Ordinary Differential Equations

Muchang Bahng

Fall 2021

Contents

1   Systems

The theory of ordinary differential equations allows us to study all evolutionary processes satisfying three properties.

1. Deterministic; that is, the entire past and future is determined by its present state.

2. Finite-dimensionality; that is, the number of parameters needed to describe any state is finite.

3. Differentiability; that is, the phase space of has the structure of a differentiable manifold.

The motion of a system in classical mechanics can be described using ordinary differential equations. However, quantum mechanics, impact theory, fluid mechanics, and heat transfer do not satisfy all three properties.

1.1   Phase Spaces, Phase Flows

Definition 1.1

Given a system that can be represented by a finite number of 1-dimensional parameters $x_1, x_2, ..., x_n$, each in their respective spaces $X_1, X_2, ..., X_n$, respectively, the *phase space* of the system is the set

$$S \equiv \prod_{i=1}^{n} X_i$$

Notice that each element $s \in S$ represents one specific state of the system. While obvious, it should explicitly be stated that
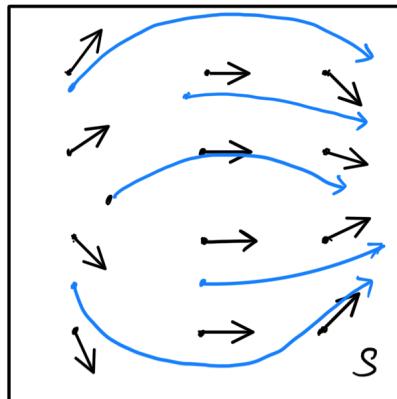
$$\dim S = n$$

The abstract notion of a phase space is extremely useful, since we can now model states of the system as points in S. As stated before, we can form a general theory and treat S as a differentiable manifold, but for simplicity, we will treat $S = \mathbb{R}^n$.

The motion of the entire system can be described by the motion of a point over a curve in the phase space. Since the system is deterministic, the entire motion over this curve is determined by the point itself.

Definition 1.2

Therefore, we can assign a velocity vector at each point $s \in S$ that models the motion of the point, called the *phase velocity vector*. The set of all phase velocity vectors in S is called the *phase velocity vector field* in the phase space S. This vector field defines the differential equation of the process.

Since we are trying to model the motion of a point through S, it is often convenient to paramaterize

---
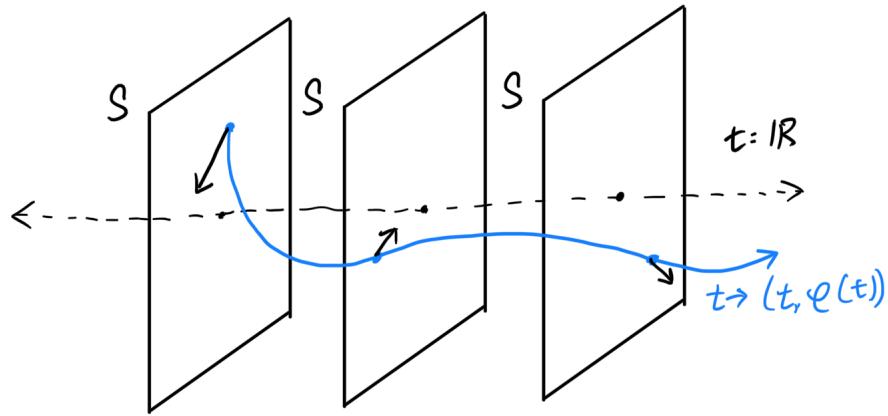
the motion as a curve with the time parameter t. This creates a well-defined curve (blue curves in the figure above) in a new space that now has a time-axis.

Definition 1.3

Given a phase space S, the *extended phase space* of a system is the space

$$S \oplus \mathbb{R}$$

where $\mathbb{R}$ represents the new time axis, with parameter t. Note that within this theory, negative values of t are well-defined, and $\dim(S \oplus \mathbb{R}) = n + 1$.
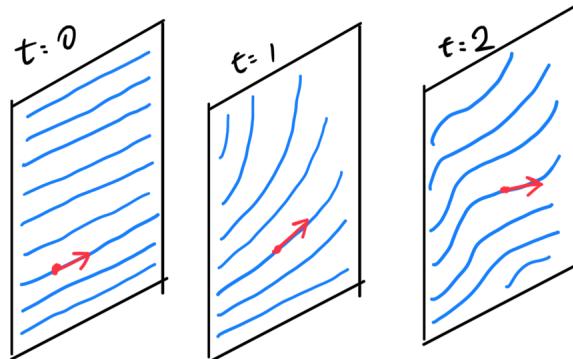
The parameterized curve, called the *integral curve*, is defined with the mapping

$$t \mapsto (t, \varphi(t))$$

where φ is the path of the particle through S.

But rather than focusing on the integral curve itself, it is useful to visualize the entire phase velocity field continuously (since we've assumed differentiability) "morphing" with respect to time. This represents a change in the entire system as time passes by.
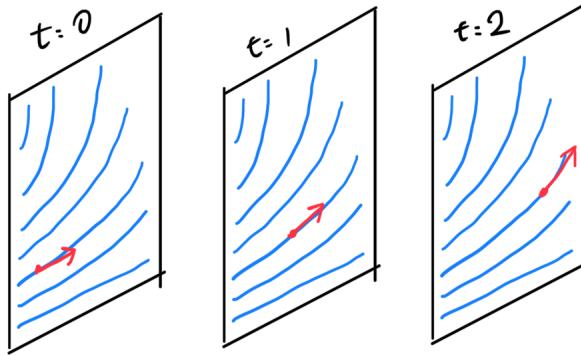
With this visual in mind, we can see that there may be multiple integral curves that can go through $S \oplus \mathbb{R}$. To determine a unique integral curve, it may be necessary to define an *initial point* that represents an initial condition.

However, there may be some systems that are invariant under time. For example, swinging a pendulum or the movement of a particle is only dependent on the position and velocity of the pendulum and particle. Whether we let go of the pendulum at $t = 0$ or $t = 1$ does not matter. This is what we call an *autonomous* system.
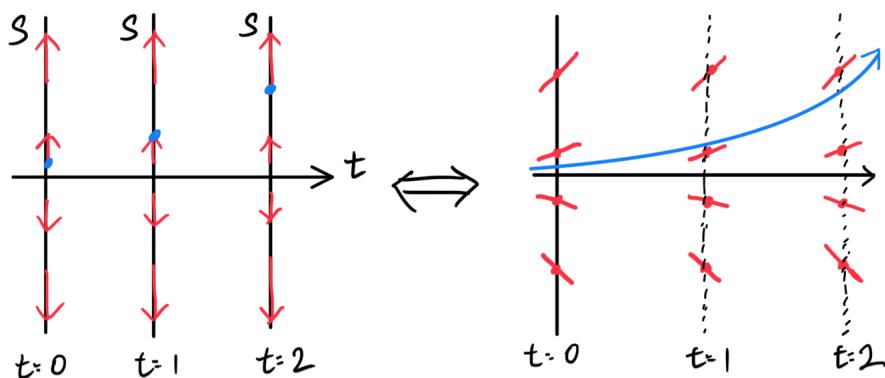
Definition 1.4

An *autonomous*, or *time-homogeneous*, system is a system where the phase space does not have a time parameter. We can visualize this type of system as one where the phase velocity vector does not morph at all.
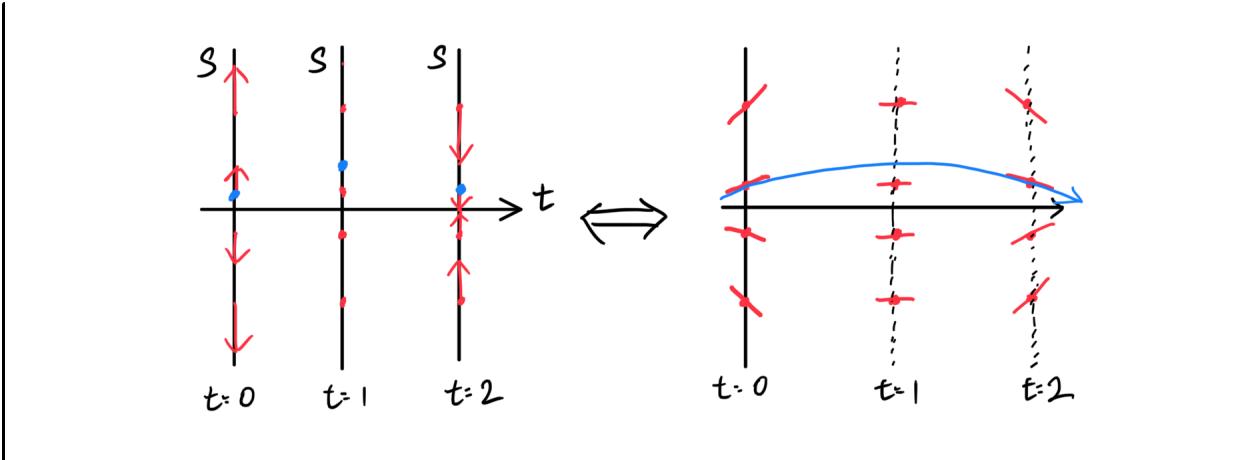
Example 1.1 (1-Dimensional Autonomous System)

Another special type of system is one that produces a one dimensional *constant* phase field that changes in the same way everywhere with respect to time. We can visualize it as a vector field in $\mathbb{R}$ that stays constant as time passes (left). Since it may be hard to visualize vectors in $\mathbb{R}$, we can view them as slopes.
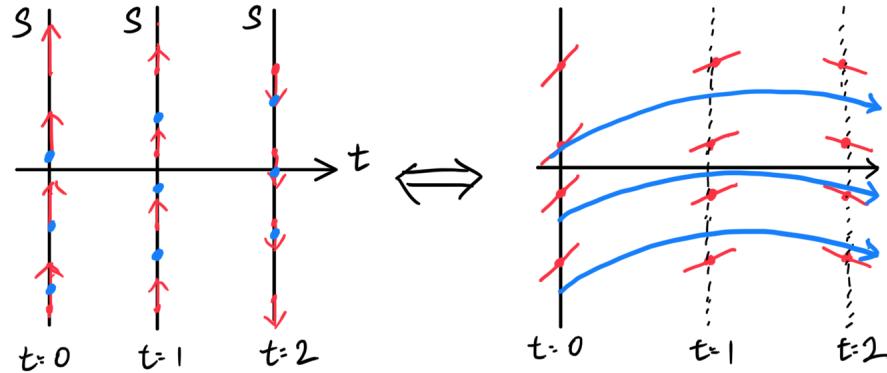
Example 1.2 (1-Dimensional General System)

A one dimensional system can be visualized as the initial vector field of $\mathbb{R}$ at $t = 0$ gradually morphing as time passes.

Example 1.3 (1-Dimensional Non-Autonomous Constant System)

Let ν be a constant, one-dimensional non-autonomous phase velocity field that changes with respect to time t.
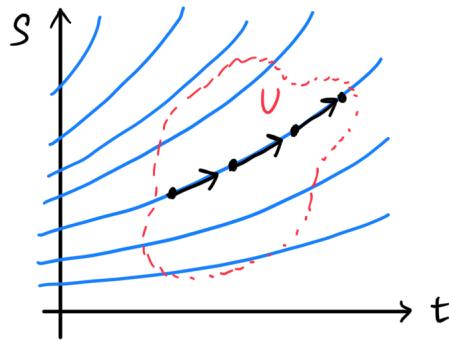
Then, the function φ in which its graph $\big(t, \varphi(t)\big)$ precisely overlaps the integral curves of ν can be computed by *Barrow's formula*

$$\varphi(t) = x_0 + \int_{t_0}^{t} \nu(\tau)\, d\tau$$

Where (t_0, x_0) is the initial condition of the system, and τ is a dummy variable representing time.

It is also worthwhile to note that we can define the function to exist within a certain subset of the phase space and to be defined over a certain time interval within $\mathbb{R}$, rather than requiring it to be defined on the whole space itself. We will denote the time interval $I \subset \mathbb{R}$ and the subset $U \subset S$.

1.1.1 Dimensionality vs Order

Definition 1.5 (Dimensionality of a System)

The *dimensionality* of the phase space of a system is dependent upon many things:
1. the dimension of the space that we are working in
2. the number of particles or bodies that we observe
3. other factors, like orientation, that may add extra parameters

But colloquially, the more things we have to keep track in the system, the higher the dimensionality.
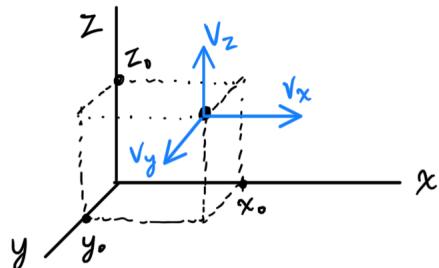
Definition 1.6 (Order of a System)

Formally, the *order* of a differential equation is the highest order derivative that is contained within the equation. It has nothing to do with how many moving parts there are (i.e. dimensionality) and more with the rule that each moving part follows.
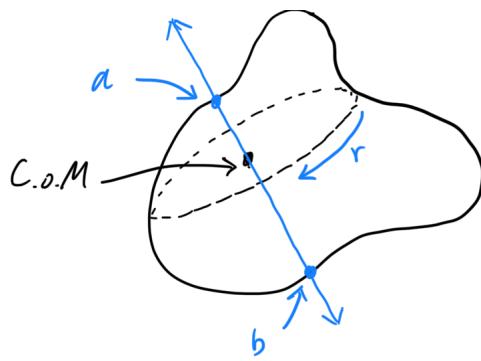
Example 1.4

We will state a couple examples of how many dimensions certain systems would have:
1. A point particle moving in three dimensional space in a constant force field would have 3 spatial dimensions (each represented by a second order differential equation). Note that a constant force field means that acceleration is also constant, so acceleration is not taken into account.
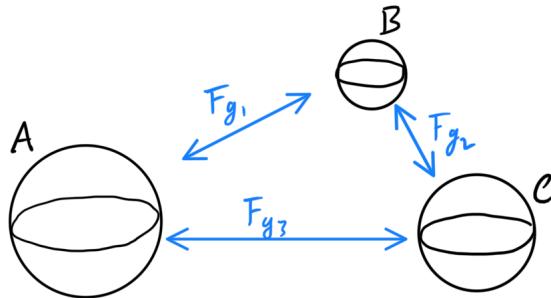
2. A rigid body in 3-space has 6 parameters (3 for its center of mass and 3 for its orientation).

Therefore, a system of 3 rigid bodies in 3-space has a total of $6 \times 3 = 18$ dimensions, where each body is individually following $F = ma$ where the force F is given by the gravitational force. This gives a second order differential equation for each body in terms of the positions of the other bodies.

1.2　First Order Differential Equations

The simplest types of differential equations are systems that contain parameters displacement and first derivative (velocity). Note that this does not necessarily mean that the dimension of the phase space is 1.

Definition 1.7 (n-dimensional First Order Differential Equation)

An *n-dimensional first order differential equation* is an equation of the form

$$y' = f(t, y)$$

where y is an n-dimensional vector representing the displacement parameters of the system, t is the time parameter, and

$$f(t_0, \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is an n-dimensional vector field that determines the phase velocity field of S at time $t = t_0$. It can also be written explicitly as a system as

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} f_1(t, y) \\ f_2(t, y) \\ \vdots \\ f_n(t, y) \end{pmatrix}$$

This equation defines an n-dimensional phase velocity field that morphs as time passes, which is used to model the movement of the n-dimensional system. Simply put, the velocity vector that determines the evolution of the system is completely dependent on the current state of the system (value of y) and the time (value of t).

The solution to this system is a smooth map $\varphi : I \subset \mathbb{R} \longrightarrow U \subset \mathbb{R}^n$ defined by the integral curve $t \mapsto (t, \varphi(t))$ in the extended phase space that satisfies the equation
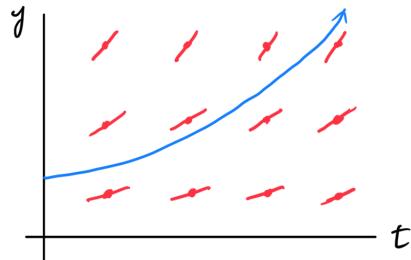
$$\varphi'(t) = f\big(t, \varphi(t)\big)$$

1.2.1 Basic One-Dimensional Examples

Example 1.5 (Normal Reproduction)

Assume that the size of a biological population is y and that the rate of reproduction is proportional to the number of organisms present. This is expressed by the first order differential equation of dimension 1

$$y' = ky, \quad k > 0$$

The solution to the equation, given initial conditions (t_0, y_0) is
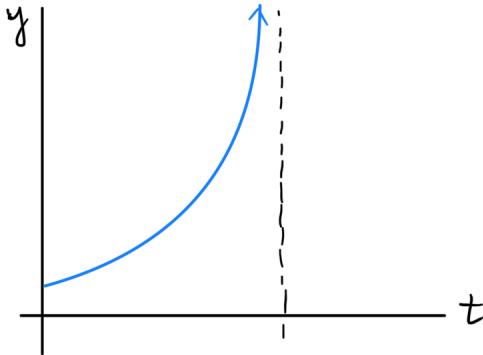
$$\varphi(t) = e^{k(t-t_0)} y_0$$

Example 1.6 (Explosion Equation)

If we assume that the rate of reproduction is proportional to the number of pairs of individuals, we get the differential equation

$$y' = ky^2$$

The integral curves of this equation shoot off to infinity in a finite time, while the integral curves of the normal reproduction system reaches infinite when t is infinite.
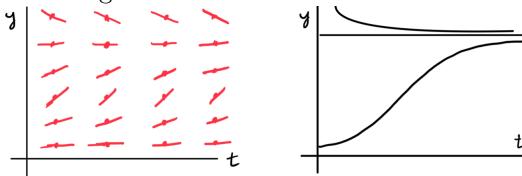
Example 1.7 (Logistic Curve)

Due to limiting factors that may restrict the rate of growth of a population, we modify the equation to the following *logistic equation.*

$$x' = (1 - x)x$$
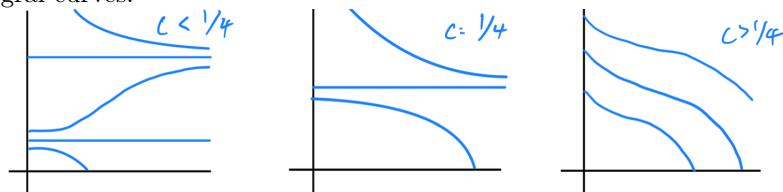
Both the direction field and the integral curves are shown.

Example 1.8 (Harvest Quotas)

Now, we model a situation where we harvest a part of the population. Let us first assume that the rate of harvest is constant. This leads to the differential equation

$$x' = (1 - x)x - c$$
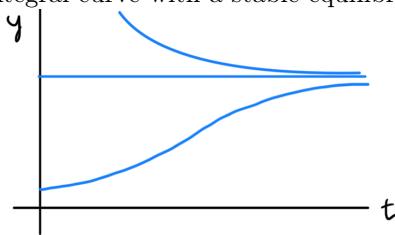
The quantity c represents the rate of harvesting and is called the *quota*. For different values of c, this results in the integral curves.

However, if we harvest with a relative quota, represented by the equation

$$x' = (1 - x)x - px$$

for $0 < p < 1$, this leads us to an integral curve with a stable equilibrium point.

1.2.2   Additional Examples

Example 1.9 (Lotka-Volterra Model)
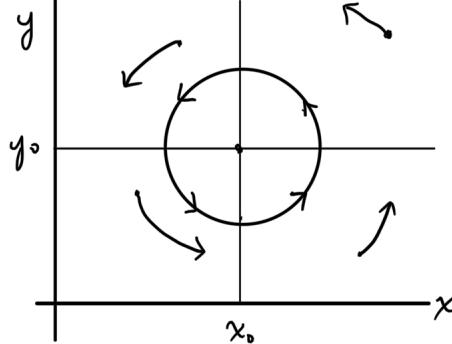
The simplest model describing the predator-prey relationship between two species is described by the system of differential equations (also a 2-dimensional first order differential equation).
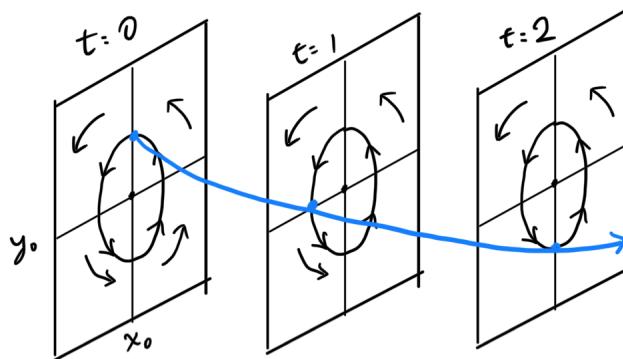
$$x' = kx - axy$$
$$y' = -ly + bxy$$

The phase space with its velocity vector field can be defined by defining the vector

$$\begin{pmatrix} kx - axy \\ -ly + bxy \end{pmatrix}$$

at each point in $\mathbb{R}^2$, labeled with the x and y axes.

Note that we have graphed the model such that the two parameters x, y are shown and not the time, but note that the time parameter t very much exists in this autonomous model.

The blue curve is the solution to this system, which describes the evolution of the autonomous system as time passes.

Example 1.10 (Pendulums)

Small oscillations of a pendulum can be approximated, using Taylor approximations, to get the differential equation

$$x'' = -kx$$
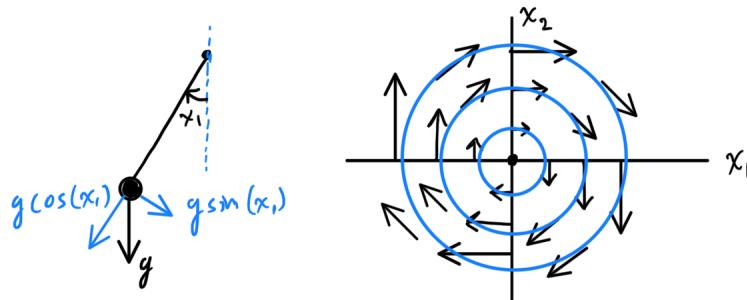
By scaling the coefficient k, we can make it equal to 1 to get

$$x'' = -x$$

The phase space of this system is 2 dimensional, with parameters $x_1 = x$ and $x_2 = x'$. This creates the autonomous system of first order differential equations

$$x_1' = x_2, \ x_2' = -x_1$$
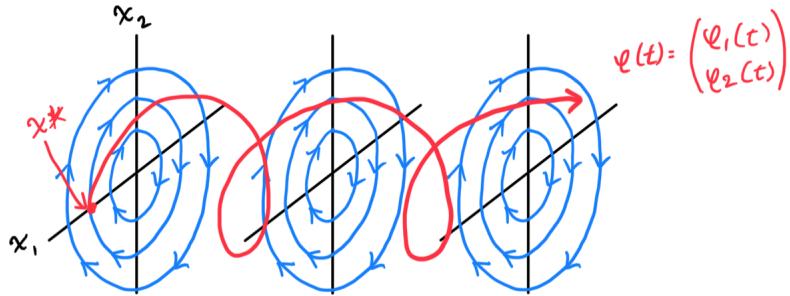
This leads to the phase velocity vector field shown below.

A more accurate equation for the undampened pendulum system (by scaling k again) is

$$\theta'' = -\sin\theta$$
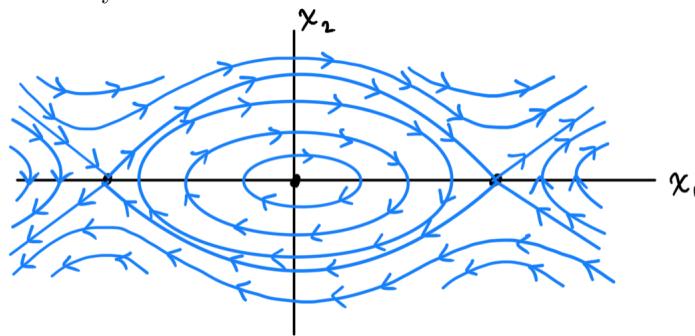
Clearly, the solution to this system, given initial value of (x_1^*, x_2^*), is a helix.

This equation can be represented as the autonomous system

$$x_1' = x_2, \ x_2' = -\sin x_1$$

which creates the phase velocity vector field

Example 1.11 (Small Oscillations of Spherical Pendulums)

A spherical pendulum adds another dimension of movement (another dimension of angle and angular velocity), therefore inducing a system where the phase space is 4 dimensional. For small oscillations, the differential equations

$$x'' = -x, \ y'' = -y$$

is represented with the system

$$x_1' = x_2, \ x_2' = -x_1, \ x_3' = x_4, \ x_4' = -x_3$$

where $x_1 = x, x_2 = x', x_3 = y, x_4 = y'$.

1.3 Existence Theory

Before we introduce methods to solve differential equations, we will state and prove conditions for existence and uniqueness of solutions. We will only prove existence in the simplest case.

Theorem 1.1 (Existence of Solutions in a First-Order System with Phase Space Dimensionality of 1)

Let f and $\partial f / \partial y$ be C^0 functions in a given region $U \in \mathbb{R}^2$, with $(t_0, y_0) \in D$. This means that Then, there exists an interval I containing t_0 and exactly 1 solution φ of the differential equation

$$y' = f(t, y)$$

passing through (t_0, y_0). The solution exists for values of t for all points $(t, \varphi(t)) \in U$. φ is also a continuous function with respect to (t, t_0, y_0).

In order to state the proof, we will introduce some lemmas.

Lemma 1.2

φ is a solution to $y' = f(t, y)$ with $\varphi(t_0) = y_0$ on interval I if and only if φ satisfies the solution y of the equation

$$y(t) = y_0 + \int_{t_0}^{t} f\big(s, y(s)\big) \, ds, \ \ t \in I \tag{1}$$

Proof. ($\rightarrow$) φ is a solution of $y' = f(t, y)$, with $\varphi(t_0) = y_0 \implies \varphi'(t) = f\big(t, \varphi(t)\big)$

$$\implies \int_{t_0}^{t} \varphi'(s) \, ds = \int_{t_0}^{t} f\big(s, \varphi(s)\big) \, ds$$

$$\implies \varphi(t) = \varphi(t_0) + \int_{t_0}^{t} f\big(s, \varphi(s)\big) \, ds = y_0 + \int_{t_0}^{t} f\big(s, \varphi(s)\big) \, ds$$

($\leftarrow$) $y(t)$ is a solution of (1) $\implies y(t)$ is continuous by the fundamental theorem of calculus. Since f is also continuous, $y(t)$ is differentiable.

$$\implies y'(t) = \frac{d}{dt} \left(y_0 + \int_{t_0}^{t} f\big(s, y(s)\big) \, ds \right) = f\big(t, y(t)\big)$$

Putting $t = t_0, y(t_0) = y_0$, this implies that y is a solution of $\varphi'(t) = f(t, y)$.

