# 3 Lecture 3: ODE review

### 3.1 Ordinary differential equation on manifold

As we have already discussed in **2.6** an ordinary differential equation on a manifold $M$ is

$$\dot{x} = X(x), \tag{3.1}$$

where $X$ is a section of $TM$. You may understand this as an ordinary $n$-vector ODE in a flat space $\mathbb{R}^n$ (or its subset).

More generally, an ODE is a functional relation among a function and its derivatives. Thus, (3.1) is not the most general form (see **3.2**), but is the most natural object to study what can happen for a time-evolving system whose phase space is $M$.

### 3.2 General ODE

Let $y$ be a $n$-times differentiable function of $t \in \boldsymbol{R}$. A functional relation

$$f(t, y(t), y'(t), \cdots, y^{(n)}(t)) = 0 \tag{3.2}$$

among $t$, $y(t)$, $y'(t)$, $\cdots$, $y^{(n)}(t)$ is called an *ordinary differential equation* (ODE) for $y(t)$, and $n$ is called its *order*, where the domain of $f$ is assumed to be appropriate. Such $y(t)$ that satisfies $f = 0$ is called a *solution* to the ODE.

If the highest order derivative of $y$ is explicitly solved as

$$y^{(n)}(t) = F(t, y, y', \cdots, y^{(n-1)}) \tag{3.3}$$

from $f = 0$, we say the ODE is in the *normal form*.[37]

### 3.3 Normal form ODE is essentially first order.

Let $y_j \equiv y^{(j-1)}$ ($j = 1, \cdots, n$). Then (3.3) can be rewritten as

$$\begin{aligned} \frac{dy_1}{dt} &= y_2, \\ &\cdots \\ \frac{dy_{n-1}}{dt} &= y_n, \\ \frac{dy_n}{dt} &= F(y, y_1, y_2, \cdots, y_n). \end{aligned} \tag{3.4}$$

[37] Notice that not normal ODE's may have many pathological phenomena, but we will not pay any attention to the non-normal form cases henceforth.

That is, (3.3) has been converted into a first order ODE for a vector $y = (y_1, y_2, \cdots, y_n)^T$.[38] Any normal form $n$-th order scalar ODE can be converted into the $n$-vector first order ODE of the form

$$\frac{dy}{dt} = X(t, y). \tag{3.5}$$

Any solution $y(t)$ can be understood as an orbit (or trajectory) parametrized with 'time' $t$ in the $n$-space (= phase space) in which $y$ lives.

### 3.4 Autonomous vs nonautonomous

For (3.5) the vector field $X$ explicitly depends on time. This means that there is a certain 'external' agent modifying the vector field. Thus, generally, we are not interested in such a system that is 'not self-contained.'[39] Such a system is called a non-autonomous system. We are interested in autonomous systems as described by (3.1).

### 3.5 When does ODE define dynamical system?

If (3.1) has a unique solution for any intial condition $x \in M$, we may define a continuous-time dynamical system. We know the following:
(1) Peano's theorem: If $X$ is continuous, (3.1) has a solution.
(2) Cauchy-Lipshitz uniqueness theorem: If $X$ satisfies a Lipshitz condition $\|X(x) - X(x')\| < L\|x - x'\|$ (see **3.6**), (3.1) has a unique solution.
(3) If $X$ is not Lipshitz, then the uniqueness of the solution is not guaranteed (counterexamples exist).

Thus, we confine our attention to differentiable vector fields that are automatically Lipshitz.[40] However, we should understand why (1)-(3). This is the purpose of the rest of the lecture. To prove (1) an approximate solution sequence is constructed (via the Euler approximation), and then we prove the existence of a limit. This requires a knowledge of functional analysis, but at least the gist of the demonstration or its delicate point should be recognized. (2) can be understood almost geometrically (the rectifiability theorem **3.19**).

### 3.6 Lipschitz condition.

[38] You may prefer $\boldsymbol{y}$ for $y$, I will maximally avoid explicit vector notation throughout the lecture notes.

