UCLA Math 135, Winter 2015
Ordinary Differential Equations

8. Theory for First-Order Systems

C. David Levermore
Department of Mathematics
University of Maryland

February 5, 2015

Contents

Return to Main Webpage

8. Theory for First-Order Systems

8.1. **Normal Forms and Solutions.** We will now consider first-order systems of n ordinary differential equations for functions $x_j(t)$, $j = 1, 2, \cdots, n$ that can be put into the normal form

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= f_1(t, x_1, x_2, \cdots, x_n)\,,\\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= f_2(t, x_1, x_2, \cdots, x_n)\,,\\
&\ \,\vdots\\
\frac{\mathrm{d}x_n}{\mathrm{d}t} &= f_n(t, x_1, x_2, \cdots, x_n)\,.
\end{aligned} \tag{8.1}$$

We say that n is the *dimension* of this system.

System (8.1) can be expressed more compactly in *vector notation* as

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x})\,, \tag{8.2}$$

where $\mathbf{x}$ and $\mathbf{f}(t, \mathbf{x})$ are given by the n-dimensional *column vectors*

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{f}(t, \mathbf{x}) = \begin{pmatrix} f_1(t, x_1, x_2, \cdots, x_n) \\ f_2(t, x_1, x_2, \cdots, x_n) \\ \vdots \\ f_n(t, x_1, x_2, \cdots, x_n) \end{pmatrix}.$$

We thereby express the system of n equations (8.1) as the single vector equation (8.2). We say x_1, x_2, $\cdots$, x_n are the *entries* of the vector $\mathbf{x}$. Similarly, we say that the functions $f_1(t, x_1, x_2, \cdots, x_n)$, $f_2(t, x_1, x_2, \cdots, x_n)$, $\cdots$, $f_n(t, x_1, x_2, \cdots, x_n)$ are the entries of the vector-valued function $\mathbf{f}(t, \mathbf{x})$.

Remark. We will use boldface, lowercase letters like $\mathbf{x}$ and $\mathbf{f}$ to denote column vectors. Other common notations include an underline like $\underline{x}$ and $\underline{f}$, or an arrow like $\vec{x}$ and $\vec{f}$. Some advanced books do not use any special notation for vectors, but expect the reader to recall what each letter represents from when it was introduced.

Here we recall from multi-variable calculus what it means for a vector-valued function $\mathbf{u}(t)$ to be either continuous or differentiable at a point.

- We say $\mathbf{u}(t)$ is *continuous at* time t if *every entry* of $\mathbf{u}(t)$ is continuous at t.
- We say $\mathbf{u}(t)$ is *differentiable at* time t if *every entry* of $\mathbf{u}(t)$ is differentiable at t.

Given these definitions, we define what it means for a vector-valued function $\mathbf{u}(t)$ to be either continuous, differentiable, or continuously differentiable over a time interval.

- We say $\mathbf{u}(t)$ is *continuous over* a time interval (t_L, t_R) if it is continuous at every t in (t_L, t_R).
- We say $\mathbf{u}(t)$ is *differentiable over* a time interval (t_L, t_R) if it is differentiable at every t in (t_L, t_R).
- We say $\mathbf{u}(t)$ is *continuously differentiable over* a time interval (t_L, t_R) if it is differentiable over (t_L, t_R) and its derivative is continuous over (t_L, t_R).

We are now ready to define what we mean by a solution of system (8.2).

Definition. We say that $\mathbf{x}(t)$ is a *solution* of system (8.2) over a time interval $(t_{_L}, t_{_R})$ when

1. $\mathbf{x}(t)$ is differentiable at every t in $(t_{_L}, t_{_R})$;
2. $\mathbf{f}(t, \mathbf{x}(t))$ is defined for every t in $(t_{_L}, t_{_R})$;
3. equation (8.2) holds at every t in $(t_{_L}, t_{_R})$.

Remark. This definition is similar to definitions of solutions to single differential equations that we gave earlier. The first point states that the left-hand side of the equation makes sense. The second point states that the right-hand side of the equation makes sense. The third point states that the two sides are equal.

8.2. **Initial-Value Problems.** We will consider initial-value problems of the form

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t, \mathbf{x}), \qquad \mathbf{x}(t_I) = \mathbf{x}^I. \tag{8.3}$$

Here t_t is the *initial time*, $\mathbf{x}^I$ is the *initial value* or *initial data*, and $\mathbf{x}(t_I) = \mathbf{x}^I$ is the *initial condition*. Below we will give conditions on $\mathbf{f}(t, \mathbf{x})$ that insure this problem has a unique solution that exists over some time interval that contains t_I. We begin with a definition.

