

Various lecture notes for 18385.

R. R. Rosales.

Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, Massachusetts, MA 02139.

September 17, 2012

Abstract

Notes, both complete and/or incomplete, for MIT's "18.385 Nonlinear Dynamics and Chaos".

Contents

1	Lecture Notes Fall 2012.	2
1.1	Lecture # 01, Thu. Sep. 06.	2
	Introduce the notion of phase space.	2
	Introduce the notion of a well posed problem.	3
	Well posed versus chaotic.	4
1.1.1	Classifications of models.	5
	Notion of phase portrait.	5
1.1.2	Canonical forms.	6
	Delay-difference and integro-differential equations.	6
1.2	Lecture # 02, Tue. Sep. 11.	7
	Inverse function theorem.	7
	Implicit function theorem.	7
	Existence and uniqueness for the I.V. problem for ode.	7
	Lipschitz continuity.	8
	Example of an ode with a discontinuity.	8
1.2.1	The linearization metatheorem.	8

List of Figures

1 Lecture Notes Fall 2012.

[vln385:LN12]

This is a short summary of (some of) the lectures from the fall of 2012.

1.1 Lecture # 01, Thu. Sep. 06.

[vln385:LN12:L01]

Students are expected to read, and be familiar, with the contents of chapter #1 in the textbook (by Strogatz). The material here is meant as a supplement to that in the book.

The notion of a *phase space* \mathcal{P} is a crucial one in mathematical modeling in science. To make a mathematical model of some physical system, one needs to identify parameter/quantities that can be used to “fully describe” the system.¹ The set of all possible values that can be taken by these parameters is the **phase space** \mathcal{P} — this set can often be given some geometrical “structure”, hence the “space” in the name. Some examples are:

1. Coin tossing. $\mathcal{P} = \{u, d\}$, where $u =$ coin up and $d =$ coin down.
2. Pendulum. $\mathcal{P} =$ *cylinder*.
3. Classical Mechanical system of unconstrained point particles in a d -dimensional Euclidean space. $\mathcal{P} =$ *set of all the possible positions and velocities* $= \mathcal{R}^{2d}$.
4. Quantum Mechanical system of unconstrained point particles in a d -dimensional Euclidean space. $\mathcal{P} =$ *set of all the complex valued functions defined on R^d , with unit square integral = unit sphere in $L_2(R^d)$.*

When “fully describing” a system the aim is not to necessarily describe the “full” system, but whatever is “relevant” (e.g.: in item 2 we ignore the color of the pendulum, the air motion around it, etc). We also may ignore “effects” that are not important to the level of accuracy to which we care to describe the system (e.g.: in item 2 we ignore the fact that the pendulum bar is not rigid — hence its length is not a constant). In fact, the same system may be associated with more than one \mathcal{P} , depending on how much detail is included (e.g.: compare items 3 and 4).

What we mean by “fully describe” the system is a rather vague, unless one introduces some mathematics. Then one can make it precise by stating that it means that: *The resulting mathematical model is well posed* (what well posed means is defined below). One cannot tell if a phase space is appropriate without simultaneously having the mathematical model that goes with it. This may seem a little circular and not very satisfactory, but the making of a mathematical model for a science or engineering problem is not a “plug into recipe and turn the crank” process. It is hard,

¹Below we “define” what “fully describe” means more precisely. For now, we leave this at the “intuitive” level.

and requires originality and a well oiled human brain to work. Think of how long it took to arrive at Newtonian Mechanics, or of how hard was the “phase space switch” when formulating Quantum Mechanics.

A mathematical problem is **well posed** if it has a unique solution that depends continuously on any parameters present in the problem. Examples:

5. Initial value problem for a linear homogeneous and constant coefficients ode system.² That is: $\dot{\vec{y}} = A\vec{y}$, with $\vec{y}(0) = \vec{y}_0$ — where \vec{y} is vector valued and A is a constant square matrix.

The (unique) solution to this problem is $\vec{y} = e^{At}\vec{y}_0$, which clearly depends continuously (in fact, analytically) on all the parameters in the problem: the entries of A and \vec{y}_0 .

