ORDINARY DIFFERENTIAL EQUATIONS

L. COLLINS

We cover the following first/second order ordinary differential equations (ODEs). The *order* of an ODE is determined by the highest order derivative present in the equation. By *ordinary*, we mean that the derivatives are of a function (y) of only one independent variable (x).

First Order ODEs

(1) **Separable**
These are equations which can be brought to the form
$$f(y)\,dy = g(x)\,dx.$$
Integrating both sides gives the general solution.

(2) **Exact Equations**
These are equations whose left hand side is an exact differential, i.e., the result of applying the product rule to some function of x and y. In general, they have the form
$$f'(x)\,g(y) + f(x)\,g'(y)\,\frac{dy}{dx} = h(x),$$
or, in terms of differentials,
$$f'(x)\,g(y)\,dx + f(x)\,g'(y)\,dy = h(x)\,dx,$$
which can be transformed to
$$d(f(x)\,g(y)) = h(x)\,dx.$$
Integrating both sides gives the general solution.

(3) **Linear Equations**
First order equations of the form
$$\frac{dy}{dx} + f(x)\,y = g(x)$$
are said to be *linear*. They can be reduced to exact equations by multiplying throughout by
$$\mu(x) = \exp\Big(\int f(x)\,dx\Big),$$
known as the *integrating factor* (where $\exp(x) \stackrel{\text{def}}{=} e^x$).

Date: 13th May, 2016.

Second Order ODEs

(1) **Homogeneous with Constant Coefficients**
A second order ODE with constant coefficients is *homogeneous* when it equals zero. In other words, we consider the equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

where $a, b, c \in \mathbb{R}$ are constants. First, we solve the **auxiliary equation**

$$ak^2 + bk + c = 0$$

whose solutions are $k = k_1$ and $k = k_2$. The general solution is then given by

$$y(x) = \begin{cases} c_1 e^{k_1 x} + c_2 e^{k_2 x} & \text{if } k_1 \neq k_2 \\ c_1 e^{kx} + c_2 x e^{kx} & \text{if } k = k_1 = k_2 \\ e^{\alpha x}(c_1 \cos \beta x + c_2 \sin \beta x) & \text{if } k = \alpha \pm \beta i \in \mathbb{C}, \end{cases}$$

where c_1, c_2 are arbitrary constants.

(2) **Inhomogeneous with Constant Coefficients**
A differential equation is *inhomogeneous* if it is not homogeneous. Here we consider the equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) \neq 0$$

where $a, b, c \in \mathbb{R}$ are constants. We solve by following these steps:

(i) Solve the **homogeneous equation**

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

to obtain the **complementary function** $\mathrm{cf}(x)$.

(ii) Guess a **trial solution**, i.e., a function $\mathrm{ts}(x)$ which, when substituted in the left-hand side of the equation, is likely to result in $f(x)$. Table 1 suggests trial solutions for common elementary functions f. Note that even if some of the constants $a, b, \dots$ in f are zero, the corresponding constants $\lambda, \mu, \dots$ in the trial solution should not be assumed zero. For example, if $f(x) = x^2$, then we still take the trial solution to be $\lambda x^2 + \mu x + \eta$. Similarly, the function $f(x) = x^2 + \sin 2x$ has trial solution $\lambda x^2 + \mu x + \eta + \zeta \cos 2x + \tau \sin 2x$.

(iii) Determine the derivatives $\mathrm{ts}'(x)$ and $\mathrm{ts}''(x)$, and substitute them into the LHS of the original ODE. Compare coefficients to determine correct values for the constants so that the result will equal $f(x)$.

The trial solution with the constant(s) found is called the **particular integral**, $\mathrm{pi}(x)$.

(iv) The general solution is given by $y(x) = \mathrm{cf}(x) + \mathrm{pi}(x)$.

	$f(x)$	Trial Solution, $\mathrm{ts}(x)$
Polynomials	a $ax+b$ ax^2+bx+c $\vdots$	λ $\lambda x+\mu$ $\lambda x^2+\mu x+\eta$ $\vdots$
Exponentials[†]	$ae^{\alpha x}$	$\lambda e^{\alpha x}$ if $k_1 \neq \alpha \neq k_2$ $\lambda x e^{\alpha x}$ if $k_1 = \alpha \neq k_2$ $\lambda x^2 e^{\alpha x}$ if $k_1 = \alpha = k_2$
Trigonometric	$a\cos\alpha x + b\sin\alpha x$	$\lambda\cos\alpha x + \mu\sin\alpha x$

[†] Note that k_1, k_2 are the solutions to the auxiliary equation solved in part (i).

TABLE 1. Trial solutions of common elementary functions.

