A Painless Introduction to Differential Equations I

Anqiao Ouyang

December 6, 2023

1 Introduction

The title of this article is inspired by many others I've read, such as *xxx Without the Agonizing Pain*. I was initially tempted to force it or even go with something like "Differential Equations Without the Agonizing Pain," but I restrained myself.

In Part I, we will explore fundamental concepts, common methods for solving first-order ordinary differential equations, techniques for solving constant coefficient linear differential equations, and an introduction to homogeneous linear systems. Additionally, we will cover numerical methods commonly used in fields such as computer science. Building on the content from my previous blog, *Introduction to Calculus V2*, this article organizes the basic concepts, solutions, and simple applications of differential equations in the form of an essay. This article assumes that readers have the foundational knowledge from previous posts on calculus basics + linear algebra P1.

Complex analysis is recommended, as it's hard to avoid if you plan to study further, but it's not a prerequisite for this article.

1.1 Scope of Content

Since there is quite a lot of material even for the basic concepts, I've decided to write differential equations as a sub-series within the calculus blog. The topics covered in this Part I include:

- Fundamental concepts of differential equations
- Common methods for solving first-order ordinary differential equations
- Methods for solving constant coefficient linear differential equations that can be extended to higher orders
- An introduction to homogeneous linear systems of ordinary differential equations
- Numerical methods primarily applied in fields such as computer science

Some conclusions that I feel need further explanation will include corollaries and proofs. For more tedious topics that require extensive listing (such as the properties of unilateral Laplace transforms), I will provide links to complete and formal listings.

2 Basic Concepts

2.1 Differential Equations and Their Significance

A differential equation is an equation that involves an unknown function and its derivatives. It describes the relationship between a function and its derivatives, and solving a differential equation essentially involves finding a function that satisfies the given equation for further analysis.

The order of a differential equation is indicated by the highest derivative present in the equation. Differential equations are divided into Ordinary Differential Equations (ODE) and Partial Differential Equations (PDE). An ODE is a differential equation in which the unknown function is a function of a single independent variable. For an unknown function y with the independent variable x, an n-th order ODE takes the following form, where Leibniz notation is used to avoid ambiguity.

$$f(x, \frac{\mathrm{d}y}{\mathrm{d}x}, \cdots, \frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}}, \frac{\mathrm{d}y}{\mathrm{d}x}, y) = 0$$

Correspondingly, a PDE involves an unknown function that depends on multiple variables. Defining $\mathbf{x} = [x_1, x_2, \cdots, x_m]$, the function is represented as $u(\mathbf{x})$. An n-th order PDE takes the following form

$$f(\mathbf{x}, u, \cdots, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \cdots, \frac{\partial u}{\partial x_n}, \frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \cdots, \frac{\partial^n u}{\partial x_1, \partial x_2, \cdots, x_n}) = 0$$

2.1.1 Initial Value Problems Boundary Value Problems

To determine a unique solution to a differential equation, we need an initial condition that provides the value of the function at a specific point. This concept is familiar even to middle school students, as it's similar to being given a specific point and then finding the equation of the function. Such problems are known as initial value problems. An initial value problem for a first-order ordinary differential equation can be generally represented as

$$y'(x) = f(x, y(x)),\ y(x_0) = y_0$$

Here, $f(x, y(x))$ describes the rate of change of the dependent variable $y(x)$ with respect to the independent variable x, i.e., it represents the derivative of $y(x)$ at each point x. Our goal is to find an appropriate $y(x)$ that satisfies the equation.

In contrast to initial value problems, boundary value problems are defined by conditions on the domain of the independent variable. A boundary value problem for a first-order ordinary differential equation can be generally represented as

$$y'(x) = f(x, y(x)),\ a \leq x \leq b,\ y(a) = \alpha,\ y(b) = \beta$$

2.1.2 Solutions of Differential Equations

In the context of differential equations, we use the terms general solution and particular solution to describe different types of solutions. The general solution is a set that contains all possible

solutions of the differential equation. In simple terms, it's a universal form that does not include any specific initial conditions, representing all potential solutions of the differential equation. It typically takes the form $y(x) = \varphi(x, C)$, where C is an undetermined constant. By assigning specific values to C, we can obtain a particular solution.

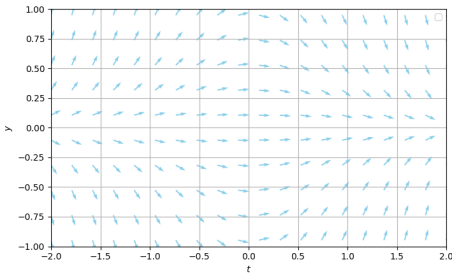
Correspondingly, a particular solution is a solution that satisfies a specific initial or boundary condition. If we have a specific pair of values x_0 and y_0, such that $y(x_0) = y_0$, then $y(x)$ is the particular solution of the differential equation at x_0.

If $y_1(x), y_2(x), \cdots, y_n(x)$ are solutions of the differential equation and are linearly independent, they are referred to as linearly independent solutions. We can use the Wronskian determinant to determine whether a set of solutions is linearly independent, defined as

$$W(y_1, y_2, \ldots, y_k) = \begin{vmatrix} y_1 & y_2 & \cdots & y_k \\ y_1' & y_2' & \cdots & y_k' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(k-1)} & y_2^{(k-1)} & \cdots & y_k^{(k-1)} \end{vmatrix}$$

If the Wonskian determinant is $x_0 \neq 0$ at some point, then the solutions are linearly independent at that point. If $W(y_1, y_2, \ldots, y_k) \neq 0$ for all x, then the set of solutions is linearly independent over the entire domain.