The previous lemma allows us to establish the existence of a solution using (1), which is generally easier. We now define the Lipshitz inequality for functions of two variables.

Definition 1.8

Let $f(t, y)$ be a continuous, bounded function in region D. For all $(t, y_1), (t, y_2) \in D$, if f satisfies the inequality

$$|f(t, y_2) - f(t, y_1)| \leq K|y_2 - y_1|$$

then f is said to satisfy the *Lipshitz condition* in D.

Now, we define a series of successive approximations

$$\varphi_0(t) = y_0, \ \varphi_{j+1}(t) = y_0 + \int_{t_0}^{t} f\big(s, \varphi_j(s)\big) \, ds$$

We hope that this sequence will eventually converge onto the actual φ, but we must properly define the φ_j's such that the points remain in the region D.

Lemma 1.3

Let $\alpha = \min\{a, \frac{b}{M}\}$. Then, the successive approximations φ_j are defined on the interval I given by $|t - t_0| < \alpha$, and on this interval

$$|\varphi_j(t) - y_0| \le M|t - t_0| < b, \ \ j = 0, 1, 2, ... \tag{2}$$

Proof. Clearly, $\varphi_0(t)$ is defined on I and satisfies (2). Now assuming that for $n \ge 1$, φ_n is defined and satisfies (2) $\implies$ the point $\big(t, \varphi(t)\big) \in D$ for $t \in I$. To prove that $\big(t, \varphi_{n+1}(t)\big) \in D$, we already know $t \in I$, so we must show that $|\varphi_{n+1}(t) - y_0| < b$. Using the intermediate value theorem,

$$|\varphi_{n+1}(t) - y_0| = \left| \int_{t_0}^{t} f\big(s, \varphi_j(s)\big) \, ds \right|$$

$$\le \left| \int_{t_0}^{t} |f\big(s, \varphi_j(s)\big)| \, ds \right|$$

$$\le M|t - t_0| < M\alpha < \le b$$

We specifically assign α because first, (2) implies that the successive approximations also have slopes bounded by the cone of the lines with slope M and $-M$ through points (t_0, y_0). The length α of I depends on where these lines meet D. Either way, the φ_j's remain in the rectangles.

Now, we can prove an existence theorem.

Theorem 1.4

Suppose f and $\partial f / \partial y$ are continuous and bounded on the rectangle D satisfying the Lipshitz condition. That is,

$$|f(t, y)| \le M, \ \ \left| \frac{\partial f}{\partial y} \right| \le K$$

Then, the sucessive approximations φ_j (defined previously) converge uniformly on the interval $I = (t_0 - \alpha, t_0 + \alpha)$ to a solution φ of the differential equation $y' = f(t, y)$ with $\varphi(t_0) = y_0$.

Proof. The second lemma shows that φ_j are defined on the interval I. Now, we define the difference between φ_j adn φ_{j+1} in the interval $[t_0, t_0 + \alpha]$ and $[t_0 - \alpha, t_0]$. Using the Lipshitz condition,

$$r_j(t) = |\varphi_{j+1}(t) - \varphi_j(t)| = \left| \int_{t_0}^{t} f\big(s, \varphi_j(s)\big) - f\big(s, \varphi_{j-1}(s)\big) \, ds \right|$$

$$\le \int_{t_0}^{t} |f\big(s, \varphi_j(s)\big) - f\big(s, \varphi_{j-1}(s)\big)| \, ds$$

$$\le K \int_{t_0}^{t} \varphi_j(s) - \varphi_{j-1}(s) \, ds$$

$$= K \int_{t_0}^{t} r_{j-1}(s) \, ds, \ \ j = 1, 2, ...$$

When $j = 0$,

$$r_0(t) = |\varphi_1(t) - \varphi_0(t)| = \left| \int_{t_0}^t f(s, \varphi(s)) \, ds \right|$$

$$\leq \int_{t_0}^t |f(s, \varphi(s)) \, ds \leq M(t - t_0)$$

Now, using induction, we prove that

$$r_j(t) \leq \frac{MK^j(t - t_0)^{j+1}}{(j+1)!}, \quad j = 0, 1, 2, ...; t_0 \leq t \leq t_0 + \alpha$$

The base case suffices. When $j = 0$,

$$r_0(t) \leq M(t - t_0)$$

Now assume that the statement is true for $j = p - 1$, for $p \geq 1$. Then,

$$r_p(t) \leq K \int_{t_0}^t r_{p-1}(s) \, ds \leq K \int_{t_0}^t \frac{MK^{p-1}(s - t_0)^p}{p!} \, ds$$

$$= \frac{MK^p}{p!} \cdot \left(\frac{(s - t_0)^{p+1}}{p+1} \bigg|_{t_0}^t \right)$$

$$= \frac{MK^p(s - t_0)^{p+1}}{(p+1)!}, \quad t_0 \leq t \leq t_0 + \alpha$$

Integrating this fact into the main proof, we have

$$r_j(t) \leq \frac{MK^j|t - t_0|^{j+1}}{(j+1)!} = \frac{M(K|t - t_0|)^{j+1}}{K(j+1)!}$$

$$< \frac{M\alpha^{j+1}K^{j+1}}{K(j+1)!}$$

$$= \frac{M(K\alpha)^{j+1}}{K(j+1)!}$$

This implies that

$$\sum_{j=0}^\infty r_j(t) = \frac{M}{K} \sum_{j=0}^\infty \frac{(K\alpha)^{j+1}}{(j+1)!}$$

$$= \frac{M}{K} e^{\alpha K}$$

This implies that

$$\sum_{j=0}^\infty \left(\varphi_{j+1}(t) - \varphi_j(t) \right)$$

absolutely and uniformly converges in the interval I where $|t - t_0| < \alpha$ to some function of t, determined arbitrarily by $\varphi(t)$.

To see the rate and error bound of convergence, we already defined

$$\varphi(t) = \varphi_0(t) + \sum_{n=0}^{\infty} \big(\varphi_{n+1}(t) - \varphi_n(t)\big)$$

$$\implies \varphi(t) - \varphi_j(t) = \sum_{n=0}^{\infty} \big(\varphi_{n+1}(t) - \varphi_n(t)\big)$$

$$\implies |\varphi(t) - \varphi_j(t)| \le \sum_{n=j}^{\infty} |\varphi_{n+1}(t) - \varphi_n(t)|$$

$$\le \sum_{n=j}^{\infty} r_n(t) \le \frac{M}{K} \sum_{n=j}^{\infty} \frac{(K\alpha)^{n+1}}{(n+1)!}$$

$$\le \frac{M}{K} \frac{(K\alpha)^{j+1}}{(j+1)!} \sum_{n=0}^{\infty} \frac{(K\alpha)^n}{n!} = \frac{M}{K} \frac{(K\alpha)^{j+1}}{(j+1)} e^{K\alpha}$$

It is clear that

$$\epsilon_j = \frac{(K\alpha)^{j+1}}{(j+1)!} \to 0 \text{ as } j \to \infty$$

To prove continuity of $\varphi(t)$ on I, we let $\epsilon > 0$ and assign

$$\varphi(t+h) - \varphi(t) = \varphi(t+h) - \varphi_j(t+h) + \varphi_j(t+h) - \varphi_j(t) + \varphi_j(t) - \varphi(t)$$

which, by the triangle inequality, implies that

$$\implies |\varphi(t+h) - \varphi(t)| \le |\varphi(t+h)\varphi_j(t+h)| + |\varphi_j(t+h) - \varphi_j(t)| + |\varphi_j(t) - \varphi(t)|$$
$$\le 2\varepsilon_j + |\varphi_j(t+h) - \varphi_j(t)|$$

Choosing j sufficiently large and $|h|$ sufficiently small, using

$$\lim_{j \to \infty} \varepsilon_j = 0$$

and continuity of $\varphi_j(t)$, we can make

$$|\varphi(t+h) - \varphi(t)| \le |\varphi_j(t+h) - \varphi_j(t)| < \varepsilon, \text{ as } h \to 0$$

$\implies \varphi(t)$ is continuous since the limit exists at t that matches the actual value.

To show that the $\varphi(t)$ satisfies the integral equation in the first lemma, we see that

$$\lim_{j \to \infty} \varphi_j(t) = \varphi(t) \implies \lim_{j \to \infty} \int_{t_0}^{t} f\big(s, \varphi_j(s)\big)\, ds = \int_{t_0}^{t} f\big(s, \varphi(s)\big)\, ds$$

$$\implies \lim_{j \to \infty} \varphi_{j+1}(t) = \lim_{j \to \infty} \left(y_0 \int_{t_0}^{t} f\big(s, \varphi_j(s)\big)\, ds \right)$$

$$\implies \varphi(t) = y_0 + \int_{t_0}^{t} f\big(s, \varphi_((s)\big)\, ds$$

Using the first lemma, we see that φ therefore satisfies the differential equation $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$.

Corollary 1.5

The error committed by $\varphi_j(t)$ satisfies the estimate

$$|\varphi(t) - \varphi_j(t)| \leq \frac{M}{K} \frac{(K\alpha)^{j+1}}{(j+1)!} e^{K\alpha} \quad \forall t \in I$$

In fact, the conditions of this theorem implies the uniqueness of solutions of $y' = f(t, y)$. However, we can prove existence without uniqueness.

Theorem 1.6

Suppose $f \in C^0$ in rectangle R with f bounded; that is, $|f(t, y)| \leq M$ for all $(t, y) \in R$. Let $\alpha = \min\{a, b/M\}$. Then, there exists a solution φ of differential equation

$$y' = f(t, y)$$

with $\varphi(t_0) = y_0$ existing on interval $|t - t_0| < \alpha$.

Now, we prove the uniqueness of solutions.

Lemma 1.7 (Gronwall Inequality)

Let $K > 0$ be a constant, and $f, g \in C^0$ be nonnegative on interval $\alpha \leq t \leq \beta$ satisfying

$$f(t) \leq K + \int_\alpha^t f(s)\, g(s)\, ds \quad \forall t \in [\alpha, \beta]$$

Then,

$$f(t) \leq K \exp\left(\int_\alpha^t g(s)\, ds \right) \quad \forall t \in [\alpha, \beta]$$

Proof. Let

$$U(t) = K + \int_\alpha^t f(s)\, g(s)\, ds \implies U(\alpha) = K$$

Then, $f(t) \leq U(t)$, which implies, by the fundamental theorem of calculus and $g(t) \geq 0$, that

$$U'(t) = f(t)\, g(t) \leq U(t)\, g(t) \quad (\alpha \leq t \leq \beta)$$
$$\implies U'(t) - U(t)\, g(t) \leq 0$$

---

which implies

$$U'(t)\left(-\exp\left(\int_\alpha^t g(s)\,ds\right)\right) - U(t)\,g(t)\left(-\exp\left(\int_\alpha^t g(s)\,ds\right)\right)$$

$$= \left(U(t)\left(-\exp\left(\int_\alpha^t g(s)\,ds\right)\right)\right)' \leq 0$$

$$\implies \int_\alpha^t \frac{d}{dr}\left(U(r)\left(-\exp\left(\int_\alpha^r g(s)\,ds\right)\right)\right)\,dr$$

$$= U(t)\exp\left(-\int_\alpha^t g(s)\,ds\right) - U(\alpha) \leq 0$$

$$\implies f(t) \leq U(t) \leq U(\alpha)\exp\left(\int_\alpha^t g(s)\,ds\right) = K\exp\left(\int_\alpha^t g(s)\,ds\right)$$

Theorem 1.8

Let $f, \partial f/\partial y \in C^0$ be bounded in rectangle $R \equiv \{(t,y) \mid |t - t_0| < \alpha, |y - y_0| < b\}$. Then there exists at most one solution of $y' = f(t,y)$ satisfying $\varphi(t_0) = y_0$.

Proof. Since the hypothesis of this theorem establishes the existence of at least one solution φ, it suffices to prove that any two solutions are equal. Assume that φ_1, φ_2 are two different solutions on a common interval J. By the first lemma, we have for all $t \in J$

$$\varphi_1(t) = y_0 + \int_{t_0}^t f\big(s, \varphi_1(s)\big)\,ds$$

$$\varphi_2(t) = y_0 + \int_{t_0}^t f\big(s, \varphi_2(s)\big)\,ds$$

which implies (using the Lipshitz condition) that

$$\varphi_2(t) - \varphi_1(t) = \int_{t_0}^t f\big(s, \varphi_2(s)\big) - f\big(s, \varphi_1(s)\big)\,ds$$

$$\implies |\varphi_2(t) - \varphi_1(t)| \leq \left|\int_{t_0}^t |f\big(s, \varphi_2(s)\big) - f\big(s, \varphi_1(s)\big)|\,ds\right|$$

$$\leq K\left|\int_{t_0}^t |\varphi_2(s) - \varphi_1(s)|\right|$$

Using Gronwall's inequality and setting $K = 0, g(s) = 1$, and $f(t) = |\varphi_2(t) - \varphi_1(t)|$, we have

$$\leq K\left|\int_{t_0}^t |\varphi_2(s) - \varphi_1(s)|\right| \leq 0 \cdot \exp\left(\int_\alpha^t 1\,ds\right) = 0$$

$$\implies |\varphi_2(t) - \varphi_1(t)| = 0 \implies \varphi_1 = \varphi_2, \text{ a contradiction.}$$

Finally, we prove continuity of f with respect to initial conditions.

Theorem 1.9

let $f, \partial f/\partial y \in C^0$ be bounded in region D and satisfy the Lipshitz condition. Let φ be the solution of $y' = f(t, y)$ with $\varphi(t_0) = y_0$ and let ψ be the solution of $y' = f(t, y)$ with $\psi(\hat{t}_0) = \hat{y}_0$. Assume that φ and ψ both exist on interval $[a, b]$. Then for each $\varepsilon > 0$, there exists a $\delta > 0$ such that if $|t - \hat{t}| < \delta$ and $|y - \hat{y}| < \delta$, then
$$|\varphi(t) - \psi(\hat{t})| < \varepsilon, \ \ a < t, \hat{t} < b$$

2　Methods of Solution

2.1　Basic Methods for First Order Scalar-Valued DEQs

2.1.1　Variables Separable

Definition 2.1 (Separable DEQs and Solution)

A differential equation in the form
$$y' = g(t)h(y)$$
is called a *variables-separable equation.*
Assume that given solution φ, $h\big(\varphi(t)\big) \neq 0$ for all $t \in I$. Then, we rearrange the equation and integrate either indefinitely or definitely, where $y = \varphi(t)$. Note that there is no difference between finding solutions using indefinite or definite integration.

$$\varphi'(t) = g(t)h(\varphi(t)) \implies \frac{y'}{h(y)} = g(t)$$

1. Indefinitely, we can use substitution $y = \varphi(t)$ to get

$$\int \frac{\varphi'(t)}{h(\varphi(t))}\, dt = \int g(t)\, dt \implies \int \frac{1}{h(y)}\, dy = \int g(t)\, dt$$

This can be remembered with the mnemonic

$$\frac{dy}{dt} = g(t)h(y) \implies \frac{1}{h(y)}dy = g(t)dt \implies \int \frac{1}{h(y)}dy = \int g(t)dt$$

2. Definitely, we can use the same substitution $y = \varphi(s)$, where $y_0 = \varphi(t_0)$, to get

$$\int_{t_0}^{t} \frac{\varphi'(s)}{h(\varphi(s))}\, ds = \int_{t_0}^{t} g(t)\, dt \implies \int_{y_0}^{y} \frac{1}{h(\tilde{y})}\, d\tilde{y} = \int_{t_0}^{t} g(\tilde{t})\, d\tilde{t}$$

where we can treat $\tilde{y}$ and $\tilde{t}$ as dummy variables.
Both methods define the solution implicitly, but it may or may not be possible to solve it explicitly as $y = \varphi(t)$. However, by the implicit function theorem, we can guarantee the existence of a function $y = \varphi(t)$ within a neighborhood of (t_0, y_0).

2.1.2　First Order Linear Equations

Definition 2.2 (First-Order Linear DEQ and Solution)

A differential equation of the form
$$y' + a_1(t)y = b(t)$$
is called a *first order linear differential equation.*
To solve this, we assume that there is a function $\mu(t)$, called the *integrating factor*, with the property

---

that
$$\mu(t)a_1(t) = \mu'(t)$$

This can be calculated with the following derivation $=$

$$\frac{\mu'(t)}{\mu(t)} = a_1(t) \implies \big(\ln \mu(t)\big)' = a_1(t)$$

$$\implies \ln \mu(t) = \int a_1(t)\,dt + k$$

$$\implies \mu(t) = e^{\int a_1(t)\,dt + k} = ke^{\int a_1(t)\,dt}$$

Okay, now what? Well multiplying this integrating factor on both sides conveniently gives us a left hand side in the form of the product rule of differentiation.

$$\mu(t)b(t) = \mu(t)y' + \mu(t)a_1(t)y$$
$$= \mu(t)y' + \mu'(t)y$$
$$= \big(\mu(t)y(t)\big)'$$

Integrating both sides gives us

$$\mu(t)y(t) = \int \mu(t)b(t)\,dt + c$$

Summarizing, the solution, if it exists, has explicit form of

$$y(t) = \frac{1}{\mu(t)}\left(\int \mu(t)b(t)\,dt + c\right), \quad \mu(t) = ke^{\int a_1(t)\,dt}$$

2.1.3 Exact Equations

Definition 2.3 (Exact DEQs and Solutions)

Let $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a C^1 function with partial derivatives ψ_t and ψ_y. A differential equation of the form

$$\psi_t(t, y) + \psi_y(t, y)y' = 0$$

is called an *exact differential equation*. The left hand side can be transformed as such:

$$\psi_t(t, y) + \psi_y(t, y)y' = \frac{d}{dt}\psi\big(t, y(t)\big) = 0$$

This implies that the solution can be written implicitly as

$$\psi\big(t, y(t)\big) = c$$

2.1.4 Bernoulli Differential Equations

Definition 2.4 (Bernoulli DEQs and Solutions)

A differential equation of the form

$$y' + p(t)\,y = q(t)\,y^n$$

is called a *Bernoulli differential equation*. To solve this, we first divide by y^n to get

$$y^{-n}y' + p(t)\,y^{1-n} = q(t)$$

---

and use the substitution $v = y^{1-n}, v' = (1-n)y^{-n}y'$ to plug into the equation and get

$$\frac{1}{1-n}v' + p(t)v = q(t)$$

This is a first order linear differential equation that we can solve for v and solve back for y.

2.2   2nd Order Equations Solvable by 1st Order Methods

A small class of 2nd order differential equations, which have the general form

$$y'' = f(t, y, y')$$

can be reduced by substitution to a system of first order differential equations. There are two types of equations:

1. Equation of form $y'' = g(t, y')$ can be solved with the substitution $p = y'$ to create the new equation

$$p' = g(t, p)$$

   which may be solved for p using any of the first order methods. We can then integrate p again to find y, adjusting the extra constant to satisfy the initial conditions $\varphi(t_0) = y_0$. That is, if $p = \psi(t)$ is the solution,

$$\varphi(t) = y_0 + \int_{t_0}^{t} \psi(s) \, ds$$

2. Equation of form $y'' = h(y, y')$ can be transformed into a pair of first order equations (i.e. a first order vector-valued equation). That is, let φ be the solution such that $\varphi(0) = y_0, \varphi'(0) = z_0$. We substitute

$$p(t) = \varphi'(t) \implies p'(t) = \varphi''(t)$$

   Assuming we found a solution $y = \varphi(t)$, suppose we can invert it to a function of y. That is, $t = s(y)$. This implies that
$$y' = \varphi'(t) = p(t) = p(s(y)) \equiv q(y)$$

   Then,
$$\varphi''(t) = p'(t) = \frac{dq}{dy}\varphi'(t) = \frac{dq}{dy}p(t) = \frac{dq}{dy}q(y)$$

   which implies that the equation can be decomposed to the following system of linear equations

$$\frac{dq}{dy}q(y) = h(y, q(y))$$
$$y' = q(y)$$

   We first find the solution of the first equation (which may not always be possible). If $z_0 = 0$, $q \equiv 0$ may be a solution $\implies \varphi(t) = y_0$. If $q \equiv 0$ is not a solution, we find $q = \psi(t)$, substitute it into the second equation to find the solution (using variables separable), and then verify it by substituting φ into the original second order differential equation.

2.3   Numerical Methods: Euler's Algorithm
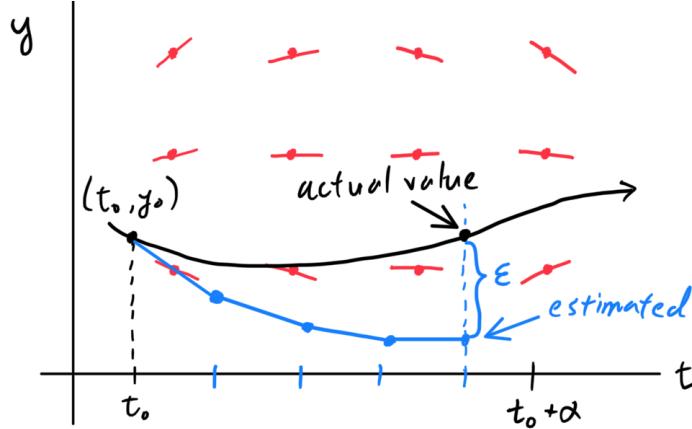
Definition 2.5 (Euler's Algorithm)

Given a differential equation $y' = f(t, y)$, we follow these steps to create an approximation graph passing through (t_0, y_0).

1. Let the interval be $[t_0, t_0 + \alpha]$. Divide the interval into n (not necessarily equally spaced) sub-intervals $[t_i, t_{i+1}]$ (for simplicity, assume equally spaced).
2. At each point (t_i, y_i), calculate the slope of the graph m_i by plugging the values into the equation $f(t_i, y_i)$.
3. From each point (t_i, y_i), calculate (t_{i+1}, y_{i+1}) using the straight-line graph

$$y_{i+1} = y_i + m_i(t_{i+1} - t_i) = y_i + f(t_i, y_i)(t_{i+1} - t_i)$$

This create a piecewise linear function that serves as an approximation for the true integral curve.

Theorem 2.1 (Error Bound for Euler's Algorithm)

The cumulative error bound for Euler's algorithm is

$$\sum_{k=1}^{n} |T_k| = \frac{1}{2} M h^2 n = \frac{1}{2} \alpha M h$$

Proof. Let $\varphi(t)$ be the exact solution. At (t_k, y_k) and (t_{k+1}, y_{k+1}),

$$\varphi(t_{k+1}) = \varphi(t_k) + \int_{t_k}^{t_{k+1}} f\big(t, \varphi(t)\big) \, dt$$

The formula for the approximate solution is

$$y_{k+1} = y_k + (t_{k+1} - t_k) f(t_k, y_k)$$

Therefore, the local error per point is

$$T_k \equiv |\varphi(t_{k+1}) - y_{k+1}|$$
$$= \left| \int_{t_k}^{t_{k+1}} f\big(t, \varphi(t)\big) \, dt - (t_{k+1} - t_k) f(t_k, y_k) \right|$$

We wish to establish an upper bound on T_k. For convenience of notation, let us denote $F(t) = f\big(t, \varphi(t)\big)$. This means that

$$T_k = \left| \int_{t_k}^{t_{k+1}} F(t) \, dt - (t_{k+1} - t_k) F(t_k) \right|$$

The mean value theorem says that there exists some $s_k \in (t_k, t)$ such that

$$F(t) - F(t_k) = (t - t_k)F'(s_k)$$

which implies that

$$\int_{t_k}^{t_{k+1}} F(t)\, dt = \int_{t_k}^{t_{k+1}} F(t_k) + (t - t_k)F'(s_k)\, dt$$

$$= \int_{t_k}^{t_{k+1}} (t - t_k)F'(s)\, dt + (t_{k+1} - t_k)F(t_k)$$

$$\implies T_k = \left| \int_{t_k}^{t_{k+1}} (t - t_k)F'(s_k)\, dt \right|$$

Let

$$M \equiv \max_{t_0 \le t \le t_0 + \alpha} \left| F'(t) \right|$$

Then,

$$F'(t) = \frac{\partial F}{\partial t}\big(t, \varphi(t)\big) + \left(\frac{\partial F}{\partial y}\big(t, \varphi(t)\big) \right) \cdot \varphi'(t)$$

$$= \frac{\partial F}{\partial t}\big(t, \varphi(t)\big) + \left(\frac{\partial F}{\partial y}\big(t, \varphi(t)\big) \right) f\big(t, \varphi(t)\big)$$

Using the triangle inequality, we get

$$M \le \max_{R} \left| \frac{\partial F}{\partial t}(t, y) \right| + \max_{R} \left| \frac{\partial F}{\partial y}(t, y) \right| \cdot \max_{R} \left| f(t, y) \right|$$

This means that the bound for M can be calculated explicitly with the inequality above. It can be checked that

$$M \le \int_{t_k}^{t_{k+1}} (t - t_k)\, dt = \frac{M}{2}(t_{k+1} - t_k)^2$$

Assuming equal subdivisions, we can let $t_{k+1} - t_k = h = \frac{\alpha}{n}$ and conclude that

$$T_k \le \frac{1}{2} M h^2$$

Therefore, the cumulative error bound is

$$\sum_{k=1}^{n} |T_k| = \frac{1}{2} M h^2 n = \frac{1}{2} \alpha M h$$

3 Linear Differential Equations

Definition 3.1 (nth Order Linear DEQ)

A *nth order linear differential equation* has the form

$$a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = f(t)$$

where $a_0, a_1, \ldots, a_n, f \in C^0$ over interval $I \subset \mathbb{R}$. Using the linear operator

$$\mathcal{L}_n : C^n(\mathbb{R}) \longrightarrow C^0(\mathbb{R}), \ \mathcal{L}_n(y) \equiv a_0 y^{(n)} + a_1 y^{(n-1)} + \ldots + a_{n-1}y' + a_n y$$

---

where

$$\mathcal{L}_n(y)(t) \equiv \left(\sum_{i=0}^{n} a_i y^{(n-i)} \right)(t) = \sum_{i=0}^{n} a_i(t) y^{(n-i)}(t)$$

the DEQ can be written as

$$\mathcal{L}_n(y)(t) = f(t), \text{ or } \mathcal{L}_n(y) = f$$

The DEQ is said to be *homogeneous* iff $f = 0$ and *inhomogeneous* iff $f \neq 0$.

In general linear differential equations are important since it is often the case that we reduce nonlinear differential equations to linear ones with approximations. For example, the equation for the pendulum is approximated as

$$\frac{d^2\theta}{dt^2} + \frac{g}{L} \sin \theta = 0 \implies \frac{d^2\theta}{dt^2} + \frac{g}{L} \theta = 0$$

for small θ.

