One-dimensional differential equations and phase lines

by Eric Cytrynbaum

April 12, 2008

The phase line

When dealing with differential equations, one is often faced with the problem of not being able to come up with a solution in closed form (i.e. a nice clean formula for the solution). In these cases, what information about solutions can you extract directly from the equation? To work with a concrete example, let's consider the equation:

$$\frac{dx}{dt} = x - x^2.$$

To facilitate the following explanations, let's give the function on the right hand side (RHS) of the equation a name: $f(x) = x - x^2$, so that the equation can be written as $dx/dt = f(x)$. Note that I have suppressed the explicit reference to the t dependence of the function and will do so often – you must always keep mental track of which letters are parameters (constants whose values are unspecified – none in the example above) and which are variables (values that change within the dynamic context of the problem), specifically independent variables (like t) and dependent variables (like x).

Steady states solutions

These are solutions that do not change in time so they must have time-derivative equal to zero. To find these, we force $dx/dt = 0$ which means that $x - x^2 = 0$. The steady state solutions for our example equation are therefore the constant functions of time $x(t) = 0$ and $x(t) = 1$.

Stability

Once we have found the steady states, we can ask how solutions that start close to the steady state behave. If *all* solutions that start sufficiently close to a steady state eventually approach the steady state value, then the steady state is called stable[1]. If there are solutions starting arbitrarily close to the steady state that leave the area near the steady state, then the steady state is called unstable. There are two ways to determine stability, qualitatively from the graph of the RHS, $x - x^2$, and by studying the RHS algebraically (or differentially, really). The former approach helps to explain the latter so I'll start with the former.

At any time t, the expression on the RHS, $x(t) - x(t)^2$ determines the instantaneous direction and amplitude of change in $x(t)$ at time t. So if $x(t) - x(t)^2 > 0$ for some value of t, then $x(t)$ is increasing. Similarly, if $x(t) - x(t)^2 < 0$, then $x(t)$ is decreasing. By drawing the graph of the RHS

[1]Technically, the definition of stability only requires that solutions which start close to the steady state stay close to the steady state. The notion of stability used here, requiring that all nearby solutions *approach* the steady state is referred to as asymptotic stability.

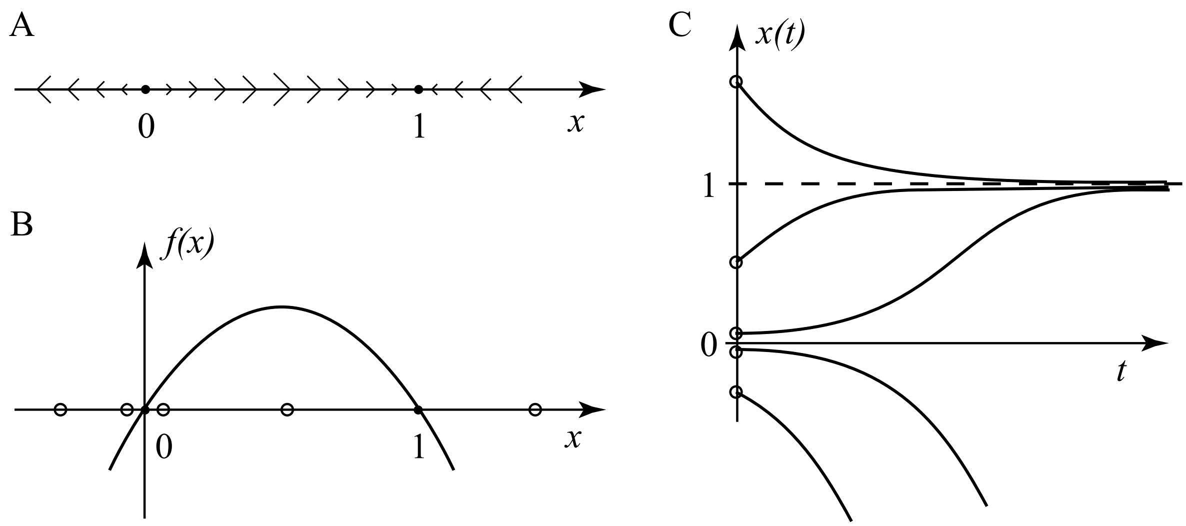

Figure 1: The phase line and interpretive aids for the equation $dx/dt = f(x)$. Note that according to my personal definition, the phase line is really just the diagram in A. It consists of an axis for the state variable (x), steady states marked with solid dots and arrows indicating how the state changes at each point on the axis. The arrows are often called a *direction field*. Panel B shows the state axis (x) annotated with a vertical axis to show the direction field as a curve instead of as arrows. The empty circles mark the points used as initial conditions for the sketch in C (also shown in C). To translate from A (or B) to C, imagine the empty circles moving in time along the axis, at every instant following the directions given by direction field (including direction and speed).

as a function of x, we can quickly identify the intervals on which $x(t)$ is increasing and those on which it is decreasing. The x axis in this graph is called the phase space[2] where by phase we mean a representation of the state of the system. Because in this case the state is entirely characterized by a single variable (x) we use the expression phase line.