2.2 Direction Fields



Here, we need to introduce a concept called **direction fields**, a graphical method for representing differential equations. It shows the slope direction at each point in the plane, which is determined by the differential equation.

The figure illustrates the overall behavior of the solution of a simple first-order ordinary differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = -2ty$ in the plane. The arrows represent the direction of change of the solution at each point, and their length indicates the rate of change at that point.

By drawing a direction field, we can gain some qualitative characteristics of the solution of the differential equation across the plane. This graphical method provides an intuitive way to understand the behavior of differential equations. At this point, you should have a basic understanding of this concept.

2.3 Exact Differentials

As mentioned earlier, the standard form of a first-order differential equation is (though not all equations follow this form):

$$y' = f(x, y)$$

If there exists a function $\phi(x, y)$ that satisfies $\frac{\partial \phi}{\partial x} = M(x, y), \frac{\partial \phi}{\partial y} = N(x, y)$ (where M and N are continuous within the domain), such that

$$d\phi = M(x,y)\mathrm{d}x + N(x,y)\mathrm{d}y$$

then this differential form is called an **exact differential**. Here, ϕ is called the **potential function**, and its existence means that the equation is an exact differential equation (or a **proper equation**), and ϕ acts as the integrating factor of the differential form.

This equation implies that the order of mixed partial derivatives can be interchanged. Referring to the supplemental section of *Basic Calculus V2*, Schwarz's theorem states that if the mixed partial derivatives of a function f exist and are continuous at a point, then the order of these mixed partial derivatives can be exchanged. For example, for a bivariate function $f(x_1, x_2)$, this can be expressed as

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

When we are concerned with the existence of the potential function ϕ, we can use the conditions of Schwarz's theorem to ensure the existence of ϕ.

2.4 Linear Differential Equations

Clearly, a linear differential equation is a differential equation that involves linear combinations of first or higher derivatives and the unknown function itself. An n-th order linear differential equation has the form:

$$a_n(x)\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1}(x)\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a_0(x)y = f(x)$$

where $a_0(x), a_1(x), \cdots, a_n(x)$ are coefficient functions. If all coefficient functions are independent of x, meaning each coefficient function is a constant, then the equation is a constant coefficient linear differential equation. If at least one coefficient function depends on x, then it is a variable coefficient linear differential equation. Different coefficient functions lead to different differential equations, so by adjusting these coefficients, we can modify the weights of the various derivative orders in the equation, thus changing the properties of the equation.

In the following chapters, we will discuss some mathematical techniques for solving this class of differential equations.

3 First-Order Ordinary Differential Equation Solutions

In the basic concepts chapter, we have gained a fundamental understanding of differential equations. Before expanding to more general differential equations, let's start with some appetizers by exploring solutions for first-order ODEs.

3.1 Variable Separation Equation

A class of differential equations that can be solved using the method of variable separation (note that this is not limited to ordinary and partial differential equations) is called a variable separation equation, generally in the form:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = g(x)h(y)$$

This means it can be expressed as the product of two variables, which transforms the equation into:

$$\frac{1}{h(y)} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = g(x)$$
$$\int \frac{1}{h(y)}\,\mathrm{d}y = \int g(x)\,\mathrm{d}x + C$$

where C is the constant of integration. As we previously mentioned, integration is the inverse operation of differentiation, so we simply need to apply integration to solve for this function. The method is straightforward, and let's consider a simple example of a differential equation:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x \cdot y^2$$

By separating and integrating, we can solve for the desired y:

$$\int \frac{1}{y^2}\,\mathrm{d}y = \int x\,\mathrm{d}x + C$$
$$-\frac{1}{y} = \frac{x^2}{2} + C$$
$$y = -\frac{1}{\frac{x^2}{2} + C}$$

3.2 Bernoulli's Equation

The equations we encounter do not necessarily have to be linear, but this makes them more challenging to handle. Many of the methods in subsequent chapters often require the equations to be linear. Therefore, we first consider how to transform nonlinear differential equations into linear ones. Starting from the standard form of the first-order linear differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)$, we add something to transform it into:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)y^\alpha$$

The unfriendly-looking y^α represents the nonlinear term of the equation, and this form is called a Bernoulli equation. We can use an appropriate variable substitution to make it linear. First, we perform a simple transformation on the equation:

$$y^{-\alpha} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} + P(x)z = Q(x)$$
$$\frac{1}{1-\alpha} \cdot \frac{\mathrm{d}z}{\mathrm{d}x} + P(x)z = Q(x)$$

Next, we introduce a new variable $z = y^{1-\alpha}$, which leads us to:

$$\frac{\mathrm{d}z}{\mathrm{d}x} = (1-\alpha)y^{-\alpha}\frac{\mathrm{d}y}{\mathrm{d}x}$$

Substituting gives us:

$$(1-\alpha)y^{-\alpha}\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)(1-\alpha)z = Q(x)(1-\alpha)$$

With this simple manipulation, it becomes a first-order linear differential equation, and we can then use standard integrating factors or other methods to solve it.

3.3 Integrating Factor Method

The integrating factor method is also commonly used to solve first-order ordinary differential equations. First, we consider the linear equation (this case is often simpler), emphasizing the following form:

$$\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y = Q(x)$$

We need to find the integrating factor $\mu(x)$ to convert it into an exact differential equation, making it easier for us to integrate. This means that the equation can be written in the form:

$$\mu(x)\left(\frac{\mathrm{d}y}{\mathrm{d}x} + P(x)y\right) = \mu(x)Q(x)$$

Regarding how to choose this integrating factor, we find that μ needs to satisfy the following equation, leading us to:

$$\begin{aligned}\mu(x)P(x) &= \frac{\mathrm{d}\mu(x)}{\mathrm{d}x} \\ \mu(x) &= e^{\int P(x)\mathrm{d}x}\end{aligned}$$

This is why we often choose $\mu(x) = e^{\int P(x)\mathrm{d}x}$ as the integrating factor in the integrating factor method; it ensures that our equation can be expressed as an exact differential form after multiplying by the integrating factor.