[39] Except perhaps the perturbation is periodic.

[40] cf. the mean-value theorem

Let $X$ be a continuous vector function whose domain is a region $D \subset \mathbb{R}^n$. For any compact[41] set $K \subset D$, if for any $y_1$ and $y_2$ both in $K$ there is a positive constant $L_K$ (which is usually dependent on $K$) such that

$$|X(y_1) - X(y_2)| \leq L_K |y_1 - y_2|, \tag{3.6}$$

then $X$ is said to satisfy a *Lipschitz condition* on $D$.

A $C^1$ function is Lipschitz continuous due to the mean value theorem. If a vector field is $C^1$, then it is Lipshitz.

### 3.7 Peano's existence theorem

Suppose $X$ in (3.1) is continuous in a bounded closed region $G \subset M$, then for any $x \in G$ there is at least one integral curve passing through it in $G$.

The proof may be obtained with the aid of Arzela's theorem **3.11** that can show the existence of the convergence of the Euler approximation sequence.[42]

### 3.8 Euler approximation

Consider

$$\dot{x} = X(x(t)) \tag{3.7}$$

for the time span $[0.T]$. if we approximate the derivative with a finite difference $\dot{x} \simeq [x(t + \Delta t) - x(t)]/\Delta t$ we may write the ODE as

$$x(t + \Delta t) = x(t) + \Delta t X(x(t)). \tag{3.8}$$

Therefore, we can make an approximate function $\varphi_i$ by making a piecewise connection of adjacent time points $\{x(n\Delta t_i)\}$, where $\Delta t_i$ is the time increment. We make an approximation sequence $\{\varphi_i\}$ for $\Delta t_i > \Delta t_{i+1} \to 0$.

Does this sequence converges to a solution (have an accumulation point corresponding to a solution)?

### 3.9 Strategy to prove Peano's theorem

First, we must show that $\{\varphi_i\}$ for $\Delta t_i > \Delta t_{i+1} \to 0$ has an accumulation point.

[41]'Compact' means in a finite dimensional space 'closed and bounded'.

[42]Kolmogorov-Fomin p102

Since the totality of continuous functions $[0,T] \to M$ is not compact (nor relative compact), to show the existence of an accumulation point is not trivial. However, Arzela's theorem tells us that $\{\varphi_i\}$ is compact. Thus, accumulation points exist (this is why we cannot prove the uniqueness).

Then, we show that the limit indeed satisfies the original ODE.

As you see, we must understand the concept of compactness in a functional space (or infinite dimensional space).

### 3.10 Review of compactness[43]

If any open covering of a set $S$ has a finite subcover,[44] $S$ is called a compact set.

If the closure of $S$ is compact, we say $S$ is relative compact.

If a space is finite-dimensional, then bounded closed set is automatically compact. It is thanks to the Bolzano-Weierstrass theorem (= bounded sequences must have an accumulation point; a finite dimensional bounded closed set is countably compact). However, if the dimension is not finite, this is not true: think of $\{e_n\}$, where $e_n = (0, \cdots, 0, \overset{n}{1}, 0, \cdots)$.

We can make a function space $C_{[0,T]}$ as a metric space by introducing a sup metric $\rho(f,g) = \sup_{t\in[0,T]} |f(t) - g(t)|$, but the space is obviously not finite dimensional. We use

**Theorem** A necessary and sufficient condition for a metric space to be compact is: (i) totally bounded and (ii) complete.

To understand this theorem we must understand:

$*$ 'totally bounded': A metric space $M$ is totally bounded if for any $\varepsilon > 0$ there is an $\varepsilon$ net $A$ consisting of finitely many points. That is $\forall y \in M\ \exists x \in A$ such that $\rho(x,y) < \varepsilon$.

