A Summary of Ordinary Differential Equations

Mathematical Modeling of Geological Processes

Matthew D. Covington
University of Arkansas

1 What is an ordinary differential equation?

An **ordinary differential equation (ODE)** is an equation that is written in terms of an unknown function and its derivatives with respect to a *single* independent variable (such as time or position). If a differential equation contains more than one independent variable (such as time *and* position), then it is a **partial differential equation**. Below are several examples of ODEs:

$$\frac{dy}{dt} = y \tag{1}$$

$$\frac{dy}{dt} = e^y \cos(2t) \tag{2}$$

$$\frac{d^2y}{dt^2} - A\frac{dy}{dt} + y = 0 \tag{3}$$

In these equations, y is the "unknown function" mentioned above, and t is the single independent variable. All other quantities in the equations are constants. The **order** of a differential equation is the highest derivative that occurs in the equation. Equations 1 and 2 are both first-order equations, containing only the first derivative of y. Equation 3 is a second-order equation.

A solution to a differential equation is a function, in the cases above $y(t)$, that results in an identity when plugged in to the differential equation. For example, a solution of Equation 1 is $y(t) = e^t$. This can be shown by calculating

$$\frac{dy}{dt} = \frac{d}{dt}(e^t) = e^t \tag{4}$$

and plugging this into Equation 1, which gives

$$e^t = e^t, \tag{5}$$

which is an identity.

2 First-order ODEs

Perhaps the simplest type of ODE is a first-order **pure time equation**. This type of equation takes the form

$$\frac{dy}{dt} = f(t), \tag{6}$$

where $f(t)$ is some arbitrary function of t. This equation can be solved by integration of both sides giving

$$y(t) = \int f(t)dt + C. \tag{7}$$

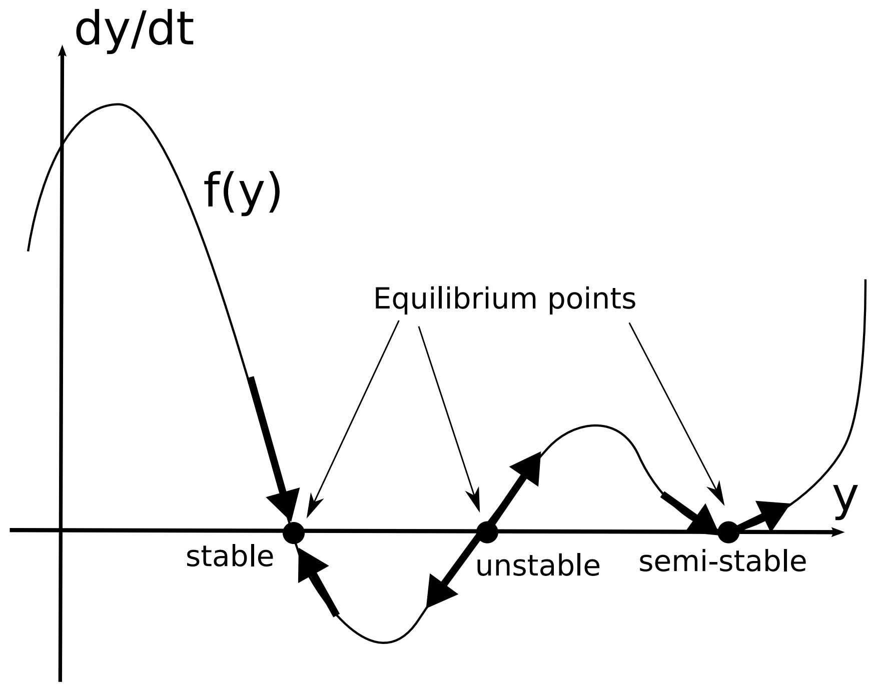

Figure 1: A phase line drawn for a first-order autonomous system, which plots the function $dy/dt = f(y)$ versus y. The three equilibrium points illustrate the three possibilities for system behavior near equilibrium.

Notice that an arbitrary constant of integration, C, appears. As a result, the equation has an infinite number of different solutions (with different values of C). C has to be determined by some additional constraint.