We simply substitute this expression in, apply the chain rule (second line), integrate simultaneously (third line), and use some basic properties of differentiation and integration to derive the final result (lines four to five):

$$
\begin{aligned}
e^{\int P(x)\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}+P(x)y\right) &= e^{\int P(x)\mathrm{d}x}Q(x)\\
\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\int P(x)\mathrm{d}x}P(x)y\right) &= e^{\int P(x)\mathrm{d}x}Q(x)\\
\int \frac{\mathrm{d}}{\mathrm{d}x}\left(e^{\int P(x)\mathrm{d}x}P(x)y\right) &= \int e^{\int P(x)\mathrm{d}x}Q(x)\\
e^{\int P(x)\mathrm{d}x}y &= \int e^{\int P(x)\mathrm{d}x}Q(x)\mathrm{d}x+C\\
y &= e^{-\int P(x)\mathrm{d}x}\left(\int e^{\int P(x)\mathrm{d}x}Q(x)\mathrm{d}x+C\right)
\end{aligned}
$$

Now we have derived a solution that also applies to non-homogeneous first-order linear differential equations. However, if it is homogeneous, meaning $Q(x) = 0$, substituting gives us a solution applicable only to homogeneous first-order linear differential equations:

$$y = Ce^{-\int P(x)\mathrm{d}x}$$

¡br¿

More generally, for first-order homogeneous nonlinear differential equations, we can often consider using the integrating factor method. Based on the definition of the linear differential equation form from the previous chapter, we can rewrite it in a form that may be more familiar to most as a first-order differential equation:

$$M(x,y)\mathrm{d}x + N(x,y)\mathrm{d}y = 0$$

First, we need to check whether it is an exact differential equation. If it is not, we need to find the integrating factor $\mu(x,y)$ to convert it into an exact differential equation, enabling easier integration, which can be expressed in the form:

$$\mu M(x,y)\mathrm{d}x + \mu N(x,y)\mathrm{d}y = 0$$

For an inexact first-order differential equation, we can find the integrating factor μ that makes it exact using this formula:

$$\mu = e^{\int \frac{My-N_x}{M}\mathrm{d}y}$$

or

$$\mu = e^{\int \frac{N_y-M_x}{M}\mathrm{d}x}$$

3.3.1 Variation of Parameters

Based on our derivation of solutions for both non-homogeneous and homogeneous first-order linear differential equations, we can introduce another method for solving. By introducing an unknown function $u(x)$ to replace the integration constant C based on the homogeneous solution $y = Ce^{-\int P(x)\mathrm{d}x}$, we can substitute back into the first-order linear differential equation to solve for u:

$$
\begin{aligned}
u'e^{-\int P(x)\mathrm{d}x} - uP(x)e^{-\int P(x)\mathrm{d}x} + P(x)ue^{-\int P(x)\mathrm{d}x} &= Q(x) \\
u'e^{-\int P(x)\mathrm{d}x} &= Q(x) \\
ue^{-\int P(x)\mathrm{d}x} &= \int Q(x)\mathrm{d}x + C_1 \\
u &= e^{-\int P(x)\mathrm{d}x}\left(\int Q(x)\mathrm{d}x + C_1\right)
\end{aligned}
$$

Then we substitute this result back into $y = Ce^{-\int P(x)\mathrm{d}x}$ and we get the general solution of the first-order nonhomogeneous linear equation

$$
y = e^{-\int P(x)\mathrm{d}x}\left(\int Q(x)\mathrm{d}x + C_1\right)e^{-\int P(x)\mathrm{d}x}
$$

The advantage of this method is that, by introducing the undetermined function, we can handle the constant term more flexibly (because C_1 includes both the integral constant C in the original homogeneous solution and the integral constant introduced in the process of solving the non-homogeneous equation), thereby making it more convenient to solve the non-homogeneous linear differential equation.

3.4 Problems

1. $\frac{\mathrm{d}y}{\mathrm{d}x} = 2x$ *Tip: Separate the variables and integrate both sides to solve the differential equation.*
2. $\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = e^{-x}$ *Tip: Rewrite the equation in standard form and find an appropriate integrating factor, then express the left side as a differential and integrate both sides of the equation.*
3. $\frac{\mathrm{d}y}{\mathrm{d}x} + (x - 2y) = 0$ *Tip: Check if it is an exact differential equation; if not, find the integrating factor μ and convert it into an exact differential form.*

4 Constant Coefficient Linear Differential Equations

Constant coefficient linear differential equations refer to equations in which all the coefficients are constants, differing from the definitions in the previous chapter. An n-th order constant coefficient linear differential equation has the form:

$$
a_n\frac{\mathrm{d}^n y}{dx^n} + a_{n-1}\frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}} + \cdots + a_1\frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = f(x)
$$

The content of this chapter includes, but is not limited to, the methods for solving higher-order constant coefficient linear differential equations.

