Lecture 11: Variation of Parameters

February 26, 2007

In this lecture we develop a very general method for solving second-order, linear, nonhomogeneous, constant coefficient ODE's

$$y'' + Py' + Qy = f(x), \tag{1}$$

where the RHS is not necessarily a function we could annihilate with a linear differential operator M with constant coefficients. (Indeed, if we could annihilate $f(x)$ with such an M, we should use the method of annihilators instead of the method discussed here.) Here $P, Q \in \mathbb{R}$ are constants, and the operator version of the equation would be $L[y] = f(x)$, where $L[y] = (D^2 + PD + Q)y$.

The method developed here is important because it can be generalized to higher-order linear ODEs, with not necessarily constant coefficients.[1] The bulk of the effort, as before, is used to find a single y_p. It is assumed that we know how to find $y_h = c_1y_1 + c_2y_2$. Of course this is no trouble for us in the case of (1), since at worst we need the quadratic formula to solve the characteristic equation $m^2 + Pm + Q = 0$.[2]

The method for finding y_p consists of assuming a certain form of y_p, based on that of y_h, and then using some clever linear algebra. In particular we will eventually employ Cramer's Rule. See Lecture 10, Footnote 6 for a quick review of Cramer's Rule.

In first section we will develop the method for (1). The development is interesting and worth reading carefully, though it ends with a formula which is very general. The formula comes from Cramer's Rule and a system of two equations in two "unknowns." We will arrive at a formula which is relatively easily employed, and leave it up to the reader whether it is better to memorize the formula, or the process that arrived at the formula. (Zill prescribes the latter approach.)

1 Variation of Parameters: Theory

We start with our second-order, linear, nonhomogeneous ODE (1) as before:

$$y'' + Py' + Qy = f(x).$$

Assume we have found a two-parameter representation of y_h:

$$y_h = c_1y_1(x) + c_2y_2(x). \tag{4}$$

---

[1] Zill at first looks at a very general solution to

$$y'' + P(x)y' + Q(x)y = f(x), \tag{2}$$

rightly pointing out that this is the second-order version of the linear equations we did back in Lecture 3 (Section 2.3 in Zill), which were of the form

$$y' + P(x)y = f(x). \tag{3}$$

He then focusses on our case (1) for most of his discussion. At the end of Section 4.6, on page 171, he discusses how one might generalize the method.

Once you are finished reading this lecture, it would be good to peruse Zill's presentation, from beginning to end to see the more general treatment. Here we will only assume (1).

[2] Finding y_h for the more general cases is a much more difficult problem, especially if the coefficients are not constant. Zill does not really address this, but his method for finding y_p—assuming we know y_h—in such cases is still valid.

Next we assume[3] that y_p can be found from the form of y_h, the difference being that we let the parameters vary (hence the name of the technique!), i.e., if we look for functions of the form[4]

$$y_p = u_1(x)y_1(x) + u_2(x)y_2(x). \tag{5}$$

Now we have to put y_p back into the original ODE and see what we can do with the expansion. To do so, we need to compute y_p' and y_p''. For the sake of simplicity we will suppress the independent variable x, but with the understanding that $y_1 = y_1(x)$, $y_2 = y_2(x)$, $u_1 = u_1(x)$, and $u_2 = u_2(x)$.

$$\begin{align}
y_p &= u_1y_1 + u_2y_2, \tag{6}\\
y_p &= u_1y_1' + u_1'y_1 + u_2y_2' + u_2'y_2, \tag{7}\\
y_p'' &= u_1y_1'' + u_1'y_1' + u_1''y_1 + u_1'y_1' + u_2y_2'' + u_2'y_2' + u_2''y_2 + u_2'y_2'. \tag{8}
\end{align}$$

Now we need to make a linear algebra argument. When we put this information into the ODE, we will have, albeit in an expanded version,

$$y_p'' + Py_p' + Qy_p = f(x). \tag{9}$$

The trouble is we then have only one equation, (9) and two unknowns, u_1 and u_2. In general function theory speak (similar to linear algebra speak), we would say the system is underdetermined. We need another equation, or constraint, to have any hope that we can solve for u_1 and u_2. Before picking another constraint, let us take a look at what (9) looks like when expanded.

