G. Graphical and Numerical Methods

In studying the first-order ODE

(1)
$$\frac{dy}{dx} = f(x, y),$$

the main emphasis is on learning different ways of finding explicit solutions. But you should realize that most first-order equations cannot be solved explicitly. For such equations, one resorts to graphical and numerical methods. Carried out by hand, the graphical methods give rough qualitative information about how the graphs of solutions to (1) look geometrically. The numerical methods then give the actual graphs to as great an accuracy as desired; the computer does the numerical work, and plots the solutions.

1. Graphical methods.

The graphical methods are based on the construction of what is called a **direction field** for the equation (1). To get this, we imagine that through each point (x, y) of the plane is drawn a little line segment whose slope is $f(x, y)$. In practice, the segments are drawn in at a representative set of points in the plane; if the computer draws them, the points are evenly spaced in both directions, forming a lattice. If drawn by hand, however, they are not, because a different procedure is used, better adapted to people.

To construct a direction field by hand, draw in lightly, or in dashed lines, what are called the **isoclines** for the equation (1). These are the one-parameter family of curves given by the equations

(2)
$$f(x, y) = c, \qquad c \text{ constant.}$$

Along the isocline given by the equation (2), the line segments all have the same slope c; this makes it easy to draw in those line segments, and you can put in as many as you want. (Note: "iso-cline" = "equal slope".)

The picture shows a direction field for the equation

$$y' = x - y .$$

The isoclines are the lines $x - y = c$, two of which are shown in dashed lines, corresponding to the values $c = 0, -1$. (Use dashed lines for isoclines).
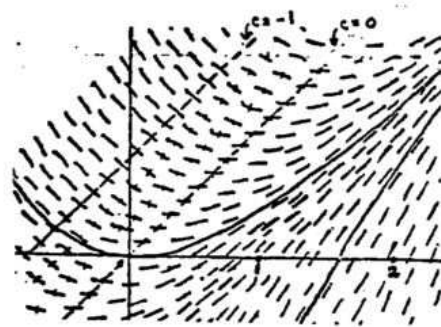
Once you have sketched the direction field for the equation (1) by drawing some isoclines and drawing in little line segments along each of them, the next step is to draw in curves which are at each point tangent to the line segment at that point. Such curves are called **integral curves** or *solution curves* for the direction field. Their significance is this:

(3)   *The integral curves are the graphs of the solutions to* $y' = f(x, y)$.

Proof. Suppose the integral curve C is represented near the point (x, y) by the graph of the function $y = y(x)$. To say that C is an integral curve is the same as saying

$$\text{slope of } C \text{ at } (x, y) \quad = \quad \text{slope of the direction field at } (x, y);$$

from the way the direction field is defined, this is the same as saying

$$y'(x) \quad = \quad f(x, y) \ .$$

But this last equation exactly says that $y(x)$ is a solution to (1). □

We may summarize things by saying, the direction field gives a picture of the first-order equation (1), and its integral curves give a picture of the solutions to (1).

Two integral curves (in solid lines) have been drawn for the equation $y' = x - y$. In general, by sketching in a few integral curves, one can often get some feeling for the behavior of the solutions. The problems will illustrate. Even when the equation can be solved exactly, sometimes you learn more about the solutions by sketching a direction field and some integral curves, than by putting numerical values into exact solutions and plotting them.

There is a theorem about the integral curves which often helps in sketching them.

Integral Curve Theorem.

(i) If $f(x, y)$ is defined in a region of the xy-plane, then integral curves of $y' = f(x, y)$ cannot cross at a positive angle anywhere in that region.

(ii) If $f_y(x, y)$ is continuous in the region, then integral curves cannot even be tangent in that region.

A convenient summary of both statements is (here "smooth"$=$ continuously differentiable):

Intersection Principle

(4)   *Integral curves of $y' = f(x, y)$ cannot intersect wherever $f(x, y)$ is smooth.*

Proof of the Theorem. The first statement (i) is easy, for at any point (x_0, y_0) where they crossed, the two integral curves would have to have the same slope, namely $f(x_0, y_0)$. So they cannot cross at a positive angle.

The second statement (ii) is a consequence of the uniqueness theorem for first-order ODE's; it will be taken up then when we study that theorem. Essentially, the hypothesis guarantees that through each point (x_0, y_0) of the region, there is a unique solution to the ODE, which means there is a unique integral curve through that point. So two integral curves cannot intersect — in particular, they cannot be tangent — at any point where $f(x, y)$ has continuous derivatives.