4.1 Characteristic Equation Method

First, we need to consider the characteristic equation method for solving the common constant coefficient linear homogeneous differential equations. This method involves transforming the differential equation into an algebraic equation known as the characteristic equation. A homogeneous constant coefficient linear differential equation has the form

$$a_n \frac{\mathrm{d}^n y}{dx^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = 0$$

For the variable r, the derivative of the exponential function e^{rx} is a constant multiple of itself, specifically $\frac{\mathrm{d}}{\mathrm{d}x} e^{rx} = re^{rx}$. Due to this favorable differential property, we can make a reasonable assumption that the solution y has an exponential form $y(x) = e^{rx}$. Substituting this back into the original equation, we arrive at what we call the characteristic equation

$$a_n r^n e^{rx} + a_{n-1} r^{n-1} e^{rx} + \cdots + a_1 r e^{rx} + a_0 e^{rx} = 0$$

Once we find the solutions to the characteristic equation, $r_1, r_2, \cdots, r_n$, and substitute them back, the general solution of the differential equation can be constructed in the form

$$y = C_1 e^{r_1 x} + C_2 e^{r_2 x} + \cdots + C_n e^{r_n x}$$

4.1.1 Discriminant of the Second-Order Characteristic Equation

Let's cover some middle school content (to prevent some of those who forget it after leaving middle school from struggling).

When we apply the characteristic equation method to a second-order constant coefficient linear homogeneous differential equation, we substitute it as follows:

$$ar^2 + br + c = 0$$

This equation is clearly a quadratic equation, and back in middle school, we learned about the discriminant $\Delta = b^2 - 4ac$, which helps us determine whether the equation has real roots. As far as I know, middle school textbooks discuss the following cases: + If $\Delta > 0$, the equation has two distinct real roots: $r = \frac{-b \pm \sqrt{\Delta}}{2a}$ + If $\Delta = 0$, the equation has one real root: $r = \frac{-b}{2a}$ + If $\Delta < 0$, the equation has no real roots.

Expanding on what middle school textbooks often overlook, when $\Delta < 0$, the roots of the characteristic equation are conjugate complex numbers. This situation frequently arises in vibrating systems, such as spring oscillators, pendulums, and mechanical vibrations. In quantum mechanics, the time evolution equation of wave functions can also be described by similar differential equations, used to represent the wave characteristics of particles. Let α and β be real numbers, and i be the imaginary unit. Then we have the conjugate pair:

$$r_1 = \alpha + i\beta r_2 = \alpha - i\beta$$

Using Euler's formula, $e^{ix} = \cos(x) + i\sin(x)$, we can directly substitute it in to obtain:

$$e^{(\alpha + i\beta)x} = e^{\alpha x}(\cos(\beta x) + i\sin(\beta x))$$

This allows us to simply combine the real and imaginary parts, and substitute to simplify the general solution:

$$
\begin{aligned}
y &= C_1 e^{r_1 x} + C_2 e^{r_2 x} \\
&= C_1 e^{\alpha x}(\cos(\beta x) + i\sin(\beta x)) + C_2 e^{\alpha x}(\cos(\beta x) - i\sin(\beta x)) \\
&= (C_1 + C_2)e^{\alpha x}(\cos(\beta x)) + i(C_1 - C_2)e^{\alpha x}(\sin(\beta x))
\end{aligned}
$$

4.2 Differential Operator Method

For higher-order constant coefficient linear differential equations, the differential operator method is often a very effective approach. We consider an n-th order constant coefficient linear differential equation of the following form

$$
a_n \frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = f(x)
$$

We introduce the operator $\mathcal{D}$ as follows

$$
\mathcal{D} = \frac{\mathrm{d}}{\mathrm{d}x}, \quad \mathcal{D}^2 = \frac{\mathrm{d}^2}{\mathrm{d}x^2}, \quad \cdots, \quad \mathcal{D}^n = \frac{\mathrm{d}^n}{\mathrm{d}x^n}
$$

Substituting this into our equation yields

$$
\begin{aligned}
a_n \mathcal{D}^n y + a_{n-1} \mathcal{D}^{n-1} y + \cdots + a_1 \mathcal{D} y + a_0 y &= f(x) \\
(a_n \mathcal{D}^n + a_{n-1} \mathcal{D}^{n-1} + \cdots + a_1 \mathcal{D} + a_0) y &= f(x)
\end{aligned}
$$

We define $\mathcal{D}^0 = I$, where I is the identity operator, indicating that no differentiation has been performed. We can then consistently define the polynomial $P(\mathcal{D}) = \sum_{i=0}^{n} a_i \mathcal{D}^i$, which allows us to express the equation as

$$
P(\mathcal{D}) y = f(x)
$$

From this, we can clearly obtain its particular solution, where $\frac{1}{P(\mathcal{D})}$ is referred to as the inverse operator of $P(\mathcal{D})$.

$$
y^* = \frac{1}{P(\mathcal{D})} f(x)
$$

4.2.1 Exponential Input Theorem

We can introduce a theorem related to the differential operator method, known as the Exponential Input Theorem, specifically designed for solving cases where the input is an exponential function. When the equation is $f(x) = e^{ax}$, i.e., $P(\mathcal{D})y = e^{ax}$, its particular solution is given by:

$$
y^* = \frac{e^{ax}}{P(a)} \quad (P(a) \neq 0)
$$

The proof for differential equations over the real numbers is straightforward, so we won't elaborate on that. Now, we also have to consider differential equations in the complex domain,

which is simple as well. We can express a complex function $f(\ddagger) = e^{a\ddagger}$ and write it in terms of its real and imaginary parts:

$$f(\ddagger) = P(\mathcal{D})y = \Re(f(\ddagger)) + i \cdot \Im(f(\ddagger))$$

Since the differential operator preserves the linearity of addition and scalar multiplication, we can handle the real and imaginary parts independently:

$$P(\mathcal{D})y = \Re(f(\ddagger))$$

$$P(\mathcal{D})y = \Im(f(\ddagger))$$

Thus, the Exponential Input Theorem also applies to differential equations in the complex domain, enabling us to solve trigonometric functions as well.