$$\begin{multline}
(u_1y_1'' + u_1'y_1' + u_1''y_1 + u_1'y_1' + u_2y_2'' + u_2'y_2' + u_2''y_2 + u_2'y_2')\\
+ P(u_1y_1' + u_1'y_1 + u_2y_2' + u_2'y_2) + Q(u_1y_1 + u_2y_2) = f(x). \tag{10}
\end{multline}$$

This looks like a terrible mess, but there is a clever choice for a second functional constraint which will clear up most of it. Our second constraint will be

$$u_1'y_1 + u_2'y_2 = 0. \tag{11}$$

Under this new assumption (11) that $u_1'y_1 + u_2'y_2 = 0$, let us revisit our computations (6)–(10).

$$\begin{align}
y_p &= u_1y_1 + u_2y_2, \tag{12}\\
\implies y_p' &= u_1y_1' + u_1'y_1 + u_2y_2' + u_2'y_2\\
&= u_1y_1' + u_2y_2' + \underbrace{u_1'y_1 + u_2'y_2}_{0}\\
\implies y_p' &= u_1y_1' + u_2y_2', \tag{13}\\
\implies y_p'' &= u_1y_1'' + u_1'y_1' + u_2y_2'' + u_2'y_2', \tag{14}
\end{align}$$

Now we put this information back into our original ODE and see, after some rearrangement, that

$$\begin{align*}
& y_p'' + Py_p' + Qy_p = f(x)\\
\iff & (u_1y_1'' + u_1'y_1' + u_2y_2'' + u_2'y_2') + P(u_1y_1' + u_2y_2') + Q(u_1y_1 + u_2y_2) = f(x)\\
\iff & u_1(\underbrace{y_1'' + Py_1' + Qy_1}_{0}) + u_2(\underbrace{y_2'' + Py_2' + Qy_2}_{0}) + u_1'y_1' + u_2'y_2' = f(x).
\end{align*}$$

---

[3] And later see if it works out. It does.

[4] Intuitively this has some hope for success, since when we plug this into the original ODE, with product rules everywhere, there might be parts which "disappear" because of the y_h-terms, and the rest may satisfy an equation we can actually solve. Still, we will have to investigate carefully as always.

The reason the terms over the braces sum to zero is that y_1 and y_2 were solutions to the homogeneous equation $y'' + Py' + Qy = 0$. Along with the contstraint (11) from before (listed first), we have

$$\begin{aligned} u_1' y_1 + u_2' y_2 &= 0, \\ u_1' y_1' + u_2' y_2' &= f(x). \end{aligned}$$

We are almost ready to make use of Cramer's Rule, except that y_1, y_2 and their derivatives are the "knowns" and u_1', u_2' are unknowns, so we write

$$\begin{array}{ccccc} y_1 u_1' & + & y_2 u_2' & = & 0 \\ y_1' u_1' & + & y_2' u_2' & = & f(x). \end{array} \tag{15}$$

So we will employ a little bit of cleverness, by treating the u_1', u_2' as the "unknowns" and y_1, y_2, y_1', y_2' as the coefficients, of the "linear system" (15). Cramer's Rule then gives us (borrowing notation from Zill, p. 168)

$$u_1' = \frac{W_1}{W}, \qquad u_2' = \frac{W_2}{W}, \tag{16}$$

where

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}, \qquad W_1 = \begin{vmatrix} 0 & y_2 \\ f(x) & y_2' \end{vmatrix}, \qquad W_2 = \begin{vmatrix} y_1 & 0 \\ y_1' & f(x) \end{vmatrix}. \tag{17}$$

Notice that W is our old friend the Wronskian (which comes up increasingly if we delve deeper into the linear theory). More importantly, notice that we know y_1, y_2 and $f(x)$, and can easily compute y_1', y_2', so Cramer's Rule gives us simple formulas for u_1' and u_2'. From these we need only integrate to get u_1 and u_2:

$$\begin{aligned} u_1' &= \frac{-y_2 f(x)}{y_1 y_2' - y_2 y_1'} &&\Longrightarrow u_1 = \int \frac{-y_2 f(x)}{y_1 y_2' - y_2 y_1'}\, dx, \\ u_2' &= \frac{y_1 f(x)}{y_1 y_2' - y_2 y_1'} &&\Longrightarrow u_2 = \int \frac{y_1 f(x)}{y_1 y_2' - y_2 y_1'}\, dx. \end{aligned}$$

Putting this together with our general theory of linear ODEs gives us then

$$y = y_p + y_h = (u_1 y_1 + u_2 y_2) + c_1 y_1 + c_2 y_2, \tag{18}$$

into which we must plug y_1, y_2, u_1, u_2.