Exercises: Section 1C

2. The ODE of a family. Orthogonal trajectories.

The solution to the ODE (1) is given analytically by an xy-equation containing an arbitrary constant c; either in the explicit form (5a), or the implicit form (5b):

$$(5) \qquad\qquad (a) \quad y = g(x, c) \qquad\qquad\qquad (b) \quad h(x, y, c) = 0 \ .$$

In either form, as the parameter c takes on different numerical values, the corresponding graphs of the equations form a one-parameter family of curves in the xy-plane.

We now want to consider the inverse problem. Starting with an ODE, we got a one-parameter family of curves as its integral curves. Suppose instead we start with a one-parameter family of curves defined by an equation of the form (5a) or (5b), can we find a first-order ODE having these as its integral curves, i.e. the equations (5) as its solutions?

The answer is yes; the ODE is found by differentiating the equation of the family (5) (using implicit differentiation if it has the form (5b)), and then using (5) to eliminate the arbitrary constant c from the differentiated equation.

Example 1. Find a first-order ODE whose general solution is the family

$$(6) \qquad\qquad\qquad y = \frac{c}{x - c} \qquad (c \text{ is an arbitrary constant}).$$

Solution. We differentiate both sides of (6) with respect to x, getting $y' = -\dfrac{c}{(x - c)^2}$. We eliminate c from this equation, in steps. By (6), $x - c = c/y$, so that

$$(7) \qquad\qquad\qquad y' = -\frac{c}{(x - c)^2} = -\frac{c}{(c/y)^2} = -\frac{y^2}{c} \ ;$$

To get rid of c, we solve (6) algebraically for c, getting $c = \dfrac{yx}{y + 1}$; substitute this for the c on the right side of (7), then cancel a y from the top and bottom; you get as the ODE having the solution (6)

$$(8) \qquad\qquad\qquad y' = -\frac{y(y + 1)}{x} \ .$$

Remark. The c must not appear in the ODE, since then we would not have a single ODE, but rather a one-parameter family of ODE's — one for each possible value of c. Instead, we want just one ODE which has each of the curves (5) as an integral curve, regardless of the value of c for that curve; thus the ODE cannot itself contain c.

Orthogonal trajectories.

Given a one-parameter family of plane curves, its **orthogonal trajectories** are another one-parameter family of curves, each one of which is perpendicular to all the curves in the original family. For instance, if the original family consisted of all circles having center at the origin, its orthogonal trajectories would be all rays (half-lines) starting at the origin.

Orthogonal trajectories arise in different contexts in applications. For example, if the original family represents the **lines of force** in a gravitational or electrostatic field, its orthogonal trajectories represent the **equipotentials**, the curves along which the gravitational or electrostatic potential is constant.

In a temperature map of the U.S., the original family would be the **isotherms**, the curves along which the temperature is constant; their orthogonal trajectories would be the **temperature gradients**, the curves along which the temperature is changing most rapidly.

More generally, if the original family is of the form $h(x, y) = c$, it represents the **level curves** of the function $h(x, y)$; its orthogonal trajectories will then be the **gradient curves** for this function — curves which everywhere have the direction of the gradient vector ∇h. This follows from the 18.02 theorem which says the gradient ∇h at any point (x, y) is perpendicular to the level curve of $h(x, y)$ passing through that point.

To find the orthogonal trajectories for a one-parameter family (5):

1. Find the ODE $y' = f(x, y)$ satisfied by the family.
2. The new ODE $y' = -\dfrac{1}{f((x, y)}$ will have as its integral curves the orthogonal trajectories to the family (5); solve it to find the equation of these curves.

The method works because at any point (x, y), the orthogonal trajectory passing through (x, y) is perpendicular to the curve of the family (5a) passing through (x, y). Therefore the slopes of the two curves are negative reciprocals of each other. Since the slope of the original curve at (x, y) is $f(x, y)$, the slope at (x, y) of the orthogonal trajectory has to be $-1/f(x, y)$. The ODE for the orthogonal trajectories then gives their slope at (x, y), thus it is

$$(9) \qquad\qquad y' = -\frac{1}{f(x, y)} \qquad \textit{ODE for orthogonal trajectories to (5a) .}$$

More generally, if the equation of the original family is given implicitly by (5b), and its ODE is also in implicit form, the procedure and its justification are essentially the same:

1. Find the ODE in implicit form $F(x, y, y') = 0$ satisfied by the family (5).
2. Replace y' by $-1/y'$; solve the new ODE $F(x, y, -1/y') = 0$ to find the orthogonal trajectores of the original family.

