Differential Equations, Part V: Further Topics

Jay Havaldar

0.0.1 Table of Contents

- Series Solutions for Second Order Linear Homogeneous Equations
- Cauchy-Euler Equations
- The Laplace Transform
- The Matrix Exponential
- Nonlinear Systems

In this chapter, we'll take a look at various additional topics which are generally covered in a differential equations course, in no particular order.

0.1 Series Solutions for Second Order Linear Homogeneous Equations

We look at the equations of the form:

$$P(x)y'' + Q(x)y' + R(x)y = 0$$

Where P, Q, R are functions have convergent power series around a point x_0. For simplicity sake, P, Q, R are generally restricted to polynomials in an intro differential equations course, and we assume that there is no common divisor for P, Q, R.

Definition: x_0 is called an **ordinary point** if $P(x_0) \neq 0$.

From continuity, we know that there is a small interval where $x_0 \neq 0$. In that interval, we divide through to get the differential equation:

$$y'' + p(x)y' + q(x)y = 0$$

And the coefficients are then continuous. Therefore, there exists a unique solution for given initial conditions in the interval.

Definition: x_0 is called an **singular point** if $P(x_0) = 0$.

At a singular point, the coefficients p, q are not continuous, and solutions are not guaranteed.

We want to find solutions of the form:

$$y = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

It may be helpful to review some basic facts about power series. We assume that there is some **radius of convergence** $\rho > 0$ so that if $|x - x_0| < rho$, the series converges at x. In the interval of convergence, we can deal with power series term by term -- treating them essentially like polynomials.

Example Let's try to solve the differential equation:

$$y'' + y = 0$$

At the point $t = 0$, which we know has solutions $\cos t, \sin t$. Assume our solution is of the form $y = \sum_{n=0}^{\infty} a_n x^n$, and substitute into the equation. Note that power series can be differentiated term by term:

$$y'' + y = 0$$

We can take derivatives to get:

$$y'' = \sum_{n=0}^{\infty} n(n-1)a_n x^{n-1}$$

And renumbering the indices, we can rewrite:

$$y'' = \sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2} x^n$$

Adding together the series for $y'' + y$, we arrive at the following equation:

$$y'' + y = \sum_{n=0}^{\infty} \left[(n+1)(n+2)a_{n+2} + a_n\right] x^n$$

It is not hard to see that for this power series to be 0, each coefficient must be 0 (to prove this, set $x = 0$ and thus the first coefficient is zero, take the derivative, and repeat this process). Therefore, we have the **recurrence relations** given by:

$$a_{n+2} = \frac{-a_n}{(n+1)(n+2)}$$

We can solve for a_0, a_1 by giving initial conditions for y, y'. Now, let's for example, look at the first few even terms:

$$\begin{aligned} a_2 &= \frac{(-1)a_0}{2 \cdot 1} \\ a_4 &= \frac{(-1)^2 a_0}{4 \cdot 3 \cdot 2 \cdot 1} \\ a_6 &= \frac{(-1)^3 a_0}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} \end{aligned}$$

In general, this pattern gives:

$$a_{2n} = a_0 \frac{(-1)^{2n}}{(2n)!}$$

But these are exactly the coefficients for the power series of $a_0 \cos x$! Looking at the odd coefficients, we similarly get the coefficients for the power series of $a_1 \sin x$. So we can indeed write our series solution:

$$y = a_0 \cos x + a_1 \sin x$$

Just as we expected. Note that this sort of thing doesn't always happen. However, in many cases it may be useful to compute quite a few coefficients to get a reasonable approximation of solutions. In the non-homogeneous case, this is essentially the technique used in Fourier series, but that topic is generally covered in a partial differential equations course.

Generalizations We considered the problem of equations of the form:

$$P(x)y'' + Q(x)y' + R(x)y = 0$$

Where P, Q, R are polynomials, $P(x_0) \neq 0$, which we defined as an ordinary point. We wish to find more general conditions under which a solution exists.