$*$ complete: any Cauchy sequence converges.[45]

[Demo]

If a metric space $A$ is compact, then it is totally bounded: If not, then there is $e_0 > 0$ such that there is no finite $\varepsilon_0$-net for $A$. Thus, we can have an infinite point set $\{a_i \in A\}$ such that $\rho(a_i, a_j) > \varepsilon_0$. Thus, $\{a_i\}$ is an infinite point set without an accumulation point. Thus, $A$ is not compact. The necessity of completeness is obvious.

To show (i)+(ii) implying compactness, we have only to show that any bounded infinite set has an accumulation point. Consider an infinite sequence

---

[43] https://www.dropbox.com/home/ApplMath?preview=AMII-ElementaryCheckList.pdf may be useful as your analysis rudiment checklist.

[44] a cover consisting of a subset of the original cover.

[45] If the distance is $L^2$,then $C$ is not complete. KF3.1

$\{x_i\}$. Take a 1-net. Then in a ball $B$ within distance one from at least one of the net point are infinitely many points in this sequence. $B$ is totally bounded, we can repeat the argument with distance 1/2. Repeating this, we can construct a Cauchy sequence. Thanks to the completeness, there must be a limit point.

Thanks to this theorem and the fact that a closed subset of a complete space is complete, we can conclude that in a complete metric space the total boundedness is enough to guarantee the relative compactness os any subset $M$.

### 3.11 Arzela's theorem[46]

**Theorem** [Arzela] A set of function $\Phi \subset C_{[0,T]}$ is relative compact iff $\Phi$ is uniformly bounded and equicontinuous.

To understand this theorem we must understand:
∗ Uniformly bounded: For any $f \in \Phi$ and for any $t \in [0,T]$, $|f(x)| \le K$ for some positive $K$.
∗ Equicontinuous: For any $\varepsilon > 0$ there is $\delta > 0$ such that for any $f \in \Phi$ $|f(t) - f(t')| < \varepsilon$ if $|t - t'| < \delta$.
[Demo]
Here we prove the sufficiency: $C_{[0,T]}$ is a complete metric space, since the convergence in sup norm means the uniform convergence on $[0,T]$. Therefore, we have only to check the total boundedness of $C_{[0,T]}$. Since $\Phi$ is uniformly bounded and equicontinuous, we can choose $\delta > 0$ appropriately so that for all $\varphi \in \Phi$

$$|\varphi| < K, \ |\varphi(x) - \varphi(x')| < \varepsilon \text{ if } |x - x'| < \delta. \tag{3.9}$$

We construct an $\varepsilon$-net consisting of piecewise linear functions. The idea must be intuitively grasped from the figure 3.1. Make the totality of piecewise linear functions connecting NE, E or SE arrows on the lattice. This is a finite $\varepsilon$-net. Therefore, $\Phi$ is compact.

### 3.12 Proof of Peano's theorem

We have constructed the piecewise linear continuous approximation sequence $\{\varphi_i(t)\} \subset C_{[0,T]}$. It is uniformly bounded and equicontinuous. Therefore, Arzela's theorem tells us that there is a uniformly convergent subsequence in $\{\varphi^{(i)}(t)\}$, converging to $\varphi(t)$. The remaining task is to show that for any $\varepsilon > 0$ we can choose $\Delta t$ small enough to

[46] A readable proof is given in Kolmogorov-Fomin.

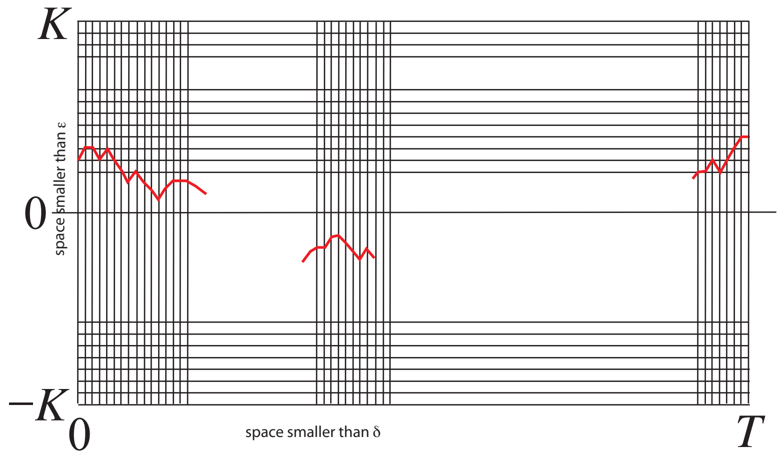


Figure 3.1: A representative piecewise linear approximate function (red) making an $\varepsilon$-net for bounded functions in $C_{[0.T]}$.

