Lecture 3: First-Order Linear ODE's

January 22, 2007

1 Some Definitions

These will not be needed to complete the problems, but it is eventually important to know the vocabulary of the theory which this section of Zill (Section 2.3) is part of. It is also interesting to see some of the general terminology we will encounter in the weeks ahead, though there will be a short detour into so-called *exact* equations before we come back to this theory with a vengeance. The actual techniques we use will be given in Section 2 of this handout.

The general nth order linear ODE, where we assume $y = y(x)$, is given by

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = g(x). \tag{1}$$

A very useful notation we will use later in the course is the following (Zill, p. 130), where $D = \frac{d}{dx}$:

$$\begin{aligned} L[y] &= a_n(x)D^n y + a_{n-1}(x)D^{n-1}y + \cdots + a_1(x)Dy + a_0(x)y \\ &= \left\{a_n(x)D^n + a_{n-1}(x)D^{n-1} + \cdots + a_1(x)D + a_0(x)\right\} y, \end{aligned} \tag{2}$$

which will let us write the linear equation in the more compact form

$$L[y] = g(x), \tag{3}$$

with the understanding that we are solving an equation, the solution being a function $y = y(x)$.

The reason that (1), and thus $L[y] = g(x)$ is called **linear** is because L is a **linear differential operator**. The "differential" part means involving derivatives and the "linear" part means that

$$L[y_1 + y_2] = L[y_1] + L[y_2], \qquad \text{for all } y_1, y_2, \tag{4}$$

$$L[\beta y] = \beta L[y], \qquad \text{for all } \beta \in \mathbb{R}. \tag{5}$$

In fact (4) and (5) together form the general linear-algebraic definition of a linear operator, with the **vector space** here being some kind of function space in which L makes sense (such as a space of n-times differentiable functions, to be somewhat specific).[1]

[1] A *vector space* is a set with operations called "vector addition" and "scalar multiplication," satisfying several structural axioms. (See any text on linear algebra.) For our part later in the course, the crucial axioms are that it is *closed* under these operations, meaning if the vector space is V, then

(1) for all $\mathbf{u}, \mathbf{v} \in V$ we have $\mathbf{u} + \mathbf{v} \in V$, and

(2) for all $\mathbf{u} \in V$ and $\beta \in \mathbb{R}$, we have $\beta\mathbf{u} \in V$.

More advanced texts define vector spaces which are "function spaces," specifically

$$C^k(I) = \left\{ f : I \to \mathbb{R} \;\middle|\; f, f', f'', \ldots, f^{(k)} \text{ exist and are continuous on } I \right\},$$

where I is some interval, and $f : I \to \mathbb{R}$ means that f inputs values in I and outputs values in $\mathbb{R}$. (I will be the domain, and $\mathbb{R}$ will contain the range.) Thus $C^n(I)$ for a given interval I, or even $C^n(\mathbb{R})$ are natural domains of the operator L in (3). Note from calculus that if f^k is defined and continuous, so is f^{k-1}, and therefore f^{k-2}, etc., until we are down to f itself. $C(I)$ is the space of functions which are continuous on I, $C^1(I)$ is the set whose derivatives are also, etc. These are all vector spaces.

In linear algebra terms, an L satisfying (4) is said to "preserve (vector) addition," while an L satisfying (5) is said to "preserve scalar multiplication," the scalars here being the constants $\beta \in \mathbb{R}$.

To see that L as in (2) does indeed fit the definition of linear operator, let us first consider it in pieces. First, assume that $L[y] = a_0(x)y$ for all functions $y = y(x)$ for which $L[y]$ makes sense. Then

$$L[y_1 + y_2] = a_0(x)(y_1 + y_2) = a_0(x)y_1 + a_0(x)y_2 = L[y_1] + L[y_2];$$

$$L[\beta y] = a_0(x)(\beta y) = \beta a_0(x)y = \beta L[y].$$

Next we will look at the case $L[y] = a_1(x)\frac{dy}{dx}$. Thus

$$\begin{aligned} L[y_1 + y_2] &= a_1(x)\frac{d}{dx}[y_1 + y_2] = a_1(x)\left\{\frac{dy_1}{dx} + \frac{dy_2}{dx}\right\} \\ &= a_1(x)\frac{dy_1}{dx} + a_1(x)\frac{dy_2}{dx} = L[y_1] + L[y_2]; \end{aligned}$$

$$L[\beta y] = a_1(x)\frac{d(\beta y)}{dx} = a_1(x)\beta\frac{dy}{dx} = \beta a_1(x)\frac{dy}{dx} = \beta L[y].$$