Another simple type of ODE is a first-order autonomous equation. An **autonomous differential equation** is an ODE that does not depend explicitly on the independent variable. Therefore, a first-order autonomous equation has the form

$$\frac{dy}{dt} = f(y), \tag{8}$$

where $f(y)$ is some function of y and does not depend on t. Such equations can also be solved directly by integration, with

$$\int \frac{1}{f(y)} dy = t + C. \tag{9}$$

A useful class of techniques, called **qualitative analysis**, can be illustrated with this equation and used to demonstrate a very general result about first-order autonomous equations. Qualitative analysis allows us to study the general behavior of a differential equation without actually solving it. It may allow us to understand much of the dynamics of the system without having to derive a solution! As an example, consider a function $f(y)$ shown in Figure 2, which plots dy/dt versus y. This type of plot is called a **phase line**. Points of equilibrium occur whenever $dy/dt = 0$, meaning y does not change with time. However, these points are precisely where $f(y) = 0$. We can take the analysis even further. When $f(y) > 0$, then y must be increasing over time, and when $f(y) < 0$ then y must be decreasing in time. We can use these conclusions to draw arrows along the phase line that tell us the direction that the solution evolves over time if started at a particular value of y. Using these arrows, we can see that no matter what function $f(y)$ is, there are three types of system behavior: 1) y can start at an equilibrium point and stay there forever, 2) y can start away from an equilibrium point and approach that equilibrium point, and 3) y can wander off to positive or negative infinity. Consequently, behavior of first-order autonomous systems is monotonic in time. They cannot oscillate. This very general conclusion can be seen from simple phase line analysis.

In general, a first-order ODE has the form $dy/dt = f(y,t)$. We have considered the special cases where $dy/dy = f(t)$ and $dy/dt = f(y)$. Another special case of interest is

$$\frac{dy}{dt} = f(y)g(t). \tag{10}$$

Such equations are called **separable** ODEs, as $f(y,t)$ can be expressed as the product of two separate functions, each of which varies with only the dependent (y) or independent (t) variable. As for the pure time and autonomous cases, separable ODEs can be solved by integration, with

$$\int \frac{1}{f(y)} dy = \int g(t)dt + C. \tag{11}$$

3 Linear ODEs

Linear ODEs are of particular importance because of how commonly they appear in scientific applications and because of the general techniques that are available for solving them. In fact, one common way of treating non-linear equations is to approximate them in a linear form, that is, to **linearize** them. A linear ODE has the form

$$A_0 y + A_1 \frac{dy}{dt} + A_2 \frac{d^2y}{dt^2} + ... = B, \tag{12}$$

where the A_i's and B's can be either constants or functions of the independent variable, t. It is called linear because it contains only a linear function of y and its derivatives. Equations 1 and 3 are examples of linear ODEs, whereas Equation 2 is a non-linear ODE. If $B = 0$ then the ODE is called **homogeneous**, whereas if B is non-zero then the ODE is called **inhomogeneous**.

There are several general theorems concerning linear ODEs that are of substantial practical importance. These include:

- **Theorem 1.** Any Nth-order linear differential equation will have a solution that contains N independent unknown constants that must be determined by boundary conditions or initial conditions. These unknown constants are called **constants of integration**.

- **Theorem 2.** If y_1 and y_2 are both solutions of a homogeneous linear differential equation, then $y_1 + y_2$ is also a solution. This is called the **superposition principle**.

- **Theorem 3.** An Nth-order homogeneous linear differential equation will have N linearly independent solutions, y_1, y_2,...y_N. An arbitrary linear combination of these solutions will also be a solution,

$$y_g(t) = \sum_{n=1}^{N} C_n y_n(t). \tag{13}$$

  This is called the **general solution**.

- **Theorem 4.** Every inhomogeneous linear ODE has a corresponding homogeneous equation (where B is set to zero) with a general solution, y_g. This general solution is also sometimes called the **complementary function**. If one can find any **particular solution** to the inhomogeneous equation, y_p, then the **complete solution** of the inhomogenous problem is given by

$$y_c = y_g + y_p. \tag{14}$$

  The general solution represents the transient response of the system to the boundary conditions, whereas the particular solution represents the system's response to the forcing of the system by the inhomogeneous term.

3.1 Boundary conditions and initial conditions

Since the solution of an Nth-order linear ODE contains N unknown constants, we need an additional N equations to determine a particular solution of that equation. These additional equations are called boundary conditions, or initial conditions (if they are conditions at $t = 0$). An **initial value problem** is the solution of an ODE where the values of the dependent variable and its derivatives are set to constant values at the beginning of a time interval, $y|_{t=0} = A_0$, $dy/dt|_{t=0} = A_1$,...$d^N y/dt^N|_{t=0} = A_N$. For the initial value problem, these are the N additional equations used to determine the N unknown constants in the solution.

