ODEs and Boundary Value Problems

An Intuitive Introduction by Iman Ebrahimi

Ongoing Project —— Last Update: October 2024

Preliminaries

Linear Differential Equations

We begin by recalling some basic definitions relevant to the study of differential equations.

Definition. An ordinary differential equation of order n, written in *normal form*, is a relation between a variable $t \in I \subset \mathbb{R}$, a function $x : \mathbb{R} \supset I \to \mathbb{R}$, and its derivatives:

$$x^{(n)} = f\left(t, x, x', ...x^{(n-1)}\right)$$

where t, when interpreted physically, represents time in most situations.

Definition. An ODE is called *linear* if the right-hand side of the above equation can be written as:

$$f(t, x, x', ..., x^{(n-1)}) = -\left(\sum_{k=0}^{n-1} a_k(t)x^k\right) + b(t).$$

where the coefficients $a_k(t)$ can sometimes be $t-$dependent and sometimes constants. Combining these two notations, we can write:

$$x^n + a_{n-1}(t)x^{(n-1)} + ...a_1(t)x'(t) + a_0(t)x(t) = b(t) \qquad (*)$$

Definition. The *Characteristic Polynomial* associated with $(*)$ is defined by:

$$L(t, \lambda) = \lambda^n + a_{n-1}(t)\lambda^{n-1} + \ldots + a_1(t)\lambda + a_0(t)$$

Let $D = \dfrac{d}{dt}$ so $Dx = x'$, $D^2x = x''$, ... Then $(*) \iff L(t, D)x = b(t)$.
$L(t, D)$ is called a differential operator.

Definition. A linear differential equation is called **homogenenous** if $b(t) \equiv 0$. Otherwise, we have an **inhomogeneous** equation.

So how do we solve such equations? The truth is, that many differential equations, particularly those that arise in real-world applications, can be too complicated to solve analytically. As a result, we often opt for numerical approximations and investigate more efficient computational algorithms. Analytical solutions are often only possible for simple or idealized cases. However, it is a fruitful effort to focus on such idealized cases in pursuit of analytical methods, since they provide insight into the fundamental behavior of the system described by the differential equation. What follows is a rigorous study of the most common cases where analytical solutions are achievable.

Linear Homogeneous Equations with Constant Coefficients

Assume that $b(t) = 0$, and a_j is independent of t. So, $(*)$ is homogeneous with constant coefficients. In this case, L is independent of t. So, we write:

$$L(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$$

In this case, $(*)$ can be written as: $L(D)x = 0$.

Lemma. Let $\mu \in \mathbb{C}$. Then:

$$L(D)(e^{\mu t}g(t)) = e^{\mu t}L(D+\mu)g(t)$$

Proof. By the product rule, we have that:

$$\begin{aligned} D(e^{\mu t}g(t)) &= D(e^{\mu t})g(t) + e^{\mu t}Dg(t) \\ &= \mu e^{\mu t}g(t) + e^{\mu t}Dg(t) \\ &= e^{\mu t}(D+\mu)g(t) \end{aligned}$$

We observe that:

$$\begin{aligned} &D^2(e^{\mu t}g(t)) = D(e^{\mu t}(D+\mu)g(t)) \\ &= e^{\mu t}(D+\mu)^2 g(t) \end{aligned}$$

And by induction:

$$D^k(e^{\mu t}g(t)) = e^{\mu t}(D+\mu)^k g(t)$$

Since $L(D)$ is a linear combination of D^k, it follows that

$$L(D)(e^{\mu t}g(t)) = e^{\mu t}L(D+\mu)g(t) \blacksquare$$

Theorem. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the roots of $L(\lambda)$ with multiplicities $n_1, n_2, \ldots, n_m$ respectively.

For arbitrary polynomials p_k of degree $\leq n_k - 1$,

$$x(t) = \sum_{k=1}^{m} p_k(t) e^{\lambda_k t}$$

is a solution of $L(D)x = 0$; conversely, every solution of $L(D)x = 0$ is of this form.

