數學建模 MA3067-*

Chapter 1. Dimensional Analysis（量綱/因次分析）

Introduction

每個變量都有他的量綱（dimension，或譯為「因次」）。分析相關變量和他們的量綱之間的關係稱為「量綱分析」（dimensional analysis），是開始建構一個新模型時很有用的基本技術之一。量綱是一個物理量的一種基本內在性質，例如「質量」、「長度」、「時間」、「電荷量」、「溫度」、……是一些基礎的物理量綱，由這些量綱可以再結合成較複雜的量綱，例如「速度」的量綱是「長度／時間」，而速度的「計量單位」可以是「公尺／秒」、「英哩／小時」或其他各種單位。量綱分析依據的原理是：變量間的關係（方程式）必須符合「量綱齊次性」（變量間的關係不隨變量之計量單位改變而改變）。任何有意義的方程式其等號兩邊的量綱必須相同，檢查方程式有否遵循這個原則是量綱分析最基本的步驟。

Introduction

每個變量都有他的量綱（dimension，或譯為「因次」）。分析相關變量和他們的量綱之間的關係稱為「量綱分析」（dimensional analysis），是開始建構一個新模型時很有用的基本技術之一。量綱是一個物理量的一種基本內在性質，例如「質量」、「長度」、「時間」、「電荷量」、「溫度」、……是一些基礎的物理量綱，由這些量綱可以再結合成較複雜的量綱，例如「速度」的量綱是「長度／時間」，而速度的「計量單位」可以是「公尺／秒」、「英哩／小時」或其他各種單位。量綱分析依據的原理是：變量間的關係（方程式）必須符合「量綱齊次性」（變量間的關係不隨變量之計量單位改變而改變）。任何有意義的方程式其等號兩邊的量綱必須相同，檢查方程式有否遵循這個原則是量綱分析最基本的步驟。

Introduction

每個變量都有他的量綱（dimension，或譯為「因次」）。分析相關變量和他們的量綱之間的關係稱為「量綱分析」（dimensional analysis），是開始建構一個新模型時很有用的基本技術之一。量綱是一個物理量的一種基本內在性質，例如「質量」、「長度」、「時間」、「電荷量」、「溫度」、……是一些基礎的物理量綱，由這些量綱可以再結合成較複雜的量綱，例如「速度」的量綱是「長度／時間」，而速度的「計量單位」可以是「公尺／秒」、「英哩／小時」或其他各種單位。量綱分析依據的原理是：變量間的關係（方程式）必須符合「量綱齊次性」（變量間的關係不隨變量之計量單位改變而改變）。任何有意義的方程式其等號兩邊的量綱必須相同，檢查方程式有否遵循這個原則是量綱分析最基本的步驟。

Introduction

Remark: 要注意「量綱」和「單位」的不同，例如「質量」是一種「量綱」,「公斤」是度量「質量」的一種「單位」,「公克」也是另一種度量「質量」的「單位」⋯不同的「單位制度」可能對同一個量綱指定不同的單位來度量，例如「力」這個物理量的量綱是「質量 × 長度／時間^2」，他的指定度量單位在 MKS 制和 CGS 制中分別是「kg×m/s^2」和「g×cm/s^2」。

原則上，究竟哪些量綱是基本的量綱、用以表達其他物理量的量綱，這是可以有不同定義的。在一些物理模型中，動量、能量、電流都曾被選作基本量綱；有些模型不以溫度為基本量綱，因為溫度表達了粒子在每個自由度上的能量，可以以能量或質量、長度、時間的組合來表達；還有些物理學者不視電荷量為基本量綱，認為它可以以質量、長度、時間的組合來表達。

Introduction

Remark: 要注意「量綱」和「單位」的不同，例如「質量」是一種「量綱」,「公斤」是度量「質量」的一種「單位」,「公克」也是另一種度量「質量」的「單位」…不同的「單位制度」可能對同一個量綱指定不同的單位來度量，例如「力」這個物理量的量綱是「質量 × 長度／ 時間^2」，他的指定度量單位在 MKS 制和 CGS 制中分別是「$\text{kg}\times\text{m}/\text{s}^2$」和「$\text{g}\times\text{cm}/\text{s}^2$」。

原則上，究竟哪些量綱是基本的量綱、用以表達其他物理量的量綱，這是可以有不同定義的。在一些物理模型中，動量、能量、電流都曾被選作基本量綱；有些模型不以溫度為基本量綱，因為溫度表達了粒子在每個自由度上的能量，可以以能量或質量、長度、時間的組合來表達；還有些物理學者不視電荷量為基本量綱，認為它可以以質量、長度、時間的組合來表達。

Introduction

Remark: 要注意「量綱」和「單位」的不同，例如「質量」是一種「量綱」,「公斤」是度量「質量」的一種「單位」,「公克」也是另一種度量「質量」的「單位」…不同的「單位制度」可能對同一個量綱指定不同的單位來度量，例如「力」這個物理量的量綱是「質量 × 長度／ 時間^2」，他的指定度量單位在 MKS 制和 CGS 制中分別是「$\mathrm{kg}\times\mathrm{m}/\mathrm{s}^2$」和「$\mathrm{g}\times\mathrm{cm}/\mathrm{s}^2$」。

原則上，究竟哪些量綱是基本的量綱、用以表達其他物理量的量綱，這是可以有不同定義的。在一些物理模型中，動量、能量、電流都曾被選作基本量綱；有些模型不以溫度為基本量綱，因為溫度表達了粒子在每個自由度上的能量，可以以能量或質量、長度、時間的組合來表達；還有些物理學者不視電荷量為基本量綱，認為它可以以質量、長度、時間的組合來表達。

Introduction

For a given physical quantity q, we use $[q]$ to denote the dimension of q, and use L, M, T to denote the dimension of length, mass, and time, respectively. A quantity q which does not change after changing unit of every fundamental dimension is called dimensionless (無量綱／無因次) and is denoted by $[q] = 1$.

Example

Let F, v, a and p demote the force, the velocity, the acceleration and the pressure, respectively. Then

$$[F] = MLT^{-2}, \qquad [v] = LT^{-1},$$
$$[a] = LT^{-2}, \qquad [p] = ML^{-1}T^{-2}.$$

The quantity $\frac{F}{ma}$, where m is the mass, is dimensionless.

Introduction

For a given physical quantity q, we use $[q]$ to denote the dimension of q, and use L, M, T to denote the dimension of length, mass, and time, respectively. A quantity q which does not change after changing unit of every fundamental dimension is called dimensionless (無量綱／無因次) and is denoted by $[q] = 1$.

Example

Let F, v, a and p demote the force, the velocity, the acceleration and the pressure, respectively. Then

$$[F] = MLT^{-2}, \qquad [v] = LT^{-1},$$
$$[a] = LT^{-2}, \qquad [p] = ML^{-1}T^{-2}.$$

The quantity $\frac{F}{ma}$, where m is the mass, is dimensionless.

Introduction

For a given physical quantity q, we use $[q]$ to denote the dimension of q, and use L, M, T to denote the dimension of length, mass, and time, respectively. A quantity q which does not change after changing unit of every fundamental dimension is called dimensionless (無量綱／無因次) and is denoted by $[q] = 1$.

Example

Let F, v, a and p demote the force, the velocity, the acceleration and the pressure, respectively. Then

$$[F] = MLT^{-2}, \qquad [v] = LT^{-1},$$
$$[a] = LT^{-2}, \qquad [p] = ML^{-1}T^{-2}.$$

The quantity $\dfrac{F}{ma}$, where m is the mass, is dimensionless.

Introduction

For a given physical quantity q, we use $[q]$ to denote the dimension of q, and use L, M, T to denote the dimension of length, mass, and time, respectively. A quantity q which does not change after changing unit of every fundamental dimension is called dimensionless (無量綱／無因次) and is denoted by $[q] = 1$.

Example

Let F, v, a and p demote the force, the velocity, the acceleration and the pressure, respectively. Then

$$[F] = MLT^{-2}, \qquad [v] = LT^{-1},$$
$$[a] = LT^{-2}, \qquad [p] = ML^{-1}T^{-2}.$$

The quantity $\frac{F}{ma}$, where m is the mass, is dimensionless.

§1.1 Dimensional Methods

量綱分析的基石為 Buckingham's Pi 定理，它指出任一個表示多個具量綱物理量之間關係的物理定律對應著一個無量綱量之間關係的等效定律。

Question: What does it mean by a relation among several dimensioned physical quantities?

Example

The air resistance F a biker encounters appears to be related to the speed v and the cross-sectional area A, as well as the air density ρ. Therefore,

$$F = \phi(\rho, A, v)$$

or equivalently,

$$\Phi\left(F, \rho, A, v\right) = F - \phi(\rho, A, v) = 0\,.$$

§1.1 Dimensional Methods

量綱分析的基石為 Buckingham's Pi 定理，它指出任一個表示多個具量綱物理量之間關係的物理定律對應著一個無量綱量之間關係的等效定律。

Question: What does it mean by a relation among several dimensioned physical quantities?

Example

The air resistance F a biker encounters appears to be related to the speed v and the cross-sectional area A, as well as the air density ρ. Therefore,

$$F = \phi(\rho, A, v)$$

or equivalently,

$$\Phi\left(F, \rho, A, v\right) = F - \phi(\rho, A, v) = 0\,.$$

§1.1 Dimensional Methods

Example

Suppose that we want to compute the yield of the first atomic explosion after viewing photographs of the spread of the fireball. In such an explosion a large amount of energy E is released in a short time in a region small enough to be considered a point. From the point of the explosion a strong shock wave spreads outwards; the pressure behind the shock is on the order of hundreds of thousands of atmospheres, far greater than the ambient air pressure whose magnitude can be accordingly neglected in the early stages of the explosion. It is plausible that there is a relation between the radius of the blast wave front r, time t, the initial air density ρ, and the energy released E. Hence, we assume there is a physical law

$$\phi(t, r, \rho, E) = 0$$

which provides a relationship among these quantities.

§1.1 Dimensional Methods

Example

Suppose that we want to compute the yield of the first atomic explosion after viewing photographs of the spread of the fireball. In such an explosion a large amount of energy E is released in a short time in a region small enough to be considered a point. From the point of the explosion a strong shock wave spreads outwards; the pressure behind the shock is on the order of hundreds of thousands of atmospheres, far greater than the ambient air pressure whose magnitude can be accordingly neglected in the early stages of the explosion. It is plausible that there is a relation between the radius of the blast wave front r, time t, the initial air density ρ, and the energy released E. Hence, we assume there is a physical law

$$\phi(t, r, \rho, E) = 0$$

which provides a relationship among these quantities.

§1.1 Dimensional Methods

Example

Suppose that we want to compute the yield of the first atomic explosion after viewing photographs of the spread of the fireball. In such an explosion a large amount of energy E is released in a short time in a region small enough to be considered a point. From the point of the explosion a strong shock wave spreads outwards; the pressure behind the shock is on the order of hundreds of thousands of atmospheres, far greater than the ambient air pressure whose magnitude can be accordingly neglected in the early stages of the explosion. It is plausible that there is a relation between the radius of the blast wave front r, time t, the initial air density ρ, and the energy released E. Hence, we assume there is a physical law

$$\phi(t, r, \rho, E) = 0$$

which provides a relationship among these quantities.

§1.1 Dimensional Methods

Example

Suppose that we want to compute the yield of the first atomic explosion after viewing photographs of the spread of the fireball. In such an explosion a large amount of energy E is released in a short time in a region small enough to be considered a point. From the point of the explosion a strong shock wave spreads outwards; the pressure behind the shock is on the order of hundreds of thousands of atmospheres, far greater than the ambient air pressure whose magnitude can be accordingly neglected in the early stages of the explosion. It is plausible that there is a relation between the radius of the blast wave front r, time t, the initial air density ρ, and the energy released E. Hence, we assume there is a physical law

$$\phi(t, r, \rho, E) = 0$$

which provides a relationship among these quantities.

§1.1 Dimensional Methods

Example

Suppose that we want to compute the yield of the first atomic explosion after viewing photographs of the spread of the fireball. In such an explosion a large amount of energy E is released in a short time in a region small enough to be considered a point. From the point of the explosion a strong shock wave spreads outwards; the pressure behind the shock is on the order of hundreds of thousands of atmospheres, far greater than the ambient air pressure whose magnitude can be accordingly neglected in the early stages of the explosion. It is plausible that there is a relation between the radius of the blast wave front r, time t, the initial air density ρ, and the energy released E. Hence, we assume there is a physical law

$$\phi(t, r, \rho, E) = 0$$

which provides a relationship among these quantities.

§1.1 Dimensional Methods

Suppose that m quantities $q_1, q_2, \cdots, q_m$ are dimensioned quantities that are expressed in terms of certain selected fundamental dimensions $L_1, L_2, \cdots, L_n$, where $n < m$, and the dimensions of q_j can be written in terms of the fundamental dimensions as

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}$$

for some exponents $a_{1j}, a_{2j}, \cdots, a_{nj}$. The $n \times m$ matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

containing the exponents is called the ***dimension matrix*** (of $q_1, \cdots, q_m$ w.r.t. dimensions $L_1, \cdots, L_n$). The entries in the j-th column give the exponents for q_j in terms of the powers of $L_1, \cdots, L_n$.

§1.1 Dimensional Methods

Suppose that m quantities $q_1, q_2, \cdots, q_m$ are dimensioned quantities that are expressed in terms of certain selected fundamental dimensions $L_1, L_2, \cdots, L_n$, where $n < m$, and the dimensions of q_j can be written in terms of the fundamental dimensions as

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}$$

for some exponents $a_{1j}, a_{2j}, \cdots, a_{nj}$. The $n \times m$ matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

containing the exponents is called the ***dimension matrix*** (of $q_1, \cdots, q_m$ w.r.t. dimensions $L_1, \cdots, L_n$). The entries in the j-th column give the exponents for q_j in terms of the powers of $L_1, \cdots, L_n$.

§1.1 Dimensional Methods

Suppose that m quantities $q_1, q_2, \cdots, q_m$ are dimensioned quantities that are expressed in terms of certain selected fundamental dimensions $L_1, L_2, \cdots, L_n$, where $n < m$, and the dimensions of q_j can be written in terms of the fundamental dimensions as

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}$$

for some exponents $a_{1j}, a_{2j}, \cdots, a_{nj}$. The $n \times m$ matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

containing the exponents is called the ***dimension matrix*** (of $q_1, \cdots, q_m$ w.r.t. dimensions $L_1, \cdots, L_n$). The entries in the j-th column give the exponents for q_j in terms of the powers of $L_1, \cdots, L_n$.

§1.1 Dimensional Methods

Suppose that m quantities $q_1, q_2, \cdots, q_m$ are dimensioned quantities that are expressed in terms of certain selected fundamental dimensions $L_1, L_2, \cdots, L_n$, where $n < m$, and the dimensions of q_j can be written in terms of the fundamental dimensions as

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}$$

for some exponents $a_{1j}, a_{2j}, \cdots, a_{nj}$. The $n \times m$ matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$

containing the exponents is called the ***dimension matrix*** (of $q_1, \cdots, q_m$ w.r.t. dimensions $L_1, \cdots, L_n$). The entries in the j-th column give the exponents for q_j in terms of the powers of $L_1, \cdots, L_n$.

§1.1 Dimensional Methods

We note that the choices of different independent fundamental dimensions results in different dimension matrices; however, the rank of dimension matrices is well-defined.

Definition

Let $q_1, \cdots, q_m$ be dimensioned quantities.

1. A quantity π is called a ***dimensionless combinations*** of $q_1, \cdots, q_m$ if $\pi = q_1^{\alpha_1} \cdots q_m^{\alpha_m}$ for some rational numbers $\alpha_1, \cdots, \alpha_m$ and $[\pi] = 1$.
2. A collection $\{\pi_1, \cdots, \pi_k\}$ of dimensionless combinations of $q_1, \cdots, q_m$ is said to be ***maximal*** if any dimensionless quantities π formed from $q_1, \cdots, q_m$ can be expressed as $\pi = \pi_1^{c_1} \cdots \pi_k^{c_k}$ for some **unique** $c_1, \cdots, c_k$.

Remark: If $\{\pi_1, \cdots, \pi_k\}$ is maximal, then k must be the nullity of the dimension matrix.

§1.1 Dimensional Methods

We note that the choices of different independent fundamental dimensions results in different dimension matrices; however, the rank of dimension matrices is well-defined.

Definition

Let $q_1, \cdots, q_m$ be dimensioned quantities.

1. A quantity π is called a ***dimensionless combinations*** of $q_1, \cdots, q_m$ if $\pi = q_1^{\alpha_1} \cdots q_m^{\alpha_m}$ for some rational numbers $\alpha_1, \cdots, \alpha_m$ and $[\pi] = 1$.
2. A collection $\{\pi_1, \cdots, \pi_k\}$ of dimensionless combinations of $q_1, \cdots, q_m$ is said to be ***maximal*** if any dimensionless quantities π formed from $q_1, \cdots, q_m$ can be expressed as $\pi = \pi_1^{c_1} \cdots \pi_k^{c_k}$ for some **unique** $c_1, \cdots, c_k$.

Remark: If $\{\pi_1, \cdots, \pi_k\}$ is maximal, then k must be the nullity of the dimension matrix.

§1.1 Dimensional Methods

We note that the choices of different independent fundamental dimensions results in different dimension matrices; however, the rank of dimension matrices is well-defined.

Definition

Let $q_1, \cdots, q_m$ be dimensioned quantities.

1. A quantity π is called a ***dimensionless combinations*** of $q_1, \cdots, q_m$ if $\pi = q_1^{\alpha_1} \cdots q_m^{\alpha_m}$ for some rational numbers $\alpha_1, \cdots, \alpha_m$ and $[\pi] = 1$.
2. A collection $\{\pi_1, \cdots, \pi_k\}$ of dimensionless combinations of $q_1, \cdots, q_m$ is said to be ***maximal*** if any dimensionless quantities π formed from $q_1, \cdots, q_m$ can be expressed as $\pi = \pi_1^{c_1} \cdots \pi_k^{c_k}$ for some **unique** $c_1, \cdots, c_k$.

Remark: If $\{\pi_1, \cdots, \pi_k\}$ is maximal, then k must be the nullity of the dimension matrix.

§1.1 Dimensional Methods

We note that the choices of different independent fundamental dimensions results in different dimension matrices; however, the rank of dimension matrices is well-defined.

Definition

Let $q_1, \cdots, q_m$ be dimensioned quantities.

1. A quantity π is called a ***dimensionless combinations*** of $q_1, \cdots, q_m$ if $\pi = q_1^{\alpha_1} \cdots q_m^{\alpha_m}$ for some rational numbers $\alpha_1, \cdots, \alpha_m$ and $[\pi] = 1$.
2. A collection $\{\pi_1, \cdots, \pi_k\}$ of dimensionless combinations of $q_1, \cdots, q_m$ is said to be ***maximal*** if any dimensionless quantities π formed from $q_1, \cdots, q_m$ can be expressed as $\pi = \pi_1^{c_1} \cdots \pi_k^{c_k}$ for some **unique** $c_1, \cdots, c_k$.

Remark: If $\{\pi_1, \cdots, \pi_k\}$ is maximal, then k must be the nullity of the dimension matrix.

§1.1 Dimensional Methods

Any fundamental dimension L_k has the property that its units can be changed upon multiplication by the appropriate conversion factor to obtain a new value in a new system of units. Let $\{[L_1]_1, \cdots, [L_n]_1\}$ and $\{[L_1]_2, \cdots, [L_n]_2\}$ be two particular choices of **units** for fundamental dimensions. Then for each $1 \leqslant k \leqslant n$, $[L_k]_2 = \lambda_k [L_k]_1$ for some dimensionless constant $\lambda_k > 0$. The value of a quantity q then can be changed in the fashion that if

$$[q] = L_1^{b_1} L_2^{b_2} \cdots L_n^{b_n}, \tag{1}$$

and $v_1(q)$ denotes the value of q in the system of units $\{[L_k]_1\}_{k=1}^n$, then

$$v_2(q) = \lambda_1^{b_1} \lambda_2^{b_2} \cdots \lambda_n^{b_n} v_1(q) \tag{2}$$

gives the value of q in the new system of units $\{[L_k]_2\}_{k=1}^n$.

§1.1 Dimensional Methods

Any fundamental dimension L_k has the property that its units can be changed upon multiplication by the appropriate conversion factor to obtain a new value in a new system of units. Let $\{[L_1]_1, \cdots, [L_n]_1\}$ and $\{[L_1]_2, \cdots, [L_n]_2\}$ be two particular choices of **units** for fundamental dimensions. Then for each $1 \leqslant k \leqslant n$, $[L_k]_2 = \lambda_k [L_k]_1$ for some dimensionless constant $\lambda_k > 0$. The value of a quantity q then can be changed in the fashion that if

$$[q] = L_1^{b_1} L_2^{b_2} \cdots L_n^{b_n}, \tag{1}$$

and $v_1(q)$ denotes the value of q in the system of units $\{[L_k]_1\}_{k=1}^n$, then

$$v_2(q) = \lambda_1^{b_1} \lambda_2^{b_2} \cdots \lambda_n^{b_n} v_1(q) \tag{2}$$

gives the value of q in the new system of units $\{[L_k]_2\}_{k=1}^n$.

§1.1 Dimensional Methods

Any fundamental dimension L_k has the property that its units can be changed upon multiplication by the appropriate conversion factor to obtain a new value in a new system of units. Let $\{[L_1]_1, \cdots, [L_n]_1\}$ and $\{[L_1]_2, \cdots, [L_n]_2\}$ be two particular choices of **units** for fundamental dimensions. Then for each $1 \leqslant k \leqslant n$, $[L_k]_2 = \lambda_k [L_k]_1$ for some dimensionless constant $\lambda_k > 0$. The value of a quantity q then can be changed in the fashion that if

$$[q] = L_1^{b_1} L_2^{b_2} \cdots L_n^{b_n}, \tag{1}$$

and $v_1(q)$ denotes the value of q in the system of units $\{[L_k]_1\}_{k=1}^n$, then

$$v_2(q) = \lambda_1^{b_1} \lambda_2^{b_2} \cdots \lambda_n^{b_n} v_1(q) \tag{2}$$

gives the value of q in the new system of units $\{[L_k]_2\}_{k=1}^n$.

§1.1 Dimensional Methods

To justify if a given physical law $\phi(q_1, \cdots, q_m) = 0$ is true, we measure each dimensioned quantities based on a particular choice of units and check if the law holds for this particular choice of units. The fact that the validity of a physical law is independent of the choice of units induces the following

Definition

Let $q_1, q_2, \cdots, q_m$ be dimensioned quantities. The physical law

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is said to be ***unit free*** (or ***physically meaningful***) if for all positive real numbers $\lambda_1, \cdots, \lambda_n$,

$$\phi\big(v_1(q_1), \cdots, v_1(q_m)\big) = 0 \quad \Leftrightarrow \quad \phi\big(v_2(q_1), \cdots, v_2(q_m)\big) = 0\,,$$

where $v_1(q_j)$ and $v_2(q_j)$ are related by (2) if q_j obeys (1).

§1.1 Dimensional Methods

To justify if a given physical law $\phi(q_1, \cdots, q_m) = 0$ is true, we measure each dimensioned quantities based on a particular choice of units and check if the law holds for this particular choice of units. The fact that the validity of a physical law is independent of the choice of units induces the following

Definition

Let $q_1, q_2, \cdots, q_m$ be dimensioned quantities. The physical law

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is said to be ***unit free*** (or ***physically meaningful***) if for all positive real numbers $\lambda_1, \cdots, \lambda_n$,

$$\phi\big(v_1(q_1), \cdots, v_1(q_m)\big) = 0 \quad \Leftrightarrow \quad \phi\big(v_2(q_1), \cdots, v_2(q_m)\big) = 0,$$

where $v_1(q_j)$ and $v_2(q_j)$ are related by (2) if q_j obeys (1).

§1.1 Dimensional Methods

Theorem (Buckingham's Pi Theorem)

Suppose that

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is a unit free physical law that relates the dimensioned quantities $q_1, q_2, \cdots, q_m$. ~~*Let* $L_1, L_2, \cdots, L_n$, *where* $n < m$, *be fundamental dimensions with*~~

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}, \qquad j = 1, \cdots, m.$$

Then there exists a maximal collection $\{\pi_1, \pi_2, \cdots, \pi_k\}$ *of dimensionless combinations of* $q_1, \cdots, q_m$ *and the physical law above is equivalent to an equation*

$$\Phi(\pi_1, \cdots, \pi_k) = 0$$

expressed only in terms of the dimensionless quantities.

Remark: If the rank of the dimension matrix is r, then $k = m - r$.

§1.1 Dimensional Methods

Theorem (Buckingham's Pi Theorem)

Suppose that

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is a unit free physical law that relates the dimensioned quantities $q_1, q_2, \cdots, q_m$. ~~*Let* $L_1, L_2, \cdots, L_n$*, where* $n < m$*, be fundamental dimensions with*~~

$$[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}, \qquad j = 1, \cdots, m.$$

Then there exists a maximal collection $\{\pi_1, \pi_2, \cdots, \pi_k\}$ *of dimensionless combinations of* $q_1, \cdots, q_m$ *and the physical law above is equivalent to an equation*

$$\Phi(\pi_1, \cdots, \pi_k) = 0$$

expressed only in terms of the dimensionless quantities.

Remark: If the rank of the dimension matrix is r, then $k = m - r$.

§1.1 Dimensional Methods

Theorem (Buckingham's Pi Theorem)

Suppose that

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is a unit free physical law that relates the dimensioned quantities $q_1, q_2, \cdots, q_m$. ~~*Let* $L_1, L_2, \cdots, L_n$, *where* $n < m$, *be fundamental dimensions with*~~

$$\require{cancel}\cancel{[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}, \qquad j = 1, \cdots, m.}$$

Then there exists a maximal collection $\{\pi_1, \pi_2, \cdots, \pi_k\}$ *of dimensionless combinations of* $q_1, \cdots, q_m$ *and the physical law above is equivalent to an equation*

$$\Phi(\pi_1, \cdots, \pi_k) = 0$$

expressed only in terms of the dimensionless quantities.

Remark: If the rank of the dimension matrix is r, then $k = m - r$.

§1.1 Dimensional Methods

Theorem (Buckingham's Pi Theorem)

Suppose that

$$\phi(q_1, q_2, \cdots, q_m) = 0$$

is a unit free physical law that relates the dimensioned quantities $q_1, q_2, \cdots, q_m$. ~~*Let* $L_1, L_2, \cdots, L_n$*, where* $n < m$*, be fundamental dimensions with*~~

$$\require{cancel}\cancel{[q_j] = L_1^{a_{1j}} L_2^{a_{2j}} \cdots L_n^{a_{nj}}, \qquad j = 1, \cdots, m.}$$

Then there exists a maximal collection $\{\pi_1, \pi_2, \cdots, \pi_k\}$ *of dimensionless combinations of* $q_1, \cdots, q_m$ *and the physical law above is equivalent to an equation*

$$\Phi(\pi_1, \cdots, \pi_k) = 0$$

expressed only in terms of the dimensionless quantities.

Remark: If the rank of the dimension matrix is r, then $k = m - r$.

§1.1 Dimensional Methods

Proof of the Pi Theorem.

Let $D = [a_{ij}]_{n\times m}$ be the dimension matrix, $r = \text{rank}(D)$. Suppose that $\pi = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_m^{\alpha_m}$ is a dimensionless quantities. Then with $\boldsymbol{\alpha}$ denoting the column vector $[\alpha_1, \cdots, \alpha_m]^{\mathrm{T}}$, we have

$$D\boldsymbol{\alpha} = \mathbf{0},$$

where $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^n$. Since $\text{rank}(D) = r$, W.L.O.G. we can assume that the first r column of D is linearly independent; thus $\alpha_1, \cdots, \alpha_r$ can be uniquely expressed in terms of $(\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_m)$. In fact,

$$D(:, 1:r)\boldsymbol{\alpha}(1:r) = -D(:, r+1:m)\boldsymbol{\alpha}(r+1:m),$$

where $D(:, i:j)$ denotes the matrix formed by the i-th to j-th columns of D and $\boldsymbol{\alpha}(i:j)$ denotes the column vector formed by the i-th to j-th components of $\boldsymbol{\alpha}$. □

§1.1 Dimensional Methods

Proof of the Pi Theorem.

Let $D = [a_{ij}]_{n\times m}$ be the dimension matrix, $r = \text{rank}(D)$. Suppose that $\pi = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_m^{\alpha_m}$ is a dimensionless quantities. Then with $\boldsymbol{\alpha}$ denoting the column vector $[\alpha_1, \cdots, \alpha_m]^{\mathrm{T}}$, we have

$$D\boldsymbol{\alpha} = \mathbf{0},$$

where $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^n$. Since $\text{rank}(D) = r$, W.L.O.G. we can assume that the first r column of D is linearly independent; thus $\alpha_1, \cdots, \alpha_r$ can be uniquely expressed in terms of $(\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_m)$. In fact,

$$D(:, 1:r)\alpha(1:r) = -D(:, r+1:m)\alpha(r+1:m),$$

where $D(:, i:j)$ denotes the matrix formed by the i-th to j-th columns of D and $\alpha(i:j)$ denotes the column vector formed by the i-th to j-th components of α. □

§1.1 Dimensional Methods

Proof of the Pi Theorem.

Let $D = [a_{ij}]_{n\times m}$ be the dimension matrix, $r = \text{rank}(D)$. Suppose that $\pi = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_m^{\alpha_m}$ is a dimensionless quantities. Then with $\boldsymbol{\alpha}$ denoting the column vector $[\alpha_1, \cdots, \alpha_m]^{\mathrm{T}}$, we have

$$D\boldsymbol{\alpha} = \mathbf{0},$$

where $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^n$. Since $\text{rank}(D) = r$, W.L.O.G. we can assume that the first r column of D is linearly independent; thus $\alpha_1, \cdots, \alpha_r$ can be uniquely expressed in terms of $(\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_m)$. In fact,

$$D(:, 1:r)\boldsymbol{\alpha}(1:r) = -D(:, r+1:m)\boldsymbol{\alpha}(r+1:m),$$

where $D(:, i:j)$ denotes the matrix formed by the i-th to j-th columns of D and $\boldsymbol{\alpha}(i:j)$ denotes the column vector formed by the i-th to j-th components of $\boldsymbol{\alpha}$. □

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Assume that the vector $\alpha(1:r)$ is given by

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1(m-r)} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{r(m-r)} \end{bmatrix} \begin{bmatrix} \alpha_{r+1} \\ \vdots \\ \alpha_m \end{bmatrix},$$

and let $\pi_1, \cdots, \pi_{m-r}$ be given by $\pi_j = q_1^{b_{1j}} q_2^{b_{2j}} \cdots q_r^{b_{rj}} q_{r+j}$. Then $\{\pi_1, \cdots, \pi_{m-r}\}$ is a maximal collection of dimensionless combinations of $q_1, \cdots, q_r$. Define

$$\begin{aligned} & F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) \\ & = \phi\big(q_1, \cdots, q_r, \pi_1 q_1^{-b_{11}} \cdots q_r^{-b_{r1}}, \cdots, \pi_{m-r} q_1^{-b_{1(m-r)}} \cdots q_r^{-b_{r(m-r)}}\big). \end{aligned}$$

We then have $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ if and only if $\phi(q_1, \cdots, q_m) = 0$. Moreover, since $\phi(q_1, q_2, \cdots, q_m) = 0$ is unit free, $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ is unit free. □

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Assume that the vector $\boldsymbol{\alpha}(1:r)$ is given by

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1(m-r)} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{r(m-r)} \end{bmatrix} \begin{bmatrix} \alpha_{r+1} \\ \vdots \\ \alpha_m \end{bmatrix},$$

and let $\pi_1, \cdots, \pi_{m-r}$ be given by $\pi_j = q_1^{b_{1j}} q_2^{b_{2j}} \cdots q_r^{b_{rj}} q_{r+j}$. Then $\{\pi_1, \cdots, \pi_{m-r}\}$ is a maximal collection of dimensionless combinations of $q_1, \cdots, q_r$. Define

$$\begin{aligned} & F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) \\ & = \phi(q_1, \cdots, q_r, \pi_1 q_1^{-b_{11}} \cdots q_r^{-b_{r1}}, \cdots, \pi_{m-r} q_1^{-b_{1(m-r)}} \cdots q_r^{-b_{r(m-r)}}). \end{aligned}$$

We then have $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ if and only if $\phi(q_1, \cdots, q_m) = 0$. Moreover, since $\phi(q_1, q_2, \cdots, q_m) = 0$ is unit free, $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ is unit free. □

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Assume that the vector $\boldsymbol{\alpha}(1:r)$ is given by

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1(m-r)} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{r(m-r)} \end{bmatrix} \begin{bmatrix} \alpha_{r+1} \\ \vdots \\ \alpha_m \end{bmatrix},$$

and let $\pi_1, \cdots, \pi_{m-r}$ be given by $\pi_j = q_1^{b_{1j}} q_2^{b_{2j}} \cdots q_r^{b_{rj}} q_{r+j}$. Then $\{\pi_1, \cdots, \pi_{m-r}\}$ is a maximal collection of dimensionless combinations of $q_1, \cdots, q_r$. Define

$$\begin{aligned} & F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) \\ & \quad = \phi\big(q_1, \cdots, q_r, \pi_1 q_1^{-b_{11}} \cdots q_r^{-b_{r1}}, \cdots, \pi_{m-r} q_1^{-b_{1(m-r)}} \cdots q_r^{-b_{r(m-r)}}\big). \end{aligned}$$

We then have $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ if and only if $\phi(q_1, \cdots, q_m) = 0$. Moreover, since $\phi(q_1, q_2, \cdots, q_m) = 0$ is unit free, $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ is unit free. □

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Assume that the vector $\boldsymbol{\alpha}(1:r)$ is given by

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1(m-r)} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{r(m-r)} \end{bmatrix} \begin{bmatrix} \alpha_{r+1} \\ \vdots \\ \alpha_m \end{bmatrix},$$

and let $\pi_1, \cdots, \pi_{m-r}$ be given by $\pi_j = q_1^{b_{1j}} q_2^{b_{2j}} \cdots q_r^{b_{rj}} q_{r+j}$. Then $\{\pi_1, \cdots, \pi_{m-r}\}$ is a maximal collection of dimensionless combinations of $q_1, \cdots, q_r$. Define

$$\begin{aligned} &F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) \\ &\quad = \phi(q_1, \cdots, q_r, \pi_1 q_1^{-b_{11}} \cdots q_r^{-b_{r1}}, \cdots, \pi_{m-r} q_1^{-b_{1(m-r)}} \cdots q_r^{-b_{r(m-r)}}) . \end{aligned}$$

We then have $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ if and only if $\phi(q_1, \cdots, q_m) = 0$. Moreover, since $\phi(q_1, q_2, \cdots, q_m) = 0$ is unit free, $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ is unit free. □

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Now, since $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$ is unit free, for any choice of unit systems $\{[L_1]_1, \cdots, [L_n]_1\}$ and $\{[L_1]_2, \cdots, [L_n]_2\}$ with conversion factors $\lambda_1, \cdots, \lambda_n > 0$ so that

$$v_2(q_j) = \lambda_1^{a_{1j}} \lambda_2^{a_{2j}} \cdots \lambda_n^{a_{nj}} v_1(q_j), \quad 1 \leqslant j \leqslant r,$$

we must have $F\big(v_2(q_1), \cdots, v_2(q_r), \pi_1, \cdots, \pi_{m-r}\big) = 0$. Since the columns of $D(:, 1:r)$ are linearly independent and $n \geqslant r$, there exist $\lambda_1, \cdots, \lambda_n$ (might not be unique if $n > r$) such that

$$\begin{bmatrix} a_{11} & \cdots & a_{n1} \\ a_{12} & \cdots & a_{n2} \\ \vdots & & \vdots \\ a_{1r} & \cdots & a_{nr} \end{bmatrix} \begin{bmatrix} \log \lambda_1 \\ \log \lambda_2 \\ \vdots \\ \log \lambda_n \end{bmatrix} = \begin{bmatrix} -\log v_1(q_1) \\ -\log v_1(q_2) \\ \vdots \\ -\log v_1(q_r) \end{bmatrix} \qquad (3)$$

□

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Now, since $F(q_1,\cdots,q_r,\pi_1,\cdots,\pi_{m-r}) = 0$ is unit free, for any choice of unit systems $\{[L_1]_1,\cdots,[L_n]_1\}$ and $\{[L_1]_2,\cdots,[L_n]_2\}$ with conversion factors $\lambda_1,\cdots,\lambda_n > 0$ so that

$$v_2(q_j) = \lambda_1^{a_{1j}}\lambda_2^{a_{2j}}\cdots\lambda_n^{a_{nj}} v_1(q_j)\,, \quad 1 \leqslant j \leqslant r\,,$$

we must have $F\big(v_2(q_1),\cdots,v_2(q_r),\pi_1,\cdots,\pi_{m-r}\big) = 0$. Since the columns of $D(:,1:r)$ are linearly independent and $n \geqslant r$, there exist $\lambda_1,\cdots,\lambda_n$ (might not be unique if $n > r$) such that

$$\begin{bmatrix} a_{11} & \cdots & a_{n1} \\ a_{12} & \cdots & a_{n2} \\ \vdots & & \vdots \\ a_{1r} & \cdots & a_{nr} \end{bmatrix} \begin{bmatrix} \log\lambda_1 \\ \log\lambda_2 \\ \vdots \\ \log\lambda_n \end{bmatrix} = \begin{bmatrix} -\log v_1(q_1) \\ -\log v_1(q_2) \\ \vdots \\ -\log v_1(q_r) \end{bmatrix} \qquad (3)$$

□

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Choose $\lambda_1, \cdots, \lambda_n$ satisfying (3). Then in the new system of units $v_2(q_j) = 1$ for all $1 \leqslant j \leqslant r$; thus we establish that as long as $q_1, \cdots, q_r$ satisfy $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$, there exists a system of units such that $v_2(q_1) = \cdots = v_2(q_r) = 1$. This implies that F is independent of $q_1, \cdots, q_r$ and we have

$$\Phi(\pi_1, \cdots, \pi_{m-r}) \equiv F(1, \cdots, 1, \pi_1, \cdots, \pi_{m-r}) = 0\,. \qquad \square$$

Example

Reconsider the biker's air resistance problem in which the physical law is

$$\Phi(F, \rho, A, v) = 0\,,$$

where F is the air resistance, ρ is the air density, A is the cross-sectional area, and v is the velocity.

§1.1 Dimensional Methods

Proof of the Pi Theorem (cont.)

Choose $\lambda_1, \cdots, \lambda_n$ satisfying (3). Then in the new system of units $v_2(q_j) = 1$ for all $1 \leqslant j \leqslant r$; thus we establish that as long as $q_1, \cdots, q_r$ satisfy $F(q_1, \cdots, q_r, \pi_1, \cdots, \pi_{m-r}) = 0$, there exists a system of units such that $v_2(q_1) = \cdots = v_2(q_r) = 1$. This implies that F is independent of $q_1, \cdots, q_r$ and we have

$$\Phi(\pi_1, \cdots, \pi_{m-r}) \equiv F(1, \cdots, 1, \pi_1, \cdots, \pi_{m-r}) = 0\,. \qquad \square$$

Example

Reconsider the biker's air resistance problem in which the physical law is

$$\Phi(F, \rho, A, v) = 0\,,$$

where F is the air resistance, ρ is the air density, A is the cross-sectional area, and v is the velocity.

§1.1 Dimensional Methods

Example (Biker's air resistance problem (cont.))

Since $[F] = MLT^{-2}$, $[\rho] = ML^{-3}$, $[A] = L^2$ and $[v] = LT^{-1}$, the dimension matrix (with the order of dimension T, L, M) is

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

The rank of the dimension matrix above is 3; thus there is only one dimensionless quantity that can be formed from F, ρ, A, v. Suppose that $\pi = F^{\alpha_1}\rho^{\alpha_2}A^{\alpha_3}v^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = F\rho^{-1}A^{-1}v^{-2}$. Therefore, an equivalent physical law is given by $g(\pi) = 0$ which shows that $\pi = k$ (or equivalently, $F = k\rho Av^2$) for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example (Biker's air resistance problem (cont.))

Since $[F] = MLT^{-2}$, $[\rho] = ML^{-3}$, $[A] = L^2$ and $[v] = LT^{-1}$, the dimension matrix (with the order of dimension T, L, M) is

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

The rank of the dimension matrix above is 3; thus there is only one dimensionless quantity that can be formed from F, ρ, A, v. Suppose that $\pi = F^{\alpha_1}\rho^{\alpha_2}A^{\alpha_3}v^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = F\rho^{-1}A^{-1}v^{-2}$. Therefore, an equivalent physical law is given by $g(\pi) = 0$ which shows that $\pi = k$ (or equivalently, $F = k\rho Av^2$) for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example (Biker's air resistance problem (cont.))

Since $[F] = MLT^{-2}$, $[\rho] = ML^{-3}$, $[A] = L^2$ and $[v] = LT^{-1}$, the dimension matrix (with the order of dimension T, L, M) is

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

The rank of the dimension matrix above is 3; thus there is only one dimensionless quantity that can be formed from F, ρ, A, v. Suppose that $\pi = F^{\alpha_1}\rho^{\alpha_2}A^{\alpha_3}v^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = F\rho^{-1}A^{-1}v^{-2}$. Therefore, an equivalent physical law is given by $g(\pi) = 0$ which shows that $\pi = k$ (or equivalently, $F = k\rho A v^2$) for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example (Biker's air resistance problem (cont.))

Since $[F] = MLT^{-2}$, $[\rho] = ML^{-3}$, $[A] = L^2$ and $[v] = LT^{-1}$, the dimension matrix (with the order of dimension T, L, M) is

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

The rank of the dimension matrix above is 3; thus there is only one dimensionless quantity that can be formed from F, ρ, A, v. Suppose that $\pi = F^{\alpha_1}\rho^{\alpha_2}A^{\alpha_3}v^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 1 & -3 & 2 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = F\rho^{-1}A^{-1}v^{-2}$. Therefore, an equivalent physical law is given by $g(\pi) = 0$ which shows that $\pi = k$ (or equivalently, $F = k\rho A v^2$) for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example

Reconsider the atomic explosion problem in which the physical law is given by $\phi(t, r, \rho, E) = 0$, where

$$[t] = T, \quad [r] = L, \quad [\rho] = ML^{-3}, \quad [E] = ML^2T^{-2},$$

so that the dimension matrix (with the order of T, L, M) is given by

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The rank of the dimension matrix above is clearly 3; thus there is only one dimensionless quantity that can be formed from t, r, ρ, E. Suppose that $\pi = t^{\alpha_1} r^{\alpha_2} \rho^{\alpha_3} E^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = t^2 r^{-5} \rho^{-1} E$ as well as an equivalent physical law $\pi = k$ for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example

Reconsider the atomic explosion problem in which the physical law is given by $\phi(t, r, \rho, E) = 0$, where

$$[t] = T, \quad [r] = L, \quad [\rho] = ML^{-3}, \quad [E] = ML^2T^{-2},$$

so that the dimension matrix (with the order of T, L, M) is given by

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The rank of the dimension matrix above is clearly 3; thus there is only one dimensionless quantity that can be formed from t, r, ρ, E. Suppose that $\pi = t^{\alpha_1} r^{\alpha_2} \rho^{\alpha_3} E^{\alpha_4}$ is a dimensionless quantity. Then

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives a dimensionless quantity $\pi = t^2 r^{-5} \rho^{-1} E$ as well as an equivalent physical law $\pi = k$ for some (dimensionless) constant k.

§1.1 Dimensional Methods

Example

At time $t = 0$ an amount of heat energy e, concentrated at a point in space, is allowed to diffuse outward into a region with temperature zero. If r denotes the radial distance from the source and t is time, the problem is to determine the temperature θ as a function of r and t.

Clearly the temperature θ depends on t, r and e. Moreover, it is "reasonable" that the "thermal diffusivity" k with dimension length-squared per time and the "heat capacity" c of the region, with dimension energy per degree per volume, play a role. Therefore, the physical law is given by

$$\phi(t, r, \theta, e, k, c) = 0 .$$

§1.1 Dimensional Methods

Example

At time $t = 0$ an amount of heat energy e, concentrated at a point in space, is allowed to diffuse outward into a region with temperature zero. If r denotes the radial distance from the source and t is time, the problem is to determine the temperature θ as a function of r and t.

Clearly the temperature θ depends on t, r and e. Moreover, it is "reasonable" that the "thermal diffusivity" k with dimension length-squared per time and the "heat capacity" c of the region, with dimension energy per degree per volume, play a role. Therefore, the physical law is given by

$$\phi(t, r, \theta, e, k, c) = 0 .$$

§1.1 Dimensional Methods

Example

At time $t = 0$ an amount of heat energy e, concentrated at a point in space, is allowed to diffuse outward into a region with temperature zero. If r denotes the radial distance from the source and t is time, the problem is to determine the temperature θ as a function of r and t.

Clearly the temperature θ depends on t, r and e. Moreover, it is "reasonable" that the "thermal diffusivity" k with dimension length-squared per time and the "heat capacity" c of the region, with dimension energy per degree per volume, play a role. Therefore, the physical law is given by

$$\phi(t, r, \theta, e, k, c) = 0\,.$$

§1.1 Dimensional Methods

Example (cont.)

This physical law has 6 dimensioned quantities

$$[t] = T, \quad [r] = L, \qquad [\theta] = \Theta,$$
$$[e] = E, \quad [k] = L^2 T^{-1}, \quad [c] = E\Theta^{-1} L^{-3}.$$

The dimension matrix (with the order of T, L, Θ, E) is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & -3 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

It is easy to see that the dimension matrix has rank 4; thus by the Pi theorem there are 2 dimensionless quantities that can be formed from t, r, θ, e, c, k. To see how we form dimensionless quantities, we assume that the combination

$$[t^{\alpha_1} r^{\alpha_2} \theta^{\alpha_3} e^{\alpha_4} k^{\alpha_5} c^{\alpha_6}] = 1.$$

§1.1 Dimensional Methods

Example (cont.)

This physical law has 6 dimensioned quantities

$$[t] = T, \quad [r] = L, \quad [\theta] = \Theta,$$
$$[e] = E, \quad [k] = L^2 T^{-1}, \quad [c] = E\Theta^{-1} L^{-3}.$$

The dimension matrix (with the order of T, L, Θ, E) is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & -3 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

It is easy to see that the dimension matrix has rank 4; thus by the Pi theorem there are 2 dimensionless quantities that can be formed from t, r, θ, e, c, k. To see how we form dimensionless quantities, we assume that the combination

$$\left[t^{\alpha_1} r^{\alpha_2} \theta^{\alpha_3} e^{\alpha_4} k^{\alpha_5} c^{\alpha_6}\right] = 1.$$

§1.1 Dimensional Methods

Example (cont.)

In other words,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & -3 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

which shows that $\alpha_1 = \alpha_5$, $\alpha_3 = -\alpha_4 = \alpha_6$, and $\alpha_2 = -2\alpha_5 + 3\alpha_6$. Therefore, two dimensionless quantities can be formed (using $(\alpha_5, \alpha_6) = (-\frac{1}{2}, 0)$ or $(\frac{3}{2}, 1)$) as

$$\pi_1 = \frac{r}{\sqrt{kt}} \qquad \text{and} \qquad \pi_2 = \frac{\theta c}{e}(kt)^{\frac{3}{2}}$$

and an equivalent physical law is given by $\Phi(\pi_1, \pi_2) = 0$ which "implies" that $\pi_2 = u(\pi_1)$ for some function u. Therefore, the temperature θ can be expressed by $\theta = \frac{e}{c(kt)^{\frac{3}{2}}} u\left(\frac{r}{\sqrt{kt}}\right)$.

§1.1 Dimensional Methods

Example (cont.)

In other words,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 2 & -3 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

which shows that $\alpha_1 = \alpha_5$, $\alpha_3 = -\alpha_4 = \alpha_6$, and $\alpha_2 = -2\alpha_5 + 3\alpha_6$. Therefore, two dimensionless quantities can be formed $\big($using $(\alpha_5, \alpha_6) = \big(-\frac{1}{2}, 0\big)$ or $\big(\frac{3}{2}, 1\big)\big)$ as

$$\pi_1 = \frac{r}{\sqrt{kt}} \qquad \text{and} \qquad \pi_2 = \frac{\theta c}{e}(kt)^{\frac{3}{2}}$$

and an equivalent physical law is given by $\Phi(\pi_1, \pi_2) = 0$ which "implies" that $\pi_2 = u(\pi_1)$ for some function u. Therefore, the temperature θ can be expressed by $\theta = \dfrac{e}{c(kt)^{\frac{3}{2}}} u\Big(\dfrac{r}{\sqrt{kt}}\Big)$.

§1.1 Dimensional Methods

Example

In this example we determine the relation between the power P that must be applied to keep a ship of length ℓ moving at a constant speed V. Assume that P depends on the density of water ϱ, the acceleration due to gravity g, and the viscosity of water ν (in length-squared per time), as well as ℓ and V. The physical law is given by

$$\phi(P, \varrho, g, \nu, \ell, V) = 0\,.$$

Suppose that the fundamental dimension is the time T, the length L, and the mass M. Then

$$\begin{aligned} [P] &= ML^2T^{-3}\,, & [\varrho] &= ML^{-3}\,, & [g] &= LT^{-2}\,, \\ [\nu] &= L^2T^{-1}\,, & [\ell] &= L\,, & [V] &= LT^{-1}\,. \end{aligned}$$

§1.1 Dimensional Methods

Example

In this example we determine the relation between the power P that must be applied to keep a ship of length ℓ moving at a constant speed V. Assume that P depends on the density of water ϱ, the acceleration due to gravity g, and the viscosity of water ν (in length-squared per time), as well as ℓ and V. The physical law is given by

$$\phi(P, \varrho, g, \nu, \ell, V) = 0 .$$

Suppose that the fundamental dimension is the time T, the length L, and the mass M. Then

$$\begin{aligned} &[P] = ML^2T^{-3}, \quad &[\varrho] = ML^{-3}, \quad &[g] = LT^{-2}, \\ &[\nu] = L^2T^{-1}, \quad &[\ell] = L, \quad &[V] = LT^{-1}. \end{aligned}$$

§1.1 Dimensional Methods

Example

In this example we determine the relation between the power P that must be applied to keep a ship of length ℓ moving at a constant speed V. Assume that P depends on the density of water ϱ, the acceleration due to gravity g, and the viscosity of water ν (in length-squared per time), as well as ℓ and V. The physical law is given by

$$\phi(P, \varrho, g, \nu, \ell, V) = 0\,.$$

Suppose that the fundamental dimension is the time T, the length L, and the mass M. Then

$$\begin{aligned} &[P] = ML^2T^{-3}\,, \quad &&[\varrho] = ML^{-3}\,, \quad &&[g] = LT^{-2}\,, \\ &[\nu] = L^2T^{-1}\,, &&[\ell] = L\,, &&[V] = LT^{-1}\,. \end{aligned}$$

§1.1 Dimensional Methods

Example (cont.)

Therefore, the dimension matrix (in the order T, L, M) is

$$D = \begin{bmatrix} -3 & 0 & -2 & -1 & 0 & -1 \\ 2 & -3 & 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which has rank 3. By the Pi Theorem, there are three dimensionless quantities π_1, π_2 and π_3 and the physical law $\phi(P, \varrho, g, \nu, \ell, V) = 0$ is equivalent to $\Phi(\pi_1, \pi_2, \pi_3) = 0$ (or sometimes $\pi_1 = F(\pi_2, \pi_3)$).

Suppose that $\pi = P^{\alpha_1} \varrho^{\alpha_2} g^{\alpha_3} \nu^{\alpha_4} \ell^{\alpha_5} V^{\alpha_6}$ is dimensionless. Then

$$\begin{bmatrix} -3 & 0 & -2 & -1 & 0 & -1 \\ 2 & -3 & 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

§1.1 Dimensional Methods

Example (cont.)

Therefore, the dimension matrix (in the order T, L, M) is

$$D = \begin{bmatrix} -3 & 0 & -2 & -1 & 0 & -1 \\ 2 & -3 & 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which has rank 3. By the Pi Theorem, there are three dimensionless quantities π_1, π_2 and π_3 and the physical law $\phi(P, \varrho, g, \nu, \ell, V) = 0$ is equivalent to $\Phi(\pi_1, \pi_2, \pi_3) = 0$ (or sometimes $\pi_1 = F(\pi_2, \pi_3)$). Suppose that $\pi = P^{\alpha_1}\varrho^{\alpha_2}g^{\alpha_3}\nu^{\alpha_4}\ell^{\alpha_5}V^{\alpha_6}$ is dimensionless. Then

$$\begin{bmatrix} -3 & 0 & -2 & -1 & 0 & -1 \\ 2 & -3 & 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

§1.1 Dimensional Methods

Example (cont.)

Three choices of $(\alpha_1, \cdots, \alpha_6)$ are

$$(1, -1, 0, 0, -2, -3), \quad \left(0, 0, -\frac{1}{2}, 0, -\frac{1}{2}, 1\right) \text{ and } (0, 0, 0, -1, 1, 1)$$

which implies that the physical law is equivalent to

$$\frac{P}{\varrho \ell^2 V^3} = F\left(\frac{V}{\sqrt{\ell g}}, \frac{V\ell}{\nu}\right).$$

The two dimensionless quantities $\frac{V}{\sqrt{\ell g}}$ and $\frac{V\ell}{\nu}$ are called the Froude number Fr and the Reynolds number Re, respectively, so that the equality above can be rewritten as

$$\frac{P}{\varrho \ell^2 V^3} = F(\text{Fr}, \text{Re}).$$

§1.1 Dimensional Methods

Example (cont.)

Three choices of $(\alpha_1, \cdots, \alpha_6)$ are

$$(1, -1, 0, 0, -2, -3), \quad \left(0, 0, -\frac{1}{2}, 0, -\frac{1}{2}, 1\right) \quad \text{and} \quad (0, 0, 0, -1, 1, 1)$$

which implies that the physical law is equivalent to

$$\frac{P}{\varrho \ell^2 V^3} = F\Big(\frac{V}{\sqrt{\ell g}}, \frac{V\ell}{\nu}\Big).$$

The two dimensionless quantities $\dfrac{V}{\sqrt{\ell g}}$ and $\dfrac{V\ell}{\nu}$ are called the Froude number Fr and the Reynolds number Re, respectively, so that the equality above can be rewritten as

$$\frac{P}{\varrho \ell^2 V^3} = F(\mathrm{Fr}, \mathrm{Re}).$$

§1.1 Dimensional Methods

Example

Suppose that at time $t = 0$ an object of mass m is given a vertical upward velocity V from the surface of a spherical planet (with mass M and radius R). The height h of the object is a function of t that obeys

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,.$$

The gravitational acceleration g on the surface of the planet is given by $g = \dfrac{GM}{R^2}$; thus including the *initial data*,

$$\frac{d^2h}{dt^2} = -\frac{R^2 g}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(0) = V\,.$$

§1.1 Dimensional Methods

Example (cont.)

The physical law of the system above can be written as

$$\phi(t, h, R, V, g) = 0\,,$$

where the five dimensioned quantities have dimension

$$[t] = T\,,\ [h] = L\,,\ [R] = L\,,\ [V] = LT^{-1} \text{ and } [g] = LT^{-2}\,,$$

and the dimension matrix (with the order of T, L) is given by

$$\begin{bmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

If $\pi = t^{\alpha_1} h^{\alpha_2} R^{\alpha_3} V^{\alpha_4} g^{\alpha_5}$ is a dimensionless quantity, then

$$\begin{bmatrix} 1 & 0 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or equivalently, $\alpha_1 = \alpha_4 + 2\alpha_5$ and $\alpha_2 = -(\alpha_3 + \alpha_4 + \alpha_5)$.

§1.1 Dimensional Methods

Example (cont.)

Since the rank of the dimension matrix is 2 there are three dimensionless quantities that can be formed: we choose $(\alpha_3, \alpha_4, \alpha_5) = (-1, 0, 0), (-1, 1, 0)$ and $\big(-1, 2, -1\big)$ to form

$$\pi_1 = \frac{h}{R}, \qquad \pi_2 = \frac{tV}{R}, \qquad \pi_3 = \frac{V^2}{gR}.$$

Therefore, the Pi theorem "implies" that there exists a function $\widetilde{\Phi}$ such that $\widetilde{\Phi}(\pi_1, \pi_2, \pi_3) = 0$ which "implies" that $\pi_1 = \Phi(\pi_2, \pi_3)$; thus

$$\frac{h}{R} = \Phi\Big(\frac{tV}{R}, \frac{V^2}{gR}\Big).$$

§1.1 Dimensional Methods

Example (cont.)

Suppose that at $t = t_{\text{max}}$ the object reaches its maximum height. Intuitively t_{max} should depends on three dimensional quantities g, R, V. On the other hand, we have $h'(t_{\text{max}}) = 0$; thus

$$0 = h'(t_{\text{max}}) = R\frac{d}{dt}\Big|_{t=t_{\text{max}}} \Phi\Big(\frac{tV}{R}, \frac{V^2}{gR}\Big) = V\frac{\partial \Phi}{\partial \pi_2}\Big(\frac{t_{\text{max}}V}{R}, \frac{V^2}{gR}\Big) .$$

The above relation "implies" that $\frac{t_{\text{max}}V}{R}$ is a function of $\frac{V^2}{gR}$; thus

$$\frac{t_{\text{max}}V}{R} = F\Big(\frac{V^2}{gR}\Big) .$$

§1.2 Characteristic Scales and Scaling

The "characteristic scales" are some **specific chosen values of dimensions** in the problem under consideration. The use of characteristic scales helps us reduce mathematical model into dimensionless form, and a good choice of characteristic scales sometimes can even simplify complicated models into simple ones.

Example

In this example we choose characteristic time scale t_c and length scale ℓ_c to recast the ODE

$$\frac{d^2h}{dt^2} = -\frac{R^2 g}{(R+h)^2}, \qquad h(0) = 0, \quad h'(0) = V.$$

We note that in pratice we know the values of R, g and V, so we should choose characteristic scales according to these values.

§1.2 Characteristic Scales and Scaling

The "characteristic scales" are some **specific chosen values of dimensions** in the problem under consideration. The use of characteristic scales helps us reduce mathematical model into dimensionless form, and a good choice of characteristic scales sometimes can even simplify complicated models into simple ones.

Example

In this example we choose characteristic time scale t_c and length scale ℓ_c to recast the ODE

$$\frac{d^2h}{dt^2} = -\frac{R^2 g}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(0) = V\,.$$

We note that in pratice we know the values of R, g and V, so we should choose characteristic scales according to these values.

§1.2 Characteristic Scales and Scaling

The "characteristic scales" are some **specific chosen values of dimensions** in the problem under consideration. The use of characteristic scales helps us reduce mathematical model into dimensionless form, and a good choice of characteristic scales sometimes can even simplify complicated models into simple ones.

Example

In this example we choose characteristic time scale t_c and length scale ℓ_c to recast the ODE

$$\frac{d^2h}{dt^2} = -\frac{R^2 g}{(R+h)^2}, \qquad h(0) = 0, \quad h'(0) = V.$$

We note that in pratice we know the values of R, g and V, so we should choose characteristic scales according to these values.

§1.2 Characteristic Scales and Scaling

The "characteristic scales" are some **specific chosen values of dimensions** in the problem under consideration. The use of characteristic scales helps us reduce mathematical model into dimensionless form, and a good choice of characteristic scales sometimes can even simplify complicated models into simple ones.

Example

In this example we choose characteristic time scale t_c and length scale ℓ_c to recast the ODE

$$\frac{d^2h}{dt^2} = -\frac{R^2 g}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(0) = V\,.$$

We note that in pratice we know the values of R, g and V, so we should choose characteristic scales according to these values.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Define the dimensionless time $\bar{t} = t/t_c$ and dimensionless height $\bar{h} = h/\ell_c$ $\left(\text{so that } \bar{h}(\bar{t}) = \frac{h(t_c\bar{t})}{\ell_c}\right)$. With the dimensionless time $\bar{t}$ and dimensionless height $\bar{h}$, the ODE above is equivalent to the dimensionless ODE

$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{t_c^2 g}{\ell_c}\frac{1}{(1+\frac{\ell_c}{R}\bar{h})^2}, \qquad \bar{h}(0) = 0, \quad \bar{h}'(0) = \frac{t_c V}{\ell_c}.$$

Three dimensioned quantities in the ODE are

$$[R] = L, \qquad [g] = LT^{-2} \qquad \text{and} \qquad [V] = LT^{-1}.$$

Therefore, three relevant time scales are $t_c = R/V$, $t_c = \sqrt{R/g}$ or $t_c = V/g$, and two relevant length scales are $\ell_c = R$ or $\ell_c = V^2/g$.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Define the dimensionless time $\bar{t} = t/t_c$ and dimensionless height $\bar{h} = h/\ell_c$ $\Big($so that $\bar{h}(\bar{t}) = \dfrac{h(t_c\bar{t})}{\ell_c}\Big)$. With the dimensionless time $\bar{t}$ and dimensionless height $\bar{h}$, the ODE above is equivalent to the dimensionless ODE

$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{t_c^2 g}{\ell_c}\frac{1}{(1+\frac{\ell_c}{R}\bar{h})^2}, \qquad \bar{h}(0) = 0, \quad \bar{h}'(0) = \frac{t_c V}{\ell_c}.$$

Three dimensioned quantities in the ODE are

$$[R] = L, \qquad [g] = LT^{-2} \qquad \text{and} \qquad [V] = LT^{-1}.$$

Therefore, three relevant time scales are $t_c = R/V$, $t_c = \sqrt{R/g}$ or $t_c = V/g$, and two relevant length scales are $\ell_c = R$ or $\ell_c = V^2/g$.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Define a dimensionless quantity $\epsilon = \dfrac{V^2}{gR}$. Using these characteristic scales, we reach at the following dimensionless problems:

1. Let $t_c = R/V$ and $\ell_c = R$. The scaled problem becomes
$$\epsilon \frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{(1+\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1\,.$$
2. Let $t_c = R/V$ and $\ell_c = V^2/g$. The scaled problem becomes
$$\epsilon^2 \frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{(1+\epsilon\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \frac{1}{\epsilon}\,.$$
3. Let $t_c = \sqrt{R/g}$ and $\ell_c = R$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{(1+\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \sqrt{\epsilon}\,.$$

§1.2 Characteristic Scales and Scaling

Example (cont.)

4. Let $t_c = \sqrt{R/g}$ and $\ell_c = V^2/g$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{\epsilon}\frac{1}{(1+\epsilon\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \frac{1}{\sqrt{\epsilon}}\,.$$

5. Let $t_c = V/g$ and $\ell_c = R$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\epsilon\frac{1}{(1+\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \epsilon\,.$$

6. Let $t_c = V/g$ and $\ell_c = V^2/g$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{(1+\epsilon\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1\,.$$

We note that these six ODEs are equivalent; however, we look for further simplification if the parameter ϵ is very small (or very large).

§1.2 Characteristic Scales and Scaling

Example (cont.)

4. Let $t_c = \sqrt{R/g}$ and $\ell_c = V^2/g$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{\epsilon}\frac{1}{(1+\epsilon\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \frac{1}{\sqrt{\epsilon}}\,.$$
5. Let $t_c = V/g$ and $\ell_c = R$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\epsilon\frac{1}{(1+\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = \epsilon\,.$$
6. Let $t_c = V/g$ and $\ell_c = V^2/g$. The scaled problem becomes
$$\frac{d^2\bar{h}}{d\bar{t}^2} = -\frac{1}{(1+\epsilon\bar{h})^2}\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1\,.$$

We note that these six ODEs are equivalent; however, we look for further simplification if the parameter ϵ is very small (or very large).

§1.2 Characteristic Scales and Scaling

Example (cont.)

Suppose that $\epsilon \ll 1$; that is, V^2 is much smaller than gR. In this case case, we are tempted to delete the terms involving ϵ (or simply setting $\epsilon = 0$) in the scaled problem. Then only case 3, 5, 6 provide meaningful models; however, only case 6 can provide a reasonable interpretation of the real phenomena: by setting $\epsilon = 0$, the scaled problem in case 6 becomes

$$\frac{d^2\bar{h}}{d\bar{t}^2} = (\approx) - 1\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1$$

whose solution is given by $\bar{h}(\bar{t}) = \bar{t} - \frac{1}{2}\bar{t}^2$. This implies that

$$h(t) = \ell_c \bar{h}\Big(\frac{t}{t_c}\Big) = \frac{V^2}{g}\Big(\frac{gt}{V} - \frac{1}{2}\frac{g^2t^2}{V^2}\Big) = Vt - \frac{1}{2}gt^2\,,$$

the formula for projectile motion that we learned in high school.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Suppose that $\epsilon \ll 1$; that is, V^2 is much smaller than gR. In this case case, we are tempted to delete the terms involving ϵ (or simply setting $\epsilon = 0$) in the scaled problem. Then only case 3, 5, 6 provide meaningful models; however, only case 6 can provide a reasonable interpretation of the real phenomena: by setting $\epsilon = 0$, the scaled problem in case 6 becomes

$$\frac{d^2\bar{h}}{d\bar{t}^2} = (\approx) - 1\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1$$

whose solution is given by $\bar{h}(\bar{t}) = \bar{t} - \frac{1}{2}\bar{t}^2$. This implies that

$$h(t) = \ell_c \bar{h}\Big(\frac{t}{t_c}\Big) = \frac{V^2}{g}\Big(\frac{gt}{V} - \frac{1}{2}\frac{g^2t^2}{V^2}\Big) = Vt - \frac{1}{2}gt^2\,,$$

the formula for projectile motion that we learned in high school.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Suppose that $\epsilon \ll 1$; that is, V^2 is much smaller than gR. In this case case, we are tempted to delete the terms involving ϵ (or simply setting $\epsilon = 0$) in the scaled problem. Then only case 3, 5, 6 provide meaningful models; however, only case 6 can provide a reasonable interpretation of the real phenomena: by setting $\epsilon = 0$, the scaled problem in case 6 becomes

$$\frac{d^2\bar{h}}{d\bar{t}^2} = (\approx) - 1\,, \qquad \bar{h}(0) = 0\,, \quad \bar{h}'(0) = 1$$

whose solution is given by $\bar{h}(\bar{t}) = \bar{t} - \frac{1}{2}\bar{t}^2$. This implies that

$$h(t) = \ell_c \bar{h}\Big(\frac{t}{t_c}\Big) = \frac{V^2}{g}\Big(\frac{gt}{V} - \frac{1}{2}\frac{g^2 t^2}{V^2}\Big) = Vt - \frac{1}{2}gt^2\,,$$

the formula for projectile motion that we learned in high school.

§1.2 Characteristic Scales and Scaling

Example (cont.)

The lesson of the example: To simplify a complicated model, one needs to be very careful about choosing characteristic scales.

The reason why $t_c = V/g$ and $\ell_c = V^2/g$ is the correct characteristic scale when $\epsilon \ll 1$?

When V is very small, we expect that the gravity acceleration is always almost g $\left(\text{instead of } \dfrac{GM}{(R+h)^2}\right)$. If the gravity acceleration is g, the object (with mass m) takes $\dfrac{V}{g}$ time to reach its maximum height $\dfrac{V^2}{2g}$; thus $t_c = \dfrac{V}{g}$ is a good choice of the characteristic time scale and $\ell_c = \dfrac{V^2}{g}$ is a good choice of the characteristic length scale.

§1.2 Characteristic Scales and Scaling

Example (cont.)

The lesson of the example: To simplify a complicated model, one needs to be very careful about choosing characteristic scales.

The reason why $t_c = V/g$ and $\ell_c = V^2/g$ is the correct characteristic scale when $\epsilon \ll 1$?

When V is very small, we expect that the gravity acceleration is always almost g $\Big($instead of $\dfrac{GM}{(R+h)^2}\Big)$. If the gravity acceleration is g, the object (with mass m) takes $\dfrac{V}{g}$ time to reach its maximum height $\dfrac{V^2}{2g}$; thus $t_c = \dfrac{V}{g}$ is a good choice of the characteristic time scale and $\ell_c = \dfrac{V^2}{g}$ is a good choice of the characteristic length scale.

§1.2 Characteristic Scales and Scaling

Example (cont.)

The lesson of the example: To simplify a complicated model, one needs to be very careful about choosing characteristic scales.

The reason why $t_c = V/g$ and $\ell_c = V^2/g$ is the correct characteristic scale when $\epsilon \ll 1$?

When V is very small, we expect that the gravity acceleration is always almost g $\left(\text{instead of } \dfrac{GM}{(R+h)^2}\right)$. If the gravity acceleration is g, the object (with mass m) takes $\dfrac{V}{g}$ time to reach its maximum height $\dfrac{V^2}{2g}$; thus $t_c = \dfrac{V}{g}$ is a good choice of the characteristic time scale and $\ell_c = \dfrac{V^2}{g}$ is a good choice of the characteristic length scale.

§1.2 Characteristic Scales and Scaling

Example

Let $p = p(t)$ denote the population of an animal species located in a fixed region at time t. The simplest model of population growth is the classic **Malthus model** which states that the rate of change of the population $\frac{dp}{dt}$ is proportional to the population p, or equivalently

$$\frac{dp}{dt} = rp\,,$$

where r is the growth rate, given in dimensions of inverse-time. A more reasonable model, called the **logistics model**, is given by

$$\frac{dp}{dt} = rp\left(1 - \frac{p}{K}\right),$$

where $K > 0$ is called the *carring capacity* (with dimension of population).

§1.2 Characteristic Scales and Scaling

Example

Let $p = p(t)$ denote the population of an animal species located in a fixed region at time t. The simplest model of population growth is the classic **Malthus model** which states that the rate of change of the population $\dfrac{dp}{dt}$ is proportional to the population p, or equivalently

$$\frac{dp}{dt} = rp\,,$$

where r is the growth rate, given in dimensions of inverse-time. A more reasonable model, called the **logistics model**, is given by

$$\frac{dp}{dt} = rp\big(1 - \frac{p}{K}\big)\,,$$

where $K > 0$ is called the *carring capacity* (with dimension of population).

§1.2 Characteristic Scales and Scaling

Example (cont.)

To complete the system, we need to impose an initial condition so that the complete equation is

$$\frac{dp}{dt} = rp\Big(1 - \frac{p}{K}\Big)\,, \qquad p(0) = p_0\,.$$

In the logistic model above, the dimension of t is time, and the dimension of population is named "population". Let t_c and p_c denote the characteristic time scale and the characteristic population scale, respectively. Introducing the dimensionless time $\bar{t} = t/t_c$ and the dimensionless population $\bar{p} = p/p_c$ $\Big($so that $\bar{p}(\bar{t}) = \dfrac{p(t_c\bar{t})}{p_c}\Big)$, we obtain the following scaled problem

$$\frac{d\bar{p}}{d\bar{t}} = rt_c\,\bar{p}\Big(1 - \frac{p_c}{K}\bar{p}\Big)\,, \qquad \bar{p}(0) = \frac{p_0}{p_c}\,.$$

§1.2 Characteristic Scales and Scaling

Example (cont.)

To complete the system, we need to impose an initial condition so that the complete equation is

$$\frac{dp}{dt} = rp\left(1 - \frac{p}{K}\right), \qquad p(0) = p_0 .$$

In the logistic model above, the dimension of t is time, and the dimension of population is named "population". Let t_c and p_c denote the characteristic time scale and the characteristic population scale, respectively. Introducing the dimensionless time $\bar{t} = t/t_c$ and the dimensionless population $\bar{p} = p/p_c$ $\left(\text{so that } \bar{p}(\bar{t}) = \frac{p(t_c\bar{t})}{p_c}\right)$, we obtain the following scaled problem

$$\frac{d\bar{p}}{d\bar{t}} = rt_c\,\bar{p}\left(1 - \frac{p_c}{K}\bar{p}\right), \qquad \bar{p}(0) = \frac{p_0}{p_c} .$$

§1.2 Characteristic Scales and Scaling

Example (cont.)

Apparently, we should choose the characteristic time scale $t_c = 1/r$. On the other hand, two characteristic population scales can be chosen: $p_c = K$ or $p_c = p_0$. Moreover, there is a dimensionless quantity $\epsilon = \dfrac{p_0}{K}$ in the system.

1. $p_c = K$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \bar{p}), \qquad \bar{p}(0) = \epsilon.$$
2. $p_c = p_0$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \epsilon\bar{p}), \qquad \bar{p}(0) = 1.$$

If $\epsilon \ll 1$, we set $\epsilon = 0$ in the scaled problem and find that only case 2 provides a reasonable interpretation of the real phenomena.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Apparently, we should choose the characteristic time scale $t_c = 1/r$. On the other hand, two characteristic population scales can be chosen: $p_c = K$ or $p_c = p_0$. Moreover, there is a dimensionless quantity $\epsilon = \dfrac{p_0}{K}$ in the system.

1. $p_c = K$: the scaled problem becomes

$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \bar{p}), \qquad \bar{p}(0) = \epsilon.$$

2. $p_c = p_0$: the scaled problem becomes

$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \epsilon\bar{p}), \qquad \bar{p}(0) = 1.$$

If $\epsilon \ll 1$, we set $\epsilon = 0$ in the scaled problem and find that only case 2 provides a reasonable interpretation of the real phenomena.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Apparently, we should choose the characteristic time scale $t_c = 1/r$. On the other hand, two characteristic population scales can be chosen: $p_c = K$ or $p_c = p_0$. Moreover, there is a dimensionless quantity $\epsilon = \frac{p_0}{K}$ in the system.

1. $p_c = K$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \bar{p}), \qquad \bar{p}(0) = \epsilon.$$
2. $p_c = p_0$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \epsilon\bar{p}), \qquad \bar{p}(0) = 1.$$

If $\epsilon \ll 1$, we set $\epsilon = 0$ in the scaled problem and find that only case 2 provides a reasonable interpretation of the real phenomena.

§1.2 Characteristic Scales and Scaling

Example (cont.)

Apparently, we should choose the characteristic time scale $t_c = 1/r$. On the other hand, two characteristic population scales can be chosen: $p_c = K$ or $p_c = p_0$. Moreover, there is a dimensionless quantity $\epsilon = \dfrac{p_0}{K}$ in the system.

1. $p_c = K$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \bar{p})\,, \qquad \bar{p}(0) = \epsilon\,.$$
2. $p_c = p_0$: the scaled problem becomes
$$\frac{d\bar{p}}{d\bar{t}} = \bar{p}(1 - \epsilon\bar{p})\,, \qquad \bar{p}(0) = 1\,.$$

If $\epsilon \ll 1$, we set $\epsilon = 0$ in the scaled problem and find that only case 2 provides a reasonable interpretation of the real phenomena.

Chapter 2. Ordinary Differential Equations（常微分方程）

Introduction

Definition

A differential equation is a mathematical equation that relates some unknown function with its derivatives. The unknown functions in a differential equations are sometimes called ***dependent variables***, and the variables which the derivatives of the unknown functions are taken with respect to are sometimes called the ***independent variables***. A differential equation is called an ***ordinary differential equation*** (ODE) if it contains an unknown function of one independent variable and its derivatives. A differential equation is called a ***partial differential equation*** (PDE) if it contains unknown multi-variable functions and their partial derivatives.

We note that in most of the mathematical ODE models, the independent variable is the time variable t or the spatial variable x.

Introduction

Definition

A differential equation is a mathematical equation that relates some unknown function with its derivatives. The unknown functions in a differential equations are sometimes called ***dependent variables***, and the variables which the derivatives of the unknown functions are taken with respect to are sometimes called the ***independent variables***. A differential equation is called an ***ordinary differential equation*** (ODE) if it contains an unknown function of one independent variable and its derivatives. A differential equation is called a ***partial differential equation*** (PDE) if it contains unknown multi-variable functions and their partial derivatives.

We note that in most of the mathematical ODE models, the independent variable is the time variable t or the spatial variable x.

Introduction

Definition

A differential equation is a mathematical equation that relates some unknown function with its derivatives. The unknown functions in a differential equations are sometimes called ***dependent variables***, and the variables which the derivatives of the unknown functions are taken with respect to are sometimes called the ***independent variables***. A differential equation is called an ***ordinary differential equation*** (ODE) if it contains an unknown function of one independent variable and its derivatives. A differential equation is called a ***partial differential equation*** (PDE) if it contains unknown multi-variable functions and their partial derivatives.

We note that in most of the mathematical ODE models, the independent variable is the time variable t or the spatial variable x.

Introduction

Definition

A differential equation is a mathematical equation that relates some unknown function with its derivatives. The unknown functions in a differential equations are sometimes called ***dependent variables***, and the variables which the derivatives of the unknown functions are taken with respect to are sometimes called the ***independent variables***. A differential equation is called an ***ordinary differential equation*** (ODE) if it contains an unknown function of one independent variable and its derivatives. A differential equation is called a ***partial differential equation*** (PDE) if it contains unknown multi-variable functions and their partial derivatives.

We note that in most of the mathematical ODE models, the independent variable is the time variable t or the spatial variable x.

Introduction

Definition

A differential equation is a mathematical equation that relates some unknown function with its derivatives. The unknown functions in a differential equations are sometimes called ***dependent variables***, and the variables which the derivatives of the unknown functions are taken with respect to are sometimes called the ***independent variables***. A differential equation is called an ***ordinary differential equation*** (ODE) if it contains an unknown function of one independent variable and its derivatives. A differential equation is called a ***partial differential equation*** (PDE) if it contains unknown multi-variable functions and their partial derivatives.

We note that in most of the mathematical ODE models, the independent variable is the time variable t or the spatial variable x.

Introduction

Definition

The ***order*** of a differential equation is the order of the highest-order derivatives present in the equation. A differential equation of order 1 is called first order, order 2 second order, etc.

Definition

The ordinary differential equation

$$F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) = 0 \tag{4}$$

is said to be ***linear*** if

$$\begin{aligned} &F(t, cy, cy', \cdots, cy^{(n-1)}, cy^{(n)}) - F(t, 0, 0, \cdots, 0) \\ &\qquad = c\big[F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) - F(t, 0, 0, \cdots, 0)\big] \end{aligned} \quad \forall\, c \in \mathbb{R}\,.$$

The ODE (4) is said to be ***nonlinear*** if it is not linear.

Introduction

Definition

The ***order*** of a differential equation is the order of the highest-order derivatives present in the equation. A differential equation of order 1 is called first order, order 2 second order, etc.

Definition

The ordinary differential equation

$$F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) = 0 \tag{4}$$

is said to be ***linear*** if

$$\begin{aligned} &F(t, cy, cy', \cdots, cy^{(n-1)}, cy^{(n)}) - F(t, 0, 0, \cdots, 0) \\ &\quad = c\big[F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) - F(t, 0, 0, \cdots, 0)\big] \end{aligned} \qquad \forall\, c \in \mathbb{R}\,.$$

The ODE (4) is said to be ***nonlinear*** if it is not linear.

Introduction

Remark: It is commonly assumed that an ordinary differential equation of order n

$$F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) = 0 \quad \text{(if the independent variable is } t\text{)}$$

can be written as

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}).$$

Moreover, given a differential equation above, we can define a vector-valued function $\boldsymbol{z} = (y, y', y'', \cdots, y^{(n-1)})^{\mathrm{T}}$ and write the ODE above as

$$\boldsymbol{z}'(t) = \frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} = \begin{bmatrix} z_2 \\ z_3 \\ \vdots \\ z_n \\ f(t, z_1, z_2, \cdots, z_n) \end{bmatrix} = \boldsymbol{f}(t, \boldsymbol{z})$$

which is a first order ODE with a vector-valued unknown.

Introduction

Remark: It is commonly assumed that an ordinary differential equation of order n

$$F(t, y, y', \cdots, y^{(n-1)}, y^{(n)}) = 0 \quad \text{(if the independent variable is } t\text{)}$$

can be written as

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}).$$

Moreover, given a differential equation above, we can define a vector-valued function $\boldsymbol{z} = \left(y, y', y'', \cdots, y^{(n-1)}\right)^{\mathrm{T}}$ and write the ODE above as

$$\boldsymbol{z}'(t) = \frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ z_n \end{bmatrix} = \begin{bmatrix} z_2 \\ z_3 \\ \vdots \\ z_n \\ f(t, z_1, z_2, \cdots, z_n) \end{bmatrix} = \boldsymbol{f}(t, \boldsymbol{z})$$

which is a first order ODE with a vector-valued unknown.

§2.1 Initial Value Problems

In this section, the ODE under consideration is always written as

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)})\,.$$

Definition

An ***initial value problem*** (**IVP**) is a (system of) differential equation

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}) \tag{5a}$$

equipped with an initial condition

$$y(t_0) = y_0,\ y'(t_0) = y_1,\ y''(t_0) = y_2,\ \cdots\ y^{(n-1)}(t_0) = y_{n-1}\,, \tag{5b}$$

where t_0 is a given point/time, and $y_0, y_1, \cdots, y_{n-1}$ are given numbers. A solution to the IVP (5) is a function y defined on an open interval I so that $t_0 \in I$ and (5) is satisfied.

§2.1 Initial Value Problems

In this section, the ODE under consideration is always written as

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}) \,.$$

Definition

An ***initial value problem*** (**IVP**) is a (system of) differential equation

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}) \tag{5a}$$

equipped with an initial condition

$$y(t_0) = y_0, \;\; y'(t_0) = y_1, \;\; y''(t_0) = y_2, \;\; \cdots \;\; y^{(n-1)}(t_0) = y_{n-1} \,, \tag{5b}$$

where t_0 is a given point/time, and $y_0, y_1, \cdots, y_{n-1}$ are given numbers. A solution to the IVP (5) is a function y defined on an open interval I so that $t_0 \in I$ and (5) is satisfied.

§2.1 Initial Value Problems

In this section, the ODE under consideration is always written as

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}) \,.$$

Definition

An ***initial value problem*** (**IVP**) is a (system of) differential equation

$$y^{(n)}(t) = f(t, y, y', \cdots, y^{(n-2)}, y^{(n-1)}) \tag{5a}$$

equipped with an initial condition

$$y(t_0) = y_0, \;\; y'(t_0) = y_1, \;\; y''(t_0) = y_2, \;\; \cdots \;\; y^{(n-1)}(t_0) = y_{n-1} \,, \tag{5b}$$

where t_0 is a given point/time, and $y_0, y_1, \cdots, y_{n-1}$ are given numbers. A solution to the IVP (5) is a function y defined on an open interval I so that $t_0 \in I$ and (5) is satisfied.

§2.1 Initial Value Problems

Example

Before we have talked about the Malthus model

$$\frac{dp}{dt} = rp\,, \qquad p(0) = p_0$$

for the growth of population. In this model, the growth rate is assumed to be positive. However, the same differential equation can be used to model the decay of radioactive substance such as plutonium (鈽). If $p(t)$ is the total amount of such kind of substance at time t, the rate of change of the amount of the plutonium $\frac{dp}{dt}$ is proportional to the total amount p, except that the "growth" rate r is negative. In such a case, r is called the **decay rate**.

The model has linear ODE and usually is called **linear model** (for population growrth or decay of radioactive substance).

§2.1 Initial Value Problems

Example

Before we have talked about the Malthus model

$$\frac{dp}{dt} = rp\,, \qquad p(0) = p_0$$

for the growth of population. In this model, the growth rate is assumed to be positive. However, the same differential equation can be used to model the decay of radioactive substance such as plutonium (鈽). If $p(t)$ is the total amount of such kind of substance at time t, the rate of change of the amount of the plutonium $\frac{dp}{dt}$ is proportional to the total amount p, except that the "growth" rate r is negative. In such a case, r is called the **decay rate**.

The model has linear ODE and usually is called **linear model** (for population growrth or decay of radioactive substance).

§2.1 Initial Value Problems

Example

Before we have talked about the Malthus model

$$\frac{dp}{dt} = rp\,, \qquad p(0) = p_0$$

for the growth of population. In this model, the growth rate is assumed to be positive. However, the same differential equation can be used to model the decay of radioactive substance such as plutonium (鈽). If $p(t)$ is the total amount of such kind of substance at time t, the rate of change of the amount of the plutonium $\frac{dp}{dt}$ is proportional to the total amount p, except that the "growth" rate r is negative. In such a case, r is called the **decay rate**.

The model has linear ODE and usually is called **linear model** (for population growrth or decay of radioactive substance).

§2.1 Initial Value Problems

Example (Spring-mass system with or without Friction)

Consider an object of mass m attached to a spring with Hook's constant k. Let $x(t)$ denote the signed distance between the object and the equilibrium position at time t. If there is **no friction**, by the Newton second law of motion we find that x obeys the ODE

$$m\ddot{x} = -kx.$$

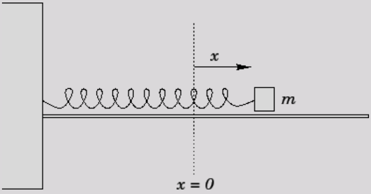


Figure 1: The spring-mass system

§2.1 Initial Value Problems

Example (Spring-mass system with or without Friction - cont.)

When the friction is under consideration, by the fact that the friction is proportional to the velocity, we find that

$$m\ddot{x} = -kx - r\dot{x}.$$

If in addition some external force $f(t)$ are exerted on the mass, the model becomes

$$m\ddot{x} = -kx - r\dot{x} + f.$$

If the initial position and the initial velocity of the object is $x(0) = x_0$ and $x'(0) = x_1$, then $x(t)$ satisfies the IVP

$$m\ddot{x} = -kx - r\dot{x} + f, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,. \tag{6}$$

The ODE in (6) is linear since the function

$$F(t, x, \dot{x}, \ddot{x}) = m\ddot{x} + r\dot{x} + kx - f(t)$$

satisfies $F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0) = c\big[F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0)\big]$.

§2.1 Initial Value Problems

Example (Spring-mass system with or without Friction - cont.)

When the friction is under consideration, by the fact that the friction is proportional to the velocity, we find that

$$m\ddot{x} = -kx - r\dot{x}.$$

If in addition some external force $f(t)$ are exerted on the mass, the model becomes

$$m\ddot{x} = -kx - r\dot{x} + f.$$

If the initial position and the initial velocity of the object is $x(0) = x_0$ and $x'(0) = x_1$, then $x(t)$ satisfies the IVP

$$m\ddot{x} = -kx - r\dot{x} + f, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,. \tag{6}$$

The ODE in (6) is linear since the function

$$F(t, x, \dot{x}, \ddot{x}) = m\ddot{x} + r\dot{x} + kx - f(t)$$

satisfies $F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0) = c\big[F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0)\big]$.

§2.1 Initial Value Problems

Example (Spring-mass system with or without Friction - cont.)

When the friction is under consideration, by the fact that the friction is proportional to the velocity, we find that

$$m\ddot{x} = -kx - r\dot{x}.$$

If in addition some external force $f(t)$ are exerted on the mass, the model becomes

$$m\ddot{x} = -kx - r\dot{x} + f.$$

If the initial position and the initial velocity of the object is $x(0) = x_0$ and $x'(0) = x_1$, then $x(t)$ satisfies the IVP

$$m\ddot{x} = -kx - r\dot{x} + f, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,. \tag{6}$$

The ODE in (6) is linear since the function

$$F(t, x, \dot{x}, \ddot{x}) = m\ddot{x} + r\dot{x} + kx - f(t)$$

satisfies $F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0) = c\big[F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0)\big]$.

§2.1 Initial Value Problems

Example (Spring-mass system with or without Friction - cont.)

When the friction is under consideration, by the fact that the friction is proportional to the velocity, we find that

$$m\ddot{x} = -kx - r\dot{x}.$$

If in addition some external force $f(t)$ are exerted on the mass, the model becomes

$$m\ddot{x} = -kx - r\dot{x} + f.$$

If the initial position and the initial velocity of the object is $x(0) = x_0$ and $x'(0) = x_1$, then $x(t)$ satisfies the IVP

$$m\ddot{x} = -kx - r\dot{x} + f, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,. \tag{6}$$

The ODE in (6) is linear since the function

$$F(t, x, \dot{x}, \ddot{x}) = m\ddot{x} + r\dot{x} + kx - f(t)$$

satisfies $F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0) = c\big[F(t, cx, c\dot{x}, c\ddot{x}) - F(t, 0, 0, 0)\big]$.

§2.1 Initial Value Problems

Example

In this example we study a closed circuit shown in the figure below.

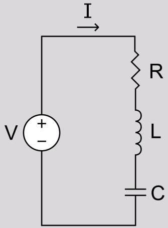


Figure 2: A closed circuit

In the figure above, V is the voltage (電壓) source powering the circuit, I is the current (電流) admitted through the circuit, R is the effective resistance (電阻) of the combined load, source, and components, L is the inductance of the inductor (電感) component, and C the capacitance of the capacitor (電容) component.

§2.1 Initial Value Problems

Example (cont.)

An electric current (電流) is the rate of flow of electric charge (電荷) past a point or region:

$$I(t) = \frac{dQ}{dt}.$$

A capacitor (電容) consists of two conductors separated by a non-conductive region which can either be a vacuum or an electrical insulator material known as a dielectric (介電質). From Coulomb's law (庫倫定律) a charge on one conductor will exert a force on the charge carriers within the other conductor, attracting opposite polarity charge and repelling like polarity charges, thus an opposite polarity charge will be induced on the surface of the other conductor. The conductors thus hold equal and opposite charges on their facing surfaces, and the dielectric develops an electric field.

§2.1 Initial Value Problems

Example (cont.)

An electric current (電流) is the rate of flow of electric charge (電荷) past a point or region:

$$\mathrm{I}(t) = \frac{d\mathrm{Q}}{dt}\,.$$

A capacitor (電容) consists of two conductors separated by a non-conductive region which can either be a vacuum or an electrical insulator material known as a dielectric (介電質). From Coulomb's law (庫倫定律) a charge on one conductor will exert a force on the charge carriers within the other conductor, attracting opposite polarity charge and repelling like polarity charges, thus an opposite polarity charge will be induced on the surface of the other conductor. The conductors thus hold equal and opposite charges on their facing surfaces, and the dielectric develops an electric field.

§2.1 Initial Value Problems

Example (cont.)

An electric current (電流) is the rate of flow of electric charge (電荷) past a point or region:

$$\mathrm{I}(t) = \frac{d\mathrm{Q}}{dt} .$$

A capacitor (電容) consists of two conductors separated by a non-conductive region which can either be a vacuum or an electrical insulator material known as a dielectric (介電質). From Coulomb's law (庫倫定律) a charge on one conductor will exert a force on the charge carriers within the other conductor, attracting opposite polarity charge and repelling like polarity charges, thus an opposite polarity charge will be induced on the surface of the other conductor. The conductors thus hold equal and opposite charges on their facing surfaces, and the dielectric develops an electric field.

§2.1 Initial Value Problems

Example (cont.)

An electric current (電流) is the rate of flow of electric charge (電荷) past a point or region:

$$\mathrm{I}(t) = \frac{d\mathrm{Q}}{dt}.$$

A capacitor (電容) consists of two conductors separated by a non-conductive region which can either be a vacuum or an electrical insulator material known as a dielectric (介電質). From Coulomb's law (庫倫定律) a charge on one conductor will exert a force on the charge carriers within the other conductor, attracting opposite polarity charge and repelling like polarity charges, thus an opposite polarity charge will be induced on the surface of the other conductor. The conductors thus hold equal and opposite charges on their facing surfaces, and the dielectric develops an electric field.

§2.1 Initial Value Problems

Example (cont.)

An ideal capacitor is characterized by a constant capacitance C which is defined as the ratio of the positive or negative charge Q on each conductor to the voltage V between them:

$$\mathrm{C} = \frac{\mathrm{Q}}{\mathrm{V}} \qquad \text{or} \qquad \mathrm{Q} = \mathrm{C}\mathrm{V}\,.$$

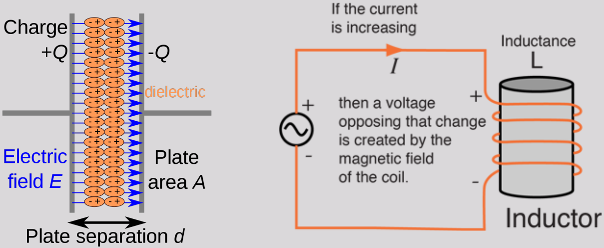


Figure 3: Left: capacitor, Right: inductance

§2.1 Initial Value Problems

Example (cont.)

Inductance (電感) is the tendency of an electrical conductor to oppose a change in the electric current flowing through it, and is defined as the ratio of the induced voltage to the rate of change of current causing it:

$$\mathrm{V}(t) = \mathrm{L}\frac{d\mathrm{I}}{dt}\,.$$

The design of inductance is based on Lenz's law (冷次定律) which states that "the current induced in a circuit due to a change in a magnetic field is directed to oppose the change in flux and to exert a mechanical force which opposes the motion" (磁通量的改變而產生的感應電流，其方向為抗拒磁通量改變的方向).

§2.1 Initial Value Problems

Example (cont.)

Inductance (電感) is the tendency of an electrical conductor to oppose a change in the electric current flowing through it, and is defined as the ratio of the induced voltage to the rate of change of current causing it:

$$\mathrm{V}(t) = \mathrm{L}\frac{d\mathrm{I}}{dt}\,.$$

The design of inductance is based on Lenz's law (冷次定律) which states that "the current induced in a circuit due to a change in a magnetic field is directed to oppose the change in flux and to exert a mechanical force which opposes the motion" (磁通量的改變而產生的感應電流，其方向為抗拒磁通量改變的方向).

§2.1 Initial Value Problems

Example (cont.)

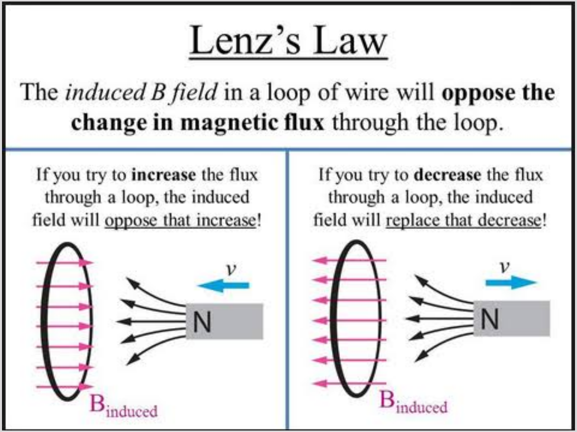


Figure 4: 冷次定律示意圖

§2.1 Initial Value Problems

Example (cont.)

In a closed circuit (a circuit without interruption, providing a continuous path through which a current can flow) shown in Figure 2, one has

$$\mathrm{V}(t) = \mathrm{I}(t)\mathrm{R} + \mathrm{L}\frac{d\mathrm{I}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q}(t)\,.$$

By the definition of I, we find that Q satisfies

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,.$$

To complete the model, initial conditions have to be imposed so that we have

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,, \qquad \mathrm{Q}(t_0) = \mathrm{Q}_0\,, \quad \mathrm{Q}'(t_0) = \mathrm{I}_0\,.$$

We note that the IVP above is essentially the same as the IVP (6) derived from studying the spring-mass system.

§2.1 Initial Value Problems

Example (cont.)

In a closed circuit (a circuit without interruption, providing a continuous path through which a current can flow) shown in Figure 2, one has

$$\mathrm{V}(t) = \mathrm{I}(t)\mathrm{R} + \mathrm{L}\frac{d\mathrm{I}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q}(t)\,.$$

By the definition of I, we find that Q satisfies

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,.$$

To complete the model, initial conditions have to be imposed so that we have

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,, \qquad \mathrm{Q}(t_0) = \mathrm{Q}_0\,, \quad \mathrm{Q}'(t_0) = \mathrm{I}_0\,.$$

We note that the IVP above is essentially the same as the IVP (6) derived from studying the spring-mass system.

§2.1 Initial Value Problems

Example (cont.)

In a closed circuit (a circuit without interruption, providing a continuous path through which a current can flow) shown in Figure 2, one has

$$\mathrm{V}(t) = \mathrm{I}(t)\mathrm{R} + \mathrm{L}\frac{d\mathrm{I}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q}(t)\,.$$

By the definition of I, we find that Q satisfies

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,.$$

To complete the model, initial conditions have to be imposed so that we have

$$\mathrm{L}\frac{d^2\mathrm{Q}}{dt^2} + \mathrm{R}\frac{d\mathrm{Q}}{dt} + \frac{1}{\mathrm{C}}\mathrm{Q} = \mathrm{V}\,, \qquad \mathrm{Q}(t_0) = \mathrm{Q}_0\,, \quad \mathrm{Q}'(t_0) = \mathrm{I}_0\,.$$

We note that the IVP above is essentially the same as the IVP (6) derived from studying the spring-mass system.

§2.1 Initial Value Problems

Example (Oscillating pendulum)

A simple pendulum consists of a mass m hanging from a string of length L and fixed at a pivot point P. When displaced to an initial angle and released, the pendulum will swing back and forth with periodic motion.

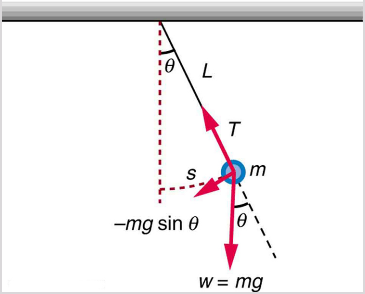


Figure 5: A simple pendulum system

§2.1 Initial Value Problems

Example (Oscillating pendulum - cont.)

Let $\theta(t)$ denote the angle, measured from the vertical dashed line (see Figure 5), at time t. By Newton's second law,

$$mL\ddot{\theta} = -mg\sin\theta\,, \qquad \theta(0) = \theta_0\,, \quad \theta'(0) = \omega_0\,.$$

The ODE in the IVP above is a **nonlinear** ODE.

When the angle of oscillation is very small; that is, $\theta \approx 0$, then by the fact that $\lim\limits_{\theta \to 0} \dfrac{\sin\theta}{\theta} = 1$ we find that in this case

$$mL\ddot{\theta} \approx -mg\theta\,;$$

thus we obtain a simplified model for simple pendulum

$$mL\ddot{\theta} = -mg\theta\,, \qquad \theta(0) = \theta_0\,, \quad \theta'(0) = \omega_0\,.$$

§2.1 Initial Value Problems

Example (Oscillating pendulum - cont.)

Let $\theta(t)$ denote the angle, measured from the vertical dashed line (see Figure 5), at time t. By Newton's second law,

$$mL\ddot{\theta} = -mg\sin\theta\,, \qquad \theta(0) = \theta_0\,, \quad \theta'(0) = \omega_0\,.$$

The ODE in the IVP above is a **nonlinear** ODE.

When the angle of oscillation is very small; that is, $\theta \approx 0$, then by the fact that $\lim\limits_{\theta\to 0}\dfrac{\sin\theta}{\theta} = 1$ we find that in this case

$$mL\ddot{\theta} \approx -mg\theta\,;$$

thus we obtain a simplified model for simple pendulum

$$mL\ddot{\theta} = -mg\theta\,, \qquad \theta(0) = \theta_0\,, \quad \theta'(0) = \omega_0\,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model)

Suppose that two different species of animals interact within the same environment or ecosystem, and suppose further that the first species eats only vegetation and the second eats only the first species. In other words, one species is a predator (掠食者) and the other is a prey (獵物).

Let $p(t)$ and $q(t)$ denote, respectively, the populations of the prey and the predator. If there is no prey, then the population of the predator should decrease/decay and follows

$$\frac{dq}{dt} = -\beta q\,, \qquad \beta > 0\,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model)

Suppose that two different species of animals interact within the same environment or ecosystem, and suppose further that the first species eats only vegetation and the second eats only the first species. In other words, one species is a predator (掠食者) and the other is a prey (獵物).

Let $p(t)$ and $q(t)$ denote, respectively, the populations of the prey and the predator. If there is no prey, then the population of the predator should decrease/decay and follows

$$\frac{dq}{dt} = -\beta q\,, \qquad \beta > 0\,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

When preys are present in the environment, it seems reasonable that the number of encounters or interactions between these two species per unit time is jointly proportional to their populations p and q; that is, proportional to the product pq. Thus when preys are present, the predator are added to the system at a rate δpq, $\delta > 0$. In other words, the population of q should follows

$$\frac{dq}{dt} = -\beta q + \delta pq\,, \qquad \beta, \delta > 0\,.$$

Here the growth rate of the population of the predator is $(\delta p - \beta)$ since

$$\frac{dq}{dt} = (\delta p - \beta) q\,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

When preys are present in the environment, it seems reasonable that the number of encounters or interactions between these two species per unit time is jointly proportional to their populations p and q; that is, proportional to the product pq. Thus when preys are present, the predator are added to the system at a rate δpq, $\delta > 0$. In other words, the population of q should follows

$$\frac{dq}{dt} = -\beta q + \delta pq\,, \qquad \beta, \delta > 0\,.$$

Here the growth rate of the population of the predator is $(\delta p - \beta)$ since

$$\frac{dq}{dt} = (\delta p - \beta) q\,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

On the other hand, if there is no predator, the population of the prey should follow the Malthus model (assuming that the supply of food is always sufficient); however, the population of the prey will decrease by the rate at which the preys are consumed during their encounters with the predator; thus

$$\frac{dp}{dt} = \alpha p - \gamma pq \,, \qquad \alpha, \gamma > 0 \,.$$

Therefore, we obtain the ***predator-prey model*** (or the ***Lotka-Volterra model***):

$$\frac{dp}{dt} = \alpha p - \gamma pq = (\alpha - \gamma q)p \,,$$
$$\frac{dq}{dt} = -\beta q + \delta pq = (-\beta + \delta p)q \,.$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

On the other hand, if there is no predator, the population of the prey should follow the Malthus model (assuming that the supply of food is always sufficient); however, the population of the prey will decrease by the rate at which the preys are consumed during their encounters with the predator; thus

$$\frac{dp}{dt} = \alpha p - \gamma pq\,, \qquad \alpha, \gamma > 0\,.$$

Therefore, we obtain the ***predator-prey model*** (or the ***Lotka-Volterra model***):

$$\begin{aligned} \frac{dp}{dt} &= \alpha p - \gamma pq = (\alpha - \gamma q)p\,, \\ \frac{dq}{dt} &= -\beta q + \delta pq = (-\beta + \delta p)q\,. \end{aligned}$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

An initial condition $p(0) = p_0$, $q(0) = q_0$ can be imposed so that it becomes an IVP.

The Lotka-Volterra model is **nonlinear** since by letting $\boldsymbol{z} = [p, q]^{\mathrm{T}}$, we can rewrite the model as

$$\dot{\boldsymbol{z}} = \boldsymbol{f}(t, \boldsymbol{z}) = \begin{bmatrix} \alpha & 0 \\ 0 & -\beta \end{bmatrix} \boldsymbol{z} + \begin{bmatrix} -\gamma z_1 z_2 \\ \delta z_1 z_2 \end{bmatrix}$$

which shows that $F(t, c\boldsymbol{z}, c\dot{\boldsymbol{z}}) - F(t, 0, 0) \neq c\big[F(t, \boldsymbol{z}, \dot{\boldsymbol{z}}) - F(t, 0, 0)\big]$ if $c \neq 1$, where

$$F(t, \boldsymbol{z}, \dot{\boldsymbol{z}}) = \dot{\boldsymbol{z}} - \boldsymbol{f}(t, \boldsymbol{z}) .$$

§2.1 Initial Value Problems

Example (Lotka-Volterra or Prey-Predator model - cont.)

An initial condition $p(0) = p_0$, $q(0) = q_0$ can be imposed so that it becomes an IVP.

The Lotka-Volterra model is **nonlinear** since by letting $\boldsymbol{z} = [p, q]^{\mathrm{T}}$, we can rewrite the model as

$$\dot{\boldsymbol{z}} = \boldsymbol{f}(t, \boldsymbol{z}) = \begin{bmatrix} \alpha & 0 \\ 0 & -\beta \end{bmatrix} \boldsymbol{z} + \begin{bmatrix} -\gamma z_1 z_2 \\ \delta z_1 z_2 \end{bmatrix}$$

which shows that $F(t, c\boldsymbol{z}, c\dot{\boldsymbol{z}}) - F(t, 0, 0) \neq c\big[F(t, \boldsymbol{z}, \dot{\boldsymbol{z}}) - F(t, 0, 0)\big]$ if $c \neq 1$, where

$$F(t, \boldsymbol{z}, \dot{\boldsymbol{z}}) = \dot{\boldsymbol{z}} - \boldsymbol{f}(t, \boldsymbol{z}) \,.$$

§2.1 Initial Value Problems

Example (SIR model for spread of diseases)

This example presents a classical model, called the SIR model, of disease transmission within a population. The total population is divided into three groups: **individuals susceptible to disease** (易感者), **infected individuals** (染病者), and **"removed" individuals** (移出者). The removed class counts those individuals who are not infected and not susceptible; in other words, immune, quarantined, or dead. Individuals may move from one class to another; for example, an individual may move from the infected class to the removed class upon recovery. Thus the model accounts for the interdependency of the different classes within the population.

§2.1 Initial Value Problems

Example (SIR model for spread of diseases)

This example presents a classical model, called the SIR model, of disease transmission within a population. The total population is divided into three groups: **individuals susceptible to disease** (易感者), **infected individuals** (染病者), and **"removed" individuals** (移出者). The removed class counts those individuals who are not infected and not susceptible; in other words, immune, quarantined, or dead. Individuals may move from one class to another; for example, an individual may move from the infected class to the removed class upon recovery. Thus the model accounts for the interdependency of the different classes within the population.

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

The fundamental relation of the SIR model is the relation

$$N = S(t) + I(t) + R(t)\,,$$

where N is the total population size, taken to be constant; $S(t)$ is the size of the susceptible population, $I(t)$ is the size of the infected population, and $R(t)$ is the size of the removed population. We note that the relationship above shows that the rate of change of S, I and R must obey the following identity

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0\,.$$

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

The fundamental relation of the SIR model is the relation

$$N = S(t) + I(t) + R(t)\,,$$

where N is the total population size, taken to be constant; $S(t)$ is the size of the susceptible population, $I(t)$ is the size of the infected population, and $R(t)$ is the size of the removed population. We note that the relationship above shows that the rate of change of S, I and R must obey the following identity

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0\,.$$

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

The derivation of the SIR model is similar to the prey-predator model: the roles of the infected group and the susceptible group are respectively similar to the predator and the prey in the prey-predator model, except that the assumption of a fixed amount of total population prohibits the growth of the susceptible group. The population of the infected group, without the presence of the susceptible group, decays due to the recovery from the disease and increases due to contact with the susceptible group. On the other hand, the only way an individual leaves the susceptible group is by becoming infected (due to contact with the infected group).

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

The derivation of the SIR model is similar to the prey-predator model: the roles of the infected group and the susceptible group are respectively similar to the predator and the prey in the prey-predator model, except that the assumption of a fixed amount of total population prohibits the growth of the susceptible group. The population of the infected group, without the presence of the susceptible group, decays due to the recovery from the disease and increases due to contact with the susceptible group. On the other hand, the only way an individual leaves the susceptible group is by becoming infected (due to contact with the infected group).

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

Therefore, we obtain the following differential equation

$$\frac{dS}{dt} = -bS(t)I(t)\,,$$
$$\frac{dI}{dt} = -\gamma I(t) + bS(t)I(t)\,,$$

where b is termed effective disease transmission, and γ is the **recovery rate**. Because of the identity

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0\,,$$

we find that

$$\frac{dR}{dt} = \gamma I(t)\,.$$

The equation above explains the term recovery rate.

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

Sometimes the system of ODEs in the previous page is written in the following form:

$$\frac{dS}{dt} = -bS(t)I(t)\,,$$
$$\frac{dI}{dt} = -\gamma I(t) + bS(t)I(t)\,,$$
$$\frac{dR}{dt} = \gamma I(t)\,.$$

where $\beta = Nb$ is called the **disease transmission rate**.

In epidemiology (流行病學), the **basic reproduction number**, denoted by R_0, of an infection is the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection. In the SIR model, $R_0 = \beta/\gamma$.

§2.1 Initial Value Problems

Example (SIR model for spread of diseases - cont.)

Sometimes the system of ODEs in the previous page is written in the following form:

$$\frac{dS}{dt} = -bS(t)I(t)\,,$$
$$\frac{dI}{dt} = -\gamma I(t) + bS(t)I(t)\,,$$
$$\frac{dR}{dt} = \gamma I(t)\,.$$

where $\beta = Nb$ is called the **disease transmission rate**.

In epidemiology (流行病學), the **basic reproduction number**, denoted by R_0, of an infection is the expected number of cases directly generated by one case in a population where all individuals are susceptible to infection. In the SIR model, $R_0 = \beta/\gamma$.

§2.1 Initial Value Problems

Example (Three springs and two Mass system)

Now we consider another spring-mass system in which there are two objects, of mass m_1 and m_2, moving on a frictionless surface under the influence of external forces $F_1(t)$ and $F_2(t)$, and they are also constrained by the three springs whose Hooke's constants are k_1, k_2 and k_3, respectively (see Figure 6).

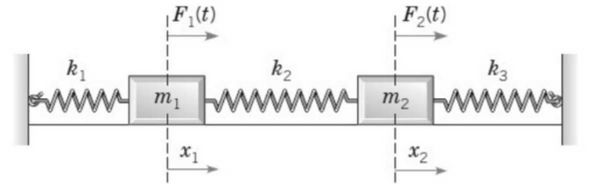


Figure 6: A two-mass, three-spring system

§2.1 Initial Value Problems

Example (Three springs and two Mass system - cont.)

Let L_1, L_2, L_3 be the length of the unconstrained springs, and ℓ_1, ℓ_2, ℓ_3 be the increment of the springs in equilibrium. Then

$$k_1\ell_1 = k_2\ell_2 = k_3\ell_3 .$$

Let $x(t)$ and $y(t)$ be the position of mass m_1 and m_2, measured from the left end, respectively. Then $x(t)$ and $y(t)$ satisfy

$$\begin{aligned}
m_1\frac{d^2x}{dt^2} &= -k_1(x-L_1)+k_2(y-x-L_2)+F_1 ,\\
m_2\frac{d^2y}{dt^2} &= -k_2(y-x-L_2)+k_3(L_1+L_2+L_3+\ell_1+\ell_2+\ell_3-y-L_3)+F_2\\
&= -k_2(y-x-L_2)+k_3(L_1+L_2+\ell_1+\ell_2+\ell_3-y)+F_2 .
\end{aligned}$$

§2.1 Initial Value Problems

Example (Three springs and two Mass system - cont.)

Let x_1, x_2 be the position of masses m_1 and m_2 measured from the equilibrium position; that is, $x_1 = x - L_1 - \ell_1$ and $x_2 = y - L_1 - \ell_1 - L_2 - \ell_2$. Then the equations for x_1 and x_2, locations of mass m_1 and m_2 measured from the equilibrium positions, are given by

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 + k_2(x_2 - x_1) + F_1 ,$$
$$m_2 \frac{d^2 x_2}{dt^2} = -k_2(x_2 - x_1) - k_3 x_2 + F_2 .$$

Note that the equation is "the same as" letting $L_1 = L_2 = L_3 = \ell_1 = \ell_2 = \ell_3 = 0$ in the equation for x, y in the previous page.

The ODE above is a second order linear ODE, and it becomes an IVP if initial conditions $x_1(t_0) = x_{10}$, $x_2(t_0) = x_{20}$, $x_1'(t_0) = x_{11}$ and $x_2'(t_0) = x_{21}$ are imposed.

§2.1 Initial Value Problems

Example (Three springs and two Mass system - cont.)

Let x_1, x_2 be the position of masses m_1 and m_2 measured from the equilibrium position; that is, $x_1 = x - L_1 - \ell_1$ and $x_2 = y - L_1 - \ell_1 - L_2 - \ell_2$. Then the equations for x_1 and x_2, locations of mass m_1 and m_2 measured from the equilibrium positions, are given by

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 + k_2(x_2 - x_1) + F_1 \,,$$
$$m_2 \frac{d^2 x_2}{dt^2} = -k_2(x_2 - x_1) - k_3 x_2 + F_2 \,.$$

Note that the equation is "the same as" letting $L_1 = L_2 = L_3 = \ell_1 = \ell_2 = \ell_3 = 0$ in the equation for x, y in the previous page.

The ODE above is a second order linear ODE, and it becomes an IVP if initial conditions $x_1(t_0) = x_{10}$, $x_2(t_0) = x_{20}$, $x_1'(t_0) = x_{11}$ and $x_2'(t_0) = x_{21}$ are imposed.

§2.1 Initial Value Problems

Example (Three springs and two Mass system - cont.)

Let x_1, x_2 be the position of masses m_1 and m_2 measured from the equilibrium position; that is, $x_1 = x - L_1 - \ell_1$ and $x_2 = y - L_1 - \ell_1 - L_2 - \ell_2$. Then the equations for x_1 and x_2, locations of mass m_1 and m_2 measured from the equilibrium positions, are given by

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 + k_2(x_2 - x_1) + F_1 \,,$$

$$m_2 \frac{d^2 x_2}{dt^2} = -k_2(x_2 - x_1) - k_3 x_2 + F_2 \,.$$

Note that the equation is "the same as" letting $L_1 = L_2 = L_3 = \ell_1 = \ell_2 = \ell_3 = 0$ in the equation for x, y in the previous page.

The ODE above is a second order linear ODE, and it becomes an IVP if initial conditions $x_1(t_0) = x_{10}$, $x_2(t_0) = x_{20}$, $x_1'(t_0) = x_{11}$ and $x_2'(t_0) = x_{21}$ are imposed.

§2.1 Initial Value Problems

Example (Planetary motion)

In this example, we consider the orbit of a planet moving around the sun in the solar system. Suppose that planet under consideration is Earth. Since Earth moves on the ecliptic plane（黃道面）, we can treat the orbit of Earth as a plane curve on the xy-plane. Let the origin of the xy-plane be the center of mass of the sun, and the location of Earth at time t be $\boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are pre-chosen but fixed directions of Cartesian coordinates. Then Newton's second law of motion implies that

$$-\frac{GMm}{\|\boldsymbol{r}(t)\|^3}\boldsymbol{r}(t) = m\boldsymbol{r}''(t)\,, \tag{7}$$

where M and m denote the mass of the sun and Earth, respectively, and $\|\boldsymbol{r}(t)\|$ is the distance from Earth to the sun at time t.

§2.1 Initial Value Problems

Example (Planetary motion)

In this example, we consider the orbit of a planet moving around the sun in the solar system. Suppose that planet under consideration is Earth. Since Earth moves on the ecliptic plane（黃道面）, we can treat the orbit of Earth as a plane curve on the xy-plane. Let the origin of the xy-plane be the center of mass of the sun, and the location of Earth at time t be $\boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are pre-chosen but fixed directions of Cartesian coordinates. Then Newton's second law of motion implies that

$$-\frac{GMm}{\|\boldsymbol{r}(t)\|^3}\boldsymbol{r}(t) = m\boldsymbol{r}''(t)\,, \tag{7}$$

where M and m denote the mass of the sun and Earth, respectively, and $\|\boldsymbol{r}(t)\|$ is the distance from Earth to the sun at time t.

§2.1 Initial Value Problems

Example (Planetary motion)

In this example, we consider the orbit of a planet moving around the sun in the solar system. Suppose that planet under consideration is Earth. Since Earth moves on the ecliptic plane（黃道面）, we can treat the orbit of Earth as a plane curve on the xy-plane. Let the origin of the xy-plane be the center of mass of the sun, and the location of Earth at time t be $\boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are pre-chosen but fixed directions of Cartesian coordinates. Then Newton's second law of motion implies that

$$-\frac{GMm}{\|\boldsymbol{r}(t)\|^3}\boldsymbol{r}(t) = m\boldsymbol{r}''(t)\,, \tag{7}$$

where M and m denote the mass of the sun and Earth, respectively, and $\|\boldsymbol{r}(t)\|$ is the distance from Earth to the sun at time t.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

We note that the two unknowns of the ODE (7) are indeed $x(t)$ and $y(t)$. To study the motion of Earth better, a polar coordinate representation of the ODE is need. We introduce a polar coordinate system in which the pole of the polar coordinate system is the sun, and the polar axis is $\mathbf{i}$. Let $(r(t), \theta(t))$ be the polar coordinate of the location of Earch at time t; that is, $\boldsymbol{r}(t) = r(t)\cos\theta(t)\mathbf{i} + r(t)\sin\theta(t)\mathbf{j}$, and define two vectors

$$\widehat{r}(t) = \cos\theta(t)\mathbf{i} + \sin\theta(t)\mathbf{j}\,,$$

$$\widehat{\theta}(t) = -\sin\theta(t)\mathbf{i} + \cos\theta(t)\mathbf{j}\,,$$

accordingly. Then $\boldsymbol{r}(t) = r(t)\widehat{r}(t)$.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

We note that the two unknowns of the ODE (7) are indeed $x(t)$ and $y(t)$. To study the motion of Earth better, a polar coordinate representation of the ODE is need. We introduce a polar coordinate system in which the pole of the polar coordinate system is the sun, and the polar axis is $\mathbf{i}$. Let $(r(t), \theta(t))$ be the polar coordinate of the location of Earch at time t; that is, $\boldsymbol{r}(t) = r(t)\cos\theta(t)\mathbf{i} + r(t)\sin\theta(t)\mathbf{j}$, and define two vectors

$$\widehat{r}(t) = \cos\theta(t)\mathbf{i} + \sin\theta(t)\mathbf{j}\,,$$
$$\widehat{\theta}(t) = -\sin\theta(t)\mathbf{i} + \cos\theta(t)\mathbf{j}\,,$$

accordingly. Then $\boldsymbol{r}(t) = r(t)\widehat{r}(t)$.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

We note that the two unknowns of the ODE (7) are indeed $x(t)$ and $y(t)$. To study the motion of Earth better, a polar coordinate representation of the ODE is need. We introduce a polar coordinate system in which the pole of the polar coordinate system is the sun, and the polar axis is $\mathbf{i}$. Let $(r(t), \theta(t))$ be the polar coordinate of the location of Earch at time t; that is, $\boldsymbol{r}(t) = r(t)\cos\theta(t)\mathbf{i} + r(t)\sin\theta(t)\mathbf{j}$, and define two vectors

$$\hat{r}(t) = \cos\theta(t)\mathbf{i} + \sin\theta(t)\mathbf{j}\,,$$

$$\hat{\theta}(t) = -\sin\theta(t)\mathbf{i} + \cos\theta(t)\mathbf{j}\,,$$

accordingly. Then $\boldsymbol{r}(t) = r(t)\hat{r}(t)$.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

By the fact that

$$\widehat{r}' = (-\sin\theta\mathbf{i} + \cos\theta\mathbf{j})\theta' = \theta'\widehat{\theta}\,,$$
$$\widehat{\theta}' = -(\cos\theta\mathbf{i} + \sin\theta\mathbf{j})\theta' = -\theta'\widehat{r},$$

we find that

$$\boldsymbol{r}'' = \frac{d}{dt}\big(r'\widehat{r} + r\theta'\widehat{\theta}\big) = r''\widehat{r} + r'\theta'\widehat{\theta} + r'\theta'\widehat{\theta} + r\theta''\widehat{\theta} - r(\theta')^2\widehat{r}$$
$$= \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}\,.$$

Therefore, (7) implies that

$$-\frac{GM}{r^2}\widehat{r} = \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}\,.$$

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

By the fact that

$$\widehat{r}' = (-\sin\theta\mathbf{i} + \cos\theta\mathbf{j})\theta' = \theta'\widehat{\theta},$$
$$\widehat{\theta}' = -(\cos\theta\mathbf{i} + \sin\theta\mathbf{j})\theta' = -\theta'\widehat{r},$$

we find that

$$\boldsymbol{r}'' = \frac{d}{dt}\big(r'\widehat{r} + r\theta'\widehat{\theta}\big) = r''\widehat{r} + r'\theta'\widehat{\theta} + r'\theta'\widehat{\theta} + r\theta''\widehat{\theta} - r(\theta')^2\widehat{r}$$
$$= \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}.$$

Therefore, (7) implies that

$$-\frac{GM}{r^2}\widehat{r} = \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}.$$

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

By the fact that

$$\widehat{r}' = (-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j})\theta' = \theta'\widehat{\theta}\,,$$
$$\widehat{\theta}' = -(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j})\theta' = -\theta'\widehat{r},$$

we find that

$$\begin{aligned}\boldsymbol{r}'' &= \frac{d}{dt}\big(r'\widehat{r} + r\theta'\widehat{\theta}\big) = r''\widehat{r} + r'\theta'\widehat{\theta} + r'\theta'\widehat{\theta} + r\theta''\widehat{\theta} - r(\theta')^2\widehat{r} \\ &= \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}\,.\end{aligned}$$

Therefore, (7) implies that

$$-\frac{GM}{r^2}\widehat{r} = \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}\,.$$

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Since $\hat{r}$ and $\hat{\theta}$ are linearly independent, we find that the polar coordinate $\big(r(t), \theta(t)\big)$ of Earch must satisfy the nonlinear ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2\,, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0\,. \tag{8b}$$

Since (8) is a second-order ODE, to make it an initial value problems we need to specify the values of $r(t_0)$, $\theta(t_0)$, $r'(t_0)$ and $\theta'(t_0)$.

Note that (8b) implies that $(r^2\theta')' = 0$; thus $r^2\theta'$ is a constant. Let ℓ be the constant angular momentum so that

$$\ell = mr^2\theta' = m\mathrm{r}_0\mathrm{v}_0\,, \tag{9}$$

where r_0 is the perihelion distance (近日點距) and v_0 is the speed at the perihelion (近日點).

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Since $\widehat{r}$ and $\widehat{\theta}$ are linearly independent, we find that the polar coordinate $\big(r(t), \theta(t)\big)$ of Earch must satisfy the nonlinear ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0. \tag{8b}$$

Since (8) is a second-order ODE, to make it an initial value problems we need to specify the values of $r(t_0)$, $\theta(t_0)$, $r'(t_0)$ and $\theta'(t_0)$.

Note that (8b) implies that $(r^2\theta')' = 0$; thus $r^2\theta'$ is a constant. Let ℓ be the constant angular momentum so that

$$\ell = mr^2\theta' = m\mathrm{r}_0\mathrm{v}_0, \tag{9}$$

where r_0 is the perihelion distance (近日點距) and v_0 is the speed at the perihelion (近日點).

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Since $\widehat{r}$ and $\widehat{\theta}$ are linearly independent, we find that the polar coordinate $\big(r(t), \theta(t)\big)$ of Earch must satisfy the nonlinear ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2\,, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0\,. \tag{8b}$$

Since (8) is a second-order ODE, to make it an initial value problems we need to specify the values of $r(t_0)$, $\theta(t_0)$, $r'(t_0)$ and $\theta'(t_0)$.

Note that (8b) implies that $(r^2\theta')' = 0$; thus $r^2\theta'$ is a constant. Let ℓ be the constant angular momentum so that

$$\ell = mr^2\theta' = m\mathrm{r}_0\mathrm{v}_0\,, \tag{9}$$

where r_0 is the perihelion distance (近日點距) and v_0 is the speed at the perihelion (近日點).

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Since $\widehat{r}$ and $\widehat{\theta}$ are linearly independent, we find that the polar coordinate $\big(r(t), \theta(t)\big)$ of Earch must satisfy the nonlinear ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2\,, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0\,. \tag{8b}$$

Since (8) is a second-order ODE, to make it an initial value problems we need to specify the values of $r(t_0)$, $\theta(t_0)$, $r'(t_0)$ and $\theta'(t_0)$.

Note that (8b) implies that $(r^2\theta')' = 0$; thus $r^2\theta'$ is a constant. Let ℓ be the constant angular momentum so that

$$\ell = mr^2\theta' = m\mathrm{r}_0\mathrm{v}_0\,, \tag{9}$$

where r_0 is the perihelion distance (近日點距) and v_0 is the speed at the perihelion (近日點).

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Note that (9) shows that θ' is sign-definite (unless $\ell = 0$), so θ is one-to-one. Let $t_1 < t_2$. The area swept out in the time interval $[t_1, t_2]$ is given by

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\, dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\, dt = \frac{\ell(t_2 - t_1)}{2m} = \frac{\mathrm{r}_0 \mathrm{v}_0}{2}(t_2 - t_1);$$

thus we conclude **Kepler's second law of planetary motion**:

> A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

Remark: Kepler's first and third laws of planetary motion will be discussed later.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Note that (9) shows that θ' is sign-definite (unless $\ell = 0$), so θ is one-to-one. Let $t_1 < t_2$. The area swept out in the time interval $[t_1, t_2]$ is given by

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\, dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\, dt = \frac{\ell(t_2 - t_1)}{2m} = \frac{\mathrm{r}_0 \mathrm{v}_0}{2}(t_2 - t_1)\,;$$

thus we conclude **Kepler's second law of planetary motion**:

> A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

Remark: Kepler's first and third laws of planetary motion will be discussed later.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Note that (9) shows that θ' is sign-definite (unless $\ell = 0$), so θ is one-to-one. Let $t_1 < t_2$. The area swept out in the time interval $[t_1, t_2]$ is given by

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\, dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\, dt = \frac{\ell(t_2 - t_1)}{2m} = \frac{\mathrm{r}_0 \mathrm{v}_0}{2}(t_2 - t_1)\,;$$

thus we conclude **Kepler's second law of planetary motion**:

> A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

Remark: Kepler's first and third laws of planetary motion will be discussed later.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

Note that (9) shows that θ' is sign-definite (unless $\ell = 0$), so θ is one-to-one. Let $t_1 < t_2$. The area swept out in the time interval $[t_1, t_2]$ is given by

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\, dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\, dt = \frac{\ell(t_2 - t_1)}{2m} = \frac{\mathrm{r}_0 \mathrm{v}_0}{2}(t_2 - t_1)\,;$$

thus we conclude **Kepler's second law of planetary motion**:

> A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

Remark: Kepler's first and third laws of planetary motion will be discussed later.

§2.1 Initial Value Problems

Example (Planetary motion - cont.)

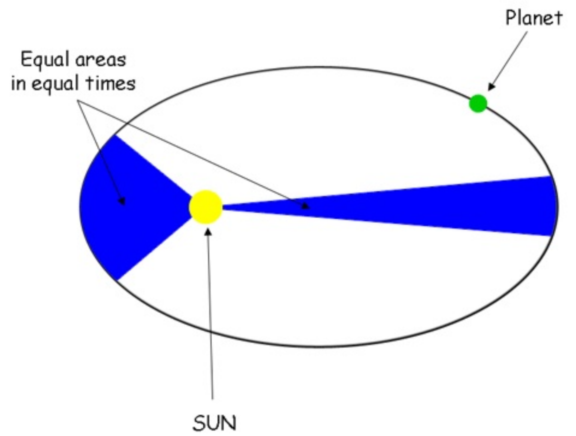


Figure 7: Kepler's second law of planetary motion

§2.1 Initial Value Problems

Remark: The angular momentum of a moving object relative to a point is the cross product of the particle's position vector $\boldsymbol{r}$ (relative to the point) and its momentum vector $\boldsymbol{p}$ (relative to the point as well). Therefore, the angular momentum of the planet relative to the Sun is

$$\boldsymbol{r} \times m\boldsymbol{r}' = mr\widehat{\mathrm{r}} \times (r'\widehat{\mathrm{r}} + r\theta'\widehat{\theta}) = mr^2\theta'\widehat{\mathrm{r}} \times \widehat{\theta} = mr^2\theta'\mathbf{k}\,;$$

thus the quantity $mr^2\theta'$ is the angular momentum of the planet relative to the sun. (8b) (or (9)) then implies that the angular momentum is a constant, so-called the ***conservation of angular momentum***（角動量守恆）.

§2.1 Initial Value Problems

Remark: The angular momentum of a moving object relative to a point is the cross product of the particle's position vector $\boldsymbol{r}$ (relative to the point) and its momentum vector $\boldsymbol{p}$ (relative to the point as well). Therefore, the angular momentum of the planet relative to the Sun is

$$\boldsymbol{r} \times m\boldsymbol{r}' = mr\widehat{r} \times (r'\widehat{r} + r\theta'\widehat{\theta}) = mr^2\theta'\widehat{r} \times \widehat{\theta} = mr^2\theta'\mathbf{k};$$

thus the quantity $mr^2\theta'$ is the angular momentum of the planet relative to the sun. (8b) (or (9)) then implies that the angular momentum is a constant, so-called the ***conservation of angular momentum***（角動量守恆）.

§2.1 Initial Value Problems

Remark: The angular momentum of a moving object relative to a point is the cross product of the particle's position vector $\boldsymbol{r}$ (relative to the point) and its momentum vector $\boldsymbol{p}$ (relative to the point as well). Therefore, the angular momentum of the planet relative to the Sun is

$$\boldsymbol{r} \times m\boldsymbol{r}' = mr\hat{\mathrm{r}} \times (r'\hat{\mathrm{r}} + r\theta'\hat{\theta}) = mr^2\theta'\hat{\mathrm{r}} \times \hat{\theta} = mr^2\theta'\mathbf{k};$$

thus the quantity $mr^2\theta'$ is the angular momentum of the planet relative to the sun. (8b) (or (9)) then implies that the angular momentum is a constant, so-called the ***conservation of angular momentum***（角動量守恆）.

§2.1 Initial Value Problems

Example (Finding relative minimum of a function)

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function. To find a relative minimum of f, we first look for critical points of f. In general, it may not be easy to solve for zeros of f'. In this example we provide a way to "find" possible local minimum of f.

Suppose that x_0 is given. If $f'(x_0) < 0$, we expect that the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the right-hand side of x_0. Similarly, if $f'(x_0) > 0$, then the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the left-hand side of x_0. Therefore, for a given point x_0, we can localize the position of the nearest critical point where f attains a local minimum by "moving" to the right or to the left based on the sign of f'.

§2.1 Initial Value Problems

Example (Finding relative minimum of a function)

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function. To find a relative minimum of f, we first look for critical points of f. In general, it may not be easy to solve for zeros of f'. In this example we provide a way to "find" possible local minimum of f.

Suppose that x_0 is given. If $f'(x_0) < 0$, we expect that the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the right-hand side of x_0. Similarly, if $f'(x_0) > 0$, then the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the left-hand side of x_0. Therefore, for a given point x_0, we can localize the position of the nearest critical point where f attains a local minimum by "moving" to the right or to the left based on the sign of f'.

§2.1 Initial Value Problems

Example (Finding relative minimum of a function)

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function. To find a relative minimum of f, we first look for critical points of f. In general, it may not be easy to solve for zeros of f'. In this example we provide a way to "find" possible local minimum of f.

Suppose that x_0 is given. If $f'(x_0) < 0$, we expect that the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the right-hand side of x_0. Similarly, if $f'(x_0) > 0$, then the value of $f(x)$ will be smaller than $f(x_0)$ when x is close but on the left-hand side of x_0. Therefore, for a given point x_0, we can localize the position of the nearest critical point where f attains a local minimum by "moving" to the right or to the left based on the sign of f'.

§2.1 Initial Value Problems

Example (Finding relative minimum of a function - cont.)

This motivates the following IVP

$$x' = -f'(x)\,, \qquad x(0) = x_0\,.$$

In general, for a continuously differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$, we can use

$$\boldsymbol{x}' = -(\nabla f)(\boldsymbol{x})\,, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0\,,$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, to find a critical point near $\boldsymbol{x}_0$.

Remark: To avoid the speed of the motion becoming too slow when $x(t)$ is close to a relative minimum of f, sometimes we can normalize the right-hand side so that the IVP under consideration becomes

$$x' = -\frac{f'(x)}{|f'(x)|} \left(\text{or} \quad \boldsymbol{x}' = -\frac{(\nabla f)(\boldsymbol{x})}{\|(\nabla f)(\boldsymbol{x})\|}\right), \qquad x(0) = x_0\,.$$

§2.1 Initial Value Problems

Example (Finding relative minimum of a function - cont.)

This motivates the following IVP

$$x' = -f'(x)\,, \qquad x(0) = x_0\,.$$

In general, for a continuously differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}$, we can use

$$\boldsymbol{x}' = -(\nabla f)(\boldsymbol{x})\,, \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0\,,$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, to find a critical point near $\boldsymbol{x}_0$.

Remark: To avoid the speed of the motion becoming too slow when $x(t)$ is close to a relative minimum of f, sometimes we can normalize the right-hand side so that the IVP under consideration becomes

$$x' = -\frac{f'(x)}{|f'(x)|} \left(\text{or} \quad \boldsymbol{x}' = -\frac{(\nabla f)(\boldsymbol{x})}{\|(\nabla f)(\boldsymbol{x})\|}\right), \qquad x(0) = x_0\,.$$

§2.1 Initial Value Problems

Theorem (Existence and Uniqueness of the Solution of IVP)

Consider the initial value problem

$$\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x}), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \in \mathbb{R}^n,$$

where $\boldsymbol{x}$ and $\boldsymbol{f}$ functions with values in $\mathbb{R}^n$. If $\boldsymbol{f}$ and the first partial derivatives of $\boldsymbol{f}$ with respect to all its variables, possibly except t, are continuous functions in some rectangular domain $R = [a, b] \times [c_1, d_1] \times [c_2, d_2] \times \cdots \times [c_n, d_n]$ that contains the point $(t_0, \boldsymbol{x}_0)$ in the interior, then the initial value problem above has a unique solution in some interval $I = (t_0 - h, t_0 + h)$ for some positive number h. Moreover, the solution is continuously differentiable on I.

Remark: Every n-th order IVP has a unique solution provided that the right-hand side function has required properties.

§2.1 Initial Value Problems

Theorem (Existence and Uniqueness of the Solution of IVP)

Consider the initial value problem

$$\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x}), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \in \mathbb{R}^n,$$

where $\boldsymbol{x}$ *and* $\boldsymbol{f}$ *functions with values in* $\mathbb{R}^n$. *If* $\boldsymbol{f}$ *and the first partial derivatives of* $\boldsymbol{f}$ *with respect to all its variables, possibly except* t, *are continuous functions in some rectangular domain* $R = [a, b] \times [c_1, d_1] \times [c_2, d_2] \times \cdots \times [c_n, d_n]$ *that contains the point* $(t_0, \boldsymbol{x}_0)$ *in the interior, then the initial value problem above has a unique solution in some interval* $I = (t_0 - h, t_0 + h)$ *for some positive number* h. *Moreover, the solution is continuously differentiable on* I.

Remark: Every n-th order IVP has a unique solution provided that the right-hand side function has required properties.

§2.1 Initial Value Problems

To be more precise, we rewrite the IVP

$$y^{(n)} = f(t, y, \cdots, y^{(n-1)}), \quad y(t_0) = y_0, \cdots, y^{(n-1)}(t_0) = y_{n-1}$$

as $\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x})$ with $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, where

$$\boldsymbol{x} = \begin{bmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{bmatrix}, \quad \boldsymbol{x}_0 = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{f}(t, \boldsymbol{x}) = \mathrm{N}\boldsymbol{x} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(t, \boldsymbol{x}) \end{bmatrix}$$

in which $\mathrm{N} = [n_{ij}]$ is the constant matrix given by $n_{k,k+1} = 1$ for $1 \leqslant k \leqslant n-1$ and $n_{ij} = 0$ elsewhere. Then

$$\frac{\partial \boldsymbol{f}}{\partial x_k} = \mathrm{N}\mathbf{e}_k + \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial f}{\partial y^{(k)}} \end{bmatrix}^{\mathrm{T}}$$

so $\frac{\partial \boldsymbol{f}}{\partial x_k}$ is continuous if and only if $\frac{\partial f}{\partial y^{(k)}}$ is continuous. This verifies the statement in the previous page.

§2.1 Initial Value Problems

To be more precise, we rewrite the IVP

$$y^{(n)} = f(t, y, \cdots, y^{(n-1)}), \quad y(t_0) = y_0, \cdots, y^{(n-1)}(t_0) = y_{n-1}$$

as $\boldsymbol{x}' = \boldsymbol{f}(t, \boldsymbol{x})$ with $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, where

$$\boldsymbol{x} = \begin{bmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{bmatrix}, \quad \boldsymbol{x}_0 = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{f}(t, \boldsymbol{x}) = \mathrm{N}\boldsymbol{x} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(t, \boldsymbol{x}) \end{bmatrix}$$

in which $\mathrm{N} = [n_{ij}]$ is the constant matrix given by $n_{k,k+1} = 1$ for $1 \leqslant k \leqslant n-1$ and $n_{ij} = 0$ elsewhere. Then

$$\frac{\partial \boldsymbol{f}}{\partial x_k} = \mathrm{N}\mathbf{e}_k + \begin{bmatrix} 0 & \cdots & 0 & \dfrac{\partial f}{\partial y^{(k)}} \end{bmatrix}^{\mathrm{T}}$$

so $\dfrac{\partial \boldsymbol{f}}{\partial x_k}$ is continuous if and only if $\dfrac{\partial f}{\partial y^{(k)}}$ is continuous. This verifies the statement in the previous page.

§2.1 Initial Value Problems

In particular, if the ODE in IVP is linear; that is,

$$f(t, y, \cdots, y^{(n-1)}) \\ = a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y + g(t),$$

then clearly the first partial derivative of f with respect to all the "y-variables" are continuous if $a_0, a_1, \cdots, a_{n-1}$ are continuous (on an open interval containing t_0). Therefore, if the coefficients and the forcing of a linear ODE are continuous (on an open interval containing t_0), then the solution of IVP exists and is uniquely determine by the initial data $y_0, \cdots, y_{n-1}$.

§2.1 Initial Value Problems

In particular, if the ODE in IVP is linear; that is,

$$f(t, y, \cdots, y^{(n-1)}) = a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y + g(t)\,,$$

then clearly the first partial derivative of f with respect to all the "y-variables" are continuous if $a_0, a_1, \cdots, a_{n-1}$ are continuous (on an open interval containing t_0). Therefore, if the coefficients and the forcing of a linear ODE are continuous (on an open interval containing t_0), then the solution of IVP exists and is uniquely determine by the initial data $y_0, \cdots, y_{n-1}$.

§2.2 Some Basic Techniques of Solving ODEs

§2.2.1 Separation of variables

The simplest ODE takes the form $x' = g(t)h(x)$. Formally we let Φ and G be an anti-derivative of $\frac{1}{h}$ and g, respectively. Then

$$\frac{d}{dt}\Phi(x(t)) = \Phi'(x(t))x'(t) = \frac{x'(t)}{h(x(t))} = g(t)$$

which implies that $\Phi(x(t)) = G(t) + C$ for some constant C. A general solution $x(t)$ then is obtained by inverting the function Φ.

If an initial condition $x(t_0) = x_0$ is provided, then we can choose Φ and G satisfying $\Phi(x_0) = G(t_0)$ so that

$$\Phi(x(t)) - \Phi(x_0) = \int_{t_0}^{t} \frac{d}{ds}\Phi(x(s))\, ds = \int_{t_0}^{t} g(s)\, ds = G(t) - G(t_0);$$

thus $\Phi(x(t)) = G(t)$ which further shows that $x(t) = \Phi^{-1}(G(t))$ (as long as Φ has an inverse function).

§2.2 Some Basic Techniques of Solving ODEs

§2.2.1 Separation of variables

The simplest ODE takes the form $x' = g(t)h(x)$. Formally we let Φ and G be an anti-derivative of $\frac{1}{h}$ and g, respectively. Then

$$\frac{d}{dt}\Phi(x(t)) = \Phi'(x(t))x'(t) = \frac{x'(t)}{h(x(t))} = g(t)$$

which implies that $\Phi(x(t)) = G(t) + C$ for some constant C. A general solution $x(t)$ then is obtained by inverting the function Φ.

If an initial condition $x(t_0) = x_0$ is provided, then we can choose Φ and G satisfying $\Phi(x_0) = G(t_0)$ so that

$$\Phi(x(t)) - \Phi(x_0) = \int_{t_0}^{t} \frac{d}{ds}\Phi(x(s))\, ds = \int_{t_0}^{t} g(s)\, ds = G(t) - G(t_0)\,;$$

thus $\Phi(x(t)) = G(t)$ which further shows that $x(t) = \Phi^{-1}(G(t))$ (as long as Φ has an inverse function).

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the logistic equation

$$p' = rp\Big(1 - \frac{p}{K}\Big)$$

introduced in Chapter 1. Letting $h(p) = rp\Big(1 - \frac{p}{K}\Big)$, we have

$$\int \frac{dp}{h(p)} = \frac{K}{r}\int \frac{dp}{p(K-p)} = \frac{1}{r}\int \Big(\frac{1}{p} + \frac{1}{K-p}\Big)\, dp$$
$$= \frac{1}{r}\big(\ln|p| - \ln|K-p|\big) + C\,.$$

Therefore, an anti-derivative of $\frac{1}{h}$ is $\Phi(p) = \frac{1}{r}\ln\left|\frac{p}{K-p}\right| + C$ whose inverse function, when considering the case $0 < p < K$ (which is the case if $0 < p(t_0) < K$), is given by

$$\Phi^{-1}(t) = \frac{Ke^{r(t-C)}}{1 + e^{r(t-C)}} = \frac{KDe^{rt}}{1 + De^{rt}}\,,$$

where $D = e^{-Cr}$; thus $p(t) = \frac{KDe^{rt}}{1 + De^{rt}}$.

§2.2 Some Basic Techniques of Solving ODEs

§2.2.2 The method of integrating factor

Consider the first-order linear ODE

$$x'(t) + q(t)x(t) = r(t)\,,$$

where q, r are given continuous functions defined on a certain interval. Let Q denote an antiderivative of q. Note that

$$\frac{d}{dt}\big[e^{Q(t)}x(t)\big] = e^{Q(t)}q(t)x(t) + e^{Q(t)}x'(t) = e^{Q(t)}\big[x'(t) + q(t)x(t)\big]\,;$$

thus

$$\frac{d}{dt}\big[e^{Q(t)}x(t)\big] = e^{Q(t)}r(t)\,. \tag{10}$$

The equation above implies that

$$e^{Q(t)}x(t) = \int e^{Q(t)}r(t)\,dt \quad \text{or} \quad x(t) = e^{-Q(t)}\int e^{Q(t)}r(t)\,dt\,.$$

§2.2 Some Basic Techniques of Solving ODEs

§2.2.2 The method of integrating factor

Consider the first-order linear ODE

$$x'(t) + q(t)x(t) = r(t)\,,$$

where q, r are given continuous functions defined on a certain interval. Let Q denote an antiderivative of q. Note that

$$\frac{d}{dt}\big[e^{Q(t)}x(t)\big] = e^{Q(t)}q(t)x(t) + e^{Q(t)}x'(t) = e^{Q(t)}\big[x'(t) + q(t)x(t)\big]\,;$$

thus

$$\frac{d}{dt}\big[e^{Q(t)}x(t)\big] = e^{Q(t)}r(t)\,. \tag{10}$$

The equation above implies that

$$e^{Q(t)}x(t) = \int e^{Q(t)}r(t)\,dt \quad \text{or} \quad x(t) = e^{-Q(t)}\int e^{Q(t)}r(t)\,dt\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Suppose now we are given an initial condition $x(t_0) = x_0$. Then we integrate both sides of (10) from t_0 to t and obtain that

$$\int_{t_0}^{t} \frac{d}{ds}\big[e^{Q(s)}x(s)\big]\,ds = \int_{t_0}^{t} e^{Q(s)}r(s)\,ds\,.$$

The Fundamental Theorem of Calculus further implies that

$$e^{Q(t)}x(t) - e^{Q(t_0)}x(t_0) = \int_{t_0}^{t} e^{Q(s)}r(s)\,ds\,;$$

thus

$$x(t) = e^{Q(t_0)-Q(t)}x_0 + \int_{t_0}^{t} e^{Q(s)-Q(t)}r(s)\,ds\,.$$

Formula above gives the solution to the initial value problem

$$x'(t) + q(t)x(t) = r(t)\,, \qquad x(t_0) = x_0\,.$$

§2.2 Some Basic Techniques of Solving ODEs

§2.2.3 Second-order linear ODEs with constant coefficients

Consider the second-order linear ODE

$$x''(t) + b(t)x'(t) + c(t)x(t) = f(t)\,, \tag{11}$$

where b, c and f are given continuous functions.

We first consider the case $f \equiv 0$. In this case, the ODE is said to be ***homogeneous***, and the theory of differential equations shows that the solution space (that is, the collection of solutions) is two dimensional. In other words, there exist two linearly independent solutions φ_1 and φ_2 such that every solution x can be written as the linear combination of φ_1 and φ_2 or equivalently,

$$x(t) = C_1\varphi_1(t) + C_2\varphi_2(t) \quad \text{for some constantn } C_1 \text{ and } C_2.$$

§2.2 Some Basic Techniques of Solving ODEