A brief sketch of the main ODE theorems

Math 611, Fall 2017

1 Main notions

Definition 1 ***An ODE of order*** *k:*
an equation relating the values of one or more unknown functions of a **single** *variable t (we will call it "time"), their derivatives up to the order k, and the independent variable itself:*

$$\Phi(t, x_1, x_2, ...x_n, , x_1', x_2', ...x_n', x_1'', \ldots, x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}) = 0. \tag{1}$$

If more than one unknown function is involved, a **system** of such equations is usually needed. A system can be the neatly written in a vector form so that it looks like a single equation, e.g.

$$\mathbf{\Phi}(t, \mathbf{x}, \mathbf{x}', ..., \mathbf{x}^{(k)}) = \mathbf{0}, \tag{2}$$

where boldface font is used to denote vectors.

Example 2

1. $x'(t)x(t)^2 - 3t\sin(x''(t)) = 8$ *is an ODE (of what order? linear or non-linear?)*
2. $x'(t) = -5x(t+7)$ *is NOT an ODE!!*
3. $x^2(t) = -x'(t) + \int\limits_0^t \cos(x'(\tau))\, d\tau$*is NOT an ODE.*

Question. Oops!

- Why aren't the latter two examples ODEs? If you read the definition, it looks at the first glance like there is nothing wrong with these examples.

- What was missing in the wording of the definition? How should it be changed to make sure we exclude such cases?

- Do you know how the equations of the type shown in the last two examples are called?

2 Classification

- ODE vs PDE

- Order

- Linear vs non-linear

3 Order reduction

Introducing new unknown functions, an ODE or a system (2) can be **reduced to a first order** system:

$$\begin{array}{c} \mathbf{\Phi}(t, \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k, \mathbf{x}_k') = \mathbf{0} \\ {\mathbf{x}_1}' = \mathbf{x}_2 \\ ... \\ {\mathbf{x}_{k-1}}' = \mathbf{x}_k \end{array}, \tag{3}$$

so now **we can always deal with the first order systems**

$$\mathbf{\Phi}(t, \mathbf{x}, \mathbf{x}') = \mathbf{0} \tag{4}$$

Definition 3

- *The* **Normal Form** *is the best one:*

$$\mathbf{x}' = \mathbf{F}(t, \mathbf{x}). \tag{5}$$

- **Autonomous**

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}). \tag{6}$$

  and **non-autonomous** *(5) equations.*

4 Reduction of a non-autonomous equation to an autonomous one:

introduce a new time τ and consider the autonomous system

$$\begin{cases} \mathbf{x}'(\tau) = \mathbf{F}(t(\tau), \mathbf{x}(\tau)) \\ t'(\tau) = 1 \end{cases} . \tag{7}$$

See that this autonomous system is equivalent to the non-autonomous (5).

Here and further we will assume that $\mathbf{x}(t)$ is a differentiable function of $t \in (a, b)$ with values in $\mathbf{R}^n$ and $\mathbf{F}$ is a function from $\mathbf{R}^{n+1}$ to $\mathbf{R}^n$. (Complex case is possible, but we will not consider it here.)

5 What evolutionary (i.e., depending on time t) process can be described by ODEs?

Let us a have a process (mechanical, biological, etc.) whose instantaneous state can be described by some parameters $\mathbf{x}$. We call the space of these parameters the **phase space**. Since the system evolves with time, the parameters become functions of time as well: $\mathbf{x}(t)$. When can such a process be described by an ordinary differential equation? Three conditions tell you when this is the case:

- The system is **finite dimensional**, i.e. it can be described by finitely many parameters $x_1, ..., x_n$. This is not the case, for instance, in fluid dynamics, heat conduction, and quantum mechanics.

- **Smoothness**: the parameters change in a differentiable manner with time.This is not the case with shock waves.

- The process is **deterministic**: the state of the system at certain moment determines the whole future behavior of the system.

Indeed, due to the first condition, we can describe the evolution of the system by a finite dimensional vector function $\mathbf{x}(t)$. This function is differentiable, as the second condition tells us. The third condition says that if we know $\mathbf{x}(t)$ for some moment t, this determines all the future values of $\mathbf{x}(\tau)$. In particular, this determines the value of the derivative $\mathbf{x}'$ at the moment t.