Proof. We will prove that $x(t)$ is a solution. Since $L(D)$ is a linear operator, we have that:

$$L(D)\left(\sum_{k=1}^{m} p_k(t) e^{\lambda_k t}\right)$$
$$= \sum_{k=1}^{m} L(D)(p_k(t) e^{\lambda_k t})$$

So it is enough to show the following for any k:

$$L(D)(p_k(t) e^{\lambda_k t}) = 0;$$

By the lemma , we have that:

$$L(D)(p_k(t) e^{\lambda_k t}) = e^{\lambda_k t} L(D + \lambda_k) P_k(t)$$

Since λ_k is a zero of order n_k,

$$L(\lambda) = (\lambda - \lambda_k)^{n_k} q(\lambda).$$

This implies that $L(\lambda + \lambda_k) = \lambda^{n_k} q(\lambda + \lambda_k)$.
This is a polynomial in λ where each term is at least of order n_k.

So: $L(D + \lambda_k) p_k(t) = 0$, since p_k has degree $< n_k - 1$■

Inhomogeneous Equations

Theorem. Every solution of $L(t, D)x = b(t)$ can be written as:

$$x(t) = x_h(t) + x_p(t)$$

where $x_p(t)$ is a solution of $(*)$ and $x_h(t)$ is a solution of the corresponding homogeneous equation.

Remark. The above theorem means that once we encounter a linear **inhomogeneous** equation $L(t, D)x = b(t)$, we first treat it as a **homogeneous** equation $L(t, D)x = 0$. Solving this, we'd have our *homogeneous solution* x_h. The next step would be to find one solution that would satisfy $L(t, D)x = b(t)$.

We call this solution x_p a *particular* solution. Since $L(t, D)$ is a **linear** differential operator, exploiting the linearity property would allow us to get the full solution by adding x_h and x_p together. The following proof formalizes these steps:

Proof. $x(t)$ is a solution of (*) since:

$$L(t, D)(x_h + x_p) = \underbrace{L(t, D)x_h}_{=0} + \underbrace{L(t, D)x_p}_{=b(t)} = b(t)$$

Let x be a solution of (*), and define $x_h = x - x_p$. Then:

$$L(t, D)x_h = L(t, D)x - L(t, D)x_p = b(t) - b(t) = 0\blacksquare.$$

Linear System of Differential Equations

As the name suggests, a *time-dependent* differential equation models the *time evolution* of a physical system. However, even the most specific subsets of systems in the physical world, albeit elegant and beautiful, are often too complicated to fit into one particular equation describing their evolution. Many physical, biological, and engineering systems can instead be modeled by **multiple** interacting variables, each of which may depend on time (t), space (x), or other independent variables. This complexity motivates the shift of our focus towards the study of **systems** of differential equations, in particular, we will be working with **linear systems**.

Example. Consider a system of first-order ODEs:

$$\begin{cases} x_1' = f_1(t, x_1, x_2, ..., x_n) \\ \vdots \\ x_n' = f_n(t, x_1, x_2, ..., x_n) \end{cases}$$

Using our physical interpretation analogy, we can think of the system as a giant machine, and each of the functions $x_1(t), x_2(t), ..., x_n(t)$ as parts or "gears" that make this machine operate. The evolution of such system is dependent on the time evolution of each of these parts. Mathematically speaking, the derivative with respect to t of the whole system, is directly related to the time derivative $\frac{d}{dt}$ of each of these n functions x_i.

The necessity of keeping track of $x_1', ..., x_n'$ (and eventually finding a solution for each component) brings us to the realm of **linear algebra**, in particular, the study of **vector algebra** and related mathematical properties. The idea is to study the evolution of these functions by considering them as n elements of a vector $\mathbf{x}(\mathrm{t})$ in $\mathbb{R}^n$. Hence, the compact nation of a first-order linear system is the following:

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}) \qquad (**)$$

where $\mathbf{f} = (f_1, ...f_n)$, $\mathbf{x} = (x_1, ..., x_n)$, such that $x_k = x_k(t)$. Furthermore, we have the function $\mathbf{f}$: $\mathbb{R}^n \times \mathbb{R} \supset \Omega \to \mathbb{R}^n$ given that $\mathbf{x} : \mathbb{R} \supset I \to \mathbb{R}^n$.

Remark. Alternatively, one can write the above compact notations as column vectors, instead of row vectors. This would be:

$$\mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix} = \begin{bmatrix} f_1(t,\mathbf{x}) \\ f_2(t,\mathbf{x}) \\ \vdots \\ f_n(t,\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} (t,\mathbf{x}) = \mathbf{f}(t,\mathbf{x})$$

We say $\mathbf{x}$ is a solution to $(**)$ if the following three conditions are satisfied:

- $\mathbf{x}$ is differentiable on I.
- $(t, \mathbf{x}(t)) \in \Omega \ \forall t \in I$
- (**) holds $\forall t \in I$

So how many solutions will there be that would satisfy these three conditions? We would, in fact, be expecting an n-parameter **family** of solutions. Additional information is needed to specify which of these solutions is the one we seek.