For second-order equations and higher, if the constraints on y and its derivatives are set at different values of the dependent variable (i.e. on the boundarys of the interval of interest) then this is called a **boundary value problem**. For example, a second-order linear ODE for $y(x)$ could be constrained so that $y|_{x=0} = A$ and $y|_{x=L} = B$. These would be the two constraints required to determine the two unknown constants in the solution to a second-order linear ODE.

There are three types of boundary conditions:

1. **Dirichlet Conditions**. The solution itself is constrained on the boundary. For example, we might set the temperature at the boundary of a domain where we are modeling heat flow using $T|_{x=0} = T_0$.

2. **Neumann Conditions.** The derivatives of the solution are prescribed at the boundary, for example, $dy/dx|_{x=0} = A$. In many physical problems, fluxes of a quantity (e.g. heat or chemical species) are proportional to the spatial gradient (first derivative) in that quantity. Therefore, Neumann Conditions are often used to constrain the flux of some quantity at the boundary.

3. **Mixed Conditions (or Robin Conditions)**. The boundary condition combines both of the other types. For example, in solution of an equation for $y(x)$ one could require that $Ady/dx|_{x=0} + By|_{x=0} = C$.

In general, you need a boundary condition or initial condition to determine each unknown constant in the solution of an ODE.

3.2 First-order linear ODEs

A first-order linear ODE has the form

$$\frac{dy}{dt} + g(t)y = f(t), \tag{15}$$

where $g(t)$ and $f(t)$ are arbitrary functions of t. Note that no multiplying factor is needed in front of the dy/dt term because if there were one there, all other terms could simply be divded by it, and g and f would still be some function of t. As for first-order pure time, autonomous, and separable equations, there is a general solution to all first-order linear ODEs that can be obtained by integration. This solution is given by

$$y(t) = e^{-I} \int f(t)e^{I} dt + Ce^{-I}, \tag{16}$$

where

$$I = \int g(t)dt. \tag{17}$$

While this solution is general, it will only be useful in relatively simple cases where the integrals over $g(t)$ and $f(t)$ can be performed by hand.

3.3 Second-order linear ODEs

A wide variety of physical processes can be represented by second-order linear ODEs, particularly processes that involve oscillation. Newton's Law produces such an equation, as do harmonic oscillators, such as a mass on a spring.

3.3.1 The homogeneous case

We will consider the special case of a second-order linear homogeneous ODE with constant coefficients, which is given by

$$\frac{d^2y}{dx^2} + A_1\frac{dy}{dx} + A_0 y = 0, \tag{18}$$

where the A_i are constants. Again, no constant is needed before the first term, because if it were there we could simply divide by it. To understand such systems, it is useful to use the notation of a **differential operator**, D, where $Dy = dy/dx$ and $D^2y = d^2y/dx^2$. We will start with a specific example of Equation 18, where

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 4y = 0, \tag{19}$$

which can be expressed in differential operator notation as

$$D^2y + 5Dy + 4y = (D^2 + 5D + 4)y = 0. \tag{20}$$

The differential operator can be factored to give

$$(D + 1)(D + 4)y = 0 \text{ or } (D + 4)(D + 1)y = 0. \tag{21}$$

Therefore, a solution to the equation $(D + 1)y = 0$ or the equation $(D + 4)y = 0$ is also a solution to Equation 19. However, these are first-order separable equations with solutions

$$y_1(x) = C_1e^{-4x} \text{ and } y_2(x) = C_2e^{-x}. \tag{22}$$

Theorem 3 above states that an Nth-order homogeneous linear ODE should contain N linearly independent solutions, and that the sum of them is the general solution. So we expect two linearly independent solutions for our problem. Linear independence between two functions y_1 and y_2 means that there are no non-zero constants A and B such that $Ay_1 + By_2 = 0$. This is true for our two solutions, there are no such non-zero As and Bs. Therefore, the general solution to Equation 19 is

$$y(x) = C_1e^{-4x} + C_2e^{-x}. \tag{23}$$

Returning to the general form of second-order homogeneous linear ODEs (Equation 18) we can see that finding the general solution is a matter of finding the roots of the so-called **auxillary (or characteristic) equation**

$$D^2 + A_1D + A_0 = 0. \tag{24}$$

In the previous example, this equation had two real roots -4 and -1. In general, the roots of this equation, r, can be found using the quadratic formula,

$$r = \frac{-A_1 \pm \sqrt{A_1^2 - 4A_0}}{2}. \tag{25}$$

There are three different possibilities for these two roots, and they result in three different classes of solutions to Equation 18. These possibilities are