Example 3.1

Consider the equation

$$ty'' + \cos(t)y' + \left(1 - \frac{1}{1+t} \right) y = 2t$$

$1 - \frac{1}{1+t}$ is discontinuous at $t_0 = -1$, and this equation is not of second order at $t = 0$. Therefore, valid intervals I must be a subset of

$$(-\infty, -1) \cup (-1, 0) \cup (0, \infty)$$

for which there exists a unique solution.

3.1   Solving Homogeneous Linear DEQs

This entire section is grounded in this theorem.

Theorem 3.1 (The Fundamental Set)

It can be seen that the set of functions $y = \varphi(t)$ that are solutions to the homogeneous equation $\mathcal{L}_n(y) = 0$ is simply just the kernel of $\mathcal{L}_n$, which is a subset of $C^n(\mathbb{R})$. Then,

$$\dim \ker \mathcal{L} = n$$

The basis of $\ker \mathcal{L}$ is called the *fundamental set*, and given that $\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ is this basis, the general solution is a linear combination of them.

$$\varphi = \sum_i k_i \varphi_i, \quad k_i \in \mathbb{R}$$

This reduces the problem of finding a general solution to just finding n linearly independent ones.

Definition 3.2 (Wronskian)

Let $f_1, f_2, \ldots, f_n$ be of class $C^{n-1}(\mathbb{R})$ in interval $I \subset \mathbb{R}$. Then, the *Wronskian* is defined

$$W(f_1, f_2, \ldots, f_n) \equiv \det \begin{pmatrix} f_1 & \cdots & f_n \\ \vdots & \ddots & \vdots \\ f_1^{(n-1)} & \cdots & f_n^{(n-1)} \end{pmatrix}$$

A simple application of the Wronskian is shown.

Theorem 3.2

Let $\varphi_1, \varphi_2, \ldots, \varphi_n$ be any n special solutions of the nth order linear homogeneous equation $\mathcal{L}(y) = 0$. Then, the set is linearly independent if and only if

$$W(\varphi_1, \ldots, \varphi_n) \neq 0$$

Note that this theorem only works when the φ_i's are known to be solutions.

Proof. Suppose that $W(\varphi_1, \ldots, \varphi_n)(t) \neq 0$ for all $t \in I$ and that $\varphi_1, \ldots, \varphi_n$ are linearly dependent. This means that there exists a nontrivial linear combination summing up to 0, and differentiating the equation on both sides gives the following (by abuse of notation, we will denote W as the Wronskian matrix, not the determinant):

$$W(\varphi_1, \ldots, \varphi_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \iff \begin{cases} b_1\varphi_1(t) + b_2\varphi_2(t) + \ldots + b_n\varphi_n(t) = 0 \\ b_1\varphi_1'(t) + b_2\varphi_2'(t) + \ldots + b_n\varphi_n'(t) = 0 \\ \ldots \\ b_1\varphi_1^{(n-1)}(t) + \ldots + b_n\varphi_n^{(n-1)}(t) = 0 \end{cases}$$

For each fixed $t \in I$, by hypothesis W is nonsingular, meaning that the only solution satisfying $Wb = 0$ is when $b = 0$ itself. But this contradicts our assumption that $b \neq 0$, and so we have proved that

$$W(\varphi_1, \ldots, \varphi_n) \neq 0 \implies \varphi_1, \ldots, \varphi_n \text{ linearly independent}$$

Now, given that $W(\varphi_1, \ldots, \varphi_n)(t) = 0$ for some $t_0 \in I$, assume that $\varphi_1, \ldots, \varphi_n$ are linearly independent. $W(\varphi_1, \ldots, \varphi_n)(t_0)$ (interpreted as a matrix) is singular, the kernel of $W(\varphi_1, \ldots, \varphi_n)(t_0)$ is nontrivial; that is, there exists a nontrivial solution $b \neq 0$ to

$$W(\varphi_1, \ldots, \varphi_n)(t_0) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

To show linear dependence of the φ_i's, we define a solution ψ of the DEQ above as

$$\psi(t) = b_1\varphi_1(t) + b_2\varphi_2(t) + \ldots + b_n\varphi_n(t)$$

where $(b_1, \ldots, b_n)^T$ is any solution of $W(t_0)b = 0$ above. Since the system above tells us that

$$\psi(t_0) = 0, \psi'(t_0) = 0, \ldots, \psi^{(n-1)}(t_0) = 0$$

it follows that $\psi = 0$ for all $t \in I$ is a solution. This means that we have found $b_1, \ldots, b_n$ such that

$$\psi = \sum_{i=1}^{n} b_i\varphi_i(t) = 0$$

contradicting the assumption that φ_i's were linearly independent. Therefore,

$$W(\varphi_1, \ldots, \varphi_n) = 0 \text{ for some } t_0 \in I \implies \varphi_1, \ldots, \varphi_n \text{ linearly dependent}$$

3.1.1 Homogeneous Equations with Constant Coefficients

Let us have the nth order linear homogeneous equation with constant coefficients $k_i \in \mathbb{R}$.

$$\mathcal{L}_n(y) \equiv k_0 y^{(n)} + k_1 y^{(n-1)} + \ldots + k_{n-1} y' + k_n y = 0$$

Now, assume that a solution in the form e^{zt} is valid for $z \in \mathbb{C}$ (explained later). Then,

$$\begin{aligned}
\mathcal{L}(e^{zt}) &= k_0 (e^{zt})^{(n)} + k_1 (e^{zt})^{(n-1)} + \ldots + k_{n-1} (e^{zt})' + k_n (e^{zt}) \\
&= k_0 z^n e^{zt} + k_1 z^{n-1} e^{zt} + \ldots + k_{n-1} z e^{zt} + k_n e^{zt} \\
&= e^{zt} \big(k_0 z^n + k_1 z^{n-1} + \ldots + k_{n-1} z + k_n \big) = 0
\end{aligned}$$

Definition 3.3 (Characteristic Polynomial)

Given the linear DEQ $\mathcal{L}_n(y) = 0$, the *characteristic polynomial* of $\mathcal{L}_n$ is

$$\sum_{i=0}^{n} k_i z^{n-i}$$

Since $e^{zt} \neq 0$ for all values of z, we must find all zeroes of the characteristic polynomial. By the fundamental theorem of algebra, there exists n solutions in $\mathbb{C}$, but it turns out that we can turn them into real solutions using the following lemma.

Lemma 3.3

Given that we have complex valued solution φ to $\mathcal{L}(y) = 0$, where

$$\varphi = e^{z^*t} = e^{(\alpha + \beta i)t} = e^{\alpha t} \big(\cos \beta t + i \sin \beta t \big)$$

and z^* is a complex solution to the characteristic polynomial, the functions

$$\begin{aligned}
\operatorname{Re}(\varphi) &= e^{\alpha t} \cos \beta t \\
\operatorname{Com}(\varphi) &= e^{\alpha t} \sin \beta t
\end{aligned}$$

are real solutions to the DEQ.

Proof. Since $z^* = \alpha + \beta i$ is a zero of the polynomial, its complex conjugate $\overline{z^*} = \alpha - \beta i$ is also a zero, meaning that the following two functions are solutions.

$$\begin{aligned}
\varphi &= e^{(\alpha + \beta i)t} = e^{\alpha t} \big(\cos \beta t + i \sin \beta t \big) \\
\overline{\varphi} &= e^{(\alpha - \beta i)t} = e^{\alpha t} \big(\cos \beta t - i \sin \beta t \big)
\end{aligned}$$

By linearity of $\ker \mathcal{L}$, the following are also solutions

$$\operatorname{Re}(\varphi) = \frac{\varphi + \overline{\varphi}}{2}, \quad \operatorname{Com}(\varphi) = \frac{\varphi - \overline{\varphi}}{2}$$

Lemma 3.4

Given differential equation $\mathcal{L}(y) = 0$, let its characteristic polynomial have real root z^* with multiplicity $m > 1$. Then, all the following are also real solutions of the DEQ:

$$e^{z^*t}, t e^{z^*t}, t^2 e^{z^*t}, \ldots, t^{m-1} e^{z^*t}$$

If this root z^* is complex then its complex conjugate $\overline{z^*}$ must also be a root of multiplicity m, and so in addition to $e^{z^*t}, \ldots, t^{m-1}e^{z^*t}$, the following are also (complex) solutions of $\mathcal{L}(y) = 0$.

$$e^{\overline{z^*}t} = \overline{e^{z^*t}}, \ te^{\overline{z^*}t} = \overline{te^{z^*t}}, \ t^2e^{\overline{z^*}t} = \overline{t^2e^{z^*t}}, \ \ldots, t^{m-1}e^{\overline{z^*}t} = \overline{t^{m-1}e^{z^*t}}$$

and we can construct a basis of real-valued functions from these $2(m-1)$ functions.

$$\mathrm{Re}(e^{z^*t}), \mathrm{Re}(te^{z^*t}), \mathrm{Re}(t^2e^{z^*t}), \ldots, \mathrm{Re}(t^{m-1}e^{z^*t})$$
$$\mathrm{Com}(e^{z^*t}), \mathrm{Com}(te^{z^*t}), \mathrm{Com}(t^2e^{z^*t}), \ldots, \mathrm{Com}(t^{m-1}e^{z^*t})$$

We can summarize all this in the following theorem.

Theorem 3.5 (Solving Homogeneous Linear DEQs with Constant Coefficients)

Let
$$L_n(y) \equiv \sum_{i=0}^{n} a_i y^{(n-i)} = a_0 y^{(n)} + a_1 y^{(n-1)} + \ldots + a_{n-1} y' + a_n y = 0, \ a_i \in \mathbb{R}$$

be a nth order linear DEQ with constant coefficients and let

$$p_n(z) = a_0 z^n + a_1 z^{n-1} + \ldots + \ldots + a_{n-1} z + a_n$$

be its characteristic polynomial with distinct roots $z_1, z_2, \ldots, z_s (s \leq n)$, each root $z_i \in \mathbb{C}$ having multiplicity m_i. Note that $\sum_j m_j = n$ and

$$p_n(z) = \prod_{i=1}^{s} (z - z_i)^{m_i}$$

Then, the n functions form the complex fundamental set of the solutions of $\mathcal{L}_n(y) = 0$.

$$\{e^{z_1t}, te^{z_1t}, \ldots, t^{m_1-1}e^{z_1t},$$
$$e^{z_2t}, te^{z_2t}, \ldots, t^{m_2-1}e^{z_2t},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$e^{z_st}, te^{z_st}, \ldots, t^{m_s-1}e^{z_st}$$

If any function $t^w e^{z_v t}$ (for $0 \leq w \leq m_s - 1, 1 \leq v \leq s$) is complex, then its conjugate $\overline{t^w e^{z_v t}}$ is also in the fundamental set and we can change the basis of the span of complex functions into the following basis of real functions having the same span:

$$\{t^w e^{z_v t}, \overline{t^w e^{z_v t}}\} \implies \{\mathrm{Re}(t^w e^{z_v t}), \mathrm{Com}(t^w e^{z_v t})\}$$

Example 3.2

The solutions to the equation $y'' + y' + y = 0$ is

$$\varphi_1(t) = \exp\left(\frac{-1+i\sqrt{3}}{2}t\right), \quad \varphi_2(t) = \exp\left(\frac{-1-i\sqrt{3}}{2}t\right)$$

This implies, by the previous theorem, that the fundamental set of real solutions is

$$Re\big(\varphi_1(t)\big) = Re\left(e^{-\frac{1}{2}t}\left(\cos\left(\frac{\sqrt{3}}{2}\right) + i\sin\left(\frac{\sqrt{3}}{2}\right)\right)t\right)$$

$$= e^{-\frac{1}{2}t}\cos\left(\frac{\sqrt{3}}{2}t\right)$$

$$Com\big(\varphi_1(t)\big) = e^{-\frac{1}{2}t}\sin\left(\frac{\sqrt{3}}{2}t\right)$$

We state the 2nd-order case separately since it is seen often in physical phenomena.

Corollary 3.6 (Solutions to 2nd-Order Linear DEQs)

Every solution φ of $y'' + py' + qy = 0$ with $p, q \in \mathbb{R}$ and $p^2 \neq 4q$ is defined on $-\infty < t < \infty$ and has the form

$$\varphi(t) = c_1 e^{z_1 t} + c_2 e^{z_2 t}$$

where z_1, z_2 are roots of the equation $z^2 + pz + q = 0$. Furthermore,
1. If $p^2 > 4q$, then $z_1, z_2 \in \mathbb{R}$ and are distinct.
2. If $p^2 < 4q$, then $z_1, z_2 \in \mathbb{C}$ with $z_1 = \bar{z}_2 \implies$ if $z_1 = \alpha + \beta i$ is a solution, a general solution can be expressed

$$\varphi(t) = e^{\alpha t}\big(a_1 \cos\beta(t) + a_2 \sin\beta(t)\big)$$

3. If $p^2 = 4q$, then $z^2 + pq + p = 0$ has a double root at $z = -\frac{p}{2} \implies$ a solution is $e^{-\frac{p}{2}t}$. The other solution is of the form, $te^{-\frac{p}{2}t}$, making the general solution

$$(a_1 + a_2 t)e^{-\frac{p}{2}t}$$

Proof. (1) and (2) are quite trivial. For (3), we know that $\varphi(t) = e^{-\frac{p}{2}t}$ is a solution, and we will try to construct another linearly independent solution. This important process that we will employ is analogous to the *reduction of order* process that will be explained later. We assume that the second solution is of form

$$\psi(t) = e^{-\frac{p}{2}t}w(t)$$

Assume that $\psi(t)$ is a solution of $y'' + py' + qy = 0$. Then manually differentiating, we get

$$\psi(t) = e^{-\frac{p}{2}t}w(t)$$

$$\psi'(t) = e^{-\frac{p}{2}t}w'(t) - \frac{p}{2}e^{-\frac{p}{2}t}w(t)$$

$$\psi''(t) = e^{-\frac{p}{2}t}w''(t) - pe^{-\frac{p}{2}t}w'(t) + \frac{p^3}{4}e^{-\frac{p}{2}t}w(t)$$

Plugging these into the differential equation and simplifying gives us

$$\psi''(t) + p\psi'(t) + q\psi(t) = e^{-\frac{p}{2}t}\left(w''(t) + \left(q - \frac{p^2}{4}\right)w(t)\right)$$

$$= e^{-\frac{p}{2}t}w''(t) = 0$$

Note that by hypothesis, $p^2 = 4q$. Since $e^{-\frac{p}{2}t} \neq 0$,

$$w''(t) = 0 \implies w(t) = a_1 + a_2 t \implies \psi(t) = e^{-\frac{p}{2}t}(a_1 + a_2 t)$$

3.1.2   Reduction of Order

We elaborate on the method used in the previous section and introduce the reduction of order method.

Theorem 3.7 (Reduction of Order of 2nd-Order Linear DEQs)

Given the homogeneous linear differential equation with nonconstant coefficients

$$\mathcal{L}_2(y) = a_0(t)y'' + a_1(t)y' + a_2(t)y = 0$$

where a_0, a_1, a_2 are C^0 on I and $a_0(t) \neq 0$ on I, assume that we know a solution φ_1. Then, the second solution is

$$\varphi_2(t) = \varphi_1(t) \int_{t_0}^{t} \frac{1}{\varphi_1^2(\sigma)} \exp\left(-\int_{t_0}^{\sigma} \frac{a_1(s)}{a_0(s)} \, ds\right) d\sigma$$

Proof. Since we know solution φ_1, we assume that the second solution is of form $\varphi_2(t) = w(t)\varphi_1(t)$ and try to find a nonconstant function satisfying the DEQ. Since,

$$\varphi_2 = w\varphi_1$$
$$\varphi_2' = w'\varphi_1 + w\varphi_1'$$
$$\varphi_2'' = w''\varphi_1 + 2w'\varphi_1' + w\varphi_1''$$

We have

$$\mathcal{L}_2(\varphi_2) = a_0\varphi_1 w'' + (2a_0\varphi_1' + a_1\varphi_1)w' + w\mathcal{L}_2(\varphi_1) \text{ for all } t \in I$$

But since $\mathcal{L}_2(\varphi_1) = 0$,

$$\mathcal{L}_2(\varphi_2) = a_0\varphi_1 w'' + (2a_0\varphi_1' + a_1\varphi_1)w'$$

But notice that this is a first-order linear equation in w'! What we have essentially done is use the fact that $\mathcal{L}_2(\varphi_1) = 0$ to "reduce" the the DEQ from a second-order equation to a first-order one. Letting $v = w'$ gives

$$v' + \left(2\frac{\varphi_1'}{\varphi_1} + \frac{a_1}{a_0}\right)v = 0$$

Separating variables we obtain the solution

$$v(t) = \frac{1}{\varphi_1^2(t)} \exp\left(-\int_{t_0}^{t} \frac{a_1(s)}{a_0(s)} \, ds\right) \implies w(t) = \int_{t_0}^{t} \frac{1}{\varphi_1^2(\sigma)} \exp\left(-\int_{t_0}^{\sigma} \frac{a_1(s)}{a_0(s)} \, ds\right) d\sigma$$

This process can be repeated for higher order equations, we show it for 3rd-order equations.

Theorem 3.8 (Reduction of Order of 3rd-Order DEQs)

Given the homogeneous linear differential equation with nonconstant coefficients

$$\mathcal{L}_3(y) = a_0(t)y''' + a_1(t)y'' + a_2(t)y' + a_3(t)y = 0$$

where a_0, a_1, a_2, a_3 are C^0 on i and $a_0(t) \neq 0$ on I, assume that we know a solution φ_1. Then, the second and third solutions are

$$\varphi_2 = \varphi_1 \int_{t_0}^{t} \gamma_1(s) \, ds, \quad \varphi_3 = \varphi_1 \int_{t_0}^{t} \gamma_2(s) \, ds$$

where γ_1, γ_2 are the solutions of differential equation

$$a_0\varphi_1 v'' + (3a_0\varphi_1' + a_1\varphi_1)v' + (3a_0\varphi_1'' + 2a_1\varphi_1' + a_2\varphi_1)v = 0$$

Proof. We assume that the second solution is of form $\varphi_2(t) = w(t)\varphi_1(t)$ and compute the derivatives

$$\varphi_2 = w\varphi_1$$
$$\varphi_2' = w'\varphi_1 + w\varphi_1'$$
$$\varphi_2'' = w''\varphi_1 + 2w'\varphi_1' + w\varphi_1''$$
$$\varphi_2''' = w'''\varphi_1 + 3w''\varphi_1' + 3w'\varphi_1'' + w\varphi_1'''$$

Substituting these into $\mathcal{L}_3 = 0$ and simplifying gives

$$\mathcal{L}_3(\varphi_2) = a_0\varphi_1 w''' + (3a_0\varphi_1' + a_1\varphi_1)w'' + (3a_0\varphi_1'' + 2a_1\varphi_1' + a_2\varphi_1)w' + w\mathcal{L}_3(\varphi_1) \text{ for all } t \in I$$

Again, since $\mathcal{L}_3(\varphi_1) = 0$, we get

$$\mathcal{L}_3(\varphi_2) = a_0\varphi_1 w''' + (3a_0\varphi_1' + a_1\varphi_1)w'' + (3a_0\varphi_1'' + 2a_1\varphi_1' + a_2\varphi_1)w'$$

which is really just a 2nd order equation in w'. Substituting $v = w'$ gives us

$$\mathcal{L}_3(\varphi_2) = a_0\varphi_1 v'' + (3a_0\varphi_1' + a_1\varphi_1)v' + (3a_0\varphi_1'' + 2a_1\varphi_1' + a_2\varphi_1)v$$

Therefore, we have successfully reduced the order of this DEQ. Upon finding the two solutions γ_1, γ_2, the solution will be of form

$$\varphi_2 = \varphi_1 \int_{t_0}^t \gamma_1(s)\,ds, \quad \varphi_3 = \varphi_1 \int_{t_0}^t \gamma_2(s)\,ds$$

3.1.3 Cauchy-Euler's Equation

Definition 3.4

A *Cauchy-Euler equation of order n* has the form

$$t^n y^{(n)} + a_1 t^{n-1} y^{(n-1)} + a_2 t^{n-2} y^{(n-2)} + \ldots + a_{n-1} t y' + a_n y = 0, \quad a_i \in \mathbb{R}$$

There are two ways we can explicitly solve for this equation. Since the most common Cauchy-Euler equation is the second-order one, we will solve the following

$$t^2 y'' + aty' + by = 0$$

1. A trial solution $y = t^m$ can be used to directly solve for basic solutions (guess and check method). We assume a trial solution $y = t^m$. Differentiating it gives

$$y'(t) = mt^{m-1}, \quad y''(t) = m(m-1)t^{m-2}$$

   Substituting into the original equation and simplifying gives:

$$t^2\big(m(m-1)t^{m-2}\big) + at\big(mt^{m-1}\big) + b\big(t^m\big) = 0 \implies m^2 + (a-1)m + b = 0$$

   There are three possible cases:

   (a) Two distinct roots m_1, m_2:
$$y = c_1 x^{m_1} + c_2 x^{m_2}$$

   (b) One repeated root m:
$$y = c_1 x^m \ln x + c_2 x^m$$

   Note that this solution is found using the reduction of order.

(c) Complex conjugate roots $\alpha \pm \beta i$:

$$y = c_1 x^\alpha \cos(\beta \ln x) + c_2 x^\alpha \sin(\beta \ln x)$$

This can be easily derived by using Euler's formula.

2.

Consider the equation of the form

We can use substitution $|t| = e^s$ to reduce the equation to have constant coefficients.

1. $t > 0, t = e^s \implies s = \log(t), y(t) = y(e^s) = w(s) = w\big(\log(t)\big)$. This implies that

$$\frac{dy}{dt} = \frac{dw}{ds}\frac{ds}{dt} = \frac{1}{t}\frac{dw}{ds}$$
$$\implies \frac{d^2 y}{dt^2} = -\frac{1}{t^2}\frac{dw}{ds} + \frac{1}{t}\frac{d}{dt}\left(\frac{dw}{ds}\right) = \frac{1}{t^2}\left(\frac{d^2 w}{ds^2} - \frac{dw}{ds}\right)$$
$$\implies \text{substituting,} \ \frac{d^2 w}{ds^2} - \frac{dw}{ds} + a_1 \frac{dw}{ds} + a_2 w(s) = 0$$
$$\implies w'' + (a_1 - 1)w' + a_2 w = 0$$

2. When $t < 0$, the procedure is similar.

This results in the equation $L(w) = 0$ having the general solution

$$\begin{cases} c_1 e^{z_1 s} + c_2 e^{z_2 s} & z_1 \neq z_2 \text{ roots of } z^2 + (a_1 - 1)z + a_2 = 0 \\ (c_1 + c_2 s)e^{z_1 s} & z_1 = z_2 \text{ roots of "} \end{cases}$$

This means that the solution of $L(y) = 0$ is one of

$$\begin{cases} \varphi(t) = c_1 |t|^{z_1} + c_2 |t|^{z_2} \\ \varphi(t) = (c_1 + c_2 \log|t|)|t|^{z_1} \end{cases}$$

If $t \in \mathbb{C}$, $e^{z \log|t|} = |t|^2$, assuming that $z = \alpha + \beta i$. This means that

$$|t|^2 = e^{(\alpha + \beta i)\log|t|} = e^{\alpha \log|t|}\big(\cos(\beta \log|t|) + i \sin(\beta \log|t|)\big)$$
$$= |t|^\alpha \big(\cos(\beta \log|t|) + i \sin(\beta \log|t|)\big)$$

3.2 Solving Inhomogeneous Linear DEQs

Definition 3.5 (Inhomogeneous Linear DEQ)

An *inhomogeneous linear equation* is a linear differential equation in the form

$$\mathcal{L}_n(y) = f$$

where $f \neq 0$. $f \neq 0$ usually denotes that there is an external force acting on the system.

Example 3.3 (Dampened Linear Mass-Spring System with Periodic External Force)

The dampened linear mass-spring system subjected to given periodic external force $A \cos \omega t$ is represented by the linear DEQ

$$y'' + by' + \frac{k}{m}y = \frac{A}{M}\cos \omega t$$

The following theorem provides insight into the structure of the solutions for these systems.

Theorem 3.9

Suppose ψ_p is a particular solution of $\mathcal{L}_n(y) = f$ on I and suppose $\varphi_1, \varphi_2, ..., \varphi_n$ are n linearly independent solutions of the homogeneous equation

$$\mathcal{L}_n(y) = 0$$

Then, every solution ψ of $\mathcal{L}_n(y) = f$ has general form

$$\psi = \psi_p + \sum_{i=1}^{n} c_i \varphi_i$$

That is, the set of all solutions is an n dimensional affine subspace within the space of all continuous functions.

Proof. Let ψ_p be a particular solution of $\mathcal{L}_n(y) = f$ and ψ_0 be any solution of $\mathcal{L}_n(y) = 0$. Then, $\psi_0 \in \ker \mathcal{L}_n$, which means that by linearity of $\mathcal{L}_n$,

$$\mathcal{L}_n(\psi_p + k\psi_0) = \mathcal{L}_n(\psi_p) + k\mathcal{L}_n(\psi_0) = f + 0 = f$$

We have reduced the problem of solving inhomogeneous DEQs into solving the homogeneous version and finding a particular solution.

3.2.1   Variation of Constants

We now attempt to find a particular solution ψ_p for a second order linear inhomogeneous equation. The great thing about the method below is that as long as the functions $a_0 \neq 0, a_1, a_2, f$ are continuous over certain interval $I \subset \mathbb{R}$, it always provides us with a specific solution.