Mathematically speaking, this "additional info" is what we consider an "initial condition". It provides us with a pair of "points" $(t_0, \mathbf{x}_0) \in \Omega$ such that $\mathbf{x}(t_0) = \mathbf{x}_0$. Among the family of solutions found previously, there would only be one $\mathbf{x}$(t) that could satisfy $\mathbf{x}(t_0) = \mathbf{x}_0$, provided that $\mathbf{f}$ satisfies certain conditions. In the following section, we elaborate more precisely on what we mean by this.

Initial Value Problems

Definition. A problem involving a differential equation, together with an initial condition is called an **initival value problem (IVP)**

$$\begin{cases} \mathbf{x}' = \mathbf{f}(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases} \tag{‡}$$

Remark. Although the focus of these notes is mainly on **boundary value problems**, the underlying concepts between those and initial value problems are identical. The distinction between initial value problems (IVPs) and boundary value problems (BVPs) lies in the conditions imposed on the solution of the differential equation.

- **IVPs** are concerned with the behavior of a solution at a specific initial time, hence they are dependent on the time variable t.

- **BVPs** are concerned with the behavior of a solution over a specified domain, typically involving conditions at different points within that domain, hence they are dependent on the spatial variable x.

Example. (Initial Value Problem) Consider a mass-spring system where a spring is attached to a wall on one end and to a mass on the other end. If the mass is initially displaced from its equilibrium position and released, it will oscillate back and forth.

The motion of the mass can be described by a second-order differential equation, such as

$$m\frac{d^2x}{dt^2} + kx = 0$$

where m is the mass, k is the spring constant, x is the displacement from the equilibrium position, and t is time.

In this scenario, an initial value problem arises because we need to specify the initial conditions of the system at $t = 0$, such as the initial displacement $x(0)$ and the initial velocity $\left.\frac{dx}{dt}\right|_{t=0}$. These initial conditions dictate the behavior of the system over **time**.

Example. (Boundary Value Problem): Consider a vibrating string fixed at both ends and plucked at some initial time. The string vibrates in a transverse motion, producing waves along its length.

The motion of the string can be described by the wave equation, such as

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}$$

where $u(x, t)$ represents the displacement of the string at position x and time t, and c is the wave speed.

In this scenario, a boundary value problem arises because we need to specify the boundary conditions at both ends of the string. These boundary conditions could include fixing the displacement or specifying the forces at the ends of the string. For example, we might have conditions such as $u(0, t) = 0$ and $u(L, t) = 0$, where L is the length of the string. These boundary conditions ensure that the ends of the string remain fixed during the vibration, which affects the pattern of the standing waves formed on the string.

Definition. Similar to the concept of **linear** differential equations, one can have a **system** of linear differential equations. The equation (**) therefore becomes:

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}) = \begin{bmatrix} f_1(t, \mathbf{x}) \\ f_2(t, \mathbf{x}) \\ \vdots \\ f_n(t, \mathbf{x}) \end{bmatrix} = \begin{bmatrix} a_{11}(t)x_1 + a_{12}(t)x_2 + ... + a_{1n}(t)x_n + b_1(t) \\ a_{21}(t)x_1 + a_{22}(t)x_2 + ... + a_{2n}(t)x_n + b_2(t) \\ \vdots \\ a_{n1}(t)x_1 + a_{n2}(t)x_2 + ... + a_{nn}(t)x_n + b_n(t) \end{bmatrix}$$

$$\iff \mathbf{x}' = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}}_{A}(t) \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{\mathbf{b}}(t)$$

$$\iff \mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t) \qquad (\dagger)$$

Remark. Modelling the evolution of a physical system with vector-valued functions is not only motivated by nature but also by the mathematical simplicity that it provides. Indeed, solving for $\mathbf{x}(t)$ in ($\dagger$) using well-known techniques of matrix algebra is far more feasible than dealing with derivatives of order n. As a matter of fact, it is common practice to **reduce** a single n-order linear differential equation to a system of n first-order linear equations:

Regardless of the physical complexity of a phenomenon, consider an ODE of order n:

$$x^{(n)} = f(t, x, x', x'', \dots, x^{(n-1)})$$

This equation can be written as an n-dimensional first-order system of ODE's. We do this by introducing new variables:

$$x_1 = x,\ x_2 = x',\ x_3 = x'', \dots, x_n = x^{(n-1)}$$

Then:

$$\begin{cases} x_1' = x_2 \\ x_2' = x_3 \\ x_3' = x_4 \\ \vdots \\ x_n' = f(t, x_1, x_2, ..., x_n) \end{cases}$$

Example: Consider the ODE:

$$t^2 x'' - 3tx' + 4x = t^3, \quad t > 0$$

This equation can be rewritten as:

$$x'' = \underbrace{\frac{3}{t}x' - \frac{4}{t^2}x + t}_{= f(t,x,x')}$$

Set $x_1 = x$ and $x_2 = x'$; we obtain the system:

$$\begin{cases} x_1' = x_2, \\ x_2' = \underbrace{\frac{3}{t}x_2 - \frac{4}{t^2}x_1 + t}_{= f(t,x,x_2)} \end{cases}$$

Having developed so much new machinery, the main question now is how *do* we get to solve a linear system? Perhaps even a much more fundamental question is yet to be investigated. That is: When is a linear system *uniquely* solvable? Motivated by such important questions, we introduce the following theorems.

Existence and Uniqueness Theorems

Theorem. (Peano's Theorem) Assume that $\mathbf{f}$ is continuous in a neighborhood of $(t_0, \mathbf{x}_0) \in \mathbb{R} \times \mathbb{R}^n$. Then there exists a $\mathcal{C}^1$ solution to the IVP (‡) defined near t_0.

Remark. Peano's theorem is excellent in proving existence of solutions. However, it fails to detect uniqueness.

Example. Consider the IVP

$$\begin{cases} \mathbf{x}' = \sqrt{|\mathbf{x}|} \\ \mathbf{x}(0) = \mathbf{0} \end{cases}$$

We see that $\mathbf{x}_1(t) \equiv 0$ is a solution. But then:

$$\mathbf{x}_2(t) = \begin{cases} \dfrac{t^2}{4} \text{ for } t \geq 0 \\ \mathbf{0} \text{ for } t \leq 0 \end{cases}$$

is also a solution. Hence, we need a stronger condition to ensure uniqueness as well as existence of solutions.

Definition. The function $\mathbf{f}$ satisfies a **Lipschitz condition** in Ω if $\exists L \geq 0$ such that $|\mathbf{f}(t,\mathbf{x}) - \mathbf{f}(t,\mathbf{y})| \leq L|\mathbf{x} - \mathbf{y}|$ $\forall (t,\mathbf{x}), (t,\mathbf{y}) \in \Omega$.

Lemma. The Lipschitz condition is satisfied if Ω is convex and $\mathbf{f}$ is $\mathcal{C}^1$ in a neighborhood of the closure $\bar{\Omega}$.

Theorem. (Picard-Lindelöf) Assume that $\mathbf{f}$ is continuous in the neighborhood of $(t_0, \mathbf{x}_0)$ and satisfies a Lipschitz condition there. Then there exists an open interval $I \ni t_0$ in which the IVP has a unique solution.

From this point, we shall freely assume that every function $\mathbf{f}$: $\mathbb{R}^n \times \mathbb{R} \supset \Omega \to \mathbb{R}^n$ we work with is Lipschitz continuous, ensuring us that studied systems *do* have a unique and well-defined solution.

Solving Linear Systems

We consider a linear system of ODEs:

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t)$$

where $\mathbf{A}(t)$ is an $n \times n$ matrix and $\mathbf{b}(t)$, $\mathbf{x}(t)$ are $n \times 1$ matrices.

Homogeneous Equations

Assume that $\mathbf{b}(t) = \mathbf{0}$.
That is, we consider the homogeneous equation $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$ where $t \in I$.
If $\mathbf{x}_1, \mathbf{x}_2$ are solutions, then $a\mathbf{x}_1 + b\mathbf{x}_2$ is also a solution $(a, b \in \mathbb{R})$.