1. The two roots, a and b, are both real, and $a \neq b$. In this case (as in our example equation above) the general solution is

$$y_g = C_1e^{ax} + C_2e^{bx}. \tag{26}$$

2. The two roots a and b are real and equal. In this case, the general solution is

$$y_g = (C_1 x + C_2)e^{ax}. \tag{27}$$

3. The two roots, $\alpha \pm i\beta$, are unequal but have imaginary components and are the complex conjugate of each other. From Case 1, we can see that the general solution is

$$y_g = C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x} = e^{\alpha x}(C_1 e^{i\beta x} + C_2 e^{-i\beta x}). \tag{28}$$

   However, using the relation $e^{\pm i\beta x} = \cos\beta x \pm i \sin\beta x$, this can be written as

$$y_g = e^{\alpha x}(A \sin\beta x + B \cos\beta x), \tag{29}$$

   where A and B are new arbitrary constants.

Therefore, second-order homogenous linear ODEs with constant coefficients support two types of behavior, exponential growth and decay (Cases 1 and 2), and oscillation (Case 3). These two types of behavior can also be combined in Case 3 since it has exponential and oscillating components. Unlike unforced first-order systems, unforced second-order systems can wiggle! *In summary, the general procedure for solving second-order homogeneous linear ODEs with constant coefficients is: 1) Determine the characteristic equation, 2) Find its roots, 3) Determine the proper solution from the list above according to the values of the roots.* Similar techniques can be applied to higher order linear homogeneous ODEs with constant coefficients.

3.3.2 The inhomogeneous case

Now we will consider the case of a second-order linear inhomogeneous equation with constant coefficients, which is given by

$$\frac{d^2y}{dx^2} + A_1 \frac{dy}{dx} + A_0 y = f(x), \tag{30}$$

where $f(x)$ is some arbitrary function of x, often called a forcing function (particularly if the independent variable is time). By Theorem 4, we know that the complete solution, y_c will be given by a sum of the general solution of the corresponding homogeneous problem, y_g and a particular solution to the inhomogeneous problem, y_p, such that $y_c = y_g + y_p$. Therefore, the first step is to set $f(x)$ to zero and solve the resulting homogeneous equation by the techniques given above.

This reduces the problem of solving of Equation 30 to a problem of finding a particular solution and adding it to the general solution found above. How might we do this? Often we guess. Typically it is safe to assume that the form of $y_p(x)$ is the same as the form of $f(x)$, up to some undetermined coefficients. For example, if $f(x)$ is a second-order polynomial, then we can assume that $y_p = Ax^2 + Bx + C$, where A, B, and C are unknown constants. This expression can be plugged into the ODE and then we can search for values of A, B, and C that would satisfy the equation. If we can find them, then we have the correct particular solution. Often $f(x)$ has a sinusoidal form. In this case, we can guess a sinusoidal particular solution (as we did with the Pitman model forced by oscillating sea level). If the function $f(x)$ is a sum of several different components, we can search for particular solutions of each of these components separately. Then the complete particular solution will be the sum of these separated solutions.

There are a variety of techniques for finding particular solutions when $f(x)$ has different forms. In fact, there is also a general technique, called **successive integration**, that will allow determination of the particular solution of any such equation through calculation of integrals. See textbooks on differential equations for details on these other techniques.

3.3.3 Other second-order equations

Once we go to second-order linear ODEs with non-constant coefficients, where the coefficients can be a function of the independent variable, there are few available general solutions. There are some special cases that can be solved, so it is often worth looking up whether the equation you are working with fits a special form that has a known solution. Integral transform solutions can also often be used, at least allowing expression of the solution in terms of an integral. We will discuss these techniques later in the course.