Assume that we have a solution $y = f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ which converges for $|x - x_0| < \rho$. Taking the mth derivative $f^{(m)}$ and setting $x = x_0$, we get the equation:

$$f^{(m)}(x_0) = m! a_m$$

So the problem of finding the coefficients of a series solution reduces to the problem of finding $f^{(m)}$ for some solution $y = f(x)$ to the differential equation. But we know that:

$$\begin{aligned} P(x)f'' + Q(x)f' + R(x)f &= 0 \\ f'' &= -p(x)f' - q(x)f \end{aligned}$$

Where as before, $p = Q/P, q = R/P$. In particular at $x = x_0$, we have:

$$f''(x_0) = -p(x_0)f'(x_0) - q(x_0)f(x_0)$$

And therefore, if we are given initial conditions $y(x_0), y'(x_0)$, then we can solve for a_2 in our power series.

Similarly, we can differentiate the above expression to obtain:

$$f''' = -\left[q'f + (q + p')f' + pf''\right]$$

And thus, if p, q can be differentiated, then we can solve for the third coefficient a_3 since all the quantities above are known at $x = x_0$. This motivates the condition we need to guarantee the existence of the power series solution. We don't just require that p, q can be differentiated infinitely many times, we also require that their power series converge.

Definition: A function $f(x)$ is **analytic** at $f(x_0)$ if it has a Taylor series expansion which converges to $f(x)$ in some interval around x_0.

This motivates a stronger definition for ordinary points.

Definition: A point $x = x_0$ of a differential equation is **ordinary** if $p(x), q(x)$ are analytic at x_0; otherwise, x_0 is said to be a **singular** point.

Finally, the following theorem connects our new definition of an ordinary point to the existence of solutions.

Theorem

If x_0 is an ordinary point of the above differential equation:

$$P(x)y'' + Q(x)y' + R(x)y = 0$$

then, the general solution of the differential equation can be written:

$$y = \sum_{n=0}^{\infty} a_n(x - x_0)^n = a_0 y_1(x) + a_1 y_2(x)$$

Where y_1, y_2 are a fundamental set of solutions which are both analytic at x_0. The radius of convergence of y_1, y_2 is at least as large as the minimum radius at which the series p, q both converge.

This theorem tells us that at an ordinary point (one where the coefficients are analytic), two series solutions exist, which form a set fundamental solutions. Furthermore, as long as p, q both converge at a point, y_1, y_2 also converge at that point.

In general, once we have found a bound on our radius of convergence, the idea is to substitute $y = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ into the differential equation and see what kinds of conditions we can place on the coefficients.

0.2 Cauchy-Euler Equations

A Cauchy-Euler equation is of the form:

$$a_n x^n y^{(n)} + a_{n-1} x^{n-1} y^{(n-1)} + \dots a_0 y = 0$$

We will consider constant coefficient equations at first. Take a second order equation of the form:

$$x^2 y'' + axy' + by = 0$$

We assume that $x > 0$. Our guess of a solution looks like $y = x^r$ for some integer r, since in each term, we take n derivatives and multiply by x^n. In the case $x < 0$ we would similarly substitute $-x^r$, or in general substitute $|x|^r$ it In particular, we get:

$$\begin{aligned} x^2y'' + axy' + by &= 0 \\ r(r-1)x^r + arx^r + bx^r &= 0 \\ [r(r-1) + ar + b]\, x^r &= 0 \end{aligned}$$

Assuming $x \neq 0$, we now are left with a characteristic equation for r. This all should look very familiar to you, as it's an almost identical process to the constant coefficient homogeneous equations we studied earlier, where we subtituted in a trial solution $y = e^{rx}$. Since $e^{x^r} = e^{rx}$, this clues us into a substitution we can make.

In the above equation, let $x = e^t$, which we can do so long as $x > 0$; then the equation becomes a differential equation in terms of t. With some use of the chain rule, we indeed get:

$$y'' + ay' + by = 0$$

Which is a constant coefficient equation with the *same characteristic equation* which we know how to solve. Now we can look at solutions of the Cauchy Euler equation by looking at the roots of the characteristic equation.