6. Consider the boundary value problem $u'' + mu = 0$ for $0 < x < 1$, with $u(0) = u(1) = 0$. Does this have a solution u ? When a solution exists, is it unique? Does the solution, if any, depend continuously on m ?

Could this be a model for a question of the type: *what is the shape of a string hanging from two supports, one at each end?*, where the shape of the string is given by $u = u(x)$?

Note also that one could have a model in which the question is: *What are the values of m for which there is a solution?* (i.e.: an eigenvalue problem), which leads to different considerations.

7. Neumann problem for Poisson in some bounded region $\Omega \subset R^2$. That is: $\Delta T = f$ in Ω , with $\partial_n T = 0$ along $\partial\Omega$.

7.1 Problem as steady state heat distribution for an insulated thin plate, heated/cooled by some source.

7.2 Argue, physically, that there cannot be a steady state unless: $\int_{\Omega} f dV = 0$ (the net heat input vanishes).

7.3 Use Gauss theorem to mathematically derive the condition in item 7.2.

This problem is well posed, provided that one restricts the sources by $\int_{\Omega} f dV = 0$.

Note that a problem of this type arises for the pressure, at each time step, in some incompressible Navier Stokes solvers. Then numerical errors “guarantee” that the condition $\int_{\Omega} f dV = 0$ is never quite satisfied. A way around this problem is to look for a T that is “as close to a solution as possible” (in some appropriate sense), which then yields a well posed problem.

8. The initial value problem $\dot{x} = 2\sqrt{|x|}$, with $x(0) = 0$. This has at least two solutions.³ $x = 0$ and $x = t^2$. Thus, it is an ill-posed problem.

²Later we will state a theorem that generalizes this to very general initial value problems for ode systems.

³In fact, it has infinitely many.

Consider the situation of a cylindrical bucket filled with water up to a height $h = h(t)$, and being emptied through a small hole at the bottom. Assume that the bucket is empty at time $t = 0$, and write the equation for $x = x(t) = h(-t)$. This gives rise to an equation similar to the one above. This should not be a surprise, since it should be obvious that we cannot determine $h(t)$ for $t < 0$ unless the time at which the bucket finished emptying is known.

Of course, **being well posed is not enough**. The model must give results and predictions that match experiments and observations, with some reasonable⁴ degree of accuracy. **But it is needed**. An ill posed (i.e.: not well posed) model is useless!

Remark 1.1.1 Well posed versus chaotic. In connection to chaos many of you may have heard statements such as “*stepping on a butterfly in China may produce a tornado in Kansas*”. We will consider later the extent to which this can be taken seriously. For now, however, notice that this implies a situation where very small changes can produce huge outcomes! Does this contradict the “*continuous dependence on parameters*” requirement for a well posed problem? Does this mean that a problem in which chaos occurs is not well posed?

The answer is a most emphatic **NO!** When dealing with an initial value problem, well posed means that: *As the size of a perturbation to the initial conditions vanishes, the size of the resulting perturbation to the solution also vanishes at any fixed time. The solution can have arbitrarily large changes as $t \rightarrow \infty$.* Chaos is tied up to “fast” growth of the errors with time, **not** with errors going from “*infinitesimal*” to *finite in arbitrarily short times (which is what lack of continuity means)*. ♣

When a model is discretized and programmed into a computer, the following things can happen:

9. The original model is ill posed, but the computer program behaves well. This, of course, means that whatever the computer is computing is not related to the “physical reasoning” that led to the model! This is, clearly, not desirable. Hence, it is important to “understand” the model, and not rely solely on a computer implementation.
10. The original model is OK, but the computer implementation behaves badly. This means: bad discretization. Numerical analysis aims to address (avoid) this problem. Important, but not the focus of this class.
11. The original model is OK, and the computer implementation behaves well. But they have little to do with each other. Again: issue dealt with by numerical analysis.

Since it is rather impossible to do “numerical analysis” for a system that is poorly understood, “understanding” is needed in all cases. This class is aimed at providing tools for “understanding”.⁵

⁴This depends on the intended use.

⁵Of course, one can get understanding from numerical computations, and use this to correct/improve the calculations, and so on. These two aspects are not independent.