Remark: Why does this method work? Differentiation is an operator, that is, a function whose inputs and outputs are themselves functions. If we denote the differentiation of f by $D[f]$, then both f and $D[f]$ are functions, which when evaluated at x, yield the numbers $f(x)$ and $D[f](x)$ respectively. The symbol D alone denotes differentiation as a function in its own right. (In Leibniz notation, this is the difference between $\frac{dy}{dx}$, which is the function $D[f]$ whose inputs are numbers, and $\frac{d}{dx}$, which is equivalent to D and whose inputs are functions.)

In general, an operator L is *linear* if for any two functions f and g,

$$L[f+g] = L[f] + L[g] \qquad \text{and} \qquad L[\alpha f] = \alpha\, L[f],$$

where α is any constant. Indeed, the differential operator D is linear, e.g., if for all x, f and g are defined by $f(x) = \sin x$ and $g(x) = x^2$, then

$$D[2f+3g](x) = 2\cos x + 6x = 2D[f](x) + 3D[g](x),$$

i.e., $D[2f+3g] = 2D[f] + 3D[g]$.

Studying linear operators abstractly is useful. Let $\mathbf{0}$ denote the zero function, i.e., the function defined by $\mathbf{0}(x) = 0$ for all x. Note that this is different from zero; the former is a function, the latter is a number. Now if L is a linear operator, the set of functions which are mapped to $\mathbf{0}$ by L is called the *kernel*, denoted $\ker(L)$. In other words,

$$f \in \ker(L) \iff L[f] = \mathbf{0}.$$

The function $\mathbf{0}$ itself is in the kernel of any linear operator L. Indeed, since for any function f, we have $(0f)(x) = 0\,f(x) = 0 = \mathbf{0}(x)$ for all x, then $0f = \mathbf{0}$. Hence since L is linear,

$$L[\mathbf{0}] = L[0\,\mathbf{0}] = 0\,L[\mathbf{0}] = \mathbf{0},$$

so $\mathbf{0} \in \ker(L)$.

The kernel of a linear operator L can tell us a lot about it, such as whether or not L is invertible. Recall that in general, a function F has an inverse if and only if it is one-to-one, i.e., if for all x and y, $F(x) = F(y)$ implies that $x = y$. Applying this reasoning to linear operators, it is easy to see that for L to have an inverse L^{-1}, only

$\mathbf{0}$ must be in its kernel. Indeed, $\mathbf{0} \in \ker(L)$ for any L by the argument above; but if $f \in \ker(L)$ where $f \neq \mathbf{0}$, then by definition of $\ker(L)$, we have $L[f] = \mathbf{0} = L[\mathbf{0}]$, but $f \neq \mathbf{0}$. This contradicts the definition of one-to-one.

What is the the kernel $\ker(D)$ of the differentiation operator D? By now, we know that $\ker(D)$ is precisely the set of *constant functions*, such as the function $\mathbf{3}$ where $\mathbf{3}(x) = 3$ for all x.

Now we finally address the problem of solving differential equations. The simplest differential equation is the implicit one in the evaluation of an indefinite integral $\int f(x)\,dx$, since this is equivalent to finding a solution $y(x)$ for the differential equation

$$\frac{dy}{dx} = f(x);$$

or with the operator notation, $D[y] = f$. Now if D were invertible, the solution would simply be $y = D^{-1}[f]$, but unfortunately the situation is not as simple, since as we have just seen, $\ker(D) \neq \{\mathbf{0}\}$. So how do we solve this problem? What we usually do is determine a particular function y_p by the techniques of integration, and then write

$$\int f(x)\,dx = y_p + c$$

where c is an "arbitrary constant". The addition of this constants incorporates *all* solutions to the differential equation. In view of the theory of kernels we have developed, this is equivalent to doing $y_p + \mathbf{c}$ for any constant function $\mathbf{c} \in \ker(D)$. Indeed, if y_p is a solution, it makes sense that $y_p + \mathbf{c}$ is also a solution, since

$$D[y_p + \mathbf{c}] = D[y_p] + D[\mathbf{c}] = D[y_p] + \mathbf{0} = D[y_p].$$

But how does this incorporate *all* solutions? Say we want to solve $L[y] = f$ for any linear operator L. Let y_p be a particular solution we found somehow, and let y represent any other solution. By linearity,

$$L[y - y_p] = L[y] - L[y_p] = f - f = \mathbf{0},$$

so $y - y_p \in \ker(L)$, i.e., $y - y_p = k$ for some function $k \in \ker(L)$. Thus

$$y = y_p + k.$$

Hence we have shown that any solution y to the equation $L[y] = f$ can be written as the particular solution y_p plus some member of the kernel, and it follows that all solutions are given by $y = y_p + k$ for $k \in \ker(L)$.