4 Non-linear equations

While many physical processes can be expressed using linear differential equations, some processes are inherently non-linear. The toolbox available for solving non-linear equations is substantially smaller than for linear equations. In fact, some of the standard techniques may produce misleading results. Consider the following equation as an example:

$$\frac{dy}{dx} = \sqrt{1 - y^2}. \tag{31}$$

This equation is separable and we can integrate it,

$$\int \frac{dy}{\sqrt{1 - y^2}} = \int dx, \tag{32}$$

giving

$$y = \sin(x + C), \tag{33}$$

where C is a constant of integration. There are two problems with this solution. First, it can be seen from examining the differential equation that the slope of y in time should always be non-negative. This is not true for all segments of a sin function. Therefore, only parts of the solution we derived are actually valid. Furthermore, there are parts of the solution that are missing. We can see this using a qualitative technique called a slope field. This is a plot that shows vectors representing the value of dy/dt at all points in x-y space. The slope field in Figure 2 depicts how solutions will evolve if started in a particular point in space. Here we can see that there are stable solutions at $y = -1$ and $y = 1$, that were not included in our previous solution. In the region between y of -1 and 1, the solution will follow the sin function we derived, but once it reaches $y = 1$ it will stay there. Figure 3 shows a phase line plot of this system, where the two equilibrium points can be seen. Note that one is

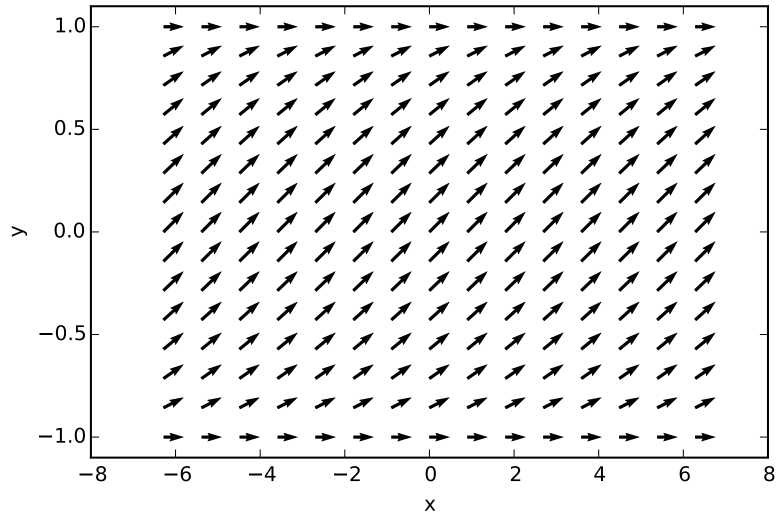

Figure 2: A slope field for the non-linear differential equation (31). Solutions will follow the arrows as they evolve.)

semistable and the other is stable. Does this make sense in light of the slope field in Figure 2? Qualitative techniques such as these are some of the most powerful techniques available for studying non-linear systems.

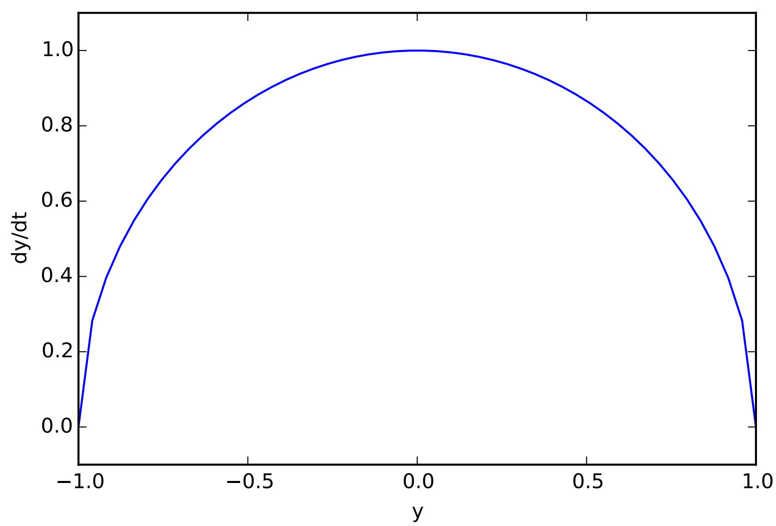

Figure 3: We can also examine the phase line of this ODE. Here we see that there are two stable points, one of which is only semi-stable.