Hence, $\mathbf{x}'(t)$ is determined by t and $\mathbf{x}(t)$. In mathematical notations we write that $\mathbf{x}'(t)$ is a function of t and $\mathbf{x}(t)$: $\mathbf{x}'(t) = \mathbf{F}(t, \mathbf{x}(t))$, which is an ODE (a system of ODEs).

6 IVP (Initial Value Problems))

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} \tag{8}$$

Definition 4 *For a differentiable function $F : \mathbf{R}^m \to \mathbf{R}^n$ the* **differential** *$DF(y)$ of F at a point y is the* **linear** *mapping from $\mathbf{R}^m$ to $\mathbf{R}^n$ with the matrix*

$$\{DF\}_{ij}(y) = \frac{\partial F_i}{\partial x_j}(y). \tag{9}$$

For a vector x of small norm, $DF(y)x$ is the linear approximation of the change of the function $F(y + x) - F(y)$. I.e.,

$$F(y + x) = F(y) + DF(y)x + o(|x|)$$

*(**Taylor formula** of first order, or **linearisation formula**)*

Theorem 5 Existence and Uniqueness Theorem
*Let $\Omega \subset \mathbf{R}^n$ be an open domain, (a, b) be an open segment of the line $\mathbf{R}_t$, and $F(t, x)$ and $D_x F(t, x)$ be continuous in $(a, b) \times \Omega$. Then, for any point $(t_0, x_0) \in (a, b) \times \Omega$ **there exists a unique** solution $x(t)$ of the IVP (8) **defined in a neighborhood** of t_0.*

Remark 6

1. *No global (i.e., on the whole (a, b)) uniqueness is guaranteed. Why?*

2. *The proof will show that another condition on F instead of differentiability suffices: it is enough that F is Lipschitz, i.e.*

   $$|F(t, x) - F(t, y)| \leq K|x - y|$$

   for some constant K and all $(t, x), (t, y)$ in our domain. Is this condition weaker than the one in the theorem?

7 Now about the proof: contraction mappings and such

7.1 Some notions and notations:

For a continuous function $x : [a, b] \to \mathbf{R}^n$ on a finite segment $[a, b]$ we denote by

$$\|x\| = \max_{t \in [a,b]} \{|x(t)|\} \tag{10}$$

its **norm** in the space of such continuous functions, where $|x|$ is the Euclidean norm of a vector in $\mathbf{R}^n$.

C_r - class of r times continuously differentiable functions (applicable both to scalar and vector valued functions of different numbers of variables, should be understandable from the context).

Definition 7 *A real valued function $A(x)$ on $\mathbf{R}$ is a* **contraction***, if it satisfies the inequality*

$$|A(x) - A(y)| \leq K|x - y|$$

for some $K < 1$ and all real x and y.

Remark 8

1. *Condition of continuous differentiability of A and estimate $|A'(x)| \leq k < 1$ guarantee that A is a contraction.*

2. *The definition of a contraction can be naturally extended to any metric space M with a metric (distance function) ρ instead of the real line, replacing $|A(x) - A(y)|$, $|x - y|$ above with $\rho(A(x), A(y))$, $\rho(x, y)$.*

7.2 A simple instant of the contraction mapping principle:

Theorem 9 *Let $A(x)$ be a contraction on $\mathbf{R}$. Then*

1. *The equation $x = A(x)$ has a unique solution x_* (called the* **fixed point of** *$A(x)$).*

2. *This fixed point can be found as $x_* = \lim_{j \to \infty} x_j$, where x_0 is arbitrary and $x_{i+1} = A(x_i)$.*

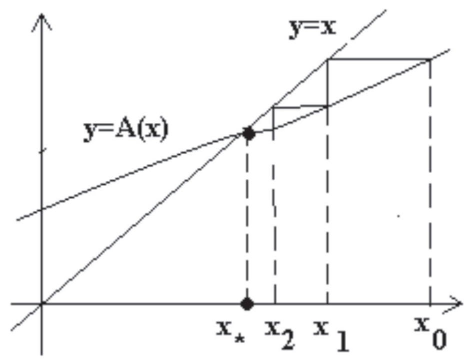

Figure 1: Geometric representation of the iterations for solving the equation $x = A(x)$ with a contraction $A(x)$.

7.3 The general contraction mapping principle:

Let X be a **metric space**: a set equipped with **metric** (= "distance") $\rho(x, y)$ with properties $\rho \geq 0$, $\rho(x, y) = \rho(y, x)$, $\rho(x, y) = 0$ only if $x = y$, triangle inequality is satisfied $\rho(x, y) \leq \rho(x, z) + \rho(y, z)$.