4.3 Laplace Transform

The Laplace transform is an integral transform, belonging to linear transformations, that converts a locally integrable function $f(t)$ defined on $[0, \infty)$ into a function of a complex variable s. It is denoted by $\mathcal{L}\{f(t)\}$. Moreover, to avoid confusion, we can also represent the Laplace transform as $F(s) = \mathcal{L}\{f(t)\}$, so we have

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

It's important to note that the Laplace transform exists only when $\int_0^\infty f(t)e^{-st}\mathrm{d}t$ converges.

Given the above equality, when we apply the Laplace transform to $f'(t)$

$$\mathcal{L}\{f'(t)\} = \int_0^\infty e^{-st} f'(t)\,\mathrm{d}t$$

By applying integration by parts, we obtain

$$\mathcal{L}\{f'(t)\} = e^{-st}f(t)\big|_0^\infty - \int_0^\infty f(t)(-se^{-st})\mathrm{d}t$$

As t approaches positive infinity, the exponential term e^{-st} will tend towards zero due to the negative exponent $-st$, leading to $\lim_{a\to\infty} e^{-st}f(a) = 0$. Therefore, the boundary term $e^{-st}f(t)|_0^\infty = -f(0)$. The other term $-\int_0^\infty f(t)(-se^{-st})\,\mathrm{d}t$ can be simplified, and we can substitute back $\int_0^\infty e^{-st}f(t)\,\mathrm{d}t$ as $F(t)$, yielding

$$\begin{aligned}
\mathcal{L}\{f'(t)\} &= e^{-st}f(t)\big|_0^\infty - \int_0^\infty f(t)(-se^{-st})\mathrm{d}t \\
&= -f(0) + s\int_0^\infty e^{-st}f(t)\,\mathrm{d}t \\
&= sF(t) - f(0)
\end{aligned}$$

Thus, we have derived the theorem for the Laplace transform of the first derivative.

This seems to diverge from the main topic of the article; I wanted to show this proof simply because I derived it myself in the past.

Now that you know what this is, let's discuss how to apply the Laplace transform to constant coefficient linear differential equations. We'll consider the following form

$$a_n \frac{\mathrm{d}^n y}{\mathrm{d}t^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y = f(t)$$

We can take the Laplace transform of both sides, converting each term into an algebraic expression in the complex frequency domain

$$a_n \mathcal{L}\{\frac{\mathrm{d}^n y}{\mathrm{d}t^n}\} + a_{n-1} \mathcal{L}\{\frac{\mathrm{d}^{n-1} y}{\mathrm{d}t^{n-1}}\} + \cdots + a_1 \mathcal{L}\{\frac{\mathrm{d}y}{\mathrm{d}t}\} + a_0 \mathcal{L}\{\dagger\} = \mathcal{L}\{\{(\sqcup)\}$$

This can sound complicated and abstract, so let's take a first-order ordinary differential equation as an example: $\frac{\mathrm{d}y}{\mathrm{d}t} + ay = 0$, where $Y(s) = \mathcal{L}\{y(t)\}$. Applying the Laplace transform gives us

$$\begin{aligned} \int_0^\infty f(t) e^{-st} \mathrm{d}t &= y e^{-st} \big|_0^\infty - \int_0^\infty y(t)(-s e^{-st}) \mathrm{d}t \\ &= sY(s) - y(0) \end{aligned}$$

Substituting back yields

$$\begin{aligned} sY(s) - y(0) - aY(s) &= 0 \\ Y(s) &= \frac{y(0)}{s-a} \end{aligned}$$

Next, we introduce the concept of the inverse Laplace transform, which is simply the reverse operation of the Laplace transform. It converts a function in the complex domain back to the time domain, in the following form, where γ is a real number containing all poles of $F(s)$, and i is the imaginary unit

$$f(t) = \mathcal{L}^{-1}\{F\}(t) = \frac{1}{2\pi i} \lim_{T \to \infty} \int_{\gamma - iT}^{\gamma + iT} e^{st} F(s) \mathrm{d}s$$

Here, $\mathcal{L}^{-1}\left\{\frac{1}{s-a}\right\} = e^{at} u(t)$ is a commonly known formula from the inverse Laplace transform theorem (since many derivations are available online, I won't write it out here, especially since I'm quite busy today), so we can apply it to obtain

$$y(t) = y(0) e^{at}$$

4.4 Problems

1. Determine the roots of the characteristic equation for $y'' + 2y' + 2y = 0$ and write its solution.
2. Solve $\frac{\mathrm{d}y}{\mathrm{d}x} + y = e^x$ using the exponential input theorem.

3. Expand the Laplace transform of the function $f(x) = e^{2x}$ on the interval $[0, \infty]$.
4. Find the antiderivative of the Laplace transform function $F(s) = \frac{1}{s^2+4s+5}$.