- For a real repeated root, $y = e^{rt}, te^{rt}, t^2e^{rt}$, etc. Substituting in $x = e^t$, we have:

$$y = x^r, (\ln x)x^r, (\ln x)^2x^r, \ldots$$

- For a complex root, $y = e^{at}\cos bt, e^{at}\sin bt$, etc. Substituting in $x = e^t$, we have:

$$y = x^r \cos(b \ln x), x^r \sin(b \ln x), x^r(\ln x)\cos(b \ln x), x^r(\ln x)\sin(b \ln x), \ldots$$

And you can check that these solutions are linearly independent. To obtain the result for $x < 0$, we can simply do a coordinate transformation $\tilde{x} = -x$ and reduce to the earlier case. You can check that in the above formulas, we need only replace x with $|x|$ to get a solution for all cases.

0.3 The Laplace Transform

The Laplace Transform is a method of transforming a problem in one domain to a problem in another. Once we have solved the transformed version of a differential equation, we can use the inverse transformation to get a set of solutions. We define:

Definition: Let $f(t)$ be defined for $t \geq 0$. Then we define the **Laplace Transform** $F(s)$, or $\mathcal{L}\{f(t)\}$, to be the improper integral:

$$\mathcal{L}\{f(t)\} = F(s) = \int_0^\infty e^{-st} f(t) dt$$

The following theorem gives us conditions for which the Laplace transform is well-defined.

Theorem

> Let f be piecewise continuous on $0 \leq t \leq A$, for any A. Furthermore, suppose there exists a positive value M so that for all $t \geq M$, we have: $|f(t)| \leq Ke^{at}$, for some $K > 0$ and a real. Then, the Laplace Transform $\mathcal{L}\{f(t)\}$ exists for all $s > a$.

This theorem essentially tells that for some large enough value, we need $f(t)$ to not overpower e^{at}, which is a reasonable condition to have. The proof is using basic calculus.

One important property of the Laplacian is that it is **linear**, meaning $\mathcal{L}[af + bg] = a\mathcal{L}[f] + b\mathcal{L}[g]$, for any constants a, b and functions f, g. This follows from the fact that integration is linear. Furthermore, we assume the inverse of the Laplace transform is well-defined, though we will not prove it.

0.3.1 Table of Laplace Transforms

The following table gives a few of the useful Laplace Transforms that we can use to solve differential equations.

f(t)	**F(s)** $= \mathcal{L}[f(t)] = F(s)$
$e^{at}f(t)$	$F(s-a)$
1	$\frac{1}{s}$
e^{at}	$\frac{1}{s-a}$
t^n	$\frac{n!}{s^{n+1}}$
$\sin bt$	$\frac{b}{s^2+b^2}$
$\cos bt$	$\frac{s}{s^2+b^2}$
$t \sin bt$	$\frac{2b}{(s^2+b^2)^2}$
$t \cos bt$	$\frac{s^2-b^2}{(s^2+b^2)^2}$
$f'(t)$	$sF(s) - f(0)$
$f''(t)$	$s^2F(s) - sf(0) - f'(0)$

A key point here is that if you look at derivatives, the Laplace transform turns a differential equation into an algebraic equation, which are in general far easier to solve.

Example Let's solve our prototypical example:

$$y'' + y = 0$$

With initial conditions $y(0) = 1, y'(0) = 0$. We already know the solution is $cosx$. Taking the Laplace transform of both sides using our table, we get:

$$s^2F(s) - sy(0) - y'(0) + F(s) = 0(s^2+1)F(s) = sF(s) = \frac{s}{1+s^2}$$

And indeed, $F(s)$ is the Laplace transform of $\cos x$, as we expected.

This example may seem fairly trivial; Laplace transform only truly becomes useful when dealing with non-continuous quantites. We take a look at two important such cases.

0.3.2 Heaviside Function

We define the **Heaviside** or **step** function:

$$u_c(t) = \begin{cases} 0 & t < c \\ 1 & t \geq c \end{cases}$$

We can use this to create a ``step up - step down'' function. For example, if we wanted a function which was zero everywhere except in the region $a < t < b$, we can simply use the difference $u_a(t) - u_b(t)$.

