we have*

$$u(x) = a\cosh\frac{x+\mu}{a}.$$

Continue to find the solutions of this problem assuming $u(0) = u(1) = \alpha > 0$. *The number of solutions will depend on* α.

Example 4 (Convexity) *In this example we assume that* $f(\xi)$ *is convex on* $\mathbf{R}$ *and* $X = \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\, u(b) = \beta,\,\}$. *We will prove existence of minimizer for*

$$I(u) = \int_a^b f(u'(x))\,dx$$

in this case and find it explicitly. By convexity of f *we know that*

$$f(x) \geq f(y) + f'(y)(x-y).$$

We can define $u_0(x) = \frac{\beta-\alpha}{b-a}(x-a) + \alpha$. *Using above inequality we see that for any* $u \in X$

$$f(u'(x)) \geq f(u_0'(x)) + f'(u_0'(x))(u'(x) - u_0'(x)).$$

It is clear that $u_0'(x) \equiv \frac{\beta-\alpha}{b-a}$ *and therefore integrating both parts over* (a,b) *we obtain*

$$\int_a^b f(u'(x))\,dx \geq \int_a^b f(u_0'(x))\,dx.$$

Therefore u_0 *is a minimizer of* $I(u)$.

Now we give several examples of nonexistence.

Example 5 (Non-convexity implies no minimizer) *We consider* $f(\xi) = e^{-\xi^2}$ *and* $X = \{u \in C^1([0,1]) \,:\, u(0) = 0,\, u(1) = 0,\}$ *and want to show that for the problem*

$$\inf_X \int_0^1 f(u'(x))\,dx$$

there is no solution.

We see that Euler-Lagrange equation is

$$\frac{d}{dx} f'(u') = 0.$$

And therefore $u' = const$ *that implies in this case (using boundary conditions) that* $u(x) \equiv 0$. *It is clear that this is a maximizer of the problem, not a minimizer. Moreover, if we take*

$$u_n(x) = n(x - \frac{1}{2})^2 - \frac{n}{4}$$

we have $u_n \in X$ *for any* $n \in \mathbb{N}$. *Calculating* $I(u_n)$ *we obtain*

$$I(u_n) = \int_0^1 e^{-4n^2(x-\frac{1}{2})^2}\,dx = \frac{1}{2n}\int_{-n}^{n} e^{-x^2}\,dx \to 0,$$

as $n \to \infty$. *Therefore* $\inf_X I(u) = 0$ *but obviously no function can deliver this infimum and so minimizer does not exist.*

Example 6 (Not smooth minimizers) *Consider* $f(\xi) = (\xi^2 - 1)^2$, $X = \{u \in C^1([0,1]) \,:\, u(0) = 0,\, u(1) = 0,\}$. *Show that there are no minimizers for this problem. However there are infinitely many minimizers that are piecewise smooth.*