Definition. We denote the linear space of all solutions on I of $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$ by V.

$$V = \{\mathbf{x} : \mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t), \forall t \in I\}$$

Note that V is a function space: $\mathbf{x} \in V$ evaluated at t_0 where $\mathbf{x}(t_0) \in \mathbb{R}^n$. We say that k elements $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in V are *linearly dependent* if there exist constants $\lambda_1, \lambda_2, \ldots, \lambda_k$ (not all equal to 0) such that:

$$\sum_{j=1}^{k} \lambda_j \mathbf{x}_j(t) = 0, \forall t \in I$$

If they are not linearly dependent, then they are linearly *independent*. For a fixed $t \in I$, we say that: $\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots \mathbf{x}_k(t)$ are linearly dependent if $\exists \lambda_1, \lambda_2, \ldots \lambda_k$ (not all equal to 0), such that:

$$\sum_{j=1}^{k} \lambda_j \mathbf{x}_j(t) = 0$$

Lemma. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ be solutions of $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$. Then the following statements are equivalent:

a) $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linearly independent in V.

b) $\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots, \mathbf{x}_k(t)$ are linearly independent in $\mathbb{R}^n$ $\forall t \in I$

c) $\mathbf{x}_1(t_0), \mathbf{x}_2(t_0), \ldots, \mathbf{x}_k(t_0)$ are linearly independent in $\mathbb{R}^n$ for some $t_0 \in I$.

Theorem. Let $\mathbf{A}(t)$ be a continuous $n \times n$ matrix on an open interval I. The solutions of $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$ form a linear space V of dimension n.

Corollary. The solution set of a linear homogeneous ODE of order n, with constant coefficients in I, is an n-dimensional linear space.

Fundamental Matrices

Definition. Let $x_1, x_2, \ldots, x_n$ be a basis of V. The matrix $\mathbf{F}(t) = \begin{bmatrix} | & | & & | \\ \mathbf{x}_1(t) & \mathbf{x}_2(t) & \ldots & \mathbf{x}_n(t) \\ | & | & & | \end{bmatrix}$ is called a **fundamental matrix** of the equation $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$.

Since $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent, $\mathbf{F}(t)$ is invertible for all $t \in I$.
$\mathbf{F}'(t) = \mathbf{A}(t)\mathbf{F}(t)$, since each column of $\mathbf{F}$ solves the equation.
Since $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is a basis of V, any solution can be written as a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$:

$$\mathbf{x}(t) = \sum_{i=1}^{n} c_i \mathbf{x}_i(t) = \mathbf{F}(t)\mathbf{c},$$

where $\mathbf{c}$ is an arbitrary column matrix.

Inhomogeneous Systems

Once again, we consider the equation:

$$\mathbf{x}'(t) = \mathbf{A}(t)x + \mathbf{b}(t), \text{ where } \mathbf{b}(t) \not\equiv 0 \qquad (***)$$

Theorem. Assuming that $\mathbf{A}(t)$, $\mathbf{b}(t)$ are continuous; let $\mathbf{F}(x)$ be the fundamental matrix of $\mathbf{x}' = \mathbf{A}(t)\mathbf{x}$. Then:

$$\mathbf{x}(t) = \mathbf{F}(t) \int_{t_0}^{t} \mathbf{F}(\tau)^{-1}\mathbf{b}(\tau) d\tau$$

is the solution of (***) with the initial condition $\mathbf{x}(t_0) = \mathbf{0}$.

Corollary. The general solution of the inhomogeneous system (1) can be written as:

$$\mathbf{x}(t) = \mathbf{F}(t)\mathbf{c} + \mathbf{F}(t) \int_{t_0}^{t} \mathbf{F}(\tau)^{-1}\mathbf{b}(\tau) d\tau$$

Where $\mathbf{c}$ is an arbitrary $n \times 1$ matrix.

Linear Equations of Order n

Earlier we indicated that reducing a linear ODE of order n to a system of n first-order ODEs provides us with a much more convenient setting. We now intend to elaborate on this reduction, and to see how we can solve them using linear systems. To simplify calculations, the order is assumed to be $n = 3$.

We consider the equation:

$$L(t, D)y = g(t), \quad L(t, \lambda) = \lambda^3 + a_2(t)\lambda^2 + a_1(t)\lambda + a_0(t). \tag{I}$$

Reduction to $n = 1$: The equation (I) is equivalent to the linear system:

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{b}(t) \tag{II}$$

where:

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) \end{pmatrix}$$

and

$$\mathbf{b}(t) = \begin{pmatrix} 0 \\ 0 \\ g(t) \end{pmatrix}.$$

Let $x_1 = y$, $x_2 = y'$, $x_3 = y''$.