We determine the Laplace transform of the Heaviside function by integration.

$$\begin{aligned} \mathcal{L}[u_c(t)] &= \int_0^\infty e^{-st} u_c(t) dt \\ &= \int_c^\infty e^{-st} dt \\ &= \frac{e^{-cs}}{s} \end{aligned}$$

One application of the Heaviside function is to translate a function. Look at the product $u_c(t)f(t-c)$:

$$u_c(t)f(t-c) = \begin{cases} 0 & t < c \\ f(t-c) & t \geq c \end{cases}$$

So we have essentially translated f over to the right by c. You can check that the Laplace transform gives:

$$\begin{aligned} \mathcal{L}[u_c(t)f(t-c)] &= \int_0^\infty e^{-st} f(t-c) dt \\ &= e^{-cs} F(s) \end{aligned}$$

So translating a function to the right by c results in multiplying its Laplace transform by e^{-cs}.

These types of shapes occur often in signal processing, which is when the Laplace transform is the most useful.

0.3.3 Dirac Delta Function

We define the **Dirac delta function** $\delta(t)$:

$$\int_{-\infty}^\infty f(t)\delta(t) = f(0)$$

We can interpret the Dirac delta function (which is not technically a function) as a function which is zero everywhere except the origin, where it is infinite. We can construct it as the limit of a step-up, step-down function:

$$\delta(t-c) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon}(u_{c-\epsilon}(t) - u_{c+\epsilon}(t))$$

Now using this definition, we can integrate and get the Laplace transform of the Dirac delta function:

$$\begin{aligned}\mathcal{L}[\delta(t-c)] &= \int_0^\infty e^{-st}\delta(t-c)dt \\ &= e^{-cs}\end{aligned}$$

And indeed, we have $\mathcal{L}[\delta(t)] = 1$.

For nonhomogeneous linear equations where the right hand side is discontinuous, we can often use the Laplace transform to make the problem easier to solve, though generally we may need to use some algebraic tricks such as partial fractions and completing the square.

0.4 The Matrix Exponential

Recall that earlier when we were dealing with the matrix system:

$$x' = Ax$$

And where A was diagonalizable, we were able to uncouple our equations and get n separate linear equations. However, in that section we mentioned there was an easier method, which involves something called the matrix exponential.

Theorem

Suppose we have the equation:

$$x' = Ax$$

With A as a constant matrix, and the initial condition given by $x(0) = x_0$. Then the solution is given by:

$$x = e^{At}x_0$$

Where e^{At} is the matrix exponential.

This can be seen by analogy with the one dimensional case. Suppose we had:

$$x' = ax$$

For some constant a, and $x(0) = x_0$. Then the solution, by integration, is $x = Ce^{at}$ for some constant C. Using the initial condition gives $C = x_0$.

0.4.1 Defining the Matrix Exponential

We use the definition of the one dimensional exponential:

$$e^{at} = 1 + at + (at)^2/2 \cdots = \sum_{k=0}^{\infty} \frac{(at)^k}{k!}$$

We can define the matrix exponential in a similar way:

$$e^{At} = I + At + (At)^2/2 \cdots = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$$

The matrix exponential enjoys many properties which we associate with the exponential. We will list a few of the useful ones here:

- $e^0 = I$.
- $e^{aX} e^{bX} = e^{a+b} X$
- $\frac{d}{dt} e^{At} = A e^{At} X$
- If A is diagonalizable, i.e. $A = SDS^{-1}$, then $e^A = S e^D S^{-1}$.
- $\det(e^A) = e^{\text{tr}(A)}$

Where tr refers to the trace of A. The first three properties can be proved using direct computation. Note that the first property tells us that when we constructed a fundamental matrix so that $X(0) = I$, we indeed had a matrix which was a matrix exponential. The final property is called **Jacobi's formula** and requires an involved proof we won't get into here.

With some knowledge of linear algebra, we can generalize the above property for diagonalizable matrices to all matrices using the Jordan normal form. The exercise of computing the matrix exponential of a Jordan normal form is left to the reader, but the key point is to compute each block separately and break each block into the sum of a diagonal matrix and another matrix.

We can use the matrix exponential to generalize the method of integrating factors.