We also assume that X is a **complete** metric space, i.e. if a sequence x_n is such that $\rho(x_n, x_m) \underset{n,m\to\infty}{\to} 0$, then there exists its limit x such that $\rho(x_n, x) \to 0$.

A mapping $A : X \mapsto X$ is a **contraction** if

$$\rho(A(x), A(y)) \leq K\rho(x, y)$$

for some $K < 1$ and all $x, y \in X$.

Theorem 10 *Let $A(x)$ be a contraction on a complete metric space X. Then*

1. *The equation $x = A(x)$ has a unique solution $x_* \in X$ (called the* **fixed point of** *$A(x)$).*

2. *This fixed point can be found as $x_* = \lim_{j\to\infty} x_j$, where $x_0 \in X$ is arbitrary and $x_{i+1} = A(x_i)$.*

7.4 An equivalent integral equation reformulation of the IVP (8):

$$x(t) = x_0 + \int_{t_0}^{t} F(\tau, x(\tau))\, d\tau. \tag{11}$$

Lemma 11 *Continuous solutions of (11) are exactly the continuously differentiable solutions of (8).*

Now the proof of Theorem 5 would be concluded if proved existence and uniqueness of continuous solutions of (11) on a small segment around t_0.

The metric space: consider the interval $[t_0 - d, t_0 + d] \subset (a, b)$ with a small d (it will be determined later on how small it should be). We also consider a ball $B = \{x \in \mathbf{R}^n \mid |x - x_0| \leq b\}$ that is entirely contained in Ω. Now define on the set X of all continuous functions $x(t)$ from $[t_0 - d, t_0 + d]$ to B the max norm (10) as before and the corresponding metric $\rho(x, y) = \|x - y\|$.

Define the following **integral operator** $x \to A(x)$

$$[A(x)](t) = x_0 + \int_{t_0}^{t} F(\tau, x(\tau))\, d\tau. \tag{12}$$

Note that this definition works for functions that map $[t_0 - d, t_0 + d]$ to B.

Lemma 12 *For a sufficiently small d, the operator $A(x)$ maps the above class of functions into itself and is a contraction, i.e.*

$$\|A(x) - A(y)\| \leq k\|x - y\| \tag{13}$$

foe some $k < 1$.

Corollary 13

1. *The integral equation (11) has a unique continuous solution in a neighborhood of t_0.*

2. *This solution can be found as the limit in the norm (10) of* **Picard iterations**

$$y_{i+1}(t) = x_0 + \int_{t_0}^{t} F(\tau, y_i(\tau))\, d\tau, \tag{14}$$

where y_0 can be chosen arbitrarily in such a way that $y_0(t_0) = x_0$, e.g. $y_0 \equiv x_0$.

3. *The Uniqueness and Existence Theorem 5 is proven.*

8 An existence theorem

Theorem 14 Peano's existence theorem. *Continuity of F alone guarantees local existence of a solution of the IVP (8).*

Remark 15 *1. In the proof of Peano's theorem, solution is also found as the limit of some sequence of functions, but rather than Picard's iterations, a sequence of Euler's piecewise linear functions (recall the Euler's method of numerical solution) is constructed.*
2. Example of the IVP problem

$$\frac{dx}{dx} = 3x^{2/3}, x(0) = 0$$

shows that conditions of Peano's theorem cannot guarantee uniqueness.

9 Geometry of ODEs: vector fields

Consider the autonomous case

$$\frac{dx}{dt} = F(x),\ x(t) \in \mathbf{R}^n. \tag{15}$$

We can think that $F(x)$ assigns to each point x a vector $F(x)$ (a vector "grows" out of any point). Then we call $F(x)$ a **vector field**. We will consider at least continuous (or smoother) functions $F(x)$ and corresponding vector fields. If F is of some class C^r, we will also say that the field is of this class.

Lemma 16 *Trajectories of solutions of (15) are exactly the curves that are tangent at each point to the vector field corresponding to this equation. Such curves are called* **phase curves** *of the field.*

Note that **vector fields are NOT defined for non-autonomous systems**.

The field $F(x)$ is said to be **non-singular** at a point x_0, if $F(x_0) \neq 0$.