Now ultimately u_1 and u_2 are given here by indefinite integrals, but we see that any "+C" terms in either would just add more scalar multiples of y_1 or y_2 to the solution (18), and so we usually use the simplest antiderivative with no additive constant.[5]

To implement the method does not require deriving the method along the way. However, memorizing the formulas for u_1 and u_2 does not seem like a reliable strategy either. A compromise, summary method is offered below.

1. Given $y'' + Py' + Qy = f(x)$, first find two linearly independent solutions which span y_h. Call them y_1 and y_2.
2. Assume y_p is of the form $y_p = u_1 y_1 + u_2 y_2$.
3. With the added constraint (11) (listed first), this will give the system (15):

$$\begin{array}{ccccc} y_1 u_1' & + & y_2 u_2' & = & 0 \\ y_1' u_1' & + & y_2' u_2' & = & f(x). \end{array}$$

---

[5] It is interesting to note that we could include the additive constant in the u_1, u_2 functions, and that would "take care" of the homogeneous part of the solution (18). Fair enough. Unfortunately this is not much of a shortcut in the method, since the method required us to know y_1 and y_2.

4. Solve the system for u_1', u_2' using Cramer's Rule. We will note here that Cramer's rule simplifies to

$$u_1' = \frac{-y_2 f(x)}{\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}} = \frac{-y_2 f(x)}{W}, \tag{19}$$

$$u_2' = \frac{y_1 f(x)}{\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}} = \frac{y_1 f(x)}{W}. \tag{20}$$

5. Integrate u_1', u_2' to get u_1, u_2. Ignore additive constants.

6. The solution to the ODE is then $y = y_p + y_h$, as always, so here we will have the two-parameter family of equations

$$y = \underbrace{u_1 y_1 + u_2 y_1}_{y_p} + \underbrace{c_1 y_1 + c_2 y_2}_{y_h}.$$

In the next section we apply the method to nonhomogeneous problems.

2 Examples

Example 1 *(Zill, p. 172, #6) Solve* $y'' + y = \sec^2 x$.

Solution: Clearly we cannot annihilate the RHS of this nonhomogeneous ODE. We first notice that it can be rewritten $(D^2+1)y = \sec^2 x$, *and so the characteristic polynomial of the operator on the LHS will be* m^2+1, *with roots* $m = \pm i$, *so two linearly independent functions spanning* y_h *will be* $y_1 = \cos x$ *and* $y_2 = \sin x$. *With these we assume*

$$y_p = u_1 y_1 + u_2 y_2 = u_1 \cos x + u_2 \sin x.$$

With the a priori assumption that $u_1' y_1 + u_2' y_2 = 0$, *we could trace through the development above and arrive at the system*

$$\begin{array}{rcrcl} u_1' y_1 & + & u_2' y_2 & = & 0 \\ u_1' y_1' & + & u_2' y_2' & = & \sec^2 x \end{array} \qquad \Longrightarrow \qquad \begin{array}{rcrcl} \cos x \cdot u_1' & + & \sin x \cdot u_2' & = & 0 \\ -\sin x \cdot u_1' & + & \cos x \cdot u_2' & = & \sec^2 x \end{array}$$

The solutions to these are as follow:

$$u_1' = \frac{\begin{vmatrix} 0 & \sin x \\ \sec^2 x & \cos x \end{vmatrix}}{\begin{vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{vmatrix}} = \frac{-\sin x \sec^2 x}{1} = -\sec x \tan x \qquad \Longrightarrow \qquad u_1 = -\sec x.$$

$$u_2' = \frac{\begin{vmatrix} \cos x & 0 \\ -\sin x & \sec^2 x \end{vmatrix}}{\begin{vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{vmatrix}} = \frac{\cos x \sec^2 x}{1} = \sec x \qquad \Longrightarrow \qquad u_2 = \ln|\sec x + \tan x|.$$

Thus $y_p = u_1 y_1 + u_2 y_2 = -\sec x \cos x + \ln|\sec x + \tan x| \cdot \sin x = -1 + \sin x \ln|\sec x + \tan x|$, *and* $y = y_p + y_h$, *giving us as the general solution the following 2-parameter family:*

$$y = \underbrace{-1 + \sin x \ln|\sec x + \tan x|}_{y_p} + \underbrace{A\cos x + B \sin x}_{y_h}.$$

In case one wishes to check the solution indeed works, it is a matter of carefully computing the effect of the differential operator (D^2+1) *on such a* y*. Now the* $y_h = A\cos x + B\sin x$ *part is clearly annihilated by our differential operator* (D^2+1)*, so we need only check that* $(D^2+1)y_p = \sec^2 x$*. In doing so it is also useful to note that* $\frac{d}{dx}\ln|\sec x + \tan x| = \sec x$*, well-known from the usual antiderivative formula for* $\sec x$ *(but worth verifying for review). Now we compute:*

$$
\begin{aligned}
& y_p = -1 + \sin x \ln|\sec x + \tan x| \\
\\
\implies\ & y_p' = \sin x \sec x + \cos x \ln|\sec x + \tan x| \\
\iff\ & y_p' = \tan x + \cos x \ln|\sec x + \tan x| \\
\\
\implies\ & y_p'' = \sec^2 x + \cos x \sec x - \sin x \ln|\sec x + \tan x| \\
\iff\ & y_p'' = \sec^2 x + 1 - \sin x \ln|\sec x + \tan x|.
\end{aligned}
$$

Thus we have $(D^2+1)y_p$ *becoming*

$$
\begin{aligned}
y_p'' + y_p &= (\sec^2 x + 1 - \sin x \ln|\sec x + \tan x|) + (-1 + \sin x \ln|\sec x + \tan x|) \\
&= \sec^2 x, \qquad \textit{q.e.d.}
\end{aligned}
$$

It should be clear that we had no hope of solving this above problem with our earlier methods. It is also clear that the Wronskian has more uses than just its utility in proving that a well-chosen set of functions is linearly independent. Indeed, the determinant in the denominators in Cramer's Rule were just the Wronskian of $\cos x, \sin x$.

It is also interesting to trace how, if we chose the functions to span y_h in the other order, i.e., $\sin x$ and then $\cos x$, though u_1 and u_2 would be switched we would still get the same y_p.

Note also how we had to integrate to get u_1 and u_2 from u_1' and u_2', respectively, and it is quite possible that no simple formula exists for the antiderivatives we need. See Zill's Example 3, p. 170.

Finally, on p. 171 Zill shows how the method generalizes to higher-order problems. For that one needs to have further constraints added so that the equation simplifies in ways analogous to the above, and so that we again have n equations in n "unknowns" $u_1', u_2', \cdots, u_n'$ if we begin with an nth order nonhomogenous linear ODE. The key is to arrive at a system

$$
\begin{array}{lllllllll}
y_1u_1' & + & y_2u_2' & + & \cdots & + & y_nu_n' & = & 0 \\
y_1'u_1' & + & y_2'u_2' & + & \cdots & + & y_n'u_n' & = & 0 \\
\vdots & & & & & & & & \vdots \\
y_1^{(n-1)}u_1' & + & y_2^{(n-1)}u_2' & + & \cdots & + & y_n^{(n-1)}u_n' & = & f(x)
\end{array}
$$

from which again Cramer's Rule can be applied to find $u_1', u_2', \cdots, u_n'$ so we can find $u_1, u_2, \cdots, u_n$.[6]

We now look at another example, but still only a second-order application.

---

[6] Linear Algebra books usually list Cramer's Rule as just a curiosity, because a system with constant coefficients—almost the only type found in such books—is usually better solved with Gaussian Elimination, in part because determinants of $n \times n$ matrices are necessarily containing $n!$ sums (in fact alternating sums and differences) of products, each with n terms. Thus the determinant of a 3×3 matrix is a sum of 6 terms, each being a product of 3 terms from the matrix, a 4×4 determinant will be a sum of 24 products of 4 terms each, and the determinant of a 5×5 matrix is a sum of 120 terms, each being a product of 5 terms from the matrix. Therefore solving a system of five equations in five unknowns, which would in general require findind six such determinants for Cramer's Rule, is usually best left to Gaussian Elimination.