make

$$\left|\frac{\varphi(t+\Delta t)-\varphi(t)}{\Delta t}-X(\varphi(t))\right|<\varepsilon. \tag{3.10}$$

To this end we have only to show for sufficiently large $k$

$$\left|\frac{\varphi^{(k)}(t+\Delta t)-\varphi^{(k)}(t)}{\Delta t}-X(\varphi^{(k)}(t))\right|<\varepsilon. \tag{3.11}$$

We can formally demonstrate this,[47] but intuitively the closeness of $\varphi^{(k)}$ to $\varphi$ and continuity of $X$ implies all the terms are close to the limits.

### 3.13 Nonuniqueness cases

Peano noted that for $X(x) = 3x^{2/3}$ $x = 0$ and $x = t^3$ are solutions satisfying $(0,0)$ as the starting point.[48] That is, continuity of $X$ is not enough for the determinacy. However, differentiability is not needed.

If $X$ is Hölder continuous with the exponent less than 1 at a point, the uniqueness is lost at the point.

[Hölder continuity].

[47] e.g., see Kolmogorov-Fomin

[48] More generally, $X = x^{1-1/n}$ ($n \in \mathbb{N}^+$).

If a function $f$ satisfies

$$|f(x) - f(y)| \le L|x - y|^\alpha \tag{3.12}$$

on its domain for constants $L$ and $\alpha \in (0, 1)$, $f$ is said to be *Hölder continuous* of order $\alpha$.

The physical reason for this nonuniqueness is 'forgetting the initial condition' due to super-ballistic acceleration.

### 3.14 Relevance to physics of nonuniqueness cases?[49]

In a fully developed turbulence the velocity field becomes not Lipshitz but only Hg̈lder wtih $\alpha \simeq 1/3$ (if we believe in the Kolmogorov spectrum). Thus, the motion of a particle advected by the flow becomes non-deterministic.

If we consider a particle in a potential for $\alpha \in (0, `)$

$$V(x) = -\frac{1}{1+\alpha}|x|^{1+\alpha} \tag{3.13}$$

the classical mechanics gives the following equation of motion:

$$\dot{x} = v, \ \ \dot{v} = |x|^\alpha \text{sign}(x). \tag{3.14}$$

Note the superballistic nature of *dotv* around $x = 0$. However, 'a physically realizable potential will however exhibit a power-law scaling as in (3.13) only over a limited range of x-values, with an inner or short-distance cutoff $\ell$ and an outer or large-distance cut-off $L$. The latter may not be very serious, but the former is serious.

Eyink and Drivas 'mollified' the potential with a smooth short-range cutoff, and then considered (1D) quantum mechanical potential (although not trapping) problem. The wave packet is not max at the origin, and even in the classical limit 'stochasticity' remains.

### 3.15 Phase flow

Let a $n$-dynamical system

$$\dot{x} = X(x) \tag{3.15}$$

be defined on a domain $U$. We can imagine a flow field on $U$ described by the vector field $X$. It is called the *phase flow* because it flows the phase space = the state space; in our case the state is specified by $x(t)$ at time $t$, so the space in which $x$ lives is

[49] G L Eyink and T D Drivas, Quantum spontaneous stochasticity arHiiv: 1509:04941 (2015).

our phase space.

We can imagine a trajectory of a point passively flowing with this flow field. It is called the phase flow.