Exercise 17 *Show that the fields arising from turning a non-autonomous system into an autonomous ones, are always non-singular at all points.*

Example of a non-singular vector field: a **constant vector field**, where $F(x)$ is a constant non-zero vector.

Question 18 *Is the existence and uniqueness theorem obvious for a constant vector field?*

A **diffeomorphism** of class C^r is a mapping G from a domain such that it is one-to-one and both G and its inverse G^{-1} are of mappings class C^r. In other words, a diffeomorphism smoothly deforms the domain. At each point x the differential $(DG)(x)$ is an invertible linear mapping of vectors in $\mathbf{R}^n$. One can act by diffeomorphisms on vector fields as well. One can come up with a right definition using the following heuristics: let $x(t)$ be a solution of the equation defined by our vector field: $x' = F(x)$. We can act on this solution by our diffeomorphism to get a new function $x_G(t) = G(x(t))$. Then the chain rule gives $x'_G = DG(x)x' = DG(x)F(x) = DG(G^{-1}(x_G))F(G^{-1}(x_G))$. In other words, the G-modified function x_G satisfies the ODE $y' = F_G(y)$ with a vector field $F_G(y) = DG(G^{-1}(y))F(G^{-1}(y))$. This leads us to the

Definition 19 *Let $F(x)$ be a vector field and G be a diffeomorphism. Then one defines a new vector field as follows: $F_G(x) = (DG)(G^{-1}(x))F(G^{-1}(x))$.*

Theorem 20 **Vector Field Rectification Theorem.** *Any vector field of class C^r in a neighborhood of any its non-singular point x_0 can be reduced to a constant field ("rectified") by applying a diffeomorphism of class C^r.*

One of the exercises is to show that the rectification theorem implies the existence and uniqueness one.

Question: Can one do the converse, i.e. get an idea of local construction of the rectifying diffeomorphism from known solution?

10 Dependence of solutions on parameters and initial data.

The solution of the IVP (8) depends on the values of t_0 and x_0. How smooth is this dependence? Another important question: Assume that the right hand side (the vector field) also depends on some parameter(s) μ:

$$\frac{d\mathbf{x}}{dt} = F(t, \mathbf{x}, \mu),\ \mathbf{x}(t_0) = x_0. \tag{16}$$

How smoothly does the solution depend on the parameter? In fact, it can be seen that dependence on the initial data reduces to dependence on parameters. Indeed, introducing new time variable $\tau = t - t_0$ and new spatial variable $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$, one reduces (8) to

$$\frac{d\mathbf{y}}{d\tau} = F(\tau + t_0, \mathbf{y} + \mathbf{x}_0),\ \mathbf{y}(0) = 0. \tag{17}$$

Now all variable parameters are in the right hand side rather than in the initial data (which become constant). So, this is the only case to handle.

Theorem 21 *Let the vector field $F(\mathbf{x}, \mu)$ (where μ belongs to an open domain of a space $\mathbf{R}^m$) be of class C^r. Let also $F(x_0, \mu_0) \neq 0$. Then the (unique) solution $x(t, t_0, x, x_0, \mu)$ of the IVP*

$$\frac{dx}{dt} = F(x, \mu),\ x(t_0) = x_0 \tag{18}$$

depends differentiably of class C^r on (t, t_0, x, x_0, μ) for sufficiently small $|t - t_0|, |x - x_0|, |\mu - \mu_0|$.

11 Extendability of local solutions.

Our theorems guaranteed existence of a **local** solution only with no guarantee of how long it will survive. Simple examples show disappearance of solutions into a singular point. Even without singular points, a solution curve can grow fast and disappear in a finite time. An example is the IVP $x' = x^2$, $x(0) = 1$ that has the solution $x = \dfrac{1}{1-t}$ that disappears at infinity when t approaches 1. Are there any other options? Answer: no.

Theorem 22 Extendability Theorem. *Let N be a compact (bounded closed) subset in Ω (the domain where the smooth field is defined). Let also F have no singular points in N. Then any local solution of (8) in $(a, b) \times \Omega$ can be extended forward (for $t > t_0$) and backward ($t < t_0$) either indefinitely, or until it reaches the boundary of N.*

12 Boundary value problems (BVPs)

- Here the conditions are imposed at **both ends** of a time interval, rather than at one end only in the IVP case.
- Important applications.
- The number of conditions should still be correct (depending on the order of system and number of unknown functions).
- No such nice existence and uniqueness theorem.