Solutions $\mathbf{x}$ of (II) are of the form $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ y'(t) \\ y''(t) \end{bmatrix}$, where y solves (I)

Let $\overline{\mathbf{x}_1}, \overline{\mathbf{x}_2}, \overline{\mathbf{x}_3}$ be three linearly independent solutions of (II). And let y_1, y_2, y_3 be the corresponding solutions of (I). Then we can write the fundamental matrix as:

$$\mathbf{F}(t) = \begin{bmatrix} | & | & | \\ \overline{\mathbf{x}_1} & \overline{\mathbf{x}_2} & \overline{\mathbf{x}_3} \\ | & | & | \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \end{bmatrix}$$

Where $\mathbf{R}_1 = (y_1, y_2, y_3)$, $\mathbf{R}_2 = (y_1', y_2', y_3')$, $\mathbf{R}_3 = (y_1'', y_2'', y_3'')$.
We write:

$$\mathbf{F}^{-1}(t) = \begin{bmatrix} | & | & | \\ \mathbf{K}_1 & \mathbf{K}_2 & \mathbf{K}_3 \\ | & | & | \end{bmatrix}$$

Using the discussed theorem, we find a solution $\overline{x}$ of (II):

$$\begin{aligned} \overline{\mathbf{x}}(t) &= \mathbf{F}(t) \int_{t_0}^{t} \mathbf{F}(\tau)^{-1} \mathbf{b}(\tau) d\tau \\ &= \mathbf{F}(t) \int_{t_0}^{t} \mathbf{K}_3(\tau) g(\tau) d\tau \end{aligned}$$

Remark. Observe that once we solve (II), the resulting solution, $\overline{\mathbf{x}}(t)$, is a vector-valued function, containing three elements $y(t)$, $y'(t)$, and $y''(t)$. But this is more than we were looking for! The reduction to $n = 1$ was meant to simplify calculations for obtaining $y(t)$. Here, however, $\overline{\mathbf{x}}(t)$ contains more information than just $y(t)$. We must therefore extract $y(t)$ from $\overline{\mathbf{x}}(t)$. That is, we take the first component of $\mathbf{F}(t) \int_{t_0}^{t} \mathbf{K}_3(\tau)g(\tau)d\tau$,

$$\begin{aligned} y(t) &= \mathbf{R}_1(t) \int_{t_0}^{t} \mathbf{K}_3(\tau)g(\tau)d\tau \\ &= \int_{t_0}^{t} \mathbf{R}_1(t)\mathbf{K}_3(\tau)g(\tau)d\tau \end{aligned}$$

Let $E(t,\tau) := \mathbf{R}_1(t)\mathbf{K}_3(\tau)$.
Then

$$y(t) = \int_{t_0}^{t} E(t,\tau)g(\tau)d\tau$$

solves (I) with initial condition $y(t_0) = y'(t_0) = y''(t_0) = 0$.

Remark. This was all very impressive, but it turns out that $E(t,\tau)$ can be characterized in a different way as well. Instead of expressing it as a product of row and column vectors, one can uniquely determine $E(t,\tau)$ as the solution of a differential equation. The following theorem formalizes this idea.

Theorem. Let $L(t,\lambda) = \lambda^n + a_{n-1}(t)\lambda^{n-1} + \ldots + a_1(t)\lambda + a_0(t)$, and let $E(t,\tau)$ be the uniquely determined solution of:

$$L(t,D)u = 0$$

$$u(\tau) = u'(\tau) = \ldots = u^{(n-2)}(\tau) = 0, \quad u^{(n-1)}(\tau) = 1$$

Then $y(t) = \int_{t_0}^{t} E(t,\tau)g(\tau)\,d\tau$ is the solution to the problem:

$$\begin{cases} L(t,D)y = g(t) \\ y(t) = y'(t) = \ldots = y^{(n-1)}(t) = 0 \end{cases}$$

Definition.
The function $E(t,\tau)$ is called the **fundamental solution** of the operator $L(t,D)$.

Boundary Value Problems

We are now well-equipped to approach the study of boundary value problems, where the solution is sought within a given domain, and the conditions are specified at different points on the boundary of this domain. These conditions can be of various types, such as specifying the values of the solution at certain points, prescribing the values of certain derivatives, or imposing integral constraints.