You can use the following software to see examples of 2-vector fields and corresponding phase flows for various initial conditions: https://media.pearsoncmg.com/aw/ide/idefiles/media/JavaTools/twoddfeq.html

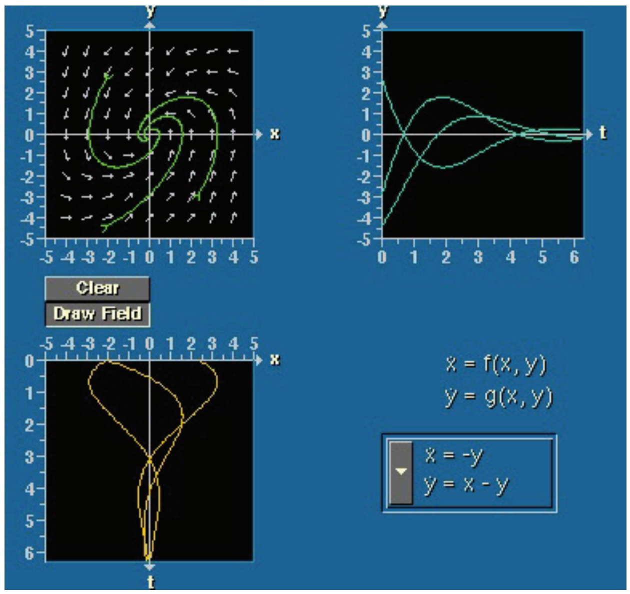


Figure 3.2: Vector-Flow demo

### 3.16 Direction field and solution curve

For (3.15) its *direction field* is a vector field $(1, X(x))$ on each $(t, x)$ in $T \times U$, where $T$ is the set of time under consideration. Again we can imagine a trajectory of a point passively flowing with this flow field. It is called the graphs of the solutions of (3.15).

You can use the following software to see examples of 1-dynamical systems (may be non-autonomous) and corresponding solution graphs for various initial conditions:https://media.pearsoncmg.com/aw/ide/idefiles/media/JavaTools/exunqtrg.html

### 3.17 Singular points

A singular point of a vector field $X$ is point where the vector field vanishes.

The essence of the general theory of ODE is that as long as the vector field is

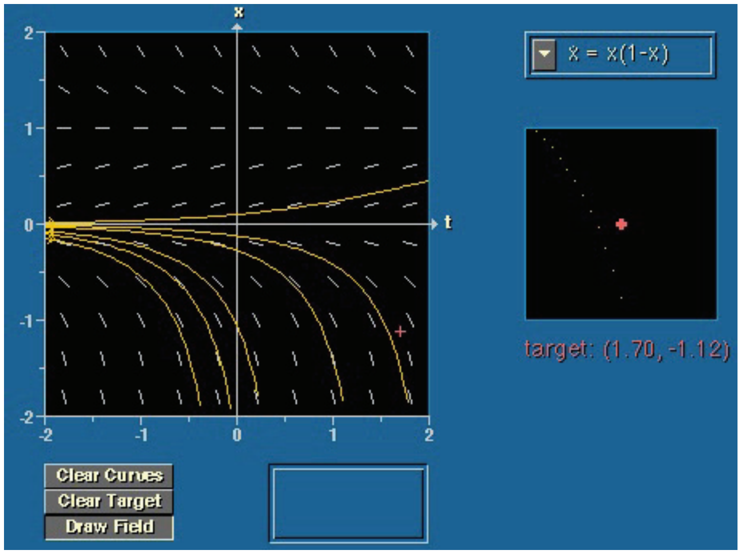


Figure 3.3: Direction field demo. You can play a shooting game.

nonsingular, unique existence of the solutions and their 'maximally' nice properties (**3.21**, **3.24**) are guaranteed. We proceed as geometrically and intuitively as possible.

### 3.18 Cauchy-Lipschitz uniqueness theorem.[50]

For (3.15), if $X$ satisfies a Lipschitz condition on $D$, and if there is a solution passing through $x_0 \in D$, it is unique.