Theorem 3.10 (Variation of Constants Formula for 2nd-Order Linear DEQs)

Given inhomogeneous linear DEQ

$$\mathcal{L}_2(y) = a_0(t)y'' + a_1(t)y' + a_2(t)y = f$$

where a_0, a_1, a_2, f are all continuous on I, assume that we have the general solution $\varphi = c_1\varphi_1 + c_2\varphi_2$ for $\mathcal{L}_2(y) = 0$. Then, a particular solution ψ_p for $\mathcal{L}_2(y) = f$ is given by the *Variation of Constants formula*

$$\psi_p = \varphi_1(t)\left(-\int_{t_0}^{t} \frac{f(s)\,\varphi_2(s)}{a_0(s)W(\varphi_1, \varphi_2)(s)}\,ds \right) + \varphi_2(t)\left(\int_{t_0}^{t} \frac{f(s)\,\varphi_1(s)}{a_0(s)W(\varphi_1, \varphi_2)(s)}\,ds \right)$$

$$= \psi_p(t) = \int_{t_0}^{t} \frac{f(s)\left(\varphi_2(t)\varphi_1(s) - \varphi_1(t)\varphi_2(s) \right)}{a_0(s)\,W(\varphi_1, \varphi_2)(s)}\,ds$$

Proof. We assume that the particular solution is of the form

$$\psi_p = u_1\varphi_1 + u_2\varphi_2$$

and that we have found these functions u_1, u_2. This assumption is quite arbitrary, but we will see why this works later. This implies that

$$\psi_p' = (u_1\varphi_1 + u_2\varphi_2)' = u_1\varphi_1' + u_2\varphi_2' + u_1'\varphi_1 + u_2'\varphi_2$$
$$\psi_p'' = u_1\varphi_1'' + u_2\varphi_2'' + 2u_1'\varphi_1 + 2u_2'\varphi_2' + u_1''\varphi_2 + u_2''\varphi_2$$

With these derivatives, we can expand out the equation $\mathcal{L}_2(\psi_p) = f$ to get

$$\mathcal{L}_2(\psi_p) = \mathcal{L}_2(u_1\varphi_1 + u_2\varphi_2) = a_0(u_1\varphi_1 + u_2\varphi_2)'' + a_1(u_1\varphi_1 + u_2\varphi_2)' + a_0(u_1\varphi_1 + u_2\varphi_2)$$

$$= \dots$$

$$= a_0\big((\varphi_1 u_1'' + \varphi_2 u_2'') + 2(\varphi_1' u_1' + \varphi_2' u_2')\big) + a_1(\varphi_1 u_1' + \varphi_2 u_2') = f$$

We would like to obtain some sort of relation from this last equation. If we assume that

$$\varphi_1 u_1' + \varphi_2 u_2' = 0$$

for all $t \in I$, then $(\varphi_1 u_1' + \varphi_2 u_2')' = \varphi_1 u_1'' + \varphi_2 u_2'' + \varphi_1' u_1' + \varphi_2' u_2' = 0$ and thus

$$a_0\big((\varphi_1 u_1'' + \varphi_2 u_2'') + 2(\varphi_1' u_1' + \varphi_2' u_2')\big) + a_1(\varphi_1 u_1' + \varphi_2 u_2') = f \implies \varphi_1' u_1' + \varphi_2' u_2' = \frac{f}{a}$$

Notice our line of reasoning so far: Assuming the existence of a solution of form $\psi_p = u_1\varphi_1 + u_2\varphi_2$ (and that $\varphi_1 u_1' + \varphi_2 u_2' = 0$) implies that $\varphi_1' u_1' + \varphi_2' u_2' = \frac{f}{a}$. However, we can reason backwards and state this: If we can find two equations u_1, u_2 satisfying $\varphi_1 u_1' + \varphi_2 u_2' = 0$ and $\varphi_1' u_1' + \varphi_2' u_2' = \frac{f}{a}$, then $\varphi_p = u_1\varphi_2 + u_2\varphi_2$ satisfies $\mathcal{L}_2(\varphi_p) = f$ on I. So assume that there exists u_1, u_2 such that (written in matrix form)

$$\begin{pmatrix} \varphi_1 & \varphi_2 \\ \varphi_1' & \varphi_2' \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 0 \\ f/a_0 \end{pmatrix}$$

Using Cramer's rule, we solve for u_1', u_2' and integrate to get u_1, u_2.

$$u_1' = \frac{1}{W(\varphi_1, \varphi_2)} \det \begin{pmatrix} 0 & \varphi_2 \\ f/a_0 & \varphi_2' \end{pmatrix} = \frac{-f\varphi_2}{a_0 W(\varphi_1, \varphi_2)} \implies u_1(t) = \int_{t_0}^{t} \frac{-f(s)\,\varphi_2(s)}{a_0(s)W(\varphi_1, \varphi_2)(s)}\,ds$$

$$u_2' = \frac{1}{W(\varphi_1, \varphi_2)} \det \begin{pmatrix} \varphi_1 & \varphi_1' \\ 0 & f/a_0 \end{pmatrix} = \frac{f\varphi_1}{a_0 W(\varphi_1, \varphi_2)} \implies u_2(t) = \int_{t_0}^{t} \frac{f(s)\,\varphi_1(s)}{a_0(s)W(\varphi_1, \varphi_2)(s)}\,ds$$

Therefore,

$$\psi_p(t) = u_1(t)\varphi_1(t) + u_2(t)\varphi_2(t) = \int_{t_0}^{t} \frac{f(s)\,\big(\varphi_2(t)\varphi_1(s) - \varphi_1(t)\varphi_2(s)\big)}{a_0(s)\,W(\varphi_1, \varphi_2)(s)}\,ds$$

Theorem 3.11 (Variation of Constants Formula)

For an nth order linear differential equation

$$\mathcal{L}_n(y) = a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_{n-1}(t)y' + a_n(t)y = f$$

Let the fundamental set of its homogeneous counterpart $\mathcal{L}_n(y) = 0$ be $\varphi_1, \dots, \varphi_n$. Then, upon solving the system

$$\begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_n \\ \varphi_1' & \varphi_2' & \cdots & \varphi_n' \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)} & \varphi_2^{(n-1)} & \cdots & \varphi_n^{(n-1)} \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \\ \vdots \\ u_n' \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f/a_0 \end{pmatrix}$$

for $u_1', u_2', \dots, u_n'$, we find ψ_p to be

$$\psi_p(t) = \sum_{i=1}^{n} u_i(t)\varphi_i(t) = \sum_{i=1}^{n} \left(\int_{t_0}^{t} u_i'(s)\,ds \right)\varphi_i(t)$$

4 Series Solutions of Linear Equations

A wider class of differential equations has solutions that can be expanded into a power series than ones that can be solved in closed form (that is, expressed in terms of elementary functions).

Example 4.1

Given the differential equation

$$y'' = -\sin(y) - y', \ \varphi(0) = \frac{\pi}{4}, \varphi'(0) = 0$$

Assuming that φ is an analytic function, we can write the Taylor series

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \ldots = \sum_{j=0}^{\infty} \frac{\varphi^{(j)}(0)}{j!} t^j$$

We can solve for each derivative manually by plugging it into the equation

$$\varphi''(0) = -\sin\left(\varphi(0)\right) - \varphi'(0) = -\frac{\sqrt{2}}{2}$$

$$\varphi'''(0) = -\cos\left(\varphi(0)\right)\varphi''(0) - \varphi''(0) = \frac{\sqrt{2}}{2}$$

$$\varphi^{(4)}(0) = \sin\left(\varphi(0)\right)\left(\varphi'(0)\right)^2 - \cos\left(\varphi(0)\right)\varphi''(0) - \varphi^{(3)}(0) = \frac{1-\sqrt{2}}{2}$$

Doing this recursively, the expansion of the solution φ begins with the terms

$$\varphi(t) = \frac{\pi}{4} - \frac{\sqrt{2}}{2}\frac{t^2}{2!} + \frac{\sqrt{2}}{2}\frac{t^3}{3!} + \frac{1-\sqrt{2}}{2}\frac{t^4}{4!} + \ldots$$

4.1 Review of Power Series

Every power series

$$\sum_{n=0}^{\infty} c_n t^n$$

has a radius of convergence $R \geq 0$ over $\mathbb{C}$. Some more properties:

1. The series converges absolutely for $|t| < R$.

2. Assigning $f(t) = \sum c_n t^n$ for $|t| < R, f \in C^{\infty}$, then

$$f^{(k)}(t) = \sum_{n=k}^{\infty} \frac{n!}{k!} c_n t^{n-k}$$

   with radius of convergence also R.

3. If $[a, b]$ is the interval of convergence, then

$$\int_a^b f(s)\, ds = \sum_{n=0}^{\infty} c_n \int_a^b s^n\, ds$$

   In particular,

$$\int_0^t f(s)\, ds = \sum_{n=0}^{\infty} \frac{c_n}{n+1} t^{n+1}$$

   which also has a radius of convergence R.

---

4. Identity theorem for power series. If $g(t) = \sum d_n t^n$ is another power series with ROC $= \tilde{R}$, and $f(t) = g(t)$ on some interval in which the series both converge, then $d_n = c_n$ for all n. In particular, if the sum of a power series is 0 for all $t \in I$, then every coefficient in the series must be 0.

5. f, g converge $\implies f + g$ converges within the ROC of both f and g.

6. f, g converges $\implies fg$ converges within the ROC of both f and g.

Definition 4.1

f is *analytic at $t = a$* if and only if it can be expanded into a Taylor series

$$f(t) = \sum_{n=0}^{\infty} c_n(t-a)^n, \quad c_n = \frac{f^{(n)}(n)}{k!}$$

with radius of convergence $R > 0$.

It is assumed that the reader is familiar with basic convergence tests such as the ratio test and the comparison test.

Lemma 4.1

Let f, g be analytic functions. That is, they can be expressed

$$f(t) = \sum_{n=p}^{\infty} c_n(t-a)^n, \quad g(t) = \sum_{n=q}^{\infty} d_n(t-a)^n$$

Without loss of generality, suppose $p < q$ and $d_q \neq 0$. Then, f and g are linearly independent on every interval I in which f, g converges.

Proof. This proof is trivial using linear algebra since f is expressed using more basis vectors.

4.2 2nd Order Linear Equations w/ Analytic Coefficients

Example 4.2

Despite what we have learned, the differential equation

$$y''(t) - ty = 0, \quad \varphi(t_0) = a, \varphi'(t_0) = b$$

cannot be solved with any of the previous methods, but there does exist unique solutions for all t. We wish to find a solution using a new method. Assume that $\varphi(t)$ is analytic at $t = t_0$. That is, it can be expanded

$$\varphi(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)^2 + \ldots = \sum_{k=0}^{\infty} c_k(t - t_0)^k$$

that converges in an interval $|(t - t_0)| < A$ (we will not worry about convergence yet). We can calculate

$$\varphi''(t) - t\varphi(t) = \sum_{k=2}^{\infty} k(k-1)c_k(t-t_0)^{k-2} - \sum_{k=0}^{\infty} c_k(t-t_0)^{k+1}$$

$$= 2c_2 + \sum_{k=3}^{\infty} \left(k(k-1)c_k - c_{k-3}\right)(t-t_0)^{k-2} = 0$$

---

if and only if $\varphi(t)$ is a solution of $y'' - ty = 0$. This implies that

$$2c_2 = 0, k(k-1)c_k - c_{k-3} = 0$$

$$\implies c_2 = 0, c_3 = \frac{1}{2\cdot 3}c_0, c_4 = \frac{1}{3\cdot 4}c_1, c_5 = \frac{1}{4\cdot 5}c_2, ...$$

$$\implies \begin{cases} c_{3m} = \frac{(1)(4)...(3m-2)}{(3m)!}c_0 \\ c_{3m+1} = \frac{(2)(5)(8)...(3m-1)}{(3m+1)!}c_1 \\ c_{3m+2} = 0 \end{cases}$$

which proves us an explicit definition of φ dependent on initial conditions $a = \varphi(0) = c_0, b = \varphi'(0) = c_1$. Therefore, a candidate for a solution is

$$\varphi(t) = a\varphi_1(t) + b\varphi_2(t)$$
$$= a\left(1 + \sum_{m=1}^{\infty}\frac{(1)(4)...(3m-2)}{(3m)!}(t-t_0)^{3m}\right) + b\left(1 + \sum_{m=1}^{\infty}\frac{(2)(5)...(3m-1)}{(3m+1)!}(t-t_0)^{3m+1}\right)$$

We can see that this infinite series converges for $|(t - t_0)| < \infty$.

Theorem 4.2 (Existence of Analytic Solutions of 2nd-Order Linear DEQs of Analytic Coefficients)

Given equation (with leading coefficient 1)

$$y'' + p(t)y + q(t)y = f(t)$$

If p, q, f are analytic at t_0, then there exists a unique solution φ satisfying $\varphi(t_0) = a, \varphi'(t_0) = b$. This solution is analytic at $t = t_0$, with expansion

$$\varphi(t) = \sum_{k=0}^{\infty} c_k(t-t_0)^k$$

converging for at least those values of t converging on p, q and f. The coefficients c_k can be determined recursively by direct substitution into the differential equation.

Notice that this series method can be very useful computationally. For values close to t_0, we can calculate the series up to a certain number of coefficients to gain a good approximation of the actual function.

Example 4.3 (Legendre Equation)

Given the equation
$$(1-t^2)y'' - 2ty' + \alpha(\alpha+1)y = 0$$

with α a given constant. In order to determine whether this differential equation has a series solution about $t = 0$, we use the previous theorem. We modify the equation to

$$y'' - \frac{2t}{1-t^2}y' + \frac{\alpha(\alpha+1)}{1-t^2}y = 0 \quad (t \neq \pm 1)$$

The coefficient functions are indeed analytic. That is,

$$-\frac{2t}{1 - t^2} = -2t(1 + t^2 + t^4 + ...) = -2t \sum_{k=0}^{\infty} t^{2k}$$

$$\frac{\alpha(\alpha + 1)}{1 - t^2} = \alpha(\alpha + 1) \sum_{k=0}^{\infty} t^{2k}$$

with a radius of convergence of 1. Therefore, the conditions of the theorem are satisfied, and the Legendre equation has a unique analytic solution satisfying $\varphi(0) = a, \varphi'(0) = b$ for all a, b.

4.3   Singular Points of Linear Equations

Many differential equations which arise in applications, the so-called equations of mathematical physics, fail to satisfy the hypotheses of the theorem stating the existence of analytic solutions. That is, there are points $t = t_0$ where the coefficients are not analytic (usually because it is unbounded). They usually pop up when we divide the entire equation by the leading coefficient $a_0(t)$, which leads to one of the non-leading coefficients to become un-analytic when it is of the form

$$\frac{a_i(t)}{a_0(t)}$$

We formalize this below

Definition 4.2 (Singular Point)

$t = t_0$ is a *singular point* of the differential equation

$$\mathcal{L}_n(y) = a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0$$

if $a_0, a_1, \ldots, a_n$ are analytic at t_0 and $a_0(t_0) = 0$, but $a_1, \ldots, a_n$ are not all zero at t_0. A point that is not a singular point is called an *ordinary point*.

In relation to the theorem on the existence of analytic solutions, we would like to clarify that given singular point t_0 and non-singular point t_1,

$$\text{At } t = t_1 \implies \text{Analytic solution exists (theorem)}$$
$$\text{At } t = t_0 \implies \text{Analytic solution could exist or not}$$

That is, the theorem tells us that given equation

$$y'' + p(t)y + q(t)y = f(t)$$

where p, q, f are analytic at t_1 (i.e. t_1 is ordinary), we are guaranteed the existence of a analytic solution in a neighborhood of t_1. However, if p, q are somehow not analytic, then there *may or may not* be an analytic solution. This will be elaborated in the following example.

Example 4.4 (Nonexistence of Analytic Solutions in 1st-Order Cauchy-Euler Equation)

Consider the first-order Cauchy-Euler equation centered around $t = t_0$
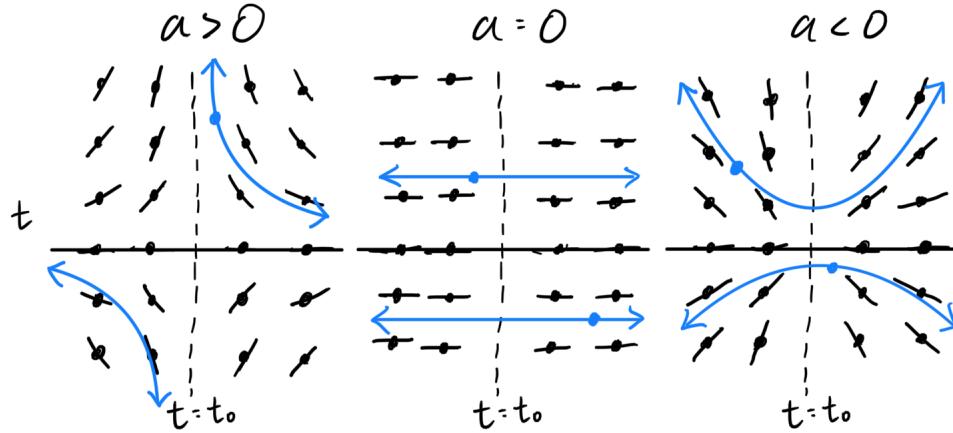
$$(t - t_0)y' + ay = 0$$

where a is a constant. Then, there cannot be an analytic solution to this DEQ at $t = t_0$ since the

---

coefficients of the equivalent, equation

$$y' = -\frac{a}{t - t_0}y$$

A special solution for this DEQ is $y = (t - t_0)^{-a}$, which is graphed (along with the phase velocity vector space/slope field) for when $a > 0, a = 0$, and $a < 0$.

Given that $a > 0$, we can see that the solution is clearly unbounded at $t = t_0$, so there does not exist an analytic function. However, if $a \leq 0$, then there does exist an analytic function.

Example 4.5 (Nonexistence of Analytic Solutions in 2nd-Order Cauchy-Euler Equation)
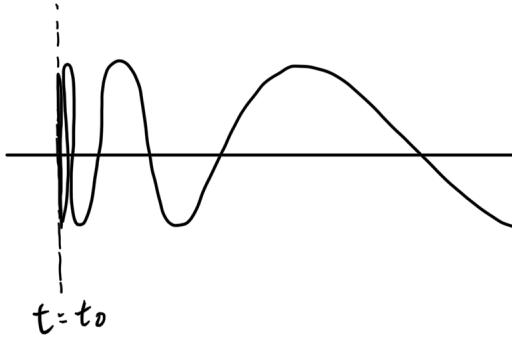
Consider the second-order Cauchy-Euler equation centered around $t = t_0$

$$(t - t_0)^2 y'' + (t - t_0)a_1 y' + a_2 y = 0$$

where a_1, a_2 are constants. Then, we can see that there cannot be analytic solutions to this DEQ at $t = t_0$ since the coefficients of the equivalent, normalized equation

$$y'' + \frac{a_1}{t - t_0}y' + \frac{a_2}{(t - t_0)^2}y = 0$$
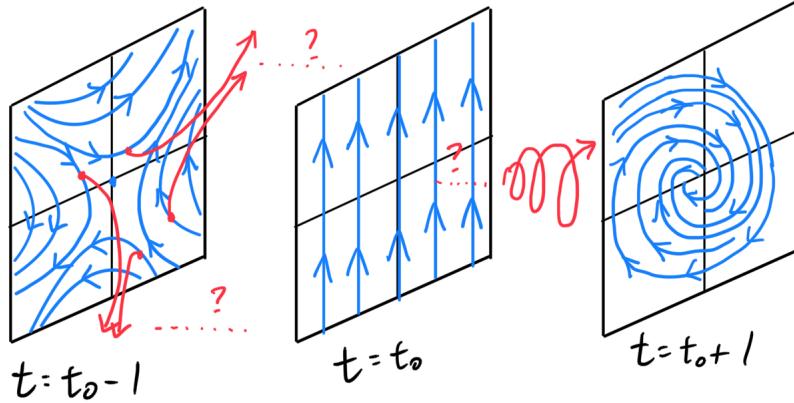
are unbounded at $t = t_0$. However, in a neighborhood of any point $t_1 \neq t_0$, there does exist analytic solutions. Assuming $a_1 = a_2 = 1$, we can see that the solution $y = \varphi(t) = \cos(\ln(t - t_0)) + \sin(\ln(t - t_0))$ oscillates infinitely as $t \to t_0$.

Furthermore, by setting this up as a system of first order equations using $y = y, x = y'$, we get

$$y' = x$$
$$x' = y'' = -\frac{a_1}{t - t_0}x - \frac{a_2}{(t - t_0)^2}y$$

Normalizing $a_1 = a_2 = 1$, we give the phase velocity vector field for the three time periods $t = t_0 - 1, t_0, t_0 + 1$.

The behavior of the solutions near the singular point t_0 (whether it is actually analytic or not) depends on how rapidly $a_0(t)$ approaches zero as $t \to t_0$. For this reason, we distinguish two different types of singular points, with the first being the regular one.

Definition 4.3 (Regular Singular Point)

The point t_0 is called a *regular singular point* of the equation

$$\mathcal{L}_n(y) = a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0$$

if it is a singular point and if

$$p_1(t) = \frac{a_1(t)}{a_0(t)}, \ p_2(t) = \frac{a_2(t)}{a_0(t)}, \ldots, p_n(t) = \frac{a_n(t)}{a_0(t)}$$

have the property that $(t-t_0)p_1(t), (t-t_0)^2 p_2(t), \ldots, (t-t_0)^n p_n(t)$ are all analytic at t_0. It immediately follows that the equation $\mathcal{L}_n(y) = 0$ has a regular singular point at $t = t_0$ if and only if it can be written in the form:

$$a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0$$
$$\implies y^{(n)} + \frac{a_1(t)}{a_0(t)}y^{(n-1)} + \ldots + \frac{a_{n-1}}{a_0(t)}(t)y' + \frac{a_n(t)}{a_0(t)}y = 0$$
$$\implies (t - t_0)^n y^{(n)} + (t - t_0)^n \frac{a_1(t)}{a_0(t)}y^{(n-1)} + \ldots + (t - t_0)^n \frac{a_{n-1}}{a_0(t)}(t)y' + (t - t_0)^n \frac{a_n(t)}{a_0(t)}y = 0$$
$$\implies (t - t_0)^n y^{(n)} + (t - t_0)^{n-1}\left((t - t_0)\frac{a_1(t)}{a_0(t)}\right)y^{(n-1)} + \ldots + \left((t - t_0)^n \frac{a_n(t)}{a_0(t)}\right)y = 0$$
$$\implies (t - t_0)^n y^{(n)} + (t - t_0)^{n-1}\alpha_1(t)y^{(n-1)} + \ldots + (t - t_0)\alpha_{n-1}(t)y' + \alpha_n(t)y = 0$$

Therefore, the equation $\mathcal{L}_n(y) = 0$ has a regular singular point if it can be written in the form

$$\sum_{i=0}^{n}(t - t_0)^{n-i}\alpha_i(t)y^{(n-i)} = 0, \quad \alpha_i(t) = (t - t_0)^i \frac{a_i(t)}{a_0(t)}$$

where the α_i's are analytic at $t = t_0$. If a singular point $t = t_0$ is not regular, then it is called an *irregular singular point*.

Example 4.6

We list three examples of equations and their corresponding singular points.
1. The Euler equation
$$(t - t_0)^2 y'' + (t - t_0)a_1 y' + a_2 y = 0$$

   is the simplest example of an equation which has a regular singular point at $t = t_0$.
2. The equation
$$t^2 y'' + \frac{3}{2}ty' + ty = 0$$

   has $t = 0$ as a regular singular point because $p(t) = \frac{3}{2t}, q(t) = \frac{1}{t}$ have the property that $tp(t) = \frac{3}{2}, t^2 q(t) = t$ are both analytic everywhere.
3. The equation
$$(t - 1)^3 y'' + 2(t - 1)^2 y' - 7ty = 0$$

   does not have a regular singular point at $t = 1$.

4.3.1   Change of Basis

To simplify the process, we can perform a change of basis

$$t \mapsto x = t - t_0$$

which allows us to transform the singular point t_0 to the origin without changing the form of the equation in any essential way. Therefore, we would let

$$\overline{\alpha}_i(x) = \alpha_i(t_0 + x) \text{ for all } i = 1, 2, \ldots, n$$

which are analytic at $x = 0$ since α_i is analytic at $t = t_0$. Furthermore, we define

$$\overline{y}(x) = y(t_0 + x)$$

and by the chain rule, we have

$$\frac{d\overline{y}}{dx} = y'(t_0 + x) = y'(t), \frac{d^2\overline{y}}{dx^2} = y''(t_0 + x) = y''(t), \ldots, \frac{d^n\overline{y}}{dx} = y^{(n)}(t_0 + x) = y^{(n)}(t)$$

Therefore, the equation simplifies as such:

$$(t - t_0)^n y^{(n)} + (t - t_0)^{n-1}\alpha_1(t)y^{(n-1)} + \ldots + (t - t_0)\alpha_{n-1}(t)y' + \alpha_n(t)y = 0 \tag{3}$$
$$\implies x^n \overline{y}^{(n)}(x) + x^{n-1}\overline{\alpha}_1(x)\overline{y}^{(n-1)}(x) + \ldots + x\overline{\alpha}_{n-1}(x)\overline{y}'(x) + \overline{\alpha}_n(x)\overline{y} = 0 \tag{4}$$

Therefore, if $\overline{y}(x)$ is a solution of the second equation with regular singular point $x = 0$, then the function $y(t) = \overline{y}(t - t_0)$ is a solution of the first equation with regular singular point $t = t_0$.

Therefore, we will assume that such a preliminary simplification has been made, and without loss of generality, we will consider a nth order homogeneous linear differential equation with regular singular point (changed accordingly to be at $t = 0$) to be of form

$$t^n y^{(n)} + t^{n-1}\alpha_1(t)y^{(n-1)} + \ldots + t^2\alpha_{n-2}(t)y'' + t\alpha_{n-1}(t)y' + \alpha_n(t)y = 0$$

---

where the α_i's are given functions analytic at $t = 0$ and having power series expansions

$$\alpha_i(t) = \sum_{k=0}^{\infty} \alpha_{ik} t^k$$

which converge in some interval $|t| < r$.

4.3.2   Examples

Example 4.7

The following DEQ can be changed as such

$$2ty'' + y' + ty = 0 \implies t^2 y'' + \frac{1}{2} ty' + \frac{1}{2} t^2 y = 0$$
$$= t^2 y'' + t\alpha_1(t) y' + \alpha_2(t) y = 0$$

where

$$\alpha_1(t) = \frac{1}{2}, \quad \alpha_2(t) = \frac{1}{2} t^2$$

which are both analytic at $t = 0$, making $t = 0$ a regular singular point. If both α_1, α_2 were constants, this following equation would be the Euler equation, making one of the solutions to be of form $|t|^z$. But since α_2 is not a constant, we attempt to find a solution of form

$$|t|^z \sum_{k=0}^{\infty} c_k t^k, \quad c_0 \neq 0$$

where the constants z, c_k are determined by substitution into the differential equation within some interval of convergence around $t = 0$. Since $t = 0$ is a singular point, we separate the cases $t > 0$ and $t < 0$. We first consider the case $t > 0$ and try the following solution of φ and its derivatives (assuming $\varphi \in C^2$):

$$\varphi(t) = t^z \sum_{k=0}^{\infty} c_k t^k = \sum_{k=0}^{\infty} c_k t^{z+k}$$

$$\varphi'(t) = \sum_{k=0}^{\infty} c_k (z+k) t^{z+k-1}$$

$$\varphi''(t) = \sum_{k=0}^{\infty} c_k (z+k)(z+k-1) t^{z+k-2}$$

We substitute this into the equation and simplify it, where we set the indicial polynomial as $f(z) = z(z - \frac{1}{2})$ (we can find this by looking at the quadratic term in z having coefficient c_0).

$$0 = t^2 \varphi''(t) + \frac{1}{2} t\varphi'(t) + \frac{1}{2} t^2 \varphi(t)$$

$$= c_0 z \left(z - \frac{1}{2} \right) t^z + c_1 (z+1) \left(z + \frac{1}{2} \right) t^{z+1} + \sum_{k=0}^{\infty} \left((z+k) \left(z+k - \frac{1}{2} \right) c_k + \frac{1}{2} c_{k-2} \right) t^{z+k}$$

$$= t^z \left(c_0 f(z) + c_1 f(z+1) t + \sum_{k=2}^{\infty} \left(f(z+k) c_k + \frac{1}{2} c_{k-2} \right) \right)$$

The equality implies that every coefficient of every power of t vanishes on the right hand side. Since we assumed $c_0 \neq 0$, we have

$$f(z) = 0$$
$$c_1 f(z+1) = 0$$
$$f(z+k) c_k + \frac{1}{2} c_{k-2} = 0 \text{ for } k = 2, 3, 4, \dots$$

$f(z) = 0 \implies z = 0$ or $z = \frac{1}{2}$.