1.1.1 Classifications of models.

[vln385:LN12:L01:CLA]

Models/problems can be classified into:

- **Static.** Example: what is the shape of a hanging cord?
- **Dynamic.** Example: vibrations of a string.

In this course we concentrate on dynamics. The time evolution in a dynamic problem can be:

- **Random.** Example: coin tossing.
- **Deterministic.** Example: orbital mechanics.

In this course we concentrate on deterministic problems. The time evolution can be:

- **Discrete.** Example: generations in a biological system.
- **Continuum.** Example: car moving along a road.

For a continuum dynamical system, the solutions can be interpreted as curves in phase space. A **main objective of this course is to describe the “topology” of these curves:** *how does the “phase portrait” look like?* That is: we are not seeking quantitative information, but mainly qualitative information. This approach was proposed by Poincaré when, towards the end of the 19-th century, it became clear that “exact solutions” were possible for only a very limited number of problems. In general, the best one could hope for (analytically) was qualitative information.⁶

Further classifications:

- **Linear.** Define. Example: harmonic oscillator.
- **Non-linear.** Define. Example: Navier-Stokes equations.

Because linear systems are vulnerable to a “divide-and-conquer” strategy, they are much better understood than non-linear systems. In fact, notice that non-linear systems are defined by “not being linear”, in other words: but what they are not. This (very poor) definition is an indication of how little we understand them, in general — though there are some types of non-linear systems for which we know more (e.g.: completely integrable Hamiltonian systems).

- **Finite dimensional.** Examples: ode.
- **Infinite dimensional.** Examples: pde.

In this course we will concentrate on finite and low [1, 2, 3] dimensional systems, governed by ode. Why only this? Because we understand 1-D and 2-D systems fairly thoroughly, and we have some understanding of 3-D (though not complete). Beyond that, most of what we know results from situations where a higher dimensional system displays behavior that can be understood in terms of low dimensional concepts. This is not intended to trivialize the techniques involved in analyzing

⁶Of course, with computers, we can now hope both for quantitative as well as qualitative information. However, this does neither make the problems trivial, nor eliminates the need for deep and intelligent thinking. When (and if) AI truly arrives, this may change. But we are not there yet.

such systems, which are covered in other courses, but we must draw the line somewhere.

Check the *Dynamical view of the world* section by Strogatz, with *the frontier* in the dimension versus nonlinearity diagram. Obviously the diagram is incomplete, and many things are missing, but it illustrates well where the boundary between known and unknown roughly is.

1.1.2 Canonical forms.

[vln385:LN12:L01:CAN]

Continuum, finite dimensional, deterministic, dynamical systems governed by ode can be written in the canonical form

$$\frac{d\vec{y}}{dt} = \vec{F}(\vec{y}), \quad (1.1.1)$$

where \vec{y} and \vec{F} are d -vector valued, and \vec{F} is some function. Then **d is the dimension of the system.** Explained in the lecture:

- How time dependent coefficients can be eliminated by adding t as an extra unknown, with the extra equation $\dot{t} = 1$.
- How higher order derivatives can be eliminated by adding derivatives as extra unknowns.
- How, after the above, one must be able to solve for $\dot{\vec{y}}$ (at least in principle) and write the equation as above: A time evolution ode problem where the time derivative is not uniquely determined by the present cannot be well posed.
- How, for a constrained system (e.g.: pendulum) the form above will apply (at least locally) once an appropriate (maybe local) coordinate system is selected incorporating the constraints. Generally, this applies to systems where the phase space \mathcal{P} is some finite dimensional manifold.

Similarly, finite dimensional, deterministic, dynamical discrete systems can be written in the form

$$\vec{y}_{n+1} = \vec{F}(\vec{y}_n), \quad (1.1.2)$$

where $\vec{y}_n \in \mathcal{P}$ and \vec{F} is some function from \mathcal{P} to \mathcal{P} . If \mathcal{P} is a d -dimensional manifold, then **d is the dimension of the system.** Explained in the lecture:

- How to reduce systems of the form $\vec{x}_{n+1} = \vec{F}(\vec{x}_n, \vec{x}_{n-1})$ to the form in (1.1.2) by letting $\vec{y}_n = (\vec{x}_n, \vec{x}_{n-1})$. Similarly for systems of the form $\vec{x}_{n+1} = \vec{F}(\vec{x}_n, \vec{x}_{n-1}, \vec{x}_{n-2})$, etc.
- How to reduce systems of the form $\vec{x}_{n+1} = \vec{F}_n(\vec{x}_n)$ to the form in (1.1.2) by introducing n as an extra unknown.