The fact that the general L is still linear follows similarly, since the $a_k(x)$ functions are **coefficients** which go along for the ride. To prove in the more general case it is easier to cite a theorem from linear algebra:

Theorem 1 *An operator is linear, i.e., satisfies* (4) *and* (5) *if and only if the following holds for all* y_1, y_2, α, β*:*

$$L[\alpha y_1 + \beta y_2] = \alpha L[y_1] + \beta L[y_2]. \tag{6}$$

The proof is a fairly quick linear algebra exercise. In short, If (6) holds, it is true for $\alpha, \beta = 1$, which gives (4), and for $\alpha = 0$, giving (5). Conversely, if (4) and (5) both hold, then

$$L[\alpha y_1 + \beta y_2] = L[\alpha y_1] + L[\beta y_2] = \alpha L[y_1] + \beta L[y_2],$$

which is (6), where we used preservation of addition first, and then preservation of scalar multiplication. That completes a proof.

The upshot of Theorem 1 is that we can prove one equation, namely (6), that L preserves **linear combinations** of functions y_1, y_2, to show (2) gives a linear operator in the sense of (4) and (5). To save space, let us just show that an operator with nth, first and zero-order terms, i.e.,

$$L[y] = a_n(x)\frac{d^n y}{dx^n} + a_1(x)\frac{dy}{dx} + a_0(x)y$$

is linear (and the middle terms would fit in the obvious way if included).

$$\begin{aligned} &L[\alpha y_1 + \beta y_2] \\ &\quad = a_n(x)\frac{d^n}{dx^n}\{\alpha y_1 + \beta y_2\} + a_1(x)\frac{d}{dx}\{\alpha y_1 + \beta y_2\} + a_0(x)\{\alpha y_1 + \beta y_2\} \\ &\quad = \alpha a_n(x)\frac{d^n y_1}{dx^n} + \beta a_n(x)\frac{d^n y_2}{dx^n} + \alpha a_1(x)\frac{dy_1}{dx} + \beta a_1(x)\frac{dy_2}{dx} + \alpha a_0(x)y_1 + \beta a_0(x)y_2 \\ &\quad = \alpha\left[a_n(x)\frac{d^n y_1}{dx^n} + a_1(x)\frac{dy_1}{dx} + a_0(x)y_1\right] + \beta\left[a_n(x)\frac{d^n y_2}{dx^n} + a_1(x)\frac{dy_2}{dx} + a_0(x)y_2\right] \end{aligned}$$

$$= \alpha L[y_1] + \beta L[y_2],$$

as we claimed.

It is important that the coefficients $a_k(x)$ are functions of x only; if they are allowed to contain y as well, we would lose linearity.

In fact, solving (1) is quite difficult, if not impossible without resorting to numerical methods, except under certain circumstances. If the coefficients $a_0(x), a_1(x), \ldots, a_n(x)$ are all constant, then there is hope, as we will see in a later lecture. Fortunately, when we have such an equation which is only order 1, even when the coefficients are nonconstant a general method is available. It is a "clever" enough method that it is best memorized and not re-invented each time it is needed. It is presented below.

2 Solving First-Order Linear ODE's

By definition, these will be of the form

$$a_1(x)\frac{dy}{dx} + a_0(x)y = g(x). \tag{7}$$

However, this is not the form that we will use to build our method upon. Instead, we will divide by $a_1(x)$, to get

$$\frac{dy}{dx} + \frac{a_0(x)}{a_1(x)}y = \frac{g(x)}{a_1(x)},$$

which we then write for convenience as

$$\frac{dy}{dx} + P(x)y = f(x). \tag{8}$$

Zill calls (8) the **standard form** of (7).

Note that we divided by $a_1(x)$, which may occasionally be zero. We have not yet discussed the topic of just *where* we can find a solution, i.e., for which x's can we solve such an equation. Thus anytime we try to solve such an equation, we must realize that our method may well break down outside of intervals on which $P(x)$ and $f(x)$ are defined and continuous. Usually it is obvious, from the form of the solution, just where the solution is valid. We will revisit this idea as we continue our development.