Remark: Even if the Lipshitz condition is not satisfied, If the variables are separable as

$$\frac{dy}{dx} = \frac{Y(y)}{X(x)}, \tag{3.16}$$

and $X$ and $Y$ are continuous and not zero near $(x_0, y_0)$, then the solution near this point is unique. However, the condition is important as we see in the next.

### 3.19 Rectifiability theorem for vector field

In a sufficiently small neighborhood of any nonsingular point a differentiable vector field is diffeomorphic to the constant field $e_1 = (1, 0, \cdots, 0)^T$.

The graph of the solution never crosses at a non-singular point.

If the original field is $C^r$, the diffeo can be $C^r$.

All the basic theorems are more or less straightforward corollaries of the funda-

[50] AMM 116 61 Does Lipschitz with Respect to $x$ Imply Uniqueness for the Differential Equation $y = f(x, y)$? Author(s): José Ángel Cid and Rodrigo López Pouso

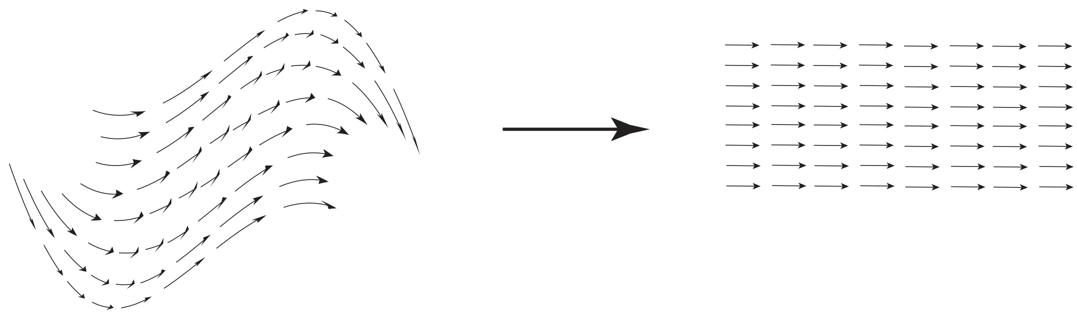

Figure 3.4: Rectification of a vector field

mental theorem.

### 3.20 Extension theorem[51]

An extension of the solution $\varphi$ is a solution which coincides with $\varphi$ on the (time) interval on which $\varphi$ is defined and which is defined on a greater (time) interval.

**Theorem** [The extension theorem] Let $K$ be a compact subset of the domain of the ODE (3.15). Then, every solution of this equation with an initial condition in $K$ can be extended to the the boundary of $K$ of infinitely in time (to $\pm\infty$).

### 3.21 Continuous dependence on initial conditions.

If the vector field is Lipschitz continuous (**3.6**), then the solution at time $t$ depends on the initial condition continuously.

Although this should be intuitively clear, since we need a more quantitative statement later, let us estimate the bounds. We need an important inequality:

### 3.22 Gronwall's inequality

Let $u, v : [a, b] \to \mathbb{R}$ be continuous nonnegative functions satisfying

$$u(t) \le \alpha + \int_a^t u(s)v(s)ds \tag{3.17}$$

for some $\alpha$ ($\ge 0$) and for $\forall t \in [a, b]$. Then,

$$u(t) \le \alpha \exp\left(\int_a^b v(s)ds\right). \tag{3.18}$$

[51] Arnold I 2.5 p17

[Demo]
If $\alpha = 0$, then $u(t) = 0$, so we assume $\alpha > 0$. Let us define $\omega(t)$ as

$$\omega(t) = \alpha + \int_a^t u(s)v(s)ds. \tag{3.19}$$

Obviously, $u(t) \leq \omega(t)$.

$\omega(a) = \alpha$ and $\omega(t) \geq \alpha > 0$. As $\omega'(t) = u(t)v(t) \leq v(t)\omega(t)$, we have

$$\omega'(t)/\omega(t) \leq v(t). \tag{3.20}$$

Integrating this, we get the inequality.