Remark 1.1.2 Delay-difference and integro-differential equations. *Some mathematical models in science and engineering involve delay-difference and integro-differential equations. In general such equations (even when they involve a finite number of unknowns which are functions of a single variable: “time”) should not be considered as finite dimensional.* ♣

1.2 Lecture # 02, Tue. Sep. 11.

[vln385:LN12:L02]

Consider an initial value problem of the form

$$\vec{G}(\dot{\vec{y}}, \vec{y}) = 0, \quad \text{with } \vec{y}(0) = \vec{y}_0. \quad (1.2.1)$$

As pointed out in § 1.1.2, for this problem to make sense, we need to be able to solve for $\dot{\vec{y}}$ in terms of \vec{y} , and write the problem in the form (1.1.1) — else there will either be no solution, or too many. The *inverse* and *implicit* function theorems deal with situations under which systems of nonlinear algebraic equations have solutions that behave “properly”. These are:⁷

- 1. Inverse function theorem.** Let $\vec{y} = \vec{G}(\vec{x})$ be a system of d functions in d variables. Assume \vec{G} is continuously differentiable. Let $\vec{y}_0 = \vec{G}(\vec{x}_0)$. Then a unique inverse function $\vec{x} = \vec{F}(\vec{y})$ — with $\vec{x}_0 = \vec{F}(\vec{y}_0)$ — is defined in a neighborhood of \vec{y}_0 , provided that the linearized problem $\delta\vec{y} = A\delta\vec{x}$ is invertible (that is, $A =$ “gradient of G at x_0 ” is an invertible matrix). ♣

Note 1: Inverse function means that $\vec{y} = \vec{G}(\vec{F}(\vec{y}))$.

Note 2: The linearized problem is obtained by substituting $\vec{y} = \vec{y}_0 + \delta\vec{y}$ and $\vec{x} = \vec{x}_0 + \delta\vec{x}$ into $\vec{y} = \vec{G}(\vec{x})$, and retaining only the linear terms.

- 2. Implicit function theorem.** Consider a system of d equations $\vec{G}(\vec{u}, \vec{y}) = 0$ in d unknowns \vec{u} , with n parameters \vec{y} — where \vec{G} is continuously differentiable. Let $\vec{G}(\vec{u}_0, \vec{y}_0) = 0$. Then the system has a unique solution $\vec{u} = \vec{F}(\vec{y})$ — with $\vec{u}_0 = \vec{F}(\vec{y}_0)$ — defined in a neighborhood of \vec{y}_0 , provided that the linearized system $A\delta\vec{u} + B\delta\vec{y} = 0$ has a unique solution (that is, $A =$ “gradient of G with respect to \vec{u} at (\vec{u}_0, \vec{y}_0) ” is an invertible matrix). ♣

These theorems are also important for bifurcation theory, which will be an important part of this course. Another important theorem is the following:

- 3. Existence and uniqueness for the I.V. problem for ode.** Consider the initial value problem:

$$\frac{d\vec{y}}{dt} = \vec{F}(\vec{y}), \quad \vec{y}(0) = \vec{y}_0, \quad \text{where } \vec{F} \text{ is Lipschitz continuous.} \quad (1.2.2)$$

This problem has a unique solution, defined on some interval $0 < t < T$. The solution depends continuously on \vec{y}_0 , as well as on any parameters on which \vec{F} depends continuously. Furthermore: If \vec{F} is differentiable, the solution has derivatives with respect to \vec{y}_0 . If a parameter dependence of \vec{F} is differentiable, so is the solution, etc ... up to analytic. ♣

No analogous theorem for infinite dimensions: (i) ode theory is fairly “complete”, but there is no complete pde theory at all. (ii) Robust “black-box” ode solvers exist, not so for pde.