Unlike initial value problems, in which one studied the change of a *time-dependent* function $x(t)$, it is common to consider a function $y(x)$ dependent on a *spatial* variable x in the study of BVPs, and therefore apply the developed machinery for initial value problems to boundary value problems.

Definition. A **boundary value problem (BVP)** is a differential equation, together with an array of boundary conditions of the form:

$$L(x, D)y = f;\ \mathbf{B}y = \mathbf{c}$$

where $L(x, D) \equiv a_n(x)D^n + a_{n-1}(x)D^{n-1} + ... + a_1(x)D + a_0(x)$ such that $D = \dfrac{d}{dx}$ and $a_n(x) \not\equiv 0\ \forall x \in [\alpha, \beta]$.

The boundary conditions $\mathbf{B}\mathrm{y} = (\mathcal{B}_1 y, \mathcal{B}_2 y, ..., \mathcal{B}_n y) = (c_1, c_2, ..., c_n) = \mathbf{c} \in \mathbb{R}^n$.

There are multiple ways one can define boundary conditions. Some examples are:

Dirichlet BCs:

$$\begin{cases} (p(x)y')' = \lambda y \\ y(\alpha) = y(\beta) = 0 \end{cases}$$

Neumann BCs:

$$y'(\alpha) = y'(\beta) = 0$$

Robin BCs:

$$b_1 y(\alpha) + b_2 y'(\alpha) = 0$$

Periodic BCs:

$$\begin{cases} y(\alpha) = y(\beta) \\ y'(\alpha) = y'(\beta) \end{cases}$$

Given the variety of ways to express boundary condition, we seek a general form to include every possible combination. For an n-th order linear ODE, where $a_0, ..., a_n, f$ are given functions such that $a_n \not\equiv 0$:

$$a_n(x)y^n + a_{n-1}(x)y^{(n-1)} + ... + a_1(x)y' + a_0(x)y = f \text{ for } \alpha < x < \beta$$

The general form of the boundary conditions is:

$$\mathcal{B}_k y = \sum_{j=1}^{n} b_{jk}^{(\alpha)} y^{(j-1)}(\alpha) + b_{jk}^{(\beta)} y^{(j-1)}(\beta)$$

Remark. The exponents (α) and (β) above the coefficients b_{jk} denote to which endpoint α, or β the coefficient belongs. They **do not** denote derivatives.

Existence and Uniqueness

Similar to linear systems, the space $\mathcal{V}$ of solutions to the homogeneous ODE $L(x, D)y = 0$ is a linear space of dimension n. Let $y_1, ..., y_n$ be a basis and form the matrix $(\mathcal{B}_k y_i)_{k,i}$, the following theorem provides us with a certain criteria to investigate uniqueness.

Theorem. The following statements are equivalent:

(a) The BVP has a unique solution $\forall f \in \mathcal{C}[\alpha, \beta]$, and $\forall \mathbf{c} \in \mathbb{R}^n$

(b) The homogeneous problem $L(x, D)y = 0$, $\mathbf{B}y = \mathbf{0}$ only has the trivial solution $y(x) \equiv 0$.

(c) $\det(\mathcal{B}_k y_i) \neq 0$

Proof. See page 178 Ordinary Differential Equations by Karl Gustav Andersson (KGA). ■

Solving Boundary Value Problems

The objective is to derive analytical methods to find solutions $y : \mathbb{R} \supset I \to \mathbb{R}$ that satisfy

$$\begin{cases} L(x, D)y = f \\ \mathbf{B}y = \mathbf{c} \end{cases} \tag{1}$$

In other words, not only do we have to make sure that y satisfies the homogeneous version of the equation $L(x, D)y = 0$, but also ensure that y fits $L(x, D)y = f$ as well as the boundary conditions. To keep track of all such requirements, we exploit the linearly of the differential operator $L(x, D)$ and break down the question into to parts.