Example 2. Find the orthogonal trajectories to the family of curves $y = c\,x^n$, where n is a fixed positive integer and c an arbitrary constant.

Solution. If $n = 1$, the curves are the family of rays from the origin, so the orthogonal trajectories should be the circles centered at the origin – this will help check our work.

We first find the ODE of the family. Differentiating the equation of the family gives $y' = nc\,x^{n-1}$; we eliminate c by using the equation of the family to get $c = y/x^n$ and substituting this into the differentiated equation, giving

$$(10) \qquad y' = \frac{n\,y}{x} \quad (\textit{ODE of family}); \qquad\qquad y' = -\frac{x}{n\,y} \quad (\textit{ODE of orthog. trajs.}) .$$

Solving the latter equation by separation of variables leads first to $n\,ydy = -xdx$, then after integrating both sides, transposing, and multiplying through by 2, to the solution

$$(11) \qquad x^2 + n\,y^2 = k, \qquad (k \geq 0 \textit{ is an arbitrary non-negative constant; n is fixed.})$$

For different k-values, the equations (11) represent the family of ellipses centered at the origin, and having x-intercepts at $\pm\sqrt{k}$ and y-intercepts at $\pm\sqrt{k/n}$.

If $n = 1$, these intercepts are equal, and the ellipses are circles centered at the origin, as predicted.

Exercises: Section 1D-2

3.   Euler's numerical method.   The graphical method gives you a quick feel for how the integral curves behave. But when they must be known accurately and the equation cannot be solved exactly, numerical methods are used. The simplest method is called **Euler's method**. Here is its geometric description.

We want to calculate the solution (integral curve) to $y' = f(x, y)$ passing through (x_0, y_0). It is shown as a curve in the picture.

We choose a step size h. Starting at (x_0, y_0), over the interval $[x_0, x_0 + h]$, we approximate the integral curve by the tangent line: the line having slope $f(x_0, y_0)$. (This is the slope of the integral curve, since $y' = f(x, y)$.)

This takes us as far as the point (x_1, y_1), which is calculated by the equations (see the picture)

$$x_1 = x_0 + h, \qquad y_1 = y_0 + h\, f(x_0, y_0) \ .$$

Now we are at (x_1, y_1). We repeat the process, using as the new approximation to the integral curve the line segment having slope $f(x_1, y_1)$. This takes us as far as the next point (x_2, y_2), where

$$x_2 = x_1 + h, \qquad y_2 = y_1 + h\, f(x_1, y_1) \ .$$

We continue in the same way. The general formulas telling us how to get from the $(n-1)$-st point to the n-th point are

(12) $$\qquad\qquad x_n = x_{n-1} + h, \qquad y_n = y_{n-1} + h\, f(x_{n-1}, y_{n-1}) \ .$$

In this way, we get an approximation to the integral curve consisting of line segments joining the points (x_0, y_0), $(x_1, y_1), \dots$.

In doing a few steps of Euler's method by hand, as you are asked to do in some of the exercises to get a feel for the method, it's best to arrange the work systematically in a table.

Example 3. For the IVP: $y' = x^2 - y^2, \quad y(1) = 0$, use Euler's method with step size .1 to find y(1.2).

Solution. We use $f(x, y) = x^2 - y^2, \quad h = .1$, and (12) above to find x_n and y_n:

n	x_n	y_n	$f(x_n, y_n)$	$hf(x_n, y_n)$
0	1	0	1	.1
1	1.1	.1	1.20	.12
2	1.2	.22		

Remarks. Euler's method becomes more accurate the smaller the step-size h is taken. But if h is too small, round-off errors can appear, particularly on a pocket calculator.

As the picture suggests, the errors in Euler's method will accumulate if the integral curve is convex (concave up) or concave (concave down). Refinements of Euler's method are aimed at using as the slope for the line segment at (x_n, y_n) a value which will correct for the convexity or concavity, and thus make the next point (x_{n+1}, y_{n+1}) closer to the true integral curve. We will study some of these. The book in Chapter 6 has numerical examples illustrating Euler' s method and its refinements.

Exercises: Section 1C

M.I.T. 18.03 Ordinary Differential Equations
18.03 Notes and Exercises