### 3.23 Initial condition dependence

We assume $X \in \mathcal{X}(M)$ is Lipschitz (usually $C^r$) with the Lipschitz constant $L$ and $M$ is compact. Make two solutions starting from $x_0, y_0 \in M$:

$$x(t) = x_0 + \int_0^t X(x(s))ds, \;\; y(t) = y_0 + \int_0^t X(y(s))ds. \tag{3.21}$$

Then,

$$x(t) - y(t) = x_0 - y_0 + \int_0^t [X(x(s)) - X(y(s))]ds. \tag{3.22}$$

This means

$$\|x(t) - y(t)\| \leq \|x_0 - y_0\| + \left\| \int_0^t [X(x(s)) - X(y(s))]ds \right\| \leq \|x_0 - y_0\| + \int_0^t \|X(x(s)) - X(y(s))\|ds. \tag{3.23}$$

Using the Lipschitz constant we have

$$\|x(t) - y(t)\| \;\; \leq \|x_0 - y_0\| + L \int_0^t \|x(s) - y(s)\|ds. \tag{3.24}$$

Now, we can apply Gronwall's inequality to obtain

$$\|x(t) - y(t)\| \leq e^{LT} \|x_0 - y_0\|. \tag{3.25}$$

That is, the solution must also be Lipschitz continuous. This explicitly proves **3.21**.

**3.24 Smooth dependence on parameter.**
If the vector field is smooth, then the solution at finite time is as smooth as the vector field. If the vector field is holomorphic, then the solution is also holomorphic. Then, we can use perturbation theory to obtain the solution in powers of the parameter. This was the idea of Poincaré.

**3.25 Picard's successive approximation method**[52]

$$x_0(t) = x(0), \tag{3.26}$$

$$x_{k+1}(t) = x(0) + \int_0^t v(s, x_k(s))ds. \tag{3.27}$$

[Demo] (may never be given explicitly)
First, we get formally

$$x(t) = x(0) + \int_0^t v(s, x(s))ds. \tag{3.28}$$

If the limit $k \to \infty$ of $x_k(t)$ exists, then obviously (3.27) gives (3.28). Therefore, we need a uniform convergence of the sequence. See Arzela **3.11**.

**3.26 History**[53]
'Differential equations' began with Leibniz, the Bernoulli brothers and others from the 1680s, not long after Newton's ' fluxional equations' in the 1670s. Applications were made largely to geometry and mechanics; isoperimetrical problems were exercises in optimisation."
According to:
**The role of the concept of construction in the transition from inverse tangent problems to differential equations.** Henk J. M. Bos p2733
Tangent problems—given a curve, to find its tangents at given points—are as old as classical Greek mathematics. 'Inverse tangent problems' was the name coined in the seventeenth century for problems of the type: given a property of tangents, find a curve whose tangents have that property. It seems that the first such problem was proposed by Florimod De Beaune in 1639. Much of the activities in the early

[52] Arnold 2.4c p16
[53] The History of Differential Equations, 1670-1950 Organised by Thomas Archibald (Wolfville) Craig Fraser (Toronto) Ivor Grattan-Guinness (Middlesex)

infinitesimal calculus (second half of the seventeenth century) were motivated by inverse tangent problems, many of them suggested by the new mechanical theory.

The transition to differential equations occurred around 1700. This transition was much more than a simple translation from figure to formula, from geometry to analytical formalism.

In the seventeenth century to solve this problem is to construct the curve required in the problem. Descartes had restricted geometry to algebraic curves. But inverse tangent problems often had non-algebraic curves as solution. Consequently mathematicians went outside the Cartesian demarcation of geometry and consequently lost a clear and shared conception of what it meant to solve a differential equation; indeed, the status of differential equations became fuzzy: were they problems? were they objects? When were their solutions satisfactory? Many puzzling developments in early analysis, and especially delays in developments expected with hindsight, can be explained by the tenacity of the older ideas on problem solving.