Essentially, this is what the method described is doing. Instead of simply having $\frac{dy}{dx} = f$ though, we have equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f,$$

where $a, b, c \in \mathbb{R}$. It is easy to see that the left-hand side is also a linear operator, since it inherits linearity from the operators $\frac{d^2}{dx^2}$, $\frac{d}{dx}$ and the identity $(I[y] = y)$.

indeed, if we define $L[y] = a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy$, then

$$\begin{aligned} L[f+g] &= a\frac{d^2}{dx^2}(f+g) + b\frac{d}{dx}(f+g) + c(f+g) \\ &= a\frac{d^2f}{dx^2} + b\frac{d^2f}{dx^2} + cf + a\frac{d^2g}{dx^2} + b\frac{d^2g}{dx^2} + cg \\ &= L[f] + L[g], \end{aligned}$$

and

$$L[\alpha\, f] = a\frac{d^2}{dx^2}(\alpha\, f) + b\frac{d}{dx}(\alpha\, f) + c(\alpha\, f) = \alpha\big(a\frac{d^2f}{dx^2} + b\frac{d^2f}{dx^2} + cf\big) = \alpha\, L[f].$$

When defining such operators, we sometimes abuse notation slightly and write $L = a\frac{d^2}{dx^2} + b\frac{d}{dx} + c$ or $L = aD^2 + bD + cI$ instead of $L[y] = a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy$ for all y. Since the given equation is equivalent to $L[y] = f$, we may also write

$$\left(a\frac{d^2}{dx^2} + b\frac{d}{dx} + c\right)[y] = f \qquad \text{or} \qquad (aD^2 + bD + cI)[y] = f.$$

Let's take an example, say,

$$\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = \cos 2x.$$

In this case, we have the operator $L = D^2 - 5D + 6I$. The first thing we need to do is to study this operator L, in particular, we need to find its kernel. In general, operators of the form $aD^2 + bD + cI$ have an exponential function $f(x) = e^{kx}$ in their kernel for some value of k.

Indeed, since $D[f](x) = ke^{kx}$ and $D^2[f](x) = k^2e^{kx}$, then

$$L[f](x) = ak^2e^{kx} + bke^{kx} + ce^{kx} = e^{kx}(ak^2 + bk + c).$$

Since $e^{kx} \neq 0$ for all $x \in \mathbb{R}$ (or $\mathbb{C}$), it follows that $L[f] = \mathbf{0}$ whenever k is a solution to the auxiliary equation $ak^2 + bk + c = 0$. There are some technical details as to why we take different general solutions depending on the multiplicity of k, or whether it is real or complex, but essentially, the first step of the solution process is determining the kernel $\ker(L)$ of the linear operator defined by the left-hand side. This is what the complementary function achieves.

The trial solution part of the method is effectively just a guess for a particular solution (hence its name), y_p. Once a correct solution is found, the general solution is given by $y = y_p + k$, as described in the general framework of linear operators.

And that's why it works.

For the unsatisfied: why the kernel comprises exponentials. If you're bothered by the fact that I didn't explain why the kernel functions look like e^{kx}, then I'll explain it briefly here.

Something cool about operators of the form $aD^2 + bD + cI$ is that we can actually "factorise" them, just like we do with quadratics. Indeed, we know that if the roots of the quadratic $ak^2 + bk + c$ are α and β, then we may write the quadratic as $a(k-\alpha)(k-\beta)$. Amazingly, we get that

$$aD^2 + bD + cI = a(D - \alpha I)(D - \beta I),$$

where for two operators L and M, their product ML means "do L, and then then do M" (just like functional composition, we could also write this as $M \circ L$). For instance,

$$(D - 2I)(D - 3I)[y] = (D - 2I)[y' - 3y]$$
$$= (y' - 3y)' - 2(y' - 3y) = y'' - 5y' + 6y.$$

Indeed, it's straightforward to verify that we can factorise in the general case. Suppose α and β are the roots of $ak^2 + bk + c$. Then

$$\begin{aligned}\big(a(D - \alpha I)(D - \beta I)\big)[y] &= \big(a(D - \alpha I)\big)[y' - \beta y] \\ &= a((y' - \beta y)' - \alpha(y' - \beta y)) \\ &= a(y'' - (\alpha + \beta)y' + \alpha\beta\, y) \\ &= a(y'' + \tfrac{b}{a}y' + \tfrac{c}{a}y) \\ &= ay'' + by' + cy \\ &= (aD^2 + bD + cI)[y].\end{aligned}$$

Ok, so we can factorise these operators. How does it help us? Well, to solve the homogeneous equation $(aD^2 + bD + cI)[y] = \mathbf{0}$, we can assume $a \neq 0$ and divide by a, and factorise the operator as $(D-\alpha I)(D-\beta I)$ (using complex roots if necessary). Thus our goal has now become to solve $(D - \alpha I)(D - \beta I)[y] = \mathbf{0}$.