1. When $z = \frac{1}{2}$, this must mean that $c_1 = 0$ in order to satisfy the equation $c_1 f(z+1) = 0$. We can rewrite the relation $f(z+k)c_k + \frac{1}{2}c_{k-2} = 0$ as

$$c_k = \frac{-1}{2f(k + \frac{1}{2})}c_{k-2} = -\frac{1}{2k(k + \frac{1}{2})}c_{k-2}$$

Remember that $c_0 \neq 0$ is arbitrary and that $c_1 = 0$, so we can manually calculate

$$c_{2m-1} = 0$$

$$c_{2m} = (-1)^m \frac{c_0}{2 \cdot 4 \cdot 6 \ldots 2m \cdot 5 \cdot 9 \ldots (4m+1)} = (-1)^m c_0 \left(\prod_{i=1}^{m} 2i(4i+1) \right)^{-1}$$

for $m = 1, 2, \ldots$. Substituting these quantities into the series $\varphi(t) = t^z \sum_{k=0}^{\infty} c_k t^k$ gives one candidate for a solution of the differential equation for $t > 0$:

$$\varphi_1(t) = t^{1/2} \left(1 + \sum_{m=1}^{\infty} (-1)^m \left(\prod_{i=1}^{m} 2i(4i+1) \right)^{-1} t^{2m} \right)$$

Note that the c_0, which is just a constant factor, has been normalized to 1.

2. When $z = 0$, we find that $c_1 = 0$ (c_0 still arbitrary), but we rewrite the relation $f(z+k)c_k + \frac{1}{2}c_{k-2} = 0$ as

$$c_k = \frac{-1}{2f(k)}c_{k-2} = -\frac{1}{2k(k - \frac{1}{2})}c_{k-2}$$

Leading to

$$c_{2m-1} = 0$$

$$c_{2m} = (-1)^m \frac{c_0}{2 \cdot 4 \cdot 6 \ldots 2m \cdot 3 \cdot 7 \ldots (4m-1)} = (-1)^m c_0 \left(\prod_{i=1}^{m} 2i(4i-1) \right)^{-1}$$

for $m = 1, 2, \ldots$. Substituting and normalizing $c_0 = 1$, we obtain the second candidate for a solution of the DEQ when $t > 0$ as

$$\varphi_2(t) = 1 + \sum_{m=1}^{\infty} (-1)^m \left(\prod_{i=1}^{m} 2i(4i-1) \right)^{-1} t^{2m}$$

Using the ratio test, we can find out that for the series φ_1, φ_2,

$$\left| \frac{a_{m+1}}{a_m} \right| \to 0 \text{ as } m \to \infty$$

where a_m is the mth term of the series. Therefore, both φ_1, φ_2 converges within $(-\infty, \infty)$. Therefore, we have found solutions of the form

$$|t|^z \sum_{k=0}^{\infty} c_k t^k$$

for when $t > 0$ (the value $t = 0$ must be omitted because the differential equation has no meaning at the singular point $t = 0$). It is easy to prove that φ_1 and φ_2 are linearly independent. The above calculations are all valid for $t < 0$ if t^z is replaced by $|t|^z = e^{z \log |t|}$, and still leads to the same solutions φ_1, φ_2 for the interval $(-\infty, 0)$.

Therefore, we can see that the assumption that a given DEQ has a solution of the form

$$|t|^z \sum_{k=0}^{\infty} c_k t^k$$

leads to a quadratic equation in z, the indicial equation, and each root of the indicial equation leads to two linearly independent solutions of the differential equation. However, an indicial polynomial may have a double root, which will require extra work to find the second solution.

Example 4.8

The differential equation $ty'' + y' + y = 0$, which may be written as

$$t^2 y'' + ty' + ty = 0$$

Clearly, $t = 0$ is a regular singular point. Assuming the existence of solution of the form $\varphi(t) = |t|^z \sum_{k=0}^{\infty} c_k t^k$ on some interval, we consider the case $t > 0$. With some steps omitted, we have

$$t^2 \varphi''(t) + t\varphi'(t) + t\varphi(t) = t^z \left(f(z)c_0 + \sum_{k=1}^{\infty} \left(f(k+z)c_k + c_{k-1} \right) t^k \right)$$

where the indicial polynomial is $f(z) = z^2$, having a double root $z = 0$. Since every term on the right hand side vanishes, we have

$$f(k+z)c_k + c_{k-1} = 0 \implies c_k = -\frac{c_{k-1}}{k^2}, \quad k = 1, 2, 3, \ldots$$

$$\implies c_k = c_0 \prod_{i=1}^{k} -\frac{1}{i^2}$$

By substituting in the c_k's, we have (for $t > 0$)

$$\varphi(t) = 1 + t^z \sum_{k=1}^{\infty} \left(\prod_{i=1}^{k} -\frac{1}{i^2} \right) t^k$$

Similar calculations show that the same solution pops up for $t < 0$. The ratio test tells us that the interval of convergence for $\varphi(t)$ is $(-\infty, \infty)$. We will talk about how to find the second solution later.

Note in the previous example that although the differential equation makes no sense at the singular point $t = 0$, the function defined by the series

$$\sum_{k=0}^{\infty} c_k t^k$$

is well defined

Example 4.9

The DEQ $ty'' + ty' - y = 0$, which can be rewritten as

$$t^2 y'' + t^2 y' - ty = 0$$

has $t = 0$ has a regular singular point. When $t > 0$, we calculate

$$^2 y'' + t^2 y' - ty == t^z \left(f(z)c_0 + \sum_{k=1}^{\infty} \left(f(k+z)c_k + (k+z-2)c_{k-1} \right) t^k \right)$$

---

where $f(z) = z(z-1)$. For every term to vanish, $z = 0$ or $z = 1$.
1. Taking $z = 1$, we have

$$c_k = -\frac{(k-1)c_{k-1}}{k(k+1)}, \quad k = 1, 2, \ldots$$

which gives $c_k = 0$ and taking $c_0 = 1$, have have solution $\varphi_1(t) = |t|$, which is clearly convergent.
2. When $z = 0$, the recursion formula becomes

$$k(k-1)c_k + (k-2)c_{k-1} = 0, \quad k = 1, 2, \ldots$$

But taking $k = 1$, c_1 must be determined from the relation

$$0 \cdot c_1 - c_0 = 0$$

But since $c_0 \neq 0$, this is impossible and there can be no solution of the assumed form corresponding to the root $z = 0$. However, a solution of a different form does exist.

4.3.3 Regular Singular Point Theorem

We can neatly summarize our theory of finding solutions of equation with singular points with the following theorem.

Theorem 4.3 (Regular Singular Point Theorem of 2nd-Order Homogeneous Linear DEQs)

Consider the differential equation

$$t^2 y'' + t\alpha_1(t)y' + \beta(t)y = 0$$

where α_1, β are analytic at regular singular point $t = 0$ and have expansions

$$\alpha(t) = \sum_{k=0}^{\infty} \alpha_k t^k, \quad \beta(t) = \sum_{k=0}^{\infty} \beta_k t^k$$

which converge for $|t| < r$ for some $r > 0$. Let z_1, z_2 be the roots of the indicial equation

$$f(z) = z(z-1) + \alpha_0 z + \beta_0 = 0$$

with $\text{Re}(z_1) \geq \text{Re}(z_2)$. Then, there is always a solution of the form

$$\varphi_1(t) = |t|^{z_1} \sum_{k=0}^{\infty} c_k t^k \quad (c_0 = 1)$$

in the punctured interval $0 < |t| < r$ whose coefficients c_k can be determined recursively from the equations

$$f(z_1 + k)c_k = -\sum_{j=0}^{k-1} \big((j+z)\alpha_{k-j} + \beta_{k-j}\big)c_j, \quad k = 1, 2, \ldots$$

As for the second solution, there are three possible cases:
1. If $z_1 \neq z_2, z_1 - z_2 \notin \mathbb{Z}$, then there is a second, linearly independent solution of the form

$$\varphi_2(t) = |t|^{z_2} \sum_{k=0}^{\infty} \hat{c}_k t^k \quad (\hat{c}_0 = 1)$$

also in the punctured interval $0 < |t| < r$. The coefficients $\hat{c}_k$ can also be solved using the recursive formula above, with z_1 replaced by z_2 and c_k replaced by $\hat{c}_k$.

---

2. If $z_1 = z_2$, then the second (linearly independent) solution is of form

$$\varphi_2(t) = |t|^{z_1+1}\left(\sum_{k=0}^{\infty} b_k t^k\right) + \varphi_1(t)\log|t|$$

valid for $0 < |t| < r$, whose coefficients b_k can be determined by direct substitution into the DEQ.

3. If $z_1 - z_2 \in \mathbb{N}$, then there is a second linearly independent solution of form

$$\varphi_2(t) = |t|^{z_2}\left(\sum_{k=0}^{\infty} b_k t^k\right) + a\varphi_1(t)\log|t|$$

valid for $0 < |t| < r$, where a is a constant (possibly 0) and the coefficients b_k can be determined by direct substitution into the DEQ.

Note that it is simpler in practice to substitute the assumed formula of the solution into the DEQ than to use the recursive formulas to solve for the coefficients.

4.3.4   The Bessel Equation

Definition 4.4 (Bessel Equation)

The **Bessel equation** arises in a natural way in many mathematical physics problems having axial/-cylindrical symmetry, and is written in the form

$$\mathcal{L}_n(y) = t^2 y'' + ty' + (t^2 - p^2)y = 0$$

where constant $p \in \mathbb{C}$ and $\text{Re}(p) \geq 0$. The point $t = 0$ is a regular singular point, and the equation is of form $t^2 y'' + t\alpha(t)y' + \beta(t)y = 0$.


Definition 4.5

Recall the gamma function
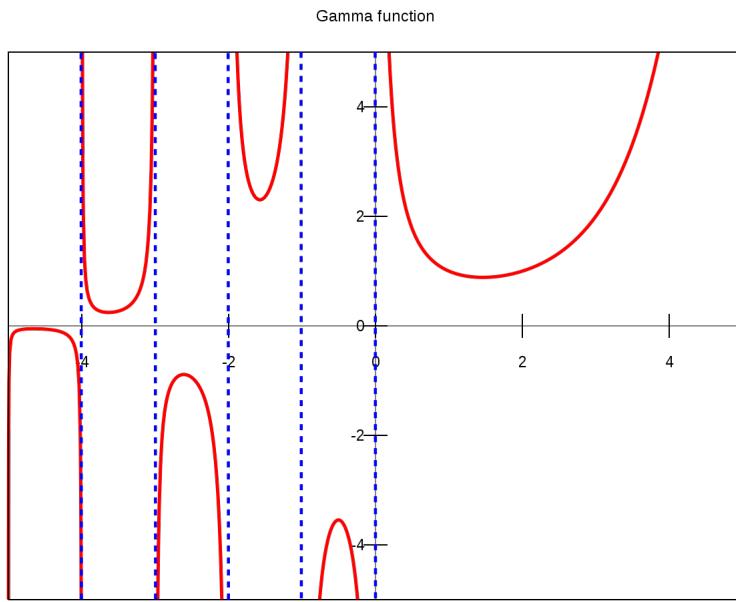
$$\Gamma(z) = \int_0^{\infty} e^{-x} x^{z-1}\, dx, \quad \text{Re}(z) > 0$$

and its properties:

1. $\Gamma(z)$ is well defined and continuous for $\text{Re}(z) > 0$.
2. $\Gamma(1) = 1, \Gamma(\frac{1}{2}) = \sqrt{\pi}$
3. $\Gamma(z) = (z-1)\Gamma(z-1)$ for $\text{Re}(z) > 1$ from integration by parts.
4. $\Gamma(z) = z!$ for $z \in \mathbb{N}$.
5. We can define Γ for negative non-integer numbers z using the recursion formula

$$\Gamma(z) = \frac{\Gamma(z+k)}{z(z+1)\ldots(z+k-1)}$$

by choosing a positive integer k such that $\text{Re}(z+k) > 0$. Therefore, we can define Γ for all complex numbers z except for when $\text{Re}(z) \in \{0, -1, -2, -3, \ldots\}$. The graph of Γ in $\mathbb{R} \times \mathbb{R}$ is shown.

Gamma function

Theorem 4.4 (The Bessel Function)

Given the Bessel equation

$$t^2 y'' + t y' + (t^2 - p^2) y = 0, \ \ p \in \mathbb{C}, \mathrm{Re}(p) \geq 0$$

where $t = 0$ is a regular singular point.

1. If $p \notin \mathbb{N} \cup \{0\}$, then the following *Bessel functions of the first kind J* (of index $p, -p$)

$$J_p(t) = \left| \frac{t}{2} \right|^p \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(p+m+1)} \left(\frac{t}{2} \right)^{2m}$$

$$J_{-p}(t) = \left| \frac{t}{2} \right|^{-p} \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m-p+1)} \left(\frac{t}{2} \right)^{2m}$$

are two linearly independent solutions for $0 < |t| < \infty$.

2. If $p = 0$, then the Bessel function of the first kind and second kind, both of index 0

$$J_0(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{(m!)^2} \left(\frac{t}{2} \right)^{2m}$$

$$K_0(t) = - \sum_{m=1}^{\infty} \left(\frac{(-1)^m}{(m!)^2} \left(1 + \frac{1}{2} + \ldots + \frac{1}{m} \right) \left(\frac{t}{2} \right)^{2m} \right) + J_0(t) \log t$$

are two linearly independent solutions for $0 < |t| < \infty$.

3. If p is a positive integer n, then the Bessel function of the first kind of index n and the *Bessel*

---

function of the second kind K (of index n)

$$J_n(t) = \left|\frac{t}{2}\right|^n \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(n+m+1)} \left(\frac{t}{2}\right)^{2m}$$

$$K_n(t) = -\frac{1}{2}\left|\frac{t}{2}\right|^{-n} \left(\sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!}\left(\frac{t}{2}\right)^{2k} + \frac{1}{n!}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right)\left(\frac{t}{2}\right)^{2n}\right)$$

$$- \frac{1}{2}\left(\frac{t}{2}\right)^n \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+n)!}\left(\left(1 + \ldots + \frac{1}{k}\right) + \left(1 + \frac{1}{2} + \ldots + \frac{1}{k+n}\right)\right)\left(\frac{t}{2}\right)^{2k}$$

$$+ J_n(t)\log|t|$$

Proof. We now derive the solutions to this Bessel equation. By the regular singular point theorem, since

$$\alpha(t) = 1, \quad \beta(t) = -p^2 + t^2$$

which both converge for $|t| < \infty$, the indicial equation is

$$f(z) = z(z-1) + z + -p^2 = z^2 - p^2$$

which has zeroes $z_1 = p, z_2 = -p$ (remember we assumed $(\text{Re}(p) \geq 0)$. If $p \neq 0$ and if $z_1 - z_2 = 2p$ is not a positive integer (i.e. p is not 0, an integer, or a half-integer), there exists two linearly independent solutions φ_1, φ_2 of the equation, valid for $0 < |t| < \infty$, of the form

$$\varphi_1(t) = |t|^p \sum_{k=0}^{\infty} c_k t^k \quad (c_0 \neq 0)$$

$$\varphi_2(t) = |t|^{-p} \sum_{k=0}^{\infty} \hat{c}_k t^k \quad (\hat{c}_0 \neq 0)$$

where coefficients $c_k, \hat{c}_k$ are determined recursively by substitution. We first compute φ_1 and assume $t > 0$. Substituting

$$\varphi_1'(t) = \sum_{k=0}^{\infty} c_k (p+k) t^{p+k-1}, \quad \varphi_1'' = \sum_{k=0}^{\infty} c_k (p+k)(p+k-1) t^{p+k-2}$$

into the equation and simplifying gives

$$\mathcal{L}_2(\varphi_1(t)) = t^p \left(f(p)c_0 + f(p+1)c_1 t + \sum_{k=0}^{\infty} (f(p+k)c_k + c_{k-2})t^k\right) = 0$$

for which we conclude that

$$f(p+1)c_1 = 2p + 1 = 0 \implies p = -\frac{1}{2} \text{ or } c_1 = 0$$

But since $\text{Re}(p) \geq 0$, this means that $c_1 = 0$. We use the recursive relations

$$f(p+k)c_k + c_{k-2} = k(2p+k)c_k + c_{k-2} = 0$$

which implies that

$$c_{2m-1} = 0$$

$$c_{2m} = \frac{(-1)^m c_0}{2^{2m}m!(p+1)(p+2)\ldots(p+m)}$$

for $m = 1, 2, \ldots$. Therefore, the solution can be written as

$$\varphi_1(t) = c_0 |t|^p \left(1 + \sum_{m=1}^{\infty} \frac{(-1)^m t^{2m}}{2^{2m} m! (p+1)(p+2) \ldots (p+m)} \right)$$

Since we can let c_0 be any nonzero constant, we define it using the Gamma function as

$$c_0 = \frac{1}{2^p \Gamma(p+1)}$$

resulting in the solution

$$J_p = \left| \frac{t}{2} \right|^p \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(p+m+1)} \left(\frac{t}{2} \right)^{2m}$$

called the *Bessel function of the first kind of index p*. It is well defined for all $t \in \mathbb{R}$ and satisfies the DEQ for $0 < |t| < \infty$.

4.3.5 Singularities at Infinity

We can also study the behavior of solutions of the equation

$$\mathcal{L}_2(y) = y'' + p(t)y' + q(t)y = 0$$

as $|t| \to \infty$ by making the change of variable $t = 1/x$ and studying the behavior of solutions of the resulting equation as $x \to 0$. Thus, we let φ be a solution of the original DEQ for $|t| > R$ and let

$$\psi(x) = \varphi(\frac{1}{x}), \quad \bar{p}(x) = p(\frac{1}{x}), \quad \bar{q}(x) = q(\frac{1}{x})$$

which should all be well-defined for $|x| < 1/R$. By the chain rule, we have

$$\varphi'(t) = -\frac{1}{t^2} \psi'(x) = -x^2 \psi'(x)$$
$$\varphi''(t) = \frac{1}{t^4} \psi''(x) + \frac{2}{t^3} \psi'(x) = x^4 \psi''(x) + 2x^3 \psi'(x)$$

Substituting this in shows that ψ satisfies the new equation

$$\bar{\mathcal{L}}_2(z) = x^4 z'' + \left(2x^3 - x^2 \bar{p}(x) \right) z' + \bar{q}(x) z = 0$$

in which x is the independent variable and z is the function. Therefore, if ψ satisfies $\bar{\mathcal{L}}_2(z) = 0$ and if $\varphi(t) = \psi(1/t)$, then φ satisfies $\mathcal{L}_2(y) = 0$. Our results lead to the following theorem.

Theorem 4.5 (Singularities at Infinity)

∞ is an ordinary point, a regular singular point, or an irregular singular point for the equation $\mathcal{L}_2(y) = 0$ if and only if 0 is respectively an ordinary point, a regular singular point, or an irregular singular point for the equation $\bar{\mathcal{L}}_2(z) = 0$.

Example 4.10

Given the equation
$$y'' + ay' + by = 0, \quad a, b \in \mathbb{R}$$
The change of variable $t = 1/x$ transforms this equation to
$$x^4 z'' + (2x^3 - ax^2)z' + bz = 0$$

which has an irregular singular point at $x = 0$. Therefore, the original equation has an irregular singular point at $t = \infty$.

5 Systems of DEQs

5.1 First-Order Systems

We expand on the theory of solving first order differential equations by studying systems of them, which can be worked on using vector algebra/calculus.

Definition 5.1 (System of 1st Order Equations)

A system of n first order equations can be put in the form
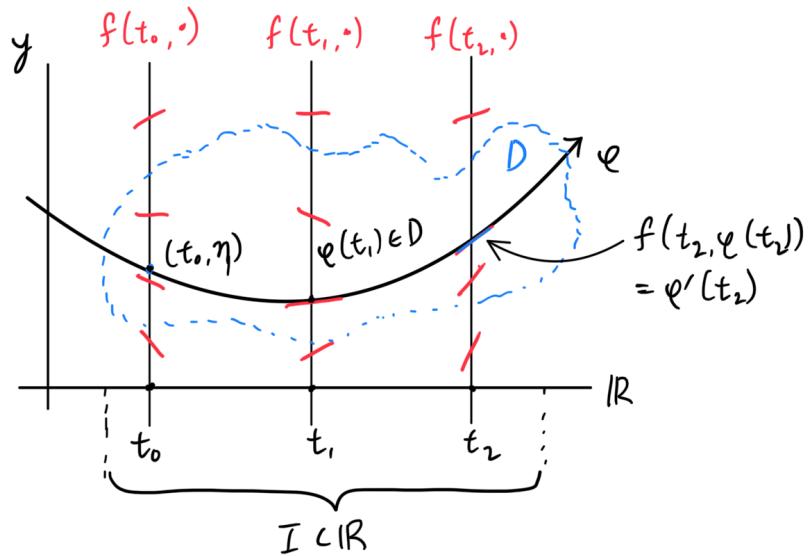
$$y' = f(t, y) \begin{cases} y_1' = f_1(t, y_1, y_2, \ldots, y_n) \\ y_2' = f_2(t, y_1, y_2, \ldots, y_n) \\ \ldots \\ y_n' = f_n(t, y_1, y_2, \ldots, y_n) \end{cases}$$

where $f : D \subset \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is defined in some $(n+1)$-dimensional region D ($\mathbb{R}$ being the time continuum and $\mathbb{R}^n$ being the n-dimensional phase space). That is, at a certain time $t = t_0$,

$$f(t_0, \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
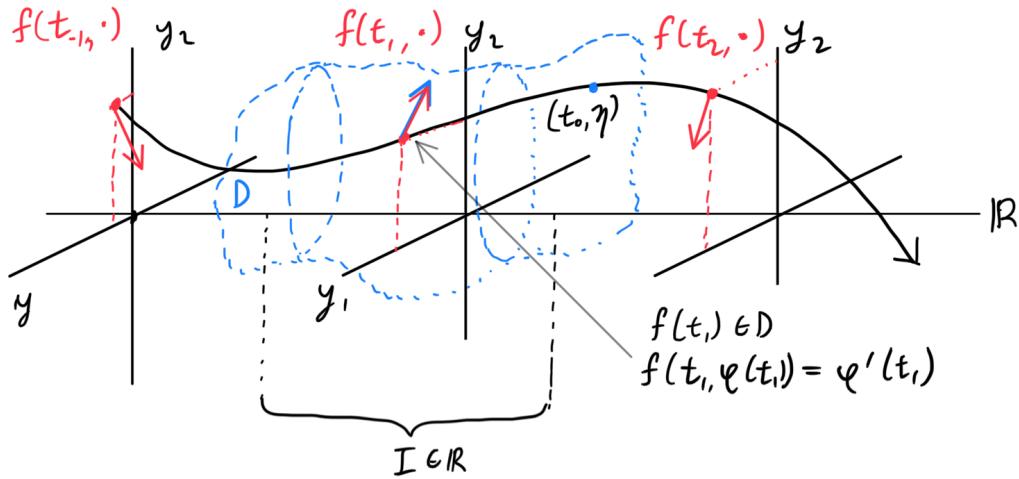
determines the phase velocity vector field of $\mathbb{R}^n$ for that instance of time. If the system is autonomous, then the vector field does not morph. As shown before, we can visualize the following.

1. System with 1-dimensional phase space (i.e. a system of one equation)
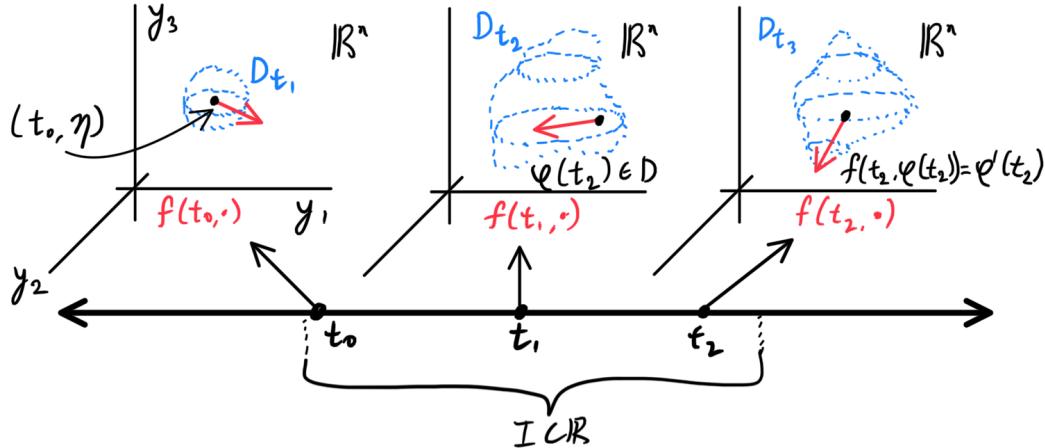
2. System with 2-dimensional phase space (two equations). The solution shown is

$$\varphi : \mathbb{R} \longrightarrow \mathbb{R}^2, \quad y = \varphi(t) \iff \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix}$$

---

3. Systems with higher-dimensional phase spaces are harder to visualize, but we just imagine a time continuum $\mathbb{R}$ where at each point $t_0 \in \mathbb{R}$, there is a vector space $\mathbb{R}^n$ with a vector field $f(t_0, \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ associated with it. In the visual below, the phase space are shown to be $\mathbb{R}^3$ (when it is really $\mathbb{R}^n$), with D being represented by its cross sections in the time axis (e.g. D_{t_0} is the cross section of D at $t = t_0$). The solution curve φ isn't explicitly shown (only the points on the curve at times $t = t_0, t_1, t_2$), but it would be of form

$$\varphi : \mathbb{R} \longrightarrow \mathbb{R}^3, y = \varphi(t) \iff \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix}(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \varphi_3(t) \end{pmatrix}$$

To find the solution of the system means to find the n equations $y_1 = \varphi_1(t), y_2 = \varphi_2(t), \ldots, y_n = \varphi_n(t)$, which is equivalent to finding the vector-valued path function

$$y = \varphi(t) = \begin{pmatrix} \varphi_1(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix}, \varphi : I \subset \mathbb{R} \longrightarrow \mathbb{R}^n$$

such that
1. the point $(t, \varphi(t)) \in D$ for each $t \in I$
2. $\varphi'(t)$ exists for each $t \in I$
3. $\varphi'(t) = f(t, \varphi(t))$ for every $t \in I$

To solve an initial value problem for the system with initial condition

$$\varphi(t_0) = \eta, \quad \eta \in \mathbb{R}^n \text{ and } (t_0, \eta) \in D$$

means to find a solution φ of the system such that $\varphi(t_0) = \eta$.