Note that in drawing $f(x)$ as a function of x we have no graphical representation of time. This means that when looking at the phase line, we must imagine that the current state is marked by a flashing dot at the current value of x. As time goes on, the flashing dot moves according to the instructions provided by $f(x)$ which specifies the rate of change of x. The phase line for the example above is illustrated in Figure 1.

The steady states of the equation are the values of x at which the RHS is zero so the graph must cross the x axis at each steady state. To determine stability for the steady state $x(t) = 1$, we are interested in what happens at values of x close to the steady state in question. The graph of $x - x^2$ is positive immediately to the left of $x = 1$, so solutions that start just below $x = 1$ (to the left of the steady state), must increase in time and must continue to do so without ever getting above $x = 1$ and so they must approach 1. The steady state at $x = 1$ is starting to look stable. But we must check all nearby starting values, in particular, those above $x = 1$. Here, $x - x^2 < 0$ so any solution that starts above $x = 1$ decreases and continues to do so without ever getting below

[2]The expression state space is often used instead of phase space and is arguably a better term because we usually talk about the state of the system rather than the phase of the system except for systems in which the state space is periodic or well characterized by plan or previous experience (phase of the moon, phase of construction, "its just a phase he's going through"). Other expressions include phase line (one state variable), phase plane (two state variables), phase portrait (arbitrary number of state variables).

$x = 1$ and so must also approach 1. Thus, we conclude that $x = 1$ is a stable steady state.

Notice that stability arose from the fact that below the steady state the RHS is positive and above the steady state the RHS is negative (i.e. the arrows on the phase line point toward the steady state). Provided the RHS is differentiable, this is equivalent to saying that the slope of the function $f(x) = x - x^2$ is negative at the steady state. Checking this condition in the example above, we see that $f'(1) = -1$.

Similarly, instability arises when $f'(x_{ss}) > 0$ which is equivalent to saying that the arrows point away from the steady state. Note that $f'(0) = 1$ so $x = 0$ is an unstable steady state.

Problems

Although phase line analysis is most useful for equations that cannot be solved in closed form, and the following questions are all focussed on phase line arguments, many of the equations below *are* nonetheless solvable. For example, separation of variables works on the quadratic ones.

1. Find all steady states of the following equations and determine stability for each one, first qualitatively by sketching the function and then quantitatively by checking the slope of the RHS at the steady state.

   (a) $dx/dt = x^2 - 2x - 3$.

   (b) $dx/dt = x^2 + 3$.

   (c) $dx/dt = x^2 - 2x + 1$.

2. Draw the phase line for the following equations for several values of the parameter a. Choose one value of a for every possible qualitatively different phase line. To define, "qualitatively different", if two values of a both give phase lines with three steady states, with the lowest and highest stable and the middle one unstable, then they have, qualitatively, the same phase line.

   (a) $dx/dt = a - x^2$.

   (b) $dx/dt = -ax + x^3$.

   (c) $dx/dt = a - x + x^3$.

   (d) $dx/dt = a + x - x^3$.

3. Consider the equation $dx/dt = -bx + x^3$. Note that for $b > 0$, $x = 0$ is a stable steady state. For values of x very close to $x = 0$, the x^3 term is insignificant compared to the $-bx$ term. With this in mind, what function does a solution $x(t)$ that approaches zero look like as it get close to zero?

4. Consider the equation $dx/dt = bx - x^3$. Note that for $b > 0$, $x = \sqrt{b}$ is a stable steady state. For values of x very close to $x = \sqrt{b}$, a phenomenon similar to the one explored in the previous problem occurs. With this in mind, what function does a solution $x(t)$ that approaches $\sqrt{b}$ look like as it get close to $\sqrt{b}$? Rewriting $f(x) = bx - x^3$ as an expansion about $\sqrt{b}$ will be useful.