5 Homogeneous Linear Differential Equation Systems

5.1 Definition

A standard linear differential equation system of order n can be compactly expressed using matrices and vectors as follows

$$\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x} = \mathbf{A}\mathbf{y}(x) + \mathbf{f}$$

where $\mathbf{y}(x)$ is the solution we are concerned with, a vector of unknown functions. $\mathbf{A}$ is the coefficient matrix, and when $\mathbf{f} = 0$, this represents a homogeneous system. The vector of unknown functions $\mathbf{y}(x)$ and the coefficient matrix $\mathbf{A}$ of a homogeneous linear differential equation system with constant coefficients is in the form of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

When the system of linear differential equations with constant coefficients is homogeneous, we can simply use the matrix exponential to represent the general solution of such a system. We introduce the definition for an n-dimensional real matrix $\mathbf{A}$, where its matrix exponential is given by

$$e^{\mathbf{A}} = \mathbf{E} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots = \sum_{n=0}^{+\infty} \frac{\mathbf{A}^n}{n!}$$

We define an identity matrix $\mathbf{E}$, and then we can define this matrix exponential function as

$$e^{\mathbf{A}x} = \mathbf{E} + \mathbf{A}x + \frac{(\mathbf{A}x)^2}{2!} + \cdots = \sum_{n=0}^{+\infty} \frac{(\mathbf{A}x)^n}{n!}$$

Next, we substitute $e^{\mathbf{A}x}$ into $\frac{\mathrm{d}\mathbf{y}(x)}{dx}$

$$\begin{aligned}
\frac{d\mathbf{y}(x)}{dx} &= \frac{\mathbf{d}}{\mathbf{d}x}\left(\mathbf{E} + \mathbf{A}x + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^i}{i!} + \cdots\right) \\
&= \mathbf{A} + \mathbf{A}^2 x + \frac{\mathbf{A}^3 x^2}{2!} + \cdots + \frac{\mathbf{A}^i x^{i-1}}{(i-1)!} + \cdots \\
&= \mathbf{A}\left(\mathbf{E} + \mathbf{A}x + \frac{(\mathbf{A}x)^2}{2!} + \cdots + \frac{(\mathbf{A}x)^{i-1}}{(i-1)!}\right) \\
&= \mathbf{A}e^{\mathrm{A}x} \\
&= \mathbf{A}\mathbf{y}(x)
\end{aligned}$$

Similarly, $e^{\mathbf{A}x}$ is the matrix form of the general solution for a system of linear differential equations with constant coefficients, known as the fundamental matrix solution. We can express it as

$$\exp \mathbf{A}x = \sum_{n=0}^{+\infty} \frac{(\mathbf{A}x)^n}{n!}$$

However, we should clearly define that we have already learned in the first chapter the concept of using the Wronskian determinant to determine the linear independence of solutions. Specifically, when the Wronskian determinant is non-zero, i.e., $W(y_1, y_2, \ldots, y_k)(x) \neq 0$, for all $x \in [a, b]$, this set of functions is linearly independent. For a system of differential equations, when the solution set $\mathbf{y}$ is linearly independent, we refer to the solution matrix $\mathbf{Y}$ it forms as the fundamental matrix solution. If there exists a point $x_0 \in [a, b]$ such that $\mathbf{Y}(x_0) = \mathbf{E}$, then $\mathbf{Y}$ is a standard fundamental matrix solution.

For a fundamental matrix solution $\mathbf{Y}$, the initial value condition given by $\mathbf{y}(x_0) = y_0$ leads to the solution that satisfies the initial value condition

$$\mathbf{y}(x) = \mathbf{Y}(x)\mathbf{Y}^{-1}(x_0)y_0$$

Next, we need to reconsider the complex solutions of $\frac{\mathrm{d}\mathbf{y}(x)}{\mathrm{d}x} = \mathbf{A}\mathbf{y}(x)$. We can separate the solution $\mathbf{y}$ into its real and imaginary parts

$$\mathbf{y}(x) = \Re(\mathbf{y}(x)) + i \cdot \Im(\mathbf{y}(x))$$

Then we substitute this back into the equation

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(\Re(\mathbf{y}(x)) + i \cdot \Im(\mathbf{y}(x))\right) &= \mathbf{A}\left(\Re(\mathbf{y}(x)) + i \cdot \Im(\mathbf{y}(x))\right) \\
\frac{\mathrm{d}}{\mathrm{d}x}\Re(\mathbf{y}(x)) + i\frac{\mathrm{d}}{\mathrm{d}x} \cdot \Im(\mathbf{y}(x)) &= \mathbf{A}\Re(\mathbf{y}(x)) + i\mathbf{A}\Im(\mathbf{y}(x))
\end{aligned}$$

Since the real and imaginary parts in the complex equation are independent of each other, we can directly compare them separately

$$\frac{\mathrm{d}\Re(\mathbf{y}(x))}{\mathrm{d}x} = \mathbf{A}\Re(\mathbf{y}(x))$$

$$\frac{\mathrm{d}\Im(\mathbf{y}(x))}{\mathrm{d}x} = \mathbf{A}\Im(\mathbf{y}(x))$$

Therefore, this applies to the complex matrix functions as well.

5.2 Eigenvectors and Eigen Decomposition

However, clearly, the matrix exponential function $e^{\mathbf{A}x}$ is an infinite series, making it quite complicated. Therefore, we need to use eigenvectors and eigen decomposition to solve the problem more intuitively. If all else fails, we could always resolve the issue with the person who raised the question (?). Although this topic has been discussed in linear algebra, it is worth mentioning again here.