0.4.2 The Non-homogeneous Case

Suppose instead we have a system which looks like:

$$x' = Ax + b$$

And again, A is constant. We attempt to multiply both sides by e^{-At} (an integrating factor) to obtain:

$$e^{-At} x' - A e^{-At} x = e^{-At} b$$

The key here is that the matrix exponential works like the regular exponential, and so the left hand side becomes an instance of the product rule:

$$\frac{d}{dt}(e^{-At}x) = e^{-At}b$$

Integrating, we have:

$$x = e^{At}\int e^{-At}b$$

Thus, the matrix exponential, if it can be easily computed, can be used to efficiently solve homogeneous or non-homogeneous matrix systems with constant coefficients.

0.5 Systems of Nonlinear Differential Equations

We begin with the one dimensional case. Suppose we are working with a non-linear differential equation of the form:

$$x' = f(x)$$

We say that x_0 is a **fixed point** or **equilibrium point** if $f(x_0) = 0$. The reason behind this term is because if we set $x(0) = x_0$, then $x'(0) = 0$; indeed the derivative never changes, and x remains constant forever.

We next consider what happens under small perturbations; will x eventually return to its stable point or will it shoot off to infinity? We let $x = x_0 + \epsilon(t)$, where $\epsilon(t)$ starts out small. Note that $x' = \epsilon'$, and $f(x_0) = 0$, so we can use Taylor series to get the first few terms of ϵ':

$$x' = \epsilon' \approx \epsilon f'(x_0) + \frac{\epsilon^2}{2}f''(x_0)$$

We ignore the second term, to get:

$$\epsilon' = \epsilon f'(x_0)$$
$$\epsilon \approx e^{f'(x_0)}$$

So we can summarize the above discussion as follows: - If $f' > 0$ at an equilbirium point, then a small perturbation will blow up exponentially, and we say such a point is **unstable**.

- If $f' < 0$ at an equilbirium point, then a small perturbation will decay exponentially, and we say such a point is **stable**.
- If $f' = 0$ at an equilbirium point, then a small perturbation will lead to neither exponential decay nor exponential growth, and we say such a point is **marginally stable**.

We can indeed try to find an approximate solution to such an equation near an equilbrium point by calculating the first few terms of the Taylor series for $f(x)$.

0.5.1 Two Dimensions

We move on to the two dimensional case, where we have:

$$x' = F(x, y)$$
$$y' = G(x, y)$$

Here, we have both x and y are functions of t, but the functions F, G only depend on x, y. We call such a system an **autonomous system**.

Again, we define an **equilibrium point** as $F(x, y) = G(x, y) = 0$.

To get a qualitative idea of what the solutions look like, we can plot x, y for various values and draw vectors indicating the direction of change at each point. This is called a **phase plane** diagram, and is particularly useful in physics, where we can plot for example position and velocity.

To find an approximation F, G near an equilibrium point, we use the nearest linear approximation to the pair of functions at that point. In one dimension, this would be simply the derivative. In two dimensions, we use the **Jacobian**:

$$A = \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{bmatrix}$$

We evaluate this matrix at a particular equilibrium point x_0, y_0. Using something like a Taylor series, we get the following differential equation

$$\begin{bmatrix} x \\ y \end{bmatrix}' = A(x_0, y_0) \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$

0.5.2 Types of Equilibrium Points

As in the one dimensional case, the behavior of our system near an equilibrium point depends on the first derivative. In particular, we look at the eigenvalues of the Jacobian matrix A. We have the following definitions:

Eigenvalues	Type of Equilbrium Point
Real, negative	Stable node
Real, positive	Unstable node
Opposite sign	Saddle
Complex with negative real part	Stable spiral
Complex with positive real part	Unstable spiral
Complex with zero real part	Stable center

These definitions come from the types of shapes the system makes in the phase plane.

For example, if the eigenvalues are both pure imaginary, then we know from Chapter 4 what the fundamental solutions to this system look like. Let the eigenvalues be called $\pm\mu$. Then we have two fundamental solutions:

$$x_1 = v_1 \cos \mu t$$
$$x_2 = v_1 \sin \mu t$$

For some vector v_1 which is the imaginary part of an eigenvector. This means every solution $\mathbf{x} = (x, y)$ looks like:

$$\mathbf{x} = (a \cos \mu t + b \sin \mu t) v_1$$

And as we plot this in the phase plane, we get nothing more than an ellipse.