Returning to (8), the following technique (trick?) was discovered over the years:

1. Given (8), i.e., $\dfrac{dy}{dx} + P(x)y = f(x)$.

2. Multiply both sides by $\eta(x) = e^{\int P(x)\,dx}$:

$$\eta(x)\frac{dy}{dx} + \eta(x)P(x)y = \eta(x)f(x), \qquad \text{i.e.,} \tag{9}$$

$$e^{\int P(x)\,dx}\frac{dy}{dx} + e^{\int P(x)\,dx}P(x)y = e^{\int P(x)\,dx}f(x). \tag{10}$$

3. Recognize that the LHS of (9) (or (10)) is a product rule. In fact, notice two things about this new equation:

   (a) The derivative of $\eta(x)$ is given by the chain rule and Fundamental Theorem of Calculus:

$$\frac{d\eta(x)}{dx} = e^{\int P(x)\,dx}\frac{d}{dx}\int P(x)\,dx = P(x)e^{\int P(x)\,dx} = P(x)\eta(x); \tag{11}$$

   (b) The RHS is a function of x alone, call it $q(x) = \eta(x)f(x)$. Thus we have

$$\eta\frac{dy}{dx} + y\frac{d\eta}{dx} = q(x). \tag{12}$$

4. Re-write the LHS as a derivative of a product:

$$\frac{d}{dx}(\eta y) = q(x). \tag{13}$$

5. This gives $\eta y = \displaystyle\int q(x)\, dx$, so that

$$y = \frac{\int q(x)\, dx}{\eta}.$$

If we would like to trace everything through based upon (8), we would get

$$y = \frac{\int e^{\int P(x)\, dx} f(x)\, dx}{e^{\int P(x)\, dx}}. \tag{14}$$

The function $\eta(x) = e^{\int P(x)\, dx}$ is called an **integrating factor**, because multiplying by this function gives a desirable form, in this case a product rule form on the LHS of (8), from which we can quickly "integrate," or solve the ODE.

A couple of remarks about constants should be made here. First, we can use any constant we would like in the integral appearing in the integrating factor. It is usually easier to just assume the arbitrary constant of integration is zero there. In fact note

$$e^{\int P(x)\, dx + C_2} = e^{\int P(x)\, dx} e^{C_2} = C_3 e^{\int P(x)\, dx},$$

so if we change the constant in $\int P(x)\, dx$ we are simply multiplying the equation in standard form (8) by a nonzero constant, which does not change anything, including the product rule form in the LHS of the new equation (12). However, the whole of the integral in the numerator of (14) will contain an arbitrary additive constant which does matter, and becomes the parameter in the one-parameter family of solutions of the original ODE.

3 The Integrating Factor in Action

One could simply memorize the solution (14) to solve these. However, the above process is usually superior because the formula is sufficiently complicated, and there are places to catch mistakes if we break it into the smaller steps. Furthermore, all we need to memorize are the forms of the ODE (8), the spirit of the process, and the integrating factor

$$\eta(x) = e^{\int P(x)\, dx}. \tag{15}$$

Example 1 *(#5 in Zill, p. 65) Solve the ODE:* $y' + 3x^2 y = x^2$.

Here $P(x) = 3x^2$, *so*

$$\eta(x) = e^{\int P(x)\, dx} = e^{\int 3x^2\, dx} = e^{x^3}.$$

Multiplying our ODE by $\eta(x)$ *gives us*

$$\begin{aligned}
y' + 3x^2 y = x^2 \implies \quad & e^{x^3} y' + 3x^2 e^{x^3} y = e^{x^3} x^2 \\
\implies \quad & \left(e^{x^3} y\right)' = x^2 e^{x^3} \\
\implies \quad & e^{x^3} y = \int x^2 e^{x^3}\, dx \\
\implies \quad & e^{x^3} y = \frac{1}{3} e^{x^3} + C \\
\implies \quad & y = \frac{\frac{1}{3} e^{x^3} + C}{e^{x^3}}.
\end{aligned}$$

Usually this process gives us a solution which can be then simplified a bit:

$$y = \frac{1}{3} + Ce^{-x^3}.$$

This is a one-parameter family of curves. In fact, the solution is valid for all $x \in \mathbb{R}$, which is the where this solution is defined and is continuous.

Example 1 is one of the simplest. In fact, it is separable (as the reader should check)! However, these can become less forgiving as they become more difficult to compute the integrals involved. It is also crucial that the equation is in the form (8), i.e., $y' + P(x)y = f(x)$.

Example 2 *Solve the ODE:* $\dfrac{dy}{dx} = x + y$.