We consider the eigenvalue problem for matrix $\mathbf{A}$ (assuming $\mathbf{A}$ is diagonalizable; otherwise, the following solution cannot be defined as a fundamental matrix solution, but we do have methods to handle such cases). There is an eigenvector matrix $\mathbf{V} = [\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_n}]$ and corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. By applying the chain rule to $\frac{\mathrm{d}}{\mathrm{d}x} e^{\lambda_i x}\mathbf{v}_i$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{\lambda_i x}\mathbf{v}_i = \lambda_i e^{\lambda_i x}\mathbf{v}_i = \mathbf{A}e^{\lambda_i x}\mathbf{v}_i$$

Thus, we can define the fundamental matrix solution as

$$\mathbf{Y} = [e^{\lambda_1 t}\mathbf{v}_1, e^{\lambda_2 t}\mathbf{v}_2, \cdots, e^{\lambda_n t}\mathbf{v}_n]$$

In Linear Algebra I, we learned that we can find the eigenvalues by solving the characteristic equation $|\mathbf{A} - \lambda I| = 0$, and in this case, we can directly solve and substitute back.

5.3 Jordan Canonical Form

For matrices $\mathbf{A}$ that are more rebellious and cannot be diagonalized, we need to use the Jordan canonical form. Here, $\mathbf{P}$ is the matrix composed of eigenvectors, allowing us to express the matrix $\mathbf{A}$ as

$$\mathbf{A} = \mathbf{PJP}^{-1}$$

where $\mathbf{J}$ is the Jordan matrix, $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \mathbf{J}_2, \cdots, \mathbf{J}_n)$. Let $\mathbf{J}_i$ denote the block corresponding to the main diagonal of $\mathbf{J}$. For instance, $\mathbf{J}_1, \mathbf{J}_2$ block the matrix $\mathbf{J}$ as follows

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \\ & \mathbf{J}_2 \end{bmatrix}$$

Each Jordan block has the form

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

We can compute each block $\mathbf{J^k}$ and the matrix exponential $e^{\mathbf{J}x}$

$$\mathbf{J^k} = \begin{bmatrix} \mathbf{J^k}_1 & \\ & \mathbf{J^k}_2 \end{bmatrix},\ e^{\mathbf{J}x} = \begin{bmatrix} e_1^{\mathbf{J}x} & \\ & e_2^{\mathbf{J}x} \end{bmatrix}$$

Combining these, we obtain $\mathbf{J^k}_i$

$$\mathbf{J}_i^k = \begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} & \frac{k(k-1)}{2}\lambda_i^{k-2} & \cdots & \frac{k!}{(k-m)!}\lambda_i^{k-m} \\ 0 & \lambda_i^k & k\lambda_i^{k-1} & \cdots & \frac{(k-1)!}{(k-m-1)!}\lambda_i^{k-m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i^k \end{bmatrix}$$

We can expand the exponential function $e^{\mathbf{J}x}$ as follows

$$e^{\mathbf{J}x} = \mathbf{E} + \mathbf{J}x + \frac{(\mathbf{J}x)^2}{2!} + \cdots = \sum_{m=0}^{+\infty} \frac{(\mathbf{J}x)^m}{m!}$$

$$e^{\mathbf{J}_i x} = \begin{bmatrix} e^{\lambda_i x} & xe^{\lambda_i x} & \frac{x^2}{2!}e^{\lambda_i x} & \cdots & \frac{x^{m-1}}{(m-1)!}e^{\lambda_i x} \\ 0 & e^{\lambda_i x} & xe^{\lambda_i x} & \cdots & \frac{x^{m-2}}{(m-2)!}e^{\lambda_i x} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_i x} \end{bmatrix}$$

In this way, we can compute the exponential form for each Jordan block and consequently obtain the entire matrix.

6 Numerical Methods for Ordinary Differential Equations

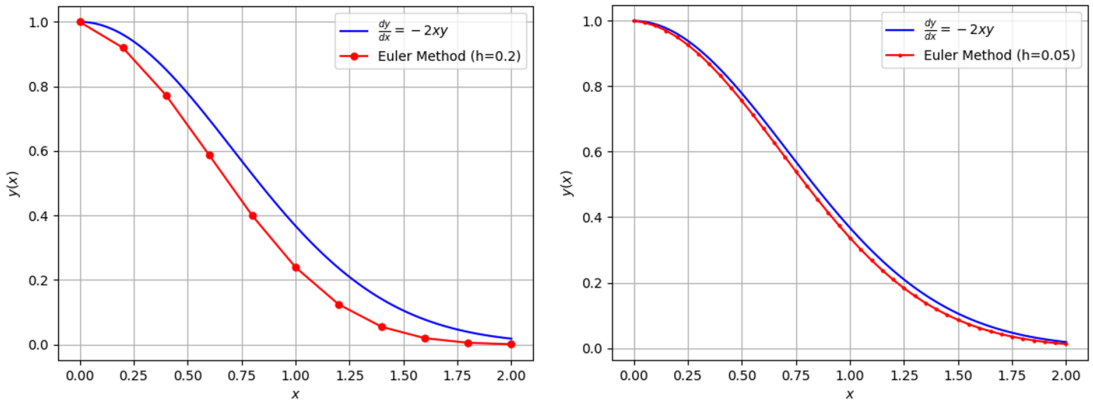
6.1 Difference Methods

The difference method is a fundamental numerical computation technique. In addition to solving differential equations, it can also be used for numerical approximation, image processing, and similar problems. The key idea is to approximate derivatives or integrals through finite interpolation, thereby discretizing a continuous problem.