⁷Proofs can be found in analysis textbooks.

Examples, etc.:

- 4. Lipschitz continuous** means that $\|\vec{F}(\vec{y}_1) - \vec{F}(\vec{y}_2)\| \leq K \|\vec{y}_1 - \vec{y}_2\|$ for some constant K . ♣
If \vec{F} is differentiable, with a bounded derivative, then it is Lipschitz continuous.

Explained in the lectures: **Definition of differentiable for vector valued functions.**

- 5.** Consider the initial value problem in item 8 in § 1.1: $\dot{x} = 2\sqrt{|x|}$, with $x(0) = 0$. Clearly $F(x) = 2\sqrt{|x|}$ is not Lipschitz continuous near zero, and sure enough uniqueness fails.⁸
- 6.** Example of a differentiable function, with derivative not bounded near a point (hence not Lipschitz):
- $$f(x) = |x|^a \sin(1/x), \text{ where } 1 < a < 2. \quad (1.2.3)$$

Then $f'(x) = -|x|^{a-2} \cos(1/x) + a|x|^{a-1} \sin(1/|x|)$ for $x \neq 0$, and $f'(0) = 0$.

Another example is: $f(x) = |x|^a \sin(1/x^2)$, where $1 < a < 3$.

In models involving an ode where \vec{F} is not Lipschitz continuous, extra care must be taken. In particular: *additional constraints/hypothesis may be needed to have a "good" model.* For example, **consider the ode with a discontinuous right hand side**

$$\dot{x} = 2 + \text{sign}(x). \quad (1.2.4)$$

One can think of this as the limit $\epsilon \downarrow 0$ of

$$\dot{x} = g(x/\epsilon), \quad (1.2.5)$$

where $g = g(z)$ is some nice smooth function — with limits at infinity, where $g(\infty) = 3$ and $g(-\infty) = 1$. For example: $g(z) = 2 + \tanh(z) + 5 \operatorname{sech}(z)$.

Unfortunately, *there are two possible limits!* One when $g(z) > 0$ for all (z) , and another when $g(z_0) = 0$ for some z_0 . Thus (1.2.4) has more than one possible interpretation! If you just write (1.2.4), and say nothing else, then you have an equation whose solutions are not properly defined.

1.2.1 The linearization metatheorem.

[vln385:LN12:L02:MTT]

Here we state a rule-of-thumb, which is useful in identifying statements that are probably true. In particular, when you can trust the results from a linearization, and when you cannot. The rule is:

*Imagine some system of equations, for which you have one solution. Imagine also that you can write a linearized system that solutions "infinitesimally close" to the one you know must satisfy.*⁹

⁸It can be shown that, for existence, continuity is enough. It is the uniqueness that requires Lipschitz continuity.

⁹It is not always possible to linearize. Examples where this happens occur, in applications.

Assume now that the linear system has some property, and that this property is structurally stable.¹⁰ Then the property is also true for the full system.

Examples:

7. Inverse and Implicit function theorems. Note that their formulation, in items 1 and 2, is exactly in the language of the metatheorem.
8. Consider now an ode problem as in (1.1.1)

$$\frac{d\vec{y}}{dt} = \vec{F}(\vec{y}),$$

and some arbitrary solution $\vec{y} = \vec{z}(t)$. Then, in order to be able to linearize the equations near \vec{z} we need \vec{F} to be differentiable. Further, for the resulting linear system to be well behaved, it seems reasonable to require that the derivative of \vec{F} be (at least) bounded. But then \vec{F} is Lipschitz continuous and (theorem in item 3) the nonlinear system is also well behaved.

Notice that the theorem in item 3 is stronger than the metatheorem, in the sense that it applies even in cases where a linearization is not possible.

On the other hand, **note that:** *properties that are not structurally stable for the linearized system are prime candidates for failure in the full system.* You must always try to investigate the effects of “small nonlinear corrections” in such situations. Of course, this is much harder than linearizing, and the subject of perturbation theory. We will cover a few tools from perturbation theory.

THE END

¹⁰Structurally stable means that the property remains true when the linear system is perturbed slightly.