Now y is a solution to $(D - \alpha I)(D - \beta I)[y]$ if the result of evaluating $(D - \beta I)[y]$ is in the kernel of $(D - \alpha I)$. The kernel of $(D - \alpha I)$ is precisely the set of all functions f which satisfy

$$(D - \alpha I)[f] = \mathbf{0},$$

i.e., the set of solutions to $f' - \alpha f = 0$, or, $f' = \alpha f$. This is a separable first order ODE, whose general solution is $f(x) = c_1 e^{\alpha x}$. Thus y is a solution to $(D-\alpha I)(D-\beta I)[y]$ precisely when it is a function of this kind, i.e., if

$$(D - \beta I)[y] = c_1 e^{\alpha x}$$

for some c_1. Now, what we have here is a linear first order differential equation, since we can write it in the usual notation as

$$y' - \beta y = c_1 e^{\alpha x}.$$

Setting $\mu(x) = e^{-\beta x}$ as our integrating factor, we multiply throughout by $\mu(x)$ to get that this equation is equivalent to

$$e^{-\beta x} y' - \beta e^{-\beta x} y = c_1 e^{(\alpha-\beta)x} \iff \frac{d}{dx}(y e^{-\beta x}) = c_1 e^{(\alpha-\beta)x}.$$

Thus, the general solution is

$$y(x) = e^{\beta x}\Big(c_1 \int e^{(\alpha-\beta)x}\, dx\Big).$$

At this point, we have three cases for the integral, depending on α and β.

(i) $\alpha \neq \beta$, both real roots. In this case, we work out the integral obtaining

$$y(x) = c_1 e^{\beta x}\Big(\frac{e^{(\alpha-\beta)x}}{\alpha - \beta} + c_2\Big) = \frac{c_1}{\alpha - \beta} e^{\alpha x} + c_1 c_2 e^{\beta x} = C_1 e^{\alpha x} + C_2 e^{\beta x},$$

where we can relabel the constants as we did, since given any $C_1, C_2 \in \mathbb{R}$, we can set $c_1 = (\alpha+\beta)C_1$ and $c_2 = C_2/c_1$ in the above. Notice that we are dividing by $\alpha - \beta$, so it is crucial that $\alpha \neq \beta$.

(ii) $\alpha = \beta$, repeated real root. In this case, the integral is simply $\int 1\,dx$, so we have

$$y(x) = c_1 e^{\beta x}(x + c_2) = e^{\beta x}(c_1 x + c_1 c_2) = C_1 e^{\beta x} + C_2 x e^{\beta x},$$

where the relabelling is justified by setting $c_1 = C_2/C_1$ and $c_2 = C_1$.

(iii) $\alpha, \beta = \sigma \pm i\tau$, complex roots. In this case, the integral is $\int e^{2\tau i x}\,dx = \int(\cos(2\tau x) + i\sin(2\tau x))\,dx$, which gives us that

$$\begin{aligned} y(x) &= c_1 e^{(\sigma - \tau i)x}\Big(\frac{\sin(2\tau x)}{2\tau} - i\,\frac{\cos(2\tau x)}{2\tau} + c_2\Big) \\ &= \frac{c_1 e^{\sigma x}}{2\tau}(\cos(\tau x) - i\sin(\tau x))(\sin(2\tau x) - i\,\cos(2\tau x) + 2\tau c_2) \\ &= \frac{c_1 e^{\sigma x}}{2\tau}(2\tau c_2\cos(\tau x) + \sin(\tau x) - (\cos(\tau x) + 2\tau c_2\sin(\tau x))i) \\ &= e^{\sigma x}\Big(\Big(c_1 c_2 - \frac{c_1}{2\tau}i\Big)\cos(\tau x) + \Big(\frac{c_1}{2\tau} - c_1 c_2 i\Big)\sin(\tau x)\Big). \end{aligned}$$

It might not be obvious, but given any constants C_1 and C_2, we can put $c_1 = C_2\tau + C_1\tau i$ and $c_2 = (2C_1C_2 - ({C_1}^2 - {C_2}^2)i)/(2\tau({C_1}^2 + {C_2}^2))$ and the above becomes

$$y(x) = e^{\sigma x}(C_1\cos(\tau x) + C_2\sin(\tau x)).$$

For our situation, want $y(x)$ to be a real-valued function, so we restrict our solutions to when $C_1, C_2 \in \mathbb{R}$ (but in truth, this function is a solution to the homogeneous equation for any $C_1, C_2 \in \mathbb{C}$).

And this is where the general solutions come from.