A solution y to (1) can be written as: $y = y_1 + y_2$, where y_1 and y_2 satisfy the BVPs:

$$\text{(II)} \quad \begin{cases} L(x, D)y = f \\ \mathbf{B}y = \mathbf{0} \end{cases} \qquad\qquad \text{(III)} \quad \begin{cases} L(x, D)y = \mathbf{0} \\ \mathbf{B}y = \mathbf{c} \end{cases}$$

Since $L(x, D)[y_1 + y_2] = L(x, D)y_1 + L(x, D)y_2 = f + \mathbf{0} = f = L(x, D)y$
And $\mathbf{B}[y_1 + y_2] = \mathbf{B}y_1 + \mathbf{B}y_2 = \mathbf{0} + \mathbf{c} = \mathbf{c} = \mathbf{B}y$

Equation II, having solution y_2 is easy to solve (That is, of course, provided that the coefficients a_j are constants). Similar to initial value problems, we have a homogeneous equation $L(x, D)y = 0$ for which we find linearly independent solutions $y_1, \ldots, y_n$ spanning the solution space $\mathcal{V}$.

$$y(x) = \sum_{k=1}^{m} p_k(x)e^{\lambda_k x}$$

where the coefficients p_k are determined by plugging-in the boundary conditions $\mathbf{B}y = \mathbf{c}$

We now shift our focus on solving the inhomogeneous system III with homogeneous boundary conditions.

Green's Function

Remark. An initial value problem (IVP) can be regarded as a boundary value problem with only one endpoint. Indeed, if one only had some information about the behavior of the function $y(x)$ at a point α in space, say $y(\alpha)$ and all its derivatives up to order $n-1$ were 0, then the equation (1) would have become:

$$\begin{cases} L(x, D)y = f \ x \geq \alpha \\ y(\alpha) = y'(\alpha) = \ldots = y^{(n-1)}(\alpha) = 0 \end{cases} \tag{2}$$

which is precisely in the form of an initial value problem. In the previous section, we found the unique solution to such a problem. Denoting the fundamental solution of L by $E(x, \xi)$, the unique solution to (2) is given by:

$$y(x) = \int_{x_0}^{x} E(x, \xi)f(\xi)d\xi$$

where $E(x, \xi)$ is a function that uniquely solves:

$$\begin{cases} L(x, D)u = 0 \\ u(\xi) = u'(\xi) = \ldots = u^{(n-2)}(\xi) = 0 \\ u^{(n-1)}(\xi) = \dfrac{1}{a_n(\xi)} \end{cases}$$

Now, in the context of boundary value problems, we are interested in the behavior and rate of change of a function $y(x)$ over a specified domain $[\alpha, \beta]$. Hence, it would make more sense to consider the fundamental solution only in that particular region and essentially disregard any contribution from the inhomogeneous term that occurs at points before our starting point. We can therefore set the following constraints on the fundamental solution:

Let

$$F(x, \xi) = \begin{cases} E(x, \xi) \text{ for } x \geq \xi \\ 0 \text{ for } x < \xi \end{cases}$$

Now $\int_\alpha^\beta F(x,\xi)f(\xi)d\xi$ would make sure we don't compute more information than we need. However, we would also want our solution to satisfy boundary conditions. Recall that the linearly independent solutions $y_1, y_2, \ldots, y_n$ to the homogeneous version of this problem $L(x,D)y = 0$ span a linear space $\mathcal{V}$ of dimension n, and that *every other* solution is a linear combination of them. Hence, adding a *suitable* linear combination of $y_1, \ldots, y_n$ to $F(x,\xi)$ would keep the resulting solution in the space $\mathcal{V}$ while ensuring that it also satisfies the boundary conditions $\mathbf{B}y = \mathbf{0}$. Formalizing this idea, we define the following function.

Theorem. Let $G(x,\xi) = F(x,\xi)+d_1(\xi)y_1(x)+\ldots+d_n(\xi)y_n(x)$ be the modified version of the previously calculated (constrained) fundamental solution $F(x,\xi)$. Then:

$$y(x) = \int_\alpha^\beta G(x,\xi)f(\xi)d\xi$$

uniquely satisfies the linear system (III), and $G(x,\xi)$ is called the **Green's Function**.

Why are the coefficients d_i functions of ξ?
The coefficients $d_1(\xi), d_2(\xi), \ldots, d_n(\xi)$ depend on ξ because they are chosen to ensure that the resulting function $G(x,\xi)$ satisfies the boundary conditions at each point ξ within the interval $[\alpha,\beta]$. Note that ξ is a *dummy variable*. We could have written it t, u, or anything else. Its job is to loop through values in the interval $[\alpha,\beta]$. One could think of ξ as a "representative" of the boundaries α and β.

In the following subsections, we introduce a powerful framework for solving BVPs with homogeneous boundary conditions.

Green's Integral Operator

Let's review the machinery and formalism that we have developed up until this point: From this point, we'd like to work with a particular class of functions.