Definition 8.1. *Let S be a set in $\mathbb{R} \times \mathbb{R}^n$. A point $(t_o, \mathbf{x}_o)$ is said to be in the* interior *of S if there exists a box $(t_L, t_R) \times (x_1^L, x_1^R) \times \cdots \times (x_n^L, x_n^R)$ that contains the point $(t_o, \mathbf{x}_o)$ and also lies within the set S.*

Our basic existence and uniqueness theorem is the following.

Theorem 8.1. *Let $\mathbf{f}(t, \mathbf{x})$ be a vector-valued function defined over a set S in $\mathbb{R} \times \mathbb{R}^n$ such that*

- *$\mathbf{f}$ is continuous over S,*
- *$\mathbf{f}$ is differentiable with respect to each x_i over S,*
- *each $\partial_{x_i}\mathbf{f}$ is continuous over S.*

Then for every inital time t_I and every initial value $\mathbf{x}^I$ such that $(t_I, \mathbf{x}^I)$ is in the interior of S there exists a unique solution $\mathbf{x}(t)$ to initial-value problem (8.3) that is defined over some time interval (a, b) such that

- *t_I is in (a, b),*
- *$\{(t, \mathbf{x}(t)) : t \in (a, b)\}$ lies within the interior of S.*

Moreover, $\mathbf{x}(t)$ extends to the largest such time interval and $\mathbf{x}'(t)$ is continuous over that time interval.

Remark. This is not the most general theorem we could state, but it applies to the first-order systems you will face in this course. It asserts that the initial-value problem (8.3) has a unique solution $\mathbf{x}(t)$ that will exist until $(t, \mathbf{x}(t))$ leaves the interior of S.

8.3. **Recasting Higher-Order Problems as First-Order Systems.** Many higher-order differential equation problems can be recast in terms of a first-order system in the normal form (8.2). For example, every n^{th}-order ordinary differential equation in the normal form

$$y^{(n)} = g\big(t, y, y', \cdots, y^{(n-1)}\big) \,,$$

can be expressed as an n-dimensional first-order system in the form (8.2) with

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t,\mathbf{x}) = \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ g(t, x_1, x_2, \cdots, x_n) \end{pmatrix} \,, \qquad \text{where} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix} \,.$$

Notice that the first-order system is expressed solely in terms of the entries of $\mathbf{x}$. The "dictionary" that relates $\mathbf{x}$ to y, y', $\cdots$, $y^{(n-1)}$ is given as a separate equation.

Example. Recast as a first-order system

$$y''' + yy' + e^t y^2 = \cos(3t) \,.$$

Solution. Because this single equation is third order, the first-order system will have dimension three. It will be

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \cos(3t) - x_1 x_2 - e^t x_1^2 \end{pmatrix} \,, \qquad \text{where} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y \\ y' \\ y'' \end{pmatrix} \,.$$

More generally, every d-dimensional m^{th}-order ordinary differential system in the normal form

$$\mathbf{y}^{(m)} = \mathbf{g}\big(t, \mathbf{y}, \mathbf{y}', \cdots, \mathbf{y}^{(n-1)}\big) \,,$$

can be expressed as an md-dimensional first-order system in the form (8.2) with

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(t,\mathbf{x}) = \begin{pmatrix} \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \\ \mathbf{g}(t, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m) \end{pmatrix} \,, \qquad \text{where} \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_m \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{y}' \\ \vdots \\ \mathbf{y}^{(m-1)} \end{pmatrix} \,.$$

Here each $\mathbf{x}_k$ is a d-dimensional vector while $\mathbf{x}$ is the md-dimensional vector constructed by stacking the vectors $\mathbf{x}_1$ through $\mathbf{x}_m$ on top of each other.

Example. Recast as a first-order system

$$q_1'' + f_1(q_1, q_2) = 0 \,, \qquad q_2'' + f_2(q_1, q_2) = 0 \,.$$

Solution. Because this two dimensional system is second order, the first-order system will have dimension four. It will be

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \\ -f_1(x_1, x_2) \\ -f_2(x_1, x_2) \end{pmatrix} \,, \qquad \text{where} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_1' \\ q_2' \end{pmatrix} \,.$$

When faced with a higher-order initial-value problem, we use the dictionary to obtain the initial values for the first-order system from those for the higher-order problem.