The study of first order equations is quite nice because we can reduce a high-order scalar differential equation of form

$$y^{(n)} = g(t, y, y', \ldots, y^{(n-1)})$$

to the following system with a change of variable $y_1 = y, y_2 = y', \ldots, y_n = y^{(n-1)}$.

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}' = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ g(t, y_1, y_2, \ldots, y_n) \end{pmatrix} \iff \begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \ldots \\ y_{n-1}' = y_n \\ y_n' = g(t, y_1, y_2, \ldots, y_n) \end{cases}$$

Example 5.1

We can write

$$2y'' - 5y' + y = 0, \quad y(3) = 6, y'(3) = -1$$

We can define the following functions to get

$$\begin{cases} x_1(t) = y(t) \\ x_2(t) = y'(t) \end{cases} \implies \begin{cases} x_1' = y' = x_2 \\ x_2' = y'' = -\frac{1}{2}x_1 + \frac{5}{2}x_2 \end{cases}$$

This gives the matrix equation

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -1/2 & 5/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(3) = \begin{pmatrix} 6 \\ -1 \end{pmatrix}$$

Solving this system is equivalent to finding a function $\varphi : \mathbb{R} \longrightarrow \mathbb{R}^2$ satisfying the equation.

It naturally extends that a system of high-order scalar differential equations can be reduced to a system of systems (creating a larger system) of first-order differential equations.

Example 5.2

The system

$$y'' = \sin(t)y' + 2y - 4$$
$$y''' = e^t y'' - 2\sin(4\pi t)y$$

can be reduced (with substitution $y_1 = y, y_2 = y', y_3 = y''$) to

$$y_1' = y_2$$
$$y_2' = y_3$$
$$y_2' = \sin(t)y_2 + 2y_1 - 4$$
$$y_3' = e^t y_3 - 2\sin(4\pi t)y_1$$

Note that even though this system consists of first-order equations, it cannot be put simply into vector form since there are two equations involving y_2'. We can attempt to solve the system excluding $y_2' = y_3$

Theorem 5.1 (Existence, Uniqueness of Solutions in a System of First-Order DEQs)
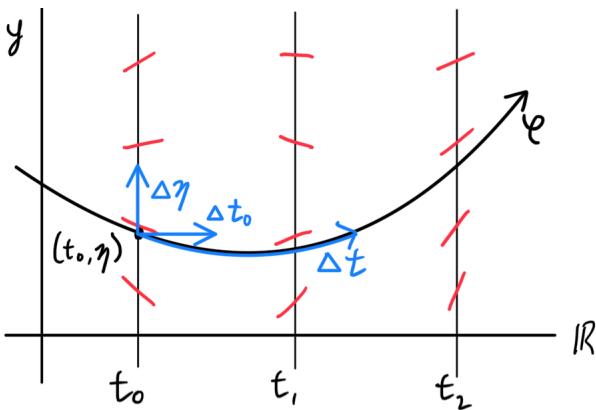
Given the system of first-order DEQs
$$y' = f(t, y)$$
Let the partial derivatives of $f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$

$$\frac{\partial f}{\partial y_k}, \quad k = 1, 2, \ldots, n$$

be continuous in D. That is, for a given $t = t_0$, the phase velocity vector field $f(t_0, \cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is C^1. Then, given any initial point $(t_0, \eta) \in D$, there exists a unique solution φ of the system $y' = f(t, y)$ satisfying the initial condition $\varphi(t_0) = \eta$. The solution φ exists for any interval I containing t_0 for which the points $(t, \varphi(t))$ lie in D. Furthermore, the solution φ is continuous with respect to t_0, t, and η.

We briefly state a useful result.

Lemma 5.2 (Gronwall Inequality)

Let $K \geq 0$ be a constant and f, g be nonnegative C^0 functions on some interval $\alpha \leq t \leq \beta$ satisfying the inequality

$$f(t) \leq K + \int_\alpha^t f(s) g(s) \, ds \text{ for } \alpha \leq t \leq \beta$$

Then,

$$f(t) \leq K \exp\left(\int_\alpha^t g(s) \, ds \right) \text{ for } \alpha \leq t \leq \beta$$

5.2 Linear Systems of DEQs

Remember that higher-order systems can be reduced to a system of first-order DEQs, so it makes sense to talk about a system of first-order linear DEQs.

Definition 5.2 (Linear Systems of DEQs)

The system $y' = f(t, y)$ linear in the components of y has the form

$$y_1' = a_{11}(t)y_1 + a_{12}(t)y_2 + \ldots + a_{1n}(t)y_n + g_1(t)$$
$$y_2' = a_{21}(t)y_1 + a_{22}(t)y_2 + \ldots + a_{2n}(t)y_n + g_2(t)$$
$$\vdots = \vdots$$
$$y_n' = a_{n1}(t)y_1 + a_{n2}(t)y_2 + \ldots + a_{nn}(t)y_n + g_n(t)$$

which can be represented as a matrix equation, where $A : \mathbb{R} \longrightarrow \mathrm{Mat}(n \times n, \mathbb{R}), g : \mathbb{R} \longrightarrow \mathbb{R}^n$.

$$y' = A(t)y + g(t) \iff \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}' = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \ldots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \ldots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \ldots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_n(t) \end{pmatrix}$$

It is actually true that there exists a unique solution to this linear system as long as η is finite. This is formalized in the theorem below.

Theorem 5.3 (Existence of Unique Solutions of Linear Systems)

Let matrix-valued function $A(t)$ and vector valued function $g(t)$ be continuous on some interval $a \leq t \leq b$. Then, if $a \leq t_0 \leq b$ and if $|\eta| < \infty$, the system of linear DEQs

$$y' = A(t)y + g(t)$$

has a unique solution $y = \varphi(t)$ satisfying initial condition $\varphi(t_0) = \eta$ and existing on the interval $a \leq t \leq b$.

5.2.1 Homogeneous Linear Systems of DEQs

Note that a first-order homogeneous linear system is of form

$$y' = A(t)y$$

Note that the scalar analogue of a homogeneous first-order linear DEQ can be changed to the above form as such

$$a_0(t)y' + a_1(t)y = 0 \implies a_0(t)y' = -a_1(t)y$$
$$\implies y' = -\frac{a_1(t)}{a_0(t)}y$$

where we treat y, y' as 1-dimensional vectors, and $-a_1(t)/a_0(t)$ as a 1×1 matrix.

Just as in the scalar case, the solutions to the linear system forms a vector space.

Theorem 5.4 (Fundamental Set of Solutions)

Given the first-order linear homogeneous system

$$y' = A(t)y$$

with $A : \mathbb{R} \longrightarrow \mathrm{Mat}(n \times n, \mathbb{C})$ being continuous over interval $I \subset \mathbb{R}$, the solutions of this system over I form an n-dimensional subspace within $C^1(\mathbb{R}, \mathbb{C}^n)$, the space of all continuously differentiable functions

mapping $\mathbb{R} \longrightarrow \mathbb{C}^n$. The basis of this space is called the *fundamental set of solutions.*

Proof. We can rearrange the system as

$$-y_1' + a_{11}(t)y_1 + a_{12}(t)y_2 + \ldots + a_{1n}(t)y_n = 0$$
$$-y_2' + a_{21}(t)y_1 + a_{22}(t)y_2 + \ldots + a_{2n}(t)y_n = 0$$
$$\vdots = \vdots$$
$$-y_n' + a_{n1}(t)y_1 + a_{n2}(t)y_2 + \ldots + a_{nn}(t)y_n = 0$$

and put it into matrix form

$$\begin{pmatrix} a_{11}(t) & \ldots & a_{1n}(t) & -1 & \ldots & 0 \\ \vdots & \ddots & \vdots(t) & \vdots & \ddots & \vdots \\ a_{n1}(t) & \ldots & a_{nn}(t) & 0 & \ldots & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ y_1' \\ \vdots \\ y_n' \end{pmatrix} = 0$$

The kernel of this linear mapping has dimension n.

Example 5.3 (4th Order DEQ into System of 1st-Order DEQs in Matrix Form)

We can write the following homogeneous 4th order differential equation

$$y^{(4)} + 3y'' - \sin(t)y' + 8y = t^2, \quad y(0) = 1, y'(0) = 2, y''(0) = 3, y^{(3)}(0) = 4$$

into the system by making the substitutions

$$x_1 = y \implies x_1' = y' = x_2$$
$$x_2 = y' \implies x_2' = y'' = x_3$$
$$x_3 = y'' \implies x_3' = y^{(3)} = x_4$$
$$x_4 = y^{(3)} \implies x_4' = y^{(4)} = -8y + \sin(t)y' - 3y'' + t^2 = -8x_1 + \sin(t)x_2 - 3x_3 + t^2$$

which leads to the matrix equation

$$x' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -8 & \sin(t) & -3 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ 0 \\ t^2 \end{pmatrix}, x(0) = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

Definition 5.3 (Solution Matrix, Fundamental Matrix)

Given that $\varphi_1, \varphi_2, \ldots, \varphi_n : \mathbb{R} \longrightarrow \mathbb{R}^n$ are n solutions to the homogeneous matrix differential equation $y' = A(t)y$, the $n \times n$ matrix whose columns are solutions is called a *solution matrix.*
Then $n \times n$ matrix with its columns being the n linearly independent solutions on I is called the

fundamental matrix.

$$\Phi = \begin{pmatrix} | & | & \cdots & | \\ \phi_1 & \phi_2 & \cdots & \phi_n \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} \phi_{11}(t) & \phi_{21}(t) & \cdots & \phi_{n1}(t) \\ \phi_{12}(t) & \phi_{22}(t) & \ddots & \phi_{n2}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{1n}(t) & \phi_{2n}(t) & \cdots & \phi_{nn}(t) \end{pmatrix}$$

Note that a fundamental matrix is not unique (that is, two different fundamental matrices $\Phi, \tilde{\Phi}$ can exist), but since the column space of Φ and $\tilde{\Phi}$ are the same, we can change any fundamental matrix into another with a linear change of basis. That is, given the passive transformation matrix P,

$$\Phi = P^{-1}\tilde{\Phi}P$$

Furthermore, the fundamental matrix $\Phi(t)$ has the property that any solution of the matrix DEQ can be expressed as a linear combination of the column space of $\Phi(t)$. That is, every solution $\psi(t)$ can be constructed as

$$\psi(t) = \Phi(t)c, \quad c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

Theorem 5.5 (Abel's Formula)

If Φ is a solution matrix of $y' = A(t)y$ on I and if t_0 is any point of I, then

$$\det \Phi(t) = \det \Phi(t_0) \exp\left(\int_{t_0}^{t} \sum_{j=1}^{n} a_{jj}(s)\, ds \right) \text{ for every } t \in I$$

It immediately follows that since t_0 is arbitrary, either
1. $\det \Phi(t) \neq 0$ for each $t \in I$, or
2. $\det \Phi(t) = 0$ for every $t \in I$

Corollary 5.6

A solution matrix Φ of the matrix equation

$$y' = A(t)y$$

on interval I is a fundamental matrix if and only if

$$\det \Phi(t) \neq 0 \text{ for every } t \in I$$

Practically, this means that to test whether a solution matrix is a fundamental matrix, it suffices to evaluate its determinant at one point!

Note that this corollary has a very close relationship with the previously mentioned theorem on how the Wronskian is used to determine the linear independence of solutions to a linear DEQ $\mathcal{L}_n(y) = 0$. More specifically, if we take a nth order linear DEQ satisfying the hypothesis of the Wronskian theorem, we can change it to a system of first-order linear equations and apply the previous corollary to determine linear independence of solutions. Both approaches are exactly the same.

5.2.2   Linear Inhomogeneous Systems

It is obvious that due to the *forcing term* $g(t)$ (representing the external force on the system), the form of solution of the inhomogeneous system

$$y' = A(t)y + g(t)$$

is

$$\varphi(t) = \psi(t) + \Phi(t)c$$

where ψ is a special solution to the DEQ over interval I. Since we know how to find the fundamental matrix Φ, all it remains is to find ψ. We can do this using the multivariate variation of constants formula.

For notational purposes, we write for vector valued functions $f : \mathbb{R} \longrightarrow \mathbb{R}^n$, where $f_1, f_2, \ldots, f_n$ are the basis functions,

$$\int_a^b f(x)\, ds = \begin{pmatrix} \int_a^b f_1(x)\, ds \\ \vdots \\ \int_a^b f_n(x)\, ds \end{pmatrix}$$

Theorem 5.7 (Multivariate Variation of Constants Formula)

Given the inhomogeneous first-order linear system

$$y' = A(t)y + g(t)$$

if Φ is the fundamental matrix of its homogeneous counterpart $y' = A(t)y$ on interval $I \subset \mathbb{R}$, then the special function given by the variation of constants formula

$$\psi(t) = \Phi(t) \int_{t_0}^t \Phi^{-1}(s)g(s)\, ds$$

is the unique solution of the original DEQ satisfying the initial condition on I

$$\psi(t_0) = 0$$

(Notice that Φ^{-1} is always well defined since Φ is nonsingular) This means that every solution ϕ of the DEQ has the form

$$\phi(t) = \psi(t) + \phi_h(t)$$

where φ_h is the solution of the homogeneous system satisfying the same initial condition at t_0 as φ (e.g. $\varphi_h(t_0) = \eta$). This makes it such that the initial conditions are met:

$$\varphi(t_0) = \psi(t_0) + \varphi_h(t_0) = 0 + \eta = \eta$$

5.3   Linear Systems with Constant Coefficients

If the homogeneous system

$$y' = A(t)y$$

has constant coefficients (i.e. A consists of constant terms), then we can obtain an explicit formula of the fundamental matrix.

Also, to reduce confusion,

1. Ax will be used to denote matrix A multiplied by vector x

2. xA will be used to denote scalar x multiplied to matrix A (row vector multiplication will not be seen in this section)

Theorem 5.8 (Fundamental Matrix of a Linear System with Constant Coefficients)

Given the system of n linear equations

$$y' = Ay, \quad A \in \mathrm{Mat}(n \times n, \mathbb{C})$$

the matrix

$$\Phi(t) = e^{(t-t_0)A} = \sum_{m=0}^{\infty} \frac{((t-t_0)A)^m}{m!}$$

is the fundamental matrix with $\Phi(t_0) = I$ on $(-\infty, \infty)$.

Proof. $\Phi(t_0) = I$ is obvious from substitution. Assuming $t_0 = 0$, we differentiate $At : \mathbb{R} \longrightarrow \mathrm{Mat}(n \times n, \mathbb{R})$ with respect to t to get

$$\left(e^{tA}\right)' = A + \frac{tA^2}{1!} + \frac{t^2 A^3}{2!} + \ldots + \frac{t^{k-1}A^k}{(k-1)!} + \ldots = Ae^{tA}$$

This means that given any constant vector $c \in \mathbb{R}^n$, we have

$$\left(e^{tA}\right)'c = Ae^{tA}c \implies \left(e^{tA}c\right)' = A(e^{tA}c)$$

and so e^{tA} is a solution matrix. Furthermore, since $\det \Phi(0) = \det I = 1$, it is a fundamental matrix (by the corollary to Abel's formula).

Note that given an arbitrary matrix A, it is stupid to calculate e^A explicitly. Rather, we can use a change of basis to put it into Jordan Canonical Form J.

$$A = P^{-1}JP$$

Then, it is clear that

$$e^A = e^{P^{-1}JP} = P^{-1}e^J P$$

Finding the eigendecomposition of a linear mapping, which may require working over the field $\mathbb{C}$ (or by introducing generalized eigenvectors over $\mathbb{R}$). This entire process is talked in detail in the linear algebra chapter.

Another trick that may work in some cases is the following lemma.

Lemma 5.9

If two $n \times n$ matrices commute, then

$$e^M \cdot e^P = e^{M+P}$$

Proof. Trivial using the Baker-Campbell-Hausdorff formula.

Finally, we mentioned a greatly simplified variation of constants formula for the case when we solve a inhomogeneous system of linear DEQs with constant coefficients.

Corollary 5.10 (Multivariate Variation of Constants Formula for Constant Coefficient Systems)

Given inhomogeneous system of N linear DEQs

$$y' = Ay + g(t), \quad A \in \mathrm{Mat}(n \times n, \mathbb{R}), g \in C^0$$

with initial conditions $\varphi(t_0) = \eta$, then a specific solution of the system is

$$\psi(t) = \int_{t_0}^{t} e^{(t-s)A} g(s) \, ds, \text{ where } \psi(t_0) = 0$$

meaning that the general solution to this equation is

$$\varphi(t) = e^{(t-t_0)A} \eta + \int_{t_0}^{t} e^{(t-s)A} g(s) \, ds \qquad -\infty < t < \infty$$

where $\varphi(t_0) = \eta$.

Example 5.4

Given the initial value problem

$$y' = Ay + g(t) \iff \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} e^{3t} \\ 1 \end{pmatrix}, \quad \varphi(0) = \eta$$

Then, we have

$$\Phi(t) = e^{tA} = e^{3t} \begin{pmatrix} 1-t & t \\ -t & 1+t \end{pmatrix}$$

$$\Phi(t)\Phi^{-1}(s) = e^{(t-s)A} = e^{3(t-s)} \begin{pmatrix} 1-(t-s) & t-s \\ -(t-s) & 1+(t-s) \end{pmatrix}$$

$$e^{(t-s)A} g(s) = e^{3t} \begin{pmatrix} 1-(t-s)+e^{-3s}(t-s) \\ -(t-s)+e^{-3s}(1+t-s) \end{pmatrix}$$

Therefore,

$$\varphi(t) = e^{3t} \begin{pmatrix} 1-t & t \\ -t & 1+t \end{pmatrix} \eta + e^{3t} \int_0^t \begin{pmatrix} 1-(t-s)+e^{-3s}(t-s) \\ -(t-s)+e^{-3s}(1+t-s) \end{pmatrix} ds$$

which can be easily evaluated.

5.3.1   Asymptotic Behavior of Solutions

In many cases, in order to apply the variation of constants formula

$$\varphi(t) = e^{(t-t_0)A} \eta + \int_{t_0}^{t} e^{(t-s)A} g(s) \, ds \qquad -\infty < t < \infty$$

and others derived from it, we must know how the matrix e^{tA} behaves. For example, in order to measure the growth of solutions of as $t \to \infty$, we need to estimate

$$\lim_{t \to \infty} \varphi(t)$$

where $|\cdot|$ is the L-1 norm. But this cannot be done without knowing a useful estimate of $|e^{tA}|$.

Theorem 5.11

Given $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A, where λ_j has multiplicity n_j and $n_1 + \ldots + n_k = n$, let ρ be any real number such that

$$\rho > \max_{j=1,\ldots,k} \big(\text{Re}(\lambda_i) \big)$$

Then, there exists a constant $K > 0$ such that

$$|e^{tA}| \leq Ke^{t\rho} \text{ for } 0 \leq t < \infty$$

Its applications really lies within its corollary.

Corollary 5.12

If all eigenvalues of A have real parts negative, then every solution $\varphi(t)$ of the system

$$y' = Ay$$

approaches 0 as $t \to +\infty$. More precisely, there exists constants $\tilde{K} > 0$, $\sigma > 0$ such that

$$|\varphi(t)| < \tilde{K}e^{-\sigma t}, \quad (0 \leq t < \infty)$$

Theorem 5.13 (Upper Bound on Growth Rate of Solutions of Nonhomogeneous Linear System with Constant Coefficients)

Suppose that in the linear inhomogeneous system

$$y' = Ay + g(t)$$

the function $g(t)$ grows no faster than an exponential function; that is, there exists real constants $M > 0, T \geq 0$, and $a \in \mathbb{R}$ such that

$$|g(t)| \leq Me^{at} \text{ for all } t \geq T$$
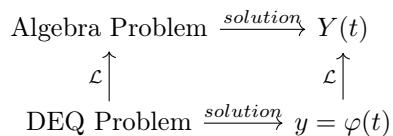
Then, every solution φ of the system grows no faster than an exponential function. That is, there exists real constants $K > 0, b$ such that
$$|\varphi(t)| \leq Ke^{bt} \text{ for all } t \geq T$$

The derivative $\varphi'(t)$ also grows no faster than an exponential function.

---

6   Laplace Transforms

Laplace transforms extends our toolkit for solving linear differential equations (and systems of them) by reducing an initial value problem into an algebra problem, which can be summarized by the diagram below, where $\mathcal{L}$ represents the transform.

$$
\begin{array}{ccc}
\text{Algebra Problem} & \xrightarrow{\text{solution}} & Y(t) \\
\mathcal{L}\uparrow & & \mathcal{L}\uparrow \\
\text{DEQ Problem} & \xrightarrow{\text{solution}} & y = \varphi(t)
\end{array}
$$

Laplace transforms also allows us to work with piecewise continuous functions.

The method of Laplace Transforms does not actually allow us to solve new types of differential equations. Rather, it is useful because it enables us to find one particular solution of the differential equation which satisfies the initial conditions directly, rather than first finding the general solution and then using the initial conditions to determine constants.

Definition 6.1 (Functions of exponential growth at infinity)

A function on $0 < t < \infty$ is said to be *of exponential growth at infinity* if it satisfies

$$|f(t)| \leq M e^{ct}$$

for some real constants $M > 0$ and c, for all sufficiently large t.

Definition 6.2 (Function Class Λ)

The *class* Λ of functions is defined as those functions on $0 < t < \infty$ which are
  1. absolutely integrable on 0
  2. piecewise continuous on $(0, \infty)$
  3. of exponential growth at infinity
Clearly, the functions $1, t, t^n (n \in \mathbb{N}), \sin t, \cos t, e^{zt} (z \in \mathbb{C})$ are in class Λ, but $e^{t^2} \notin \Lambda$.

Definition 6.3 (Laplace Transform)

Let $f \in \Lambda$. The *Laplace transform* of f is denoted $\mathcal{L}\{f(t)\}$ or $F(t)$, defined

$$F(s) \equiv \int_0^\infty f(t) e^{-st}\, dt = \lim_{A \to \infty} f(t) e^{-st}\, dt$$

Sometimes, the Laplace transform may be defined by setting the lower limit to $-\infty$.

$$F(t) \equiv \int_{-\infty}^\infty f(t) e^{-st}\, dt$$

Example 6.1

The Laplace transformation of constant function 1 is

$$\int_0^\infty e^{-st}\, dt = \lim_{A \to \infty} \int_0^A e^{-st}\, dt = \lim_{A \to \infty} \left(\frac{1}{s} - \frac{e^{-sA}}{s} \right) = \frac{1}{s} \quad (\mathrm{Re}(s) > 0)$$

Clearly, the integral does not converge for $\mathrm{Re}(s) \leq 0$. The Laplace transform of e^{zt} is

$$\mathcal{L}(e^{zt}) = \frac{1}{s - z} \quad (\mathrm{Re}(s) > \mathrm{Re}(z))$$

From the examples before and our notation of the transform, we can see and prove the following.

Theorem 6.1 (Linearity of the Laplace Transform)

The Laplace transform is a linear map with respect to its function argument. That is,

$$\mathcal{L}\{af + bg\} = a\mathcal{L}\{f\} + b\mathcal{L}\{g\}$$

This means that if $f : \mathbb{R} \longrightarrow \mathbb{C}$ is a complex-valued function

$$f(t) = u(t) + iv(t)$$

Then,

$$\mathcal{L}\big(f(t)\big) = \mathcal{L}\big(u(t) + iv(t)\big) = \mathcal{L}\big(u(t)\big) + i\mathcal{L}\big(v(t)\big)$$

That is,
$$\operatorname{Re}\big((\mathcal{L}(f(t)))=\mathcal{L}(\operatorname{Re}(f(t))), \operatorname{Com}\big((\mathcal{L}(f(t)))=\mathcal{L}(\operatorname{Com}(f(t)))$$

6.1   Common Laplace Transforms

Example 6.2

Given $z = \alpha + i\beta \in \mathbb{C}$, with $\alpha, \beta \in \mathbb{R}$, and $f(t) = e^{zt}$, we have

$$\mathcal{L}(e^{zt}) = \mathcal{L}(e^{\alpha t}e^{i\beta t}) = \mathcal{L}\big(e^{\alpha t}\cos\beta t + ie^{\alpha t}\sin\beta t\big)$$
$$= \frac{1}{s-z} = \frac{1}{z - \alpha - i\beta}$$
$$= \frac{s - \alpha + i\beta}{(s-\alpha)^2 + \beta^2}$$

By linearity, we can take the real and complex parts of this transform to get

$$\mathcal{L}\big(e^{\alpha t}\cos\beta t\big) = \frac{s - \alpha}{(s-\alpha)^2 + \beta^2} \tag{5}$$
$$\mathcal{L}\big(e^{\alpha t}\sin\beta t\big) = \frac{\beta}{(s-\alpha)^2 + \beta^2} \tag{6}$$

Taking $\alpha = 0$ gives us

$$\mathcal{L}\big(\cos\beta t\big) = \frac{s}{s^2 + \beta^2}, \quad \mathcal{L}\big(\sin\beta t\big) = \frac{\beta}{s^2 + \beta^2} \quad \operatorname{Re}(s) > 0$$

To compute the Laplace transforms functions of some other forms, we can use the following theorem.

Theorem 6.2

The Laplace transform of a function f in the class Λ has derivatives of all orders, given by

$$F^{(k)}(s) = (-1)^k \int_0^\infty t^k e^{-st} f(t)\, dt, \quad k = 1, 2, \ldots$$

Furthermore, $f(t) \in \Lambda \implies t^k f(t) \in \Lambda$ for every positive integer k, and its Laplace transform is given by

$$\mathcal{L}\big(t^k f(t)\big) = (-1)$$

Corollary 6.3

It follows immediately that

$$\mathcal{L}\big(t^k\big) = (-1)^k \frac{d^k}{ds^k}\left(\frac{1}{s}\right) = \frac{k!}{s^{k+1}} \quad (\operatorname{Re}(s) > 0)$$
$$\mathcal{L}\big(t^k e^{zt}\big) = (-1)^k \frac{d^k}{ds^k}\left(\frac{1}{s-z}\right) = \frac{k!}{(s-z)^{k+1}} \quad (\operatorname{Re}(s) > \operatorname{Re}(z))$$

Another theorem for computing Laplace transforms.

Theorem 6.4

If the function $f \in \Lambda$ has a Laplace transform F, then for any constant $a \in \mathbb{C}$,

$$\mathcal{L}\big(e^{at}f(t)\big) = F(s-a)$$

Proof. It is easy to prove that $e^{at}f(t) \in \Lambda$, since

$$|f(t)| \leq Me^{ct} \implies |e^{at}f(t)| \leq e^{at}Me^{ct} = Me^{(a+c)t}$$

and thus $e^{at}f(t)$ is of exponential growth at infinity. Calculating the Laplace transform, we get

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

We conclude this subsection by providing a table of common Laplace transforms.