First we need to get the correct form, which again was $\quad y' + P(x)y = f(x)$*:*

$$\frac{dy}{dx} - y = x.$$

This gives $P(x) = -1$, so $\eta(x) = e^{\int P(x)\,dx} = e^{\int(-1)\,dx} = e^{-x}$. Multiplying by this integrating factor gives

$$\begin{aligned} y' - y = x \implies & \quad e^{-x}y' - e^{-x}y = e^{-x}x \\ \implies & \quad \left(e^{-x}y\right)' = xe^{-x} \\ \implies & \quad e^{-x}y = \int xe^{-x}\,dx. \end{aligned}$$

Of course now we must integrate by parts.

$$\begin{aligned} u &= x & dv &= e^{-x}\,dx \\ du &= dx & v &= -e^{-x} \end{aligned}$$

This gives us

$$e^{-x}y = uv - \int v\,du = x(-e^{-x}) + \int e^{-x}\,dx = -xe^{-x} - e^{-x} + C.$$

Multiplying by e^x then gives us $y = e^x\left(-xe^{-x} - e^{-x} + C\right)$ and so

$$y = -x - 1 + Ce^x.$$

Example 3 *(From a previous edition of Zill) Solve the ODE:* $\dfrac{dy}{dx} + y\cot x = 2\cos x$.

Here $P(x) = \cot x$, and so

$$\eta(x) = e^{\int P(x)\,dx} = e^{\int \cot x\,dx} = e^{\ln|\sin x|} = |\sin x|.$$

Here we can wave our hands a bit. After all, $|\sin x| = \pm\sin x$, depending upon whether $\sin x$ is positive or negative, but we can certainly multiply both sides of our ODE by either (and check that it works, i.e., that we get a product rule form on the LHS). For simplicity we will multiply by $\sin x$:

$$\begin{aligned} y' + y\cot x = 2\cos x \implies & \quad y'\sin x + y\cot x\sin x = 2\cos x\sin x \\ \implies & \quad y'\sin x + y\cos x = 2\cos x\sin x \\ \implies & \quad (y\sin x)' = 2\cos x\sin x \\ \implies & \quad y\sin x = \int 2\cos x\sin x\,dx = \frac{1}{2}\sin^2 x + C \\ \implies & \quad y = \frac{\frac{1}{2}\sin^2 x + C}{\sin x}. \end{aligned}$$

Thus

$$y = \frac{1}{2}\sin x + C\csc x,$$

and is valid on all intervals of the form $(n\pi, (n+1)\pi)$*, i.e., except where* $\sin x = 0$.

4 Derivation of $\eta(x)$

For completeness the derivation of η is given here. The idea is that one assumes an integrating factor exists, and then attempts to use the ODE to find what the form of η must be. Recall that the whole point of multiplying by η was to make the LHS a product rule, in this case. Thus we can ignore the RHS (so long as it is a function of x). What we really need is $(\eta y)' = \text{RHS}$, i.e.,

$$(\eta y)' = \eta y' + P(x)\eta y. \tag{16}$$

Expanding the derivative on the left then gives

$$\eta y' + y\eta' = \eta y' + P(x)\eta y. \tag{17}$$

Subtracting $\eta y'$ from both sides gives

$$y\eta' = P(x)\eta y, \tag{18}$$

which, after dividing by ηy gives

$$\frac{\eta'}{\eta} = P(x). \tag{19}$$

Recalling that $\eta = \eta(x)$, and putting in the integrals gives

$$\int \frac{\eta'(x)}{\eta(x)}\,dx = \int P(x)\,dx. \tag{20}$$

The integrand on the RHS is the same as $(\ln|\eta(x)|)'$, so we have

$$\ln|\eta(x)| = \int P(x)\,dx, \tag{21}$$

giving us

$$|\eta(x)| = e^{\int P(x)\,dx} \tag{22}$$

so that

$$\eta(x) = \pm e^{\int P(x)\,dx}. \tag{23}$$

Now as we mentioned before, if η works as an integrating factor, so does $-\eta(x)$ (and in fact any nonzero multiple of η will work because multiplicative constants can go along for the ride), so we wave our hands a little and just take

$$\eta(x) = e^{\int P(x)\,dx}. \tag{24}$$

Again, this derivation is not necessary once we know the method, but it is interesting to see how one could derive the method. Furthermore the techniques used in this derivation can be attempted with other, more exotic ODE's, and in fact will appear in a little different form in Section 4.6 of Zill.

Homework 3-A

In all of the problems below, find an integrating factor $\eta(x)$ and solve using the method of Section 2 and the examples of Section 3 above.

1. Solve the ODE $y' + 2xy = x$.

2. Solve #17 in Zill: $\cos x\dfrac{dy}{dx} + y\sin x = 1$.