However, the elimination strategy necessarily relies heavily upon the particular coefficients, where Cramer's Rule does not at all. In our application our "coefficients" of our linear systems (15) are not constants so even a system of two equations in two unknowns will not so easily succumb to elimination (and a system of three equations in three unknowns would be nearly impossible, and we would find ourselves re-inventing Cramer's Rule without necessarily seeing the simple structure of the solution with determinants).

Do note how the Wronskian not being identically zero is crucial in (16).

Example 2 *(Zill, p. 172, #9, with a slight variation) Solve* $y'' - 4y = \frac{e^{2x}}{x}$*, assuming* $x > 0$*. (The interested reader could see how it would change if we assume* $x < 0$*.)*

<u>Solution</u>: We begin by noting that we have no hope of annihilating the RHS with a nontrivial linear differential operator with constant coefficients. Next we note that this ODE is of the form $(D^2 - 4)y = \frac{1}{x}e^{2x}$*, i.e.,* $(D-2)(D+2)y = \frac{1}{x}e^{2x}$*, so our functions for* y_h *can be* e^{2x} *and* e^{-2x}*, so* $y_h = Ae^{2x} + Be^{-2x}$*, and*

$$y_p = u_1y_1 + u_2y_2 = u_1e^{2x} + u_2e^{-2x}.$$

Again from the theoretical development, including the added constraint so y_p *is no longer underdetermined, we eventually get our system (15), which for this example means*

$$\begin{array}{rcrcl} u_1'y_1 & + & u_2'y_2 & = & 0 \\ u_1'y_1' & + & u_2'y_2' & = & \frac{1}{x}e^{2x} \end{array} \qquad \Longrightarrow \qquad \begin{array}{rcrcl} e^{2x}\cdot u_1' & + & e^{-2x}\cdot u_2' & = & 0 \\ 2e^{2x}\cdot u_1' & + & (-2e^{-2x})\cdot u_2' & = & \frac{1}{x}e^{2x}. \end{array}$$

The solution to these is again given by Cramer's Rule, as follows:

$$u_1' = \frac{\begin{vmatrix} 0 & e^{-2x} \\ \frac{1}{x}e^{2x} & -2e^{-2x} \end{vmatrix}}{\begin{vmatrix} e^{2x} & e^{-2x} \\ 2e^{2x} & -2e^{-2x} \end{vmatrix}} = \frac{-\frac{1}{x}}{-4} = \frac{1}{4x} \qquad \Longrightarrow \qquad u_1 = \frac{1}{4}\ln x.$$

$$u_2' = \frac{\begin{vmatrix} e^{2x} & 0 \\ 2e^{2x} & \frac{1}{x}e^{2x} \end{vmatrix}}{\begin{vmatrix} e^{2x} & e^{-2x} \\ 2e^{2x} & -2e^{-2x} \end{vmatrix}} = \frac{\frac{1}{x}e^{4x}}{-4} = -\frac{1}{4}\cdot\frac{1}{x}e^{4x} \qquad \Longrightarrow \qquad u_2 = -\frac{1}{4}\int_1^x \frac{1}{t}e^{4t}\,dt.$$

In forms we used for u_1 *and* u_2 *we made use of the assumption* $x > 0$*. The lower limit of integration for our form of* u_2 *does not matter as long as it is positive (Zill uses* $x_0 > 0$*); the difference in using another choice will in* u_2*, be an additive constant that can be absorbed into our parameters. Recalling that* $y_p = u_1y_1 + u_2y_2$*, but switching the order of multiplication for mainly aesthetic reasons, we get our 2-parameter general solution* $y = y_p + y_h$*, i.e.,*

$$y = \underbrace{\frac{1}{4}e^{2x}\ln x - \frac{1}{4}e^{-2x}\int_1^x \frac{1}{t}e^{4t}\,dt}_{y_p} + \underbrace{Ae^{2x} + Be^{-2x}}_{y_h}.$$

If $x < 0$ were desired, we would replace $\ln x$ with $\ln|x| = \ln(-x)$, and 1 with -1 or a similar negative number.