Theorem 6.5 (Table of Common Laplace Transforms)

Here we have some common transforms, where $a \in \mathbb{R}$ is a constant. Note that the gamma function

$$\Gamma(t) \equiv \int_0^\infty e^{-x}\,x^{t-1}\,dx$$

is an extension of the factorial function to the real numbers.

1. $f(t) = e^{at} \implies F(t) = \frac{1}{s-a}$
2. $f(t) = t^n, n \in \mathbb{N} \implies F(t) = \frac{n!}{s^{n+1}}$
3. $f(t) = t^p, p > -1 \implies F(t) = \frac{\Gamma(p+1)}{t^{p+1}}$
4. $f(t) = \sin(at) \implies F(t) = \frac{a}{t^2+a^2}$
5. $f(t) = \cos(at) \implies F(t) = \frac{t}{t^2+a^2}$
6. $f(t) = t\sin(at) \implies F(t) = \frac{2at}{(t^2+a^2)^2}$
7. $f(t) = t\cos(at) \implies F(t) = \frac{t^2-a^2}{(s^2+a^2)^2}$
8. $f(t) = \sin(at+b) \implies F(t) = \frac{t\sin(b)+a\cos(b)}{t^2+a^2}$
9. $f(t) = \cos(at+b) \implies F(t) = \frac{t\cos(b)-a\sin(b)}{t^2+a^2}$

6.1.1   Laplace Transforms of Derivatives

Now that we know how to calculate Laplace transforms, we must extend this to derivatives of functions in order to integrate it within differential equations.

Lemma 6.6

Let f be a differentiable function in the class Λ whose derivative also belongs to the class Λ, and let the Laplace transform of f be F. Then,

$$\mathcal{L}\big(f'(t)\big) = sF(s) - f(0)$$

Proof. We use the definition of Laplace transform and integrate by parts to get

$$\mathcal{L}\big(f'(t)\big) = \int_0^\infty e^{-st} f'(t)\, dt = \lim_{A\to\infty} e^{-st} f'(t)\, dt$$

$$= \lim_{A\to\infty} \left(e^{-st} f(t) \big|_0^A + \int_0^A s e^{-st} f(t)\, dt \right)$$

$$= -f(0) + s \int_0^\infty e^{-st} f(t)\, dt = sF(s) - f(0)$$

where we used the fact $(f \in \Lambda \implies f$ is of exponential growth at infinity) that

$$\lim_{A\to\infty} e^{-sA} f(A) = 0$$

for sufficiently large $\mathrm{Re}(s)$.

Definition 6.4 (The Class Λ^k)

For each positive integer k, define Λ^k to be the class of $C^k((0,\infty))$ (defined over $(0,\infty)$) functions in Λ whose derivatives up to order k also belong to Λ. That is,

$$\Lambda^k = \{f \in C^k((0,\infty)) \mid f, f', f'', \ldots, f^{(k)} \in \Lambda\}$$

Theorem 6.7 (Laplace Transform of Derivatives)

If $f \in \Lambda^k$ for some positive integer k, and if F is the Laplace transform of f, then for any $1 \le j \le k$,

$$\mathcal{L}\big(f^{(j)}(t)\big) = s^j F(s) - \sum_{i=1}^{j} s^{j-i} f^{(i-1)}(0)$$

$$= s^j F(s) - s^{j-1} f(0) - s^{j-2} f'(0) - \ldots - s f^{(j-2)}(0) - f^{(j-1)}(0)$$

Proof. Using the lemma, proof by induction on j for any fixed k.

Note that in order to solve a differential equation using Laplace transforms, we must know the initial values $\varphi(0), \varphi'(0), \ldots$ at $t = 0$ (must be at 0!) for us to simplify the algebraic equation that is produced. This is shown in the examples.

Example 6.3

Given the first-order differential equation with initial conditions

$$y' + ay = 0, \quad \varphi_0(0) = y_0$$

where a, y_0 are given constants. Assuming that the solution is in the class Λ^1, we try to find this using Laplace transforms. Let $Y_0(s) = \mathcal{L}(\varphi_0)$. Then,

$$\mathcal{L}\big(\varphi_0'(t) + a\varphi_0(t)\big) = \mathcal{L}\big(\varphi_0'(t)\big) + a\mathcal{L}\big(\varphi_0(t)\big)$$

$$= sY_0(s) - \varphi_0(0) + aY_0(s) = 0$$

So,

$$(s+a)Y_0(s) = y_0 \implies Y_0(s) = \frac{y_0}{s+a}$$

Now the only problem remaining is to find a function whose Laplace transform is this expression. By direct verification, we can see that $\varphi_0(t) = y_0 e^{-at}$ is indeed the solution.

Example 6.4

The Laplace transform of the nonhomogeneous first-order linear DEQ

$$y' + ay = f(t), \quad \varphi(0) = y_0$$

is (where $Y(s) = \mathcal{L}(\varphi)$)

$$sY(s) - \varphi(0) + aY(s) = F(s) \implies Y(s) = \frac{y_0}{s+a} + \frac{F(s)}{s+a}$$

Therefore, we must find a function which has this expression as its Laplace transform. By the previous example, we know that $t_0 e^{-at}$ satisfies the first term, but we have no method as yet of finding a function whose Laplace transform is

$$\frac{F(s)}{s+a}$$

This suggests that we will need a method of finding a function whose Laplace transform is the product of two given functions $F(s)$ and $1/(s+a)$.

This final example shows what would happen if we tried to take the Laplace transform of linear DEQ with nonconstant coefficients.

Example 6.5

Taking the Laplace transform of

$$y' + 2ty = 0, \quad \psi(0) = y_0$$

gives (where $\mathcal{L}(\psi) = Z(s)$)

$$sZ(s) - \psi(0) - 2Z'(s) = 0$$

While in the previous problems the equation was greatly simplified into an algebraic problem, here we have a differential equation which is no simpler than the original problem.

Clearly, this example suggests that the usefulness of the Laplace transform method is limited mainly to equations with constant coefficients.
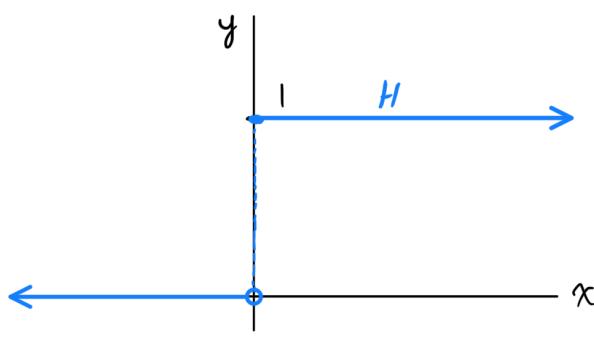
## 6.2   Heaviside Step, Dirac Delta Functions

Definition 6.5 (Heaviside Step Function)

A *Heaviside (step) function*, denoted as H, $\mathbb{I}$, or u, is defined

$$H(x) \equiv \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Clearly, it can be horizontally translated c units to result in the function $H(t-c) = u_c$ which is 1 if $t \geq c$ and 0 otherwise.

Looking back a Heaviside step functions, we can combine them to create more complicated models. For example, we can redefine the piecewise function

$$f(t) = \begin{cases} -4 & t < 6 \\ 25 & 6 \leq t < 8 \\ 16 & 8 \leq t < 30 \\ 10- & t \geq 30 \end{cases}$$

in terms of Heaviside functions as such

$$f(t) = -4 + 29u_6(t) - 9u_8(t) - 6u_{30}(t)$$

This is analogous to a "switch" that we can turn on or off. Furthermore, we can use Heaviside functions to "shift" functions horizontally a certain length c while turning them "off" for values of $t < c$ and "on" for values $t \geq c$. This can be written as

$$g(t) \equiv u_c(t) f(t - c)$$

To find the Laplace transform of $g(t)$, we can evaluate it as such.

$$\mathcal{L}\{u_c(t) f(t - c)\} = \int_0^\infty e^{-st} u_c(t) f(t - c) \, dt$$

$$= \int_c^\infty e^{-st} f(t - c) \, dt$$

Substituting $u = t - c$ gives

$$\mathcal{L}\{u_c(t) f(t - c)\} = \int_0^\infty e^{-s(u+c)} f(u) \, du$$

$$= e^{-cs} \int_0^\infty e^{-su} f(u) \, du$$

$$= e^{-cs} F(s)$$

This leads to the following formulas.

Theorem 6.8

The Laplace transforms for Heaviside functions are:

$$\mathcal{L}\{u_c(t) f(t - c)\} = e^{-ct} F(t)$$

$$\mathcal{L}\{u_c(t)\} = \frac{e^{-ct}}{t}$$
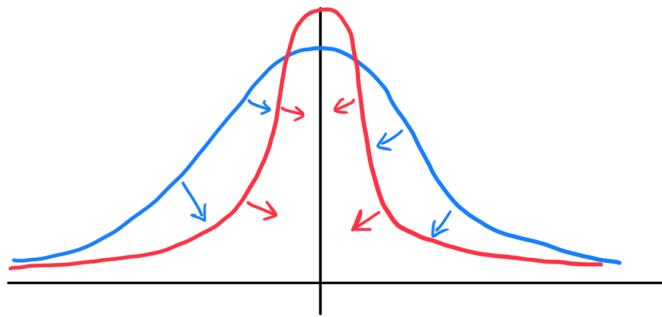
Definition 6.6 (Dirac Delta Function)

The *Dirac delta function* is a generalized function or distribution that models forces exerted over small time frames. It can be thought of as a function δ satisfying the properties

$$1. \delta(t-a) = 0, \quad t \neq a$$

$$2. \int_{a-\varepsilon}^{a+\varepsilon} \delta(t-a)\, dt, \quad \varepsilon > 0$$

$$3. \int_{a-\varepsilon}^{a+\varepsilon} f(t)\delta(t-a)\, dt = f(a), \quad \varepsilon > 0$$

Alternatively, it may be evaluated as the limit of a Gaussian distribution centered at a where $\sigma \to 0$. That is,

$$\delta(t) = \lim_{\sigma \to 0} \mathrm{Normal}(a, \sigma^2)$$

To apply the Dirac delta function in Laplace transforms, we calculate

$$\mathcal{L}\{\delta(t-a)\} = \int_0^\infty e^{-st}\delta(t-a)\, dt = e^{-as}$$

given that $a > 0$.

Theorem 6.9

More generally, we have

$$\mathcal{L}\{u_c(t)f(t-c)\} = e^{-cs}F(s)$$

where F is the Laplace transform of f.

Lemma 6.10

The derivative of the Heaviside function $u_a(t)$ is the Dirac delta function centered at a. That is,

$$u_a'(t) = \delta(t-a)$$

Proof. With a bit of background in probability, the cumulative distribution function (CDF) of the Dirac delta function is defined

$$\int_{-\infty}^{t} \delta(u-a)\, du = \begin{cases} 0 & t < a \\ 1 & t \geq a \end{cases}$$

But this is precisely the definition of the Heaviside function. By the fundamental theorem of calculus,

---

we have

$$u_a'(t) = \frac{d}{dt} \int_{-\infty}^{t} \delta(u - a) \, du = \delta(t - a)$$

6.3   Inverse Laplace Transform

Definition 6.7 (Inverse Laplace Transform)

The *Inverse Laplace transform* is the inverse of a Laplace transform. That is,

$$\mathcal{L}\big(f(t)\big) = F(s) \implies \mathcal{L}^{-1}\big(F(s)\big) = f(t)$$

Now, as we have mentioned in the previous examples, we would like to derive a certain method to find the inverse transform of a function.

Theorem 6.11 (Inverse Laplace Transforms of Products of Functions)

We assume that we are given two functions $f(t), g(t)$, with their Laplace transforms $\{ = F(s), \} = G(s)$. Then,

$$\mathcal{L}\left(\int_0^t f(t - v)g(v) \, dv \right) = F(s)G(s)$$

That is, given the function $F(s)G(s)$ (with their respective inverse transforms known), its inverse Laplace transform is

$$\mathcal{L}^{-1}\big(F(s)G(s)\big) = \int_0^t f(t - v)g(v) \, dv$$

given by the convolution integral of f and g.

Proof. We let $H(s) = F(s)G(s)$ and try to find the function $h(t)$ where $\mathcal{L}(h) = H(s)$. If there is such as function h, then

$$\mathcal{L}\big(h(t)\big) = \int_0^{\infty} e^{-st} h(t) \, dt$$

$$= F(s)G(s) = \int_0^{\infty} e^{-su} f(u) \, du \int_0^{\infty} e^{-sv} g(v) \, dv$$

where $\mathrm{Re}(s) > \sigma = \max(\alpha, \beta)$, where $F(s) = \mathcal{L}(f)$ for $\mathrm{Re}(s) > \alpha$ and $G(s) = \mathcal{L}(g)$ for $\mathrm{Re}(s) > \beta$ $(\alpha, \beta \in \mathbb{R})$. Since each integral converges absolutely for $\mathrm{Re}(s) > \sigma$, we can write the product of the two integrals as the iterated integral (using Fubini's theorem)

$$\int_0^{\infty} e^{-st} h(t) \, dt = \int_0^{\infty} \int_0^{\infty} e^{-s(u+v)} f(u)g(v) \, du \, dv$$

$$= \int_0^{\infty} g(v) \left(\int_0^{\infty} e^{-s(u+v)} f(u) \, du \right) dv$$

$$= \int_0^{\infty} g(v) \left(\int_0^{\infty} e^{-st} f(t - v) \, dt \right) dv$$

$$= \int_0^{\infty} e^{-st} \left(\int_0^t f(t - v)g(v) \, dv \right) dt$$

where the second to last step was done by making the change of variable $u + v = t$ for $\mathrm{Re}(s) > \sigma$.

Therefore,

$$F(s)G(s) = \int_0^\infty e^{-st} h(t)\, dt = \int_0^\infty e^{-st} \left(\int_0^t f(t-v)g(v)\, dv \right) dt$$

$$\implies h(t) = \int_0^t f(t-v)g(v)\, dv$$

Theorem 6.12

$\mathcal{L}$ is a linear operator.
$$\mathcal{L}^{-1}\{aF + bG\} = a\mathcal{L}^{-1}\{F\} + b\mathcal{L}^{-1}\{G\}$$

Corollary 6.13

The only function in the class Λ whose Laplace transform is identically 0 is the zero function.

Example 6.6

We compute
$$\mathcal{L}^{-1}\left(\frac{1}{s^2 - 1} \right) = \mathcal{L}^{-1}\left(\frac{1}{(s+1)(s-1)} \right)$$

Since
$$\mathcal{L}^{-1}\left(\frac{1}{s-1} \right) = e^t, \ \mathcal{L}^{-1}\left(\frac{1}{s+1} \right) = e^{-t}$$

we have
$$\mathcal{L}^{-1}\left(\frac{1}{(s+1)(s-1)} \right) = \int_0^t e^{t-u} e^{-u}\, du = e^t \int_0^t e^{-2u}\, du = \frac{1}{2}\left(e^t - e^{-t} \right)$$

We can also compute it using partial fraction decomposition and by using the linearity of $\mathcal{L}^{-1}$.

$$\mathcal{L}^{-1}\left(\frac{1}{s^2 - 1} \right) = \frac{1}{2}\left(\mathcal{L}^{-1}\left(\frac{1}{s-1} \right) - \mathcal{L}^{-1}\left(\frac{1}{s+1} \right) \right)$$
$$= \frac{1}{2}\left(e^t - e^{-t} \right)$$

Example 6.7

We compute
$$\mathcal{L}^{-1}\left(\frac{1}{s^2(s^2 + 1)} \right)$$

we know that
$$\mathcal{L}^{-1}\left(\frac{1}{s^2} \right) = t, \ \ \mathcal{L}^{-1}\left(\frac{1}{s^2 + 1} \right) = \sin t$$

So,
$$\mathcal{L}^{-1}\left(\frac{1}{s^2(s^2 + 1)} \right) = \int_0^t (t-u)\sin u\, du = t\int_0^t \sin u\, du - \int_0^t u\sin u\, du = t - \sin t$$

Using partial fractions, we have

$$\mathcal{L}^{-1}\left(\frac{1}{s^2(s^2+1)}\right) = \mathcal{L}^{-1}\left(\frac{1}{s^2}\right) - \mathcal{L}^{-1}\left(\frac{1}{s^2+1}\right) = t - \sin t$$

Example 6.8

Here is an example where we use only partial fractions. Since

$$G(s) = \frac{86s - 78}{(s+3)(s-4)(5s-1)} = -\frac{3}{s+3} + \frac{2}{s-4} + \frac{5}{5s-1}$$

we have

$$\mathcal{L}^{-1}\{G(s)\} = \mathcal{L}^{-1}\{-\frac{3}{s+3} + \frac{2}{s-4} + \frac{5}{5s-1}\}$$
$$= -3e^{-3t} + 2e^{4t} + e^{t/5}$$

Therefore, for rational functions of the form

$$\frac{N(s)}{D(s)}$$

where N, D are polynomials with the degree of N less than the degree of D, we can take advantage of partial fractions to compute its inverse Laplace transform.

Example 6.9

Solve the following differential equation.

$$2y'' + 3y' - 2y = te^{-2t}, \quad y(0) = 0, y'(0) = -2$$

We take the Laplace transforms of all the terms in the differential equation to get

$$2\big(t^2Y(t) - ty(0) - y'(0)\big) + 3\big(tY(t) - y(0)\big) - 2Y(t) = \frac{1}{(t+2)^2}$$

which, after simplification, gives

$$(2t^2 + 3t - 2)Y(t) + 4 = \frac{1}{(t+2)^2}$$

Solving for Y, we get

$$Y(t) = \frac{1}{(2t-1)(t+2)^3} - \frac{4}{(2t-1)(t+2)}$$
$$= \frac{1}{125}\left(\frac{-96}{s-\frac{1}{2}} + \frac{96}{s+2} - \frac{10}{(s+2)^2} - \frac{25}{(s+2)^3}\right)$$

The inverse transform of this gives

$$y(t) = \frac{1}{125}\left(-96e^{\frac{t}{2}} + 96e^{-2t} - 10te^{-2t} - \frac{25}{2}t^2e^{-2t}\right)$$

We now deal with Heaviside functions in this next example.

Example 6.10

Solve the following initial value problem.

$$2y'' + 10y = 3u_{12}(t) - 5\delta(t-4), \quad y(0) = -1, y'(0) = -2$$

We take the Laplace transform

$$2\big(t^2Y(t) - ty(0) - y'(0)\big) + 10Y(t) = \frac{3e^{-12t}}{t} - 5e^{-4t}$$

$$\implies (2t^2 + 10)Y(t) + 2t + 4 = \frac{3e^{-12t}}{t} - 5e^{-4t}$$

and solve for Y

$$Y(t) = 3e^{-12t}\left(\frac{1}{t(2t^2 + 10)}\right) - 5e^{-4t}\left(\frac{1}{2t^2 + 10}\right) - \frac{2t+4}{2t^2+10}$$

$$= 3e^{-12t}\left(\frac{1}{10t} - \frac{1}{10(t^2+5)}\right) - 5e^{-4t}\left(\frac{1}{2t^2+10}\right) - \frac{2t+4}{2t^2+10}$$

$$= 3e^{-12t}F(t) - 5e^{-4t}G(t) - H(t)$$

Using the rules for inverse transforming Dirac delta functions, we can get

$$y(t) = 3u_{12}(t)f(t-12) - 5u_4(t)g(t-4) - h(t)$$

where

$$f(t) = \frac{1}{10} - \frac{1}{10}\cos\left(\sqrt{5}t\right)$$

$$g(t) = \frac{1}{2\sqrt{5}}\sin\left(\sqrt{5}t\right.$$

$$h(t) = \cos\left(\sqrt{5}t\right) + \frac{2}{\sqrt{5}}\sin\left(\sqrt{5}t\right.$$

In general, the best tool for systematically solving inverse Laplace transforms is a table, experience, and luck.

Finally, we conclude by saying that not every type of function is necessarily the Laplace transform of some other function. For example, this theorem:

Theorem 6.14

If f belongs to the class Λ and if $F(s)$ is its Laplace transform, then

$$\lim_{\text{Re}(s)\to\infty} F(s) = 0$$

Furthermore, not only does $F(s) \to 0$ as $s \to \infty$, but in fact $|sF(s)|$ remains bounded as $\text{Re}(s) \to \infty$. Clearly, not all functions have this property.

Proof. Since $f \in \Lambda$, it is of exponential growth at infinity, meaning $|f(t)| \leq Me^{ct}$ for some $M > 0, c$

and sufficiently large t. This means that

$$|F(s)| \leq M \int_0^\infty |e^{-st}| e^{ct} \, dt$$

$$= M \int_0^\infty e^{-\operatorname{Re}(s)t} e^{ct} \, dt = \frac{M}{\operatorname{Re}(s) - c}, \quad (\operatorname{Re}(s) > c)$$

and it is clear that this limit approaches 0 as $\operatorname{Re}(s) \to \infty$.

In actuality, it is often impossible to find the inverse transform for a general function explicitly.

7 Numerical Methods of Solving DEQs

In many problems the only effective method for obtaining information about the solution of a differential equation is to use a numerical approximation procedure. In this section, we will talk about

1. the Euler Method and modified Euler Method

2. the Milne Method

3. Runge-Kutta Methods

For simplicity, we constrain our view to single first-order differential equations.

## 7.1 The Euler Method and Modified Euler Method
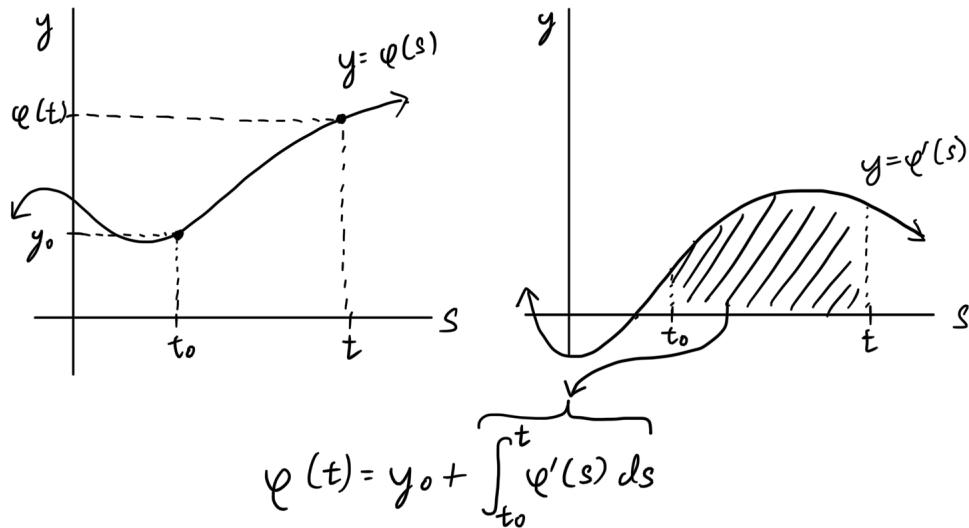
Given the first-order differential equation

$$y' = f(t, y) \text{ with initial conditions } \varphi(t_0) = y_0$$

we wish to find an approximation to the value of the solution φ at $t = t_0 + T$. By the fundamental rule of calculus, we know that assuming that $f(t, y)$ is C^1 for $t \in (t_0, t_0 + T)$,
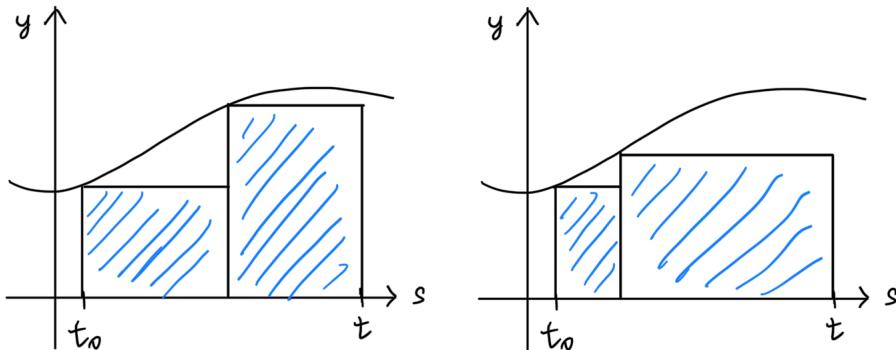
$$\varphi(t) = \varphi(t_0) + \int_{t_0}^t \varphi'(s) \, ds$$

$$= y_0 + \int_{t_0}^t f(s, \varphi(s)) \, ds$$

In visual terms, we just take y_0 and add the area under the derivative curve $\varphi'(s)$ from t_0 to whatever t we would like.

$$\varphi(t) = y_0 + \int_{t_0}^{t} \varphi'(s)\, ds$$
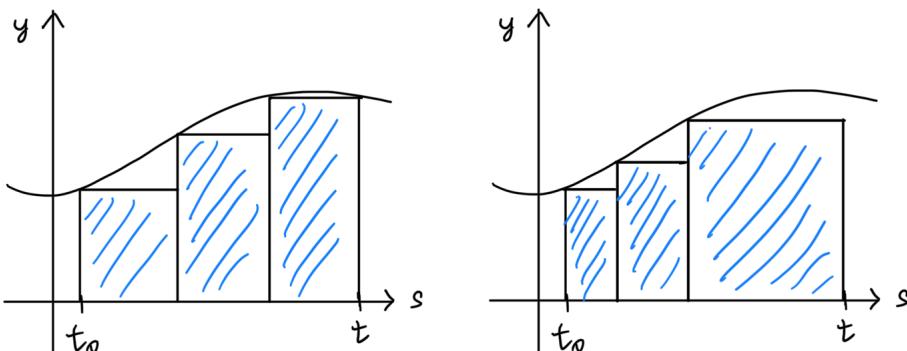
Since we know y_0, finding $f(t_0 + T)$ rests entirely on solving the integral representing the area under the velocity curve. Euler's method is simply doing this with left-hand Riemann sums. Note that while it is conventional that a the rectangles have the same width, they do not necessarily need to be.
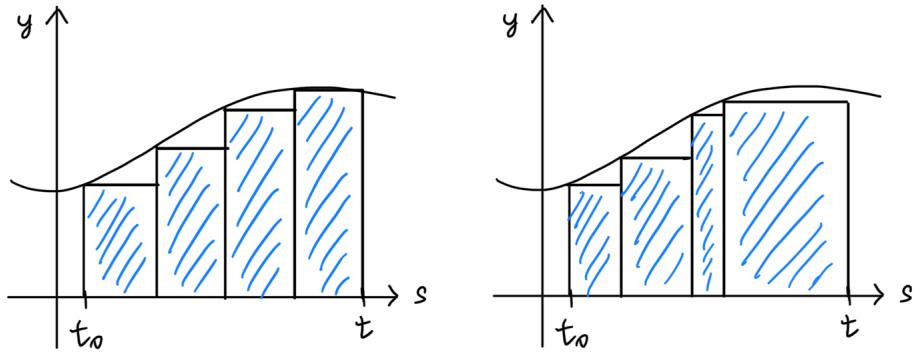
1. 2 Riemann rectangles (even and uneven).