Among these, the forward difference approximates using information from the current point and the next point. For equidistant nodes $x_k = x_0 + kh,\ (k = 0, 1, \cdots, n)$, the difference is the increment of the function value $f(x)_{n+1}$ minus the current function value. We denote it as:

$$\Delta f(x)_n = f(x)_{n+1} - f(x)_n$$

This can be directly generalized to higher orders. The k-th order difference $\Delta^k f(x)$ is defined, and the difference of differences is a higher-order difference. Thus, a k-th order forward difference formula is given by $\Delta^k f(x)_n = \Delta^{k-1}(f(x)_{n+1} - f(x)_n)$, leading to:



$$\Delta^k f(x) = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} f(x+i)$$
$$= \sum_{i=0}^{k} \binom{k}{i} (-1)^{i} f(x+k-i)$$

Similarly, the backward difference is defined as the difference between the current function value and the previous function value, i.e., $\Delta y_n = y_n - y_{n-1}$. A k-th order backward difference formula is given by $\Delta^k f(x)_n = \Delta^{k-1}(f(x)_n - f(x)_{n+1})$. Note that the above formulas also apply to backward differences because we include the power of -1, which determines whether it is a forward or backward difference.

6.2 Euler's Method

Euler's method is a commonly used numerical technique for approximating the solution to initial value problems in ordinary differential equations. Based on linear approximation, it can be considered one of the simplest numerical methods. We estimate the function value at the next point using the slope of the tangent line at the current point, assuming that over a step size, the change in the solution can be approximated by the current derivative multiplied by the time step. To represent this compactly, we define $f(x) = y$, and if the estimated step size is h, the iterative formula is

$$y_{n+1} = y_n + h \cdot f(x_n, y_n)$$

The figure below illustrates the approximations of the differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = -2xy$ using Euler's method with two different step sizes. Clearly, the smaller the step size, the more iterations are required, resulting in a higher degree of accuracy.

Note that Euler's method is a basic first-order numerical method, and its error generally increases with the step size. Therefore, for problems requiring high precision in numerical solutions, Euler's method may not be accurate enough and is not suitable for curves that change rapidly. However, it is often a good choice for simple initial value problems.

The aforementioned form is called the "explicit Euler method," and correspondingly, there is an implicit Euler method defined as

$$y_{n+1} = y_n + h \cdot f(x_{n+1}, y_{n+1})$$

Before proceeding with this, we need to solve this equation, for example, by using Newton's method to convert it into a root-finding problem. Rearranging the equation and defining a symbol leads us to

$$G(y_{n+1}) = y_{n+1} - y_n - h \cdot f(x_{n+1}, y_{n+1}) = 0$$

The formula for Newton's method is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

We can substitute this into our expression, compactly represented as

$$y_{n+1}^{k+1} = y_n - \frac{F(y_{n+1}^k)}{F'(y_{n+1}^k)}$$

We can continue the iterative computation in a similar manner.

6.3 Runge-Kutta Methods

The Runge-Kutta methods are numerical techniques for solving ordinary differential equations. They adjust the slope repeatedly to estimate the solution and use appropriate weights to average these slopes, making them generally more accurate than the Euler method.

If we modify the Euler method's formula $y_{n+1} = y_n + h \cdot f(x_n, y_n)$ to use two slopes for estimating the next segment of the solution, we can consider a step size h along with the following ordinary differential equation and initial value problem:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y), \quad y(x_0) = y_0$$

Similarly, we can compute our two target slopes:

$$k_1 = h \cdot f(x_n, y_n)$$
$$k_2 = h \cdot f(x_n + h, y_n + k_1)$$

We can then use the average of these slopes to obtain the estimated solution:

$$y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2)$$

This method, which averages two slopes, is known as the second-order Runge-Kutta method, abbreviated as RK2. In practice, RK4 is more commonly used, which takes four slopes into account and gives more weight to the midpoint slope. This is similar in concept, and both are referred to as "explicit Runge-Kutta" methods. A s-th order Runge-Kutta method (apologies for the usual use of k and n being occupied) has weights c_i and is expressed as follows:

$$
\begin{aligned}
k_1 &= hf(t_n, y_n) \\
k_2 &= hf(t_n + c_2 h, y_n + a_{21} k_1) \\
k_3 &= hf(t_n + c_3 h, y_n + a_{31} k_1 + a_{32} k_2) \\
&\vdots \\
k_s &= hf(t_n + c_s h, y_n + \sum_{i=1}^{s-1} a_{si} k_i)
\end{aligned}
$$

We define the weight coefficients b_i, giving us:

$$
y_{n+1} = y_n + \sum_{i=1}^{s} b_i k_i
$$

Note not to confuse c_i, which are the weights for local adjustments, with b_i, which are the weights adjusted according to the order. Refer to the following table:

$$
\begin{array}{c|cccc}
0 & & & & \\
c_2 & a_{21} & & & \\
c_3 & a_{31} & a_{32} & & \\
\vdots & \vdots & \vdots & \ddots & \\
c_s & a_{s1} & a_{s2} & \dots & a_{s,s-1} \\
\hline
 & b_1 & b_2 & \dots & b_s
\end{array}
$$

6.4 Problems

1. Write a function that takes a list as input and returns the difference values of each element with its subsequent element in the list.
2. Write a function that accepts a tuple representing an interval and solves the ordinary differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = x - y$ with a step size of $h = 0.1$.
3. Use the fourth-order Runge-Kutta method to solve the ordinary differential equation $\frac{\mathrm{d}y}{\mathrm{d}x} = x + y$ with a step size of $h = 0.1$.

References

1. Wikipedia contributors. (n.d.). Calculus. Wikipedia. `https://en.wikipedia.org/wiki/Calculus`
2. MacCluer, B. D., Bourdon, P. S., Kriete, T. L. (n.d.). *Differential equations: Techniques, theory, and applications.*
3. MIT OpenCourseWare. (n.d.). *Differential equations.* Massachusetts Institute of Technology. `https://ocw.mit.edu/courses/mathematics/18-03-differential-equations-spring-2010/`