Example. Recast as an initial-value problem for a first-order system

$$y'''' - e^y = 0\,, \qquad y(0) = 2\,, \quad y'(0) = -1\,, \quad y''(0) = 5\,, \quad y'''(0) = -4\,.$$

Solution. The first-order initial-value problem is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ e^{x_1} \end{pmatrix}, \qquad \begin{pmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \\ x_4(0) \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 5 \\ -4 \end{pmatrix}, \qquad \text{where} \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y \\ y' \\ y'' \\ y''' \end{pmatrix},$$

Remark. We can also find single higher-order equations that are satisfied by the entries of a first-order system. We will not discuss how this is done because it is not as useful.

8.4. **Linear First-Order Systems.** The n-dimensional first-order system (8.1) is called *linear* when it has the form

$$\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + f_1(t)\,, \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + f_2(t)\,, \\ &\vdots \\ \frac{\mathrm{d}x_n}{\mathrm{d}t} &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t)\,. \end{aligned} \tag{8.4}$$

The functions $a_{jk}(t)$ are called *coefficients* while the functions $f_j(t)$ are called *forcings*.

We can use *matrix notation* to write the linear system (8.4) compactly as

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t)\,, \tag{8.5}$$

where $\mathbf{x}$ and $\mathbf{f}(t)$ are the n-dimensional column vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{f}(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix},$$

while $\mathbf{A}(t)$ is the $n{\times}n$ matrix

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

We call $\mathbf{A}(t)$ the *coefficient matrix* and $\mathbf{f}(t)$ the *forcing vector*. System (8.5) is said to be *homogeneous* if $\mathbf{f}(t) = 0$ and *nonhomogeneous* otherwise.

The product $\mathbf{A}(t)\mathbf{x}$ appearing in system (8.5) denotes the column vector that results from the *matrix multiplication* of the matrix $\mathbf{A}(t)$ with the column vector $\mathbf{x}$. The sum appearing in (8.5) denotes the column vector that results from the *matrix addition* of the column vector $\mathbf{A}(t)\mathbf{x}$ with the column vector $\mathbf{f}(t)$.

Remark. Any n^{th}-order linear equation in the normal form

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_1(t)\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_{n-1}(t)\frac{\mathrm{d}y}{\mathrm{d}t} + a_n(t)y = f(t)\,,$$

can be recast as the n-dimensional first-order linear system

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t)\,, \qquad \text{where} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}\,,$$

with

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_n(t) & \cdots & -a_3(t) & -a_2(t) & -a_1(t) \end{pmatrix}\,, \qquad \mathbf{f}(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{pmatrix}\,.$$

Therefore the study of first-order linear systems contains the study of higher-order linear equations. Conversely, solving a first-order linear system can be reduced to solving a higher-order linear equation.

We will consider linear initial-value problems in the form

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t)\,, \qquad \mathbf{x}(t_I) = \mathbf{x}^I\,, \tag{8.6}$$

where $\mathbf{x}^I$ is called the *vector of initial values*, or simply the *initial vector*.

Many facts that we studied about higher-order linear equations have analogous facts about linear first-order systems. For example, the basic existence and uniqueness theorem is the following.

Theorem 8.2. If $\mathbf{A}(t)$ and $\mathbf{f}(t)$ are continuous over the time interval (t_L, t_R) then for every initial time t_I in (t_L, t_R) and every initial vector $\mathbf{x}^I$ the initial-value problem (8.6) has a unique solution $\mathbf{x}(t)$ that is continuously differentiable over (t_L, t_R). Moreover, if $\mathbf{A}(t)$ and $\mathbf{f}(t)$ are k-times continuously differentiable over the time interval (t_L, t_R) then $\mathbf{x}(t)$ will be is $(k+1)$-times continuously differentiable over (t_L, t_R).

The Basic Existence and Uniqueness Theorem can be used to identify the interval of definition for solutions of the initial-value problem (8.6). This is done very much like the way we identified intervals of definition for solutions of higher-order linear equations. Specifically, if $\mathbf{x}(t)$ is the solution of the initial-value problem (8.6) then its interval of definition will be (t_L, t_R) whenever:

- every entry of the coefficient matrix $\mathbf{A}(t)$ and the forcing vector $\mathbf{f}(t)$ are continuous over (t_L, t_R),
- the initial time t_I is in (t_L, t_R),
- an entry of either the coefficient matrix or the forcing vector is undefined at each of $t = t_L$ and $t = t_R$.

We can do this because the first two bullets along with the Basic Existence and Uniqueness Theorem imply that the interval of definition will be at least (t_L, t_R), while the last two bullets along with our definition of solution imply that the interval of definition can be no bigger than (t_L, t_R) because the equation breaks down at $t = t_L$ and $t = t_R$. This argument works when $t_L = -\infty$ or $t_R = \infty$.