2. 3 Riemann rectangles (even and uneven).

3. 4 Riemann rectangles (even and uneven).

In fact, the question of the best choice of unequal spacing is a major unsolved mathematical problem. We formalize this algorithm in the following steps.

Theorem 7.1 (Euler's Method)

Given the first order DEQ $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$, say that we would like to approximate the value of $\varphi(t_0 + T)$. Then, the *Euler method* gives us steps:
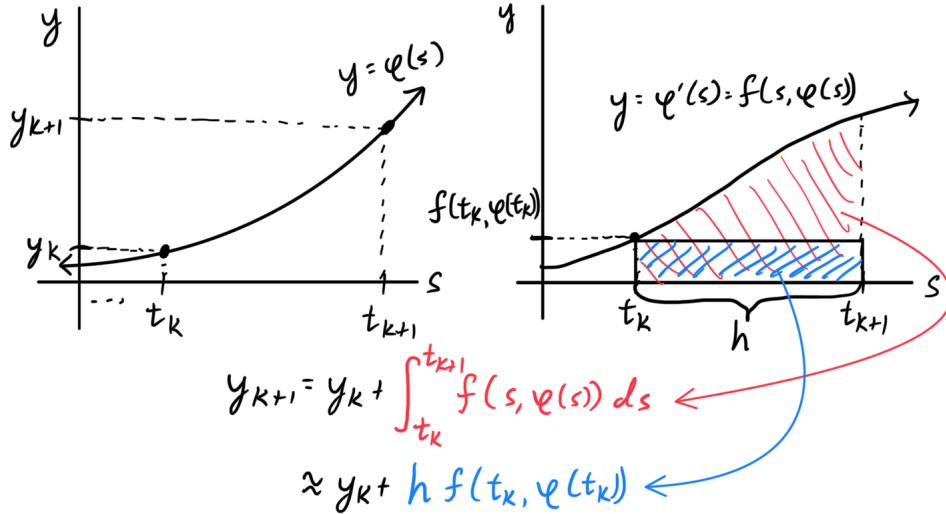
1. Divide the interval $[t_0, t_0 + T]$ into N (not necessarily equal) subintervals by specifying intermediate points

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T$$

   For simplicity of calculations later, we will assume equal spacing, with spacing $h = T/N$.

2. Looking at points t_k and t_{k+1} for $k = 1, \ldots, N-1$, we can approximate the value of y_{k+1} with y_k using Riemann rectangles.

$$y_{k+1} = y_k + \int_{t_k}^{t_{k+1}} f\big(s, \varphi(s)\big)\, ds \implies y_{k+1} = y_k + (t_{k+1} - t_k) f(t_k, y_k)$$

$$\implies y_{k+1} = y_k + h f(t_k, y_k)$$

3. Use step 2 to find y_1, then y_2, and so on until $y_N \approx \varphi(t_N) = \varphi(t_0 + T)$ is found. Then terminate.

Furthermore, for some constant M, the local truncation error of the Euler method (for each step) is

$$|T_k| \leq \frac{1}{2}Mh^2$$
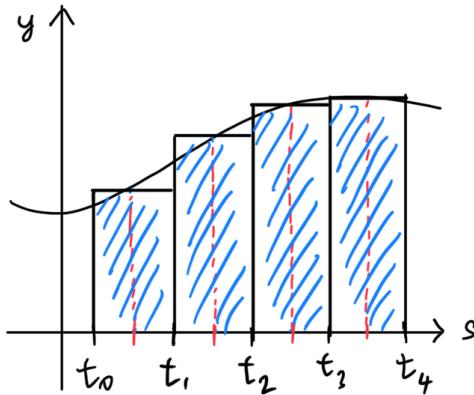
that is, no greater than a constant multiple of h^2, while the cumulative truncation error of it is

$$|\varphi(t_n) - y_n| \leq \frac{1}{2}Mh^2N = \frac{1}{2}MTh$$

that is, no greater than a constant multiple of h.

Since the cumulative error bound is bounded by a linear function of h, we can make the error as small as we want by making the h arbitrarily small. However, the truncation error for Euler's method is too large and is not used in applications. Rather, it is more efficient to use a more sophisticated method of approximation where the cumulative truncation error is no greater than a constant multiplied by some higher power of h.

An obvious improvement is to change the method of approximation of integrals from left-hand Riemann sums (taken at the start point of each interval) to Riemann rectangles taken at the midpoint of each interval. This method of approximation is called *midpoint quadrature*


Theorem 7.2 (Modified Euler's Method)
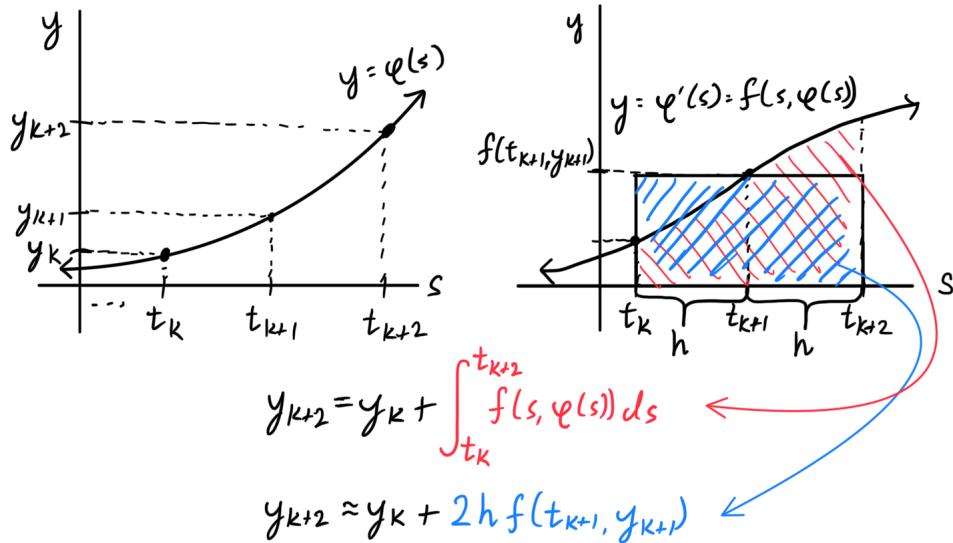
Given the first order DEQ $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$, say that we would like to approximate the value of $\varphi(t_0 + T)$. Then, the *modified Euler's method* gives us steps:
1. Divide the interval $[t_0, t_0+T]$ into N equal subintervals (of length h), where N is an even number, by specifying intermediate points

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T$$

2. Rather than integrating over one subinterval $[t_k, t_{k+1}]$, we integrate over two subintervals $[t_k, t_{k+2}]$ and approximate it using the value of the function at t_{k+1}, the midpoint of the interval. That is,

$$y_{k+2} = y_k + \int_{t_k}^{t_{k+2}} f\big(s, \varphi(s)\big)\, ds \implies y_{k+2} = y_k + (t_{k+2} - t_k)f(t_{k+1}, y_{k+1})$$

$$\implies y_{k+2} = y_k + 2hf(t_{k+1}, y_{k+1})$$

$$y_{k+2} = y_k + \int_{t_k}^{t_{k+2}} f(s, \varphi(s))\, ds$$

$$y_{k+2} \approx y_k + 2h f(t_{k+1}, y_{k+1})$$

Note that the modified Euler's method expresses y_{k+2} in terms of both y_k and y_{k+1}. This is a problem, since in the first step we know only y_0 but not y_1 in order to calculate y_2. Therefore, we must solve for y_1 with some other method, such as Euler's method or approximating using Taylor series (provided the function is analytic).

3. Use step to to find y_2, then y_4, then y_6, and so on until $y_N \approx \varphi(t_N) = \varphi(t_0 + T)$ is found. Then terminate.

Furthermore, for some constant M, the local truncation error of the modified Euler method (for each step) is

$$|T_k| \leq \frac{M}{3} h^3$$

that is, no greater than a constant multiple of h^2, while the cumulative truncation error of it is no great than a constant multiple of h^2.

Example 7.1 (Comparison of Euler Algorithm and Modified Version on Approximating e)

By setting $h = 0.1$, we estimate e where $\varphi(1) = e$. Remember the differential equation with this solution is $y' = f(t, y) = y$. Using Euler's method we have

$$y_{k+1} = y_k + 0.1y_k = 1.1y_k$$

Doing this iteratively 10 times starting at $y_0 = 1$ gives us $1.1^{10} = 2.593$. However, using the modified Euler's method, we have

$$y_{k+2} = y_k + 0.2f(t_{k+1}, y_{k+1})$$
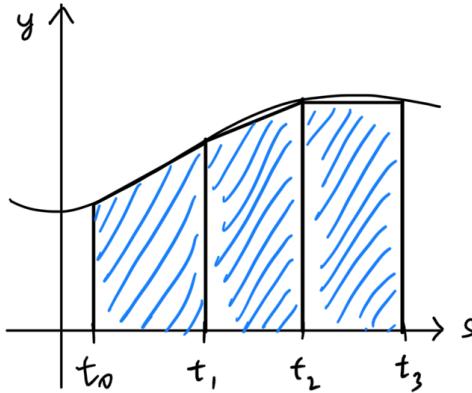
We can first approximate $y_1 = \varphi(0.1) = 1 + 0.1 + \frac{1}{2}(0.1)^2 + \ldots$ using a power series expansion, giving $y_1 \approx 1.105$ (correct to 3 decimal places). Since $f(t, y) = y$, we get

$$y_{k+2} = y_k + 0.2y_{k+1}$$

Calculating these values $y_2, y_3, \ldots$, we get $e \approx 2.713$, which is considerably better than the 2.593 approximation.

Another way to approximate definite integrals is to use trapezoidal sums as a method of approximation.

$$\int_{t_k}^{t_{k+1}} f\big(s, \varphi(s)\big)\, ds \approx \frac{h}{2}\big(f(t_k, \varphi(t_k)) + f(t_{k+1}, \varphi(t_{k+1}))\big)$$

Theorem 7.3 (Improved Euler's Method)

Given the first order DEQ $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$, say that we would like to approximate the value of $\varphi(t_0 + T)$, then the *improved Euler method* gives us steps:

1. Divide the interval $[t_0, t_0 + T]$ into N (not necessarily equal) subintervals by specifying intermediate points
$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T$$
For simplicity of calculations later, we will assume equal spacing, with spacing $h = T/N$.

2. Looking at points t_k and t_{k+1} for $k = 1, \ldots, N - 1$, we can approximate the value of y_{k+1} with y_k using trapezoids.
$$y_{k+1} = y_k + \int_{t_k}^{t_{k+1}} f\big(s, \varphi(s)\big)\, ds \implies y_{k+1} = y_k + \frac{h}{2}\big(f(t_k, y_k) + f(t_{k+1}, y_{k+1})\big)$$
Note that the improved Euler method expresses y_{k+1} implicitly rather than explicitly. There are methods for dealing with implicit formulas, but we leave it at this.

3. Use step 2 to find y_2, then y_3, and so on until $y_N \sum \varphi(t_N)$ is found.

7.2 The Milne Method

The Milne method gives us a quadratic approximation to certain integrals.

Lemma 7.4 (Simpson's Rule)
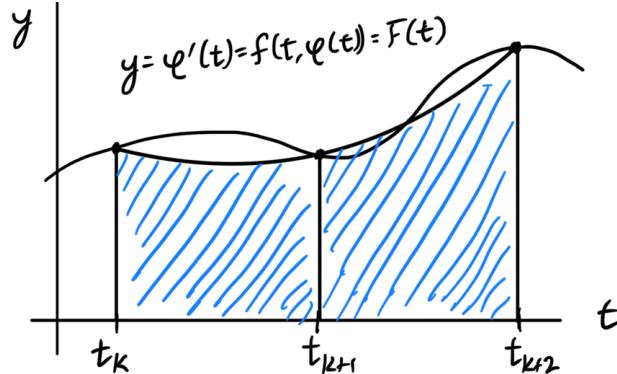
Simpson's rule gives the approximation

$$\int_{t_k}^{t_{k+2}} f\big(s, \varphi(s)\big)\, ds \approx \frac{h}{3}\big(f(t_k, \varphi(t_k)) + 4f(t_{k+1}, y_{k+1}) + f(t_{k+2}, y_{k+2})\big)$$

Proof. This approximation pops up from attempting to find a quadratic graph of best fit for arbitrary function $y = f(t, \varphi(t))$. For convenience, we assign $F(t) = f(t, \varphi(t))$. Now, we must determine constants a, b, c so that the parabola

$$y = a + b(s - t_{k+1}) + c(s - t_{k+1})^2$$

---

with vertex (t_{k+1}, a) passes through the points $(t_k, F(t_k)), (t_{k+1}, F(t_{k+1})), (t_{k+2}, F(t_{k+2}))$.

Since $h = t_{k+2} - t_{k+1} = t_{k+1} - t_k$, we have the set of conditions

$$F(t_k) = a - bh + ch^2$$
$$F(t_{k+1}) = a$$
$$F(t_{k+2}) = a + bh + ch^2$$

We can solve this system to get

$$a = F(t_{k+1}), \qquad c = \frac{F(t_{k+2}) - 2F(t_{k+1}) + F(t_k)}{2h^2}$$

We then integrate this quadratic approximation

$$\int_{t_k}^{t_{k+2}} \left(a + b(s - t_{k+1}) + c(s - t_{k+1})^2 \right) ds = \left(as - \frac{b}{2}(s - t_{k+1})^2 + \frac{c}{3}(s - t_{k+1})^3 \right)\Big|_{t_k}^{t_{k+2}}$$

$$= 2ah + \frac{2}{3}ch^3$$

Note that we do not even need to find the value of b. Substituting the solutions a, c in, we get Simpson's Rule.

Theorem 7.5 (Milne Method)

Given the first order DEQ $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$, say that we would like to approximate the value of $\varphi(t_0 + T)$. Then, the *Milne method* gives us steps:
1. Divide the interval $[t_0, t_0+T]$ into N (not necessarily equal) subintervals by specifying intermediate points

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T$$

For simplicity of calculations later, we will assume equal spacing, with spacing $h = T/N$.
2. Rather than integrating over one subinterval $[t_k, t_{k+1}]$, we integrate over two subintervals $[t_k, t_{k+2}]$ and approximate it using Milne's formula:

$$y_{k+2} = y_k + \int_{t_k}^{t_{k+2}} f\big(s, \varphi(s)\big)\, ds$$

$$\implies y_{k+2} = y_k + \frac{h}{3}\big(f(t_k, \varphi(t_k)) + 4f(t_{k+1}, y_{k+1}) + f(t_{k+2}, y_{k+2})\big)$$

Note that this solves for y_{k+2} implicitly (which isn't a problem when dealing with linear equations), and you need to know both y_k and y_{k+1} in order to calculate y_{k+2}.

3. Use step 2 to find $y_2, y_3, \ldots$ until $y_N \approx \varphi(t_N)$ is found. Then terminate.

Furthermore, the cumulative truncation error of the Milne method is no greater than a constant multiple of h^4.

We can now state a very important theorem on error bounds of one-step methods (applying to the Euler method, but not to the Milne method).

Theorem 7.6 (Truncation Error Bounds on One-Step Methods)

Suppose that $f(t, y) \in C^1$ for $t_0 \leq t \leq t_0 + T$ and all y. Suppose

$$y_1, y_2, \ldots, y_N$$

are the approximations calculated by some one-step method with step length h to the solution φ of

$$y' = f(t, y), \quad \varphi(t_0) = y_0$$

If the local truncation error is no greater than ϵ, then the cumulative truncation error is no greater than a constant multiple of ϵ/h.

This theorem shows that if the local truncation error of a one-step process is no greater than a constant multiple of h^p, then the cumulative truncation error is no grater than a constant multiple of h^{p-1}. Analogous results can be proved for two-steps methods (and for methods involving any finite number of steps). Therefore, the accuracy of approximation is improved if a method whose truncation error involves a higher power of h is used.

7.3 Stability, Consistency, and Convergence

Definition 7.1 (Round-Off Errors)

When we use a numerical method to obtain an approximation to the solution of a differential equation, we are trying to find a set of numbers

$$y_1, y_2, y_3, \ldots, y_N$$

defined by the method, where y_N is the final approximation. However, due to the floating-point nature of numbers in computers, we actually round-off all numbers calculated to a specific number of significant figures, obtaining a slightly different set of numbers

$$z_1, z_2, z_3, \ldots, z_N$$

the difference

$$r_k = |z_k - y_k|, \quad k = 1, 2, \ldots, N$$

is called the *round-off error*. The h value is also called the *mesh*.

Definition 7.2 (Finite-Difference Methods)

Finite-difference methods are a class of numerical techniques for solving differential equations by approximating derivatives with finite differences. Both the spatial domain and time interval are discretized (i.e. broken up into a finite number of steps), and the value of the solution at these discrete points is approximated by solving algebraic equations containing finite differences and values from nearby points.

Definition 7.3 (Consistency, Stability, Convergence)

We now define three common terms that describes numerical methods.
1. *Consistency*: A finite difference method is considered consistent if by reducing the mesh and time step size, the truncation error terms could be made to approach 0. In this case the solution to the difference equation would approach the true solution of the DE.
2. *Stability*: A finite difference approximation is stable if the errors (truncation, round-off, etc.) decay (i.e. remain bounded, preferably within some deducible function) as the computation proceeds from one marching step to the next.
3. *Convergence*: The solution to the finite difference approximation approaches the true solution of the DE when the mesh is refined (step size reduced). This means that the approximation y_N tend to the actual solution as $h \to 0$.

Every method mentioned in this section is convergent. Obviously, a method which is not convergent is useless for obtaining numerical approximations. It is possible to give conditions for stability and consistency of numerical methods which are easy to verify. If we do this, we can use the following theorem to prove convergence.

Theorem 7.7 (Lax Equivalence Theorem)

A finite difference approximation method satisfying consistency (meaning truncation error approaches 0 when step size and mesh size goes to 0) and stability (meaning error goes on diminishing as time step passes) is convergent.

It turns out that the Milne method may be numerically instable, which is why it is not always suitable, despite its small truncation error.

7.4   Runge-Kutta Methods

Let us revisit Taylor polynomials that serve as approximations to differentiable functions. Let φ be the solution of the differential equation

$$y' = f(t, y), \quad \varphi(t_0) = y_0$$

Assuming that φ has a continuous second derivative on the interval $[t_0, t_0 + T]$, we can use Taylor's theorem (with the Lagrange form of the remainder) to write the first order approximation centered at t_0.

$$\varphi(t) = \varphi(t_0) + (t - t_0)\varphi'(t_0) + \frac{(t - t_0)^2}{2!}\varphi''(\zeta)$$

for some $\zeta \in (t_0, t)$. If $t_1 = t_0 + h$, then

$$\varphi(t_1) = \varphi(t_0) + h\varphi'(t_0) + \frac{h^2}{2!}\varphi''(\zeta)$$
$$= \varphi(t_0) + hf(t_0, y_0) + \frac{h^2}{2!}\varphi''(\zeta)$$
$$\approx y_0 + hf(t_0, y_0)$$

where the approximation is gotten by neglecting the term $\frac{h^2}{2!}\varphi''(\zeta)$. If we divide the interval $[t_0, t_0 + T]$ into N subintervals of length h by defining the partition points $t_k = t_0 + kh$ $(k = 0, 1, \ldots, N)$, we can use this procedure to obtain an iterative formula

$$y_{k+1} = y_k + hf(t_k, y_k)$$

with local truncation error $\frac{h^2}{2!}\varphi''(\zeta)$. This first-order approximation is, of course, just Euler's algorithm.

If we look at the second degree Taylor expansion

$$\varphi(t) = \varphi(t_0) + (t - t_0)\varphi'(t_0) + \frac{(t - t_0)^2}{2!}\varphi''(t_0) + \frac{(t - t_0)^3}{3!}\varphi'''(\zeta)$$

we can ignore the error term but we still have to deal with the second order term. We can evaluate $\varphi''(t_0)$ by differentiating $\varphi'(t) = f(t, \varphi(t))$ using the chain rule, which gives

$$\varphi''(t) = f_t\big(t, \varphi(t)\big) + f_y\big(t, \varphi(t)\big)\varphi'(t)$$
$$= f_t\big(t, \varphi(t)\big) + f_y\big(t, \varphi(t)\big)f\big(t, \varphi(t)\big)$$

Where f_t and f_y represents the partial derivatives with respect to t and y, respectively. Note that when computing $f_t(t, \varphi(t))$, we should keep $\varphi(t)$ constant, and when computing $f_y(t, \varphi(t))$, we should keep t constant while deriving with respect to $y = \varphi(t)$. Do not get confused by this. This ultimately leads to the iterative approximation formula

$$y_{k+1} = y_k + hf(t_k, y_k) + \frac{h^2}{2!}\big(f_t(t_k, y_k) + f_y(t_k, y_k)f(t_k, y_k)\big)$$

with local truncation error $\frac{h^3}{3!}\varphi'''(\zeta)$. This is a plausible means of obtaining numerical approximations, but it suffers from the disadvantage of having to calculate derivatives of f, which isn't easy for a computer to do. This problem persists for procedures using higher-order Taylor approximations.

The Runge-Kutta method is an attempt to obtain formulas equivalent to Taylor approximations which do not involve derivatives of f. There are many variations of Runge-Kutta methods, but we will describe the most common version.

Theorem 7.8 (Runge-Kutta Method)

Given the first order DEQ $y' = f(t, y)$ with initial conditions $\varphi(t_0) = y_0$, say that we would like to approximate the value of $\varphi(t_0 + T)$. Then, the *Runge-Kutta method* gives us steps:
1. Divide the interval $[t_0, t_0+T]$ into N (not necessarily equal) subintervals by specifying intermediate points

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T$$

   For simplicity of calculations later, we will assume equal spacing, with spacing $h = T/N$.
2. Given the value y_k, we can approximate the value y_{k+1} using the one-step process

$$y_{k+1} = y_k + \frac{h}{6}(p_1 + 2p_2 + 2p_3 + p_4)$$

where

$$p_1 = f(t_k, y_k) \qquad\qquad p_2 = f\left(t_k + \frac{h}{2}, y_k + \frac{hp_1}{2}\right)$$

$$p_3 = f\left(t_k + \frac{h}{2}, y_k + \frac{hp_2}{2}\right) \qquad\qquad p_4 = f(t_{k+1}, y_k + hp_3)$$

The term $\frac{1}{6}(p_1 + 2p_2 + 2p_3 + p_4)$ represents an "average" slope of φ over the interval $[t_k, t_{k+1}]$. More specifically,
1. p_1 is the slope at t_1
2. p_2 is an approximation of the slops at the midpoint of the interval obtained by means of the Euler method
3. p_3 is a second approximation to the slope at the midpoint
4. p_4 is an approximation to the slope at t_{k+1} obtained by means of the Euler method with slope p_3.

Even though it requires a lot of calculations, it is an explicit one-step procedure that does not require calculation of derivative of f. Furthermore, it is equivalent to a 4th order Taylor formula, and the local truncation error of the Runge-Kutta method is no greater than a constant multiple of h^5.

Example 7.2 (Deriving a Runge-Kutta Formula equivalent to 2nd-Order Taylor Formula)

Earlier in this subsection, we have found that the second order approximation formula for φ is

$$y_{k+1} = y_k + hf(t_k, y_k) + \frac{h^2}{2!}\big(f_t(t_k, y_k) + f_y(t_k, y_k)f(t_k, y_k)\big)$$

which requires us to compute derivatives. We will demonstrate how to obtain a Runge-Kutta formula equivalent to this 2nd-order Taylor formula. We will assume a method of form

$$y_{k+1} = y_k + h\Big(af(t_k, y_k) + bf\big(t_k + \alpha h, y_k + \beta hf(t_k, t_k)\big)\Big)$$

where a, b, α, β are constants to be determined. By Taylor's theorem for functions of two variables we write

$$f\big(t_k + \alpha h, y_k + \beta hf(t_k, y_k)\big) = f(t_k, y_k) + \alpha hf_t(t_k, y_k) + \beta hf(t_k, y_k)f_y(t_k, y_k) + Rh^2$$

where R is a remainder term involving the second-order partial derivatives of f. Substituting this into the original form gives

$$y_{k+1} = y_k + h(a+b)f(t_k, y_k) + h^2\big(\alpha bf_t(t_k, y_k) + \beta bf(t_k, y_k)f_y(t_k, y_k)\big) + Rbh^3$$

Comparing this with the original approximation formula for φ (with derivatives), we can compare each term and see that if

$$a + b = 1, \quad , \alpha b = \frac{1}{2}, \quad \beta b = \frac{1}{2}$$

then the approximations obtained by these two formulas differ only by the term Rbh^3. Thus, any quadruples of constants satisfying these conditions is a Runge-Kutta formula of the desired type. For example, choosing $a = \frac{1}{2}, b = \frac{1}{2}, \alpha = 1, \beta = 1$, gives

$$y_{k+1} = y_k + \frac{h}{2}\big(f(t_k, y_k) + f(t_{k+1}, y_k + hf(t_k, y_k))\big)$$

which is equivalent in accuracy to a three-term Taylor formula. Also note that this formula is similar to the improved Euler method.

7.5   Numerical Methods for Systems and Equations of Higher Order

Everything in this section had been designed for first-order equations, but these methods are equally suitable for systems of differential equations. We can write a system of equations as a vector equation

$$y' = f(t, y) \iff \begin{cases} y_1' & = f_1(t, y_1, y_2, \ldots, y_n) \\ y_2' & = f_2(t, y_1, y_2, \ldots, y_n) \\ \vdots & = \vdots \\ y_n' & = f_n(t, y_1, y_2, \ldots, y_n) \end{cases}$$

where $y \in \mathbb{R}^n$ and $f : \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}^n$. We can apply the approximation methods developed in this chapter to the system by applying them to each component in the vector equation.

Example 7.3

Consider the system

$$u' = v, \quad v' = g(t, u, v)$$

This can be written in form

$$y' = f(t, y) \iff \begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} v \\ g(t, u, v) \end{pmatrix}$$

The Euler method applied to this system leads to a pair of iterative formulas

$$\begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} u_k \\ v_k \end{pmatrix} + h \begin{pmatrix} v_k \\ g(t_k, u_k, v_k) \end{pmatrix} \implies \begin{cases} u_{k+1} & = u_k + h v_k \\ v_{k+1} & = v_k + h g(t_k, u_k, v_k) \end{cases}$$

Therefore, once we are given initial values u_0 and v_0, and once we have found both u_k and v_k, we can use this to compute u_{k+1} and v_{k+1}.

As we have seen before, higher order equations can be treated as a system of first-order DEQs, which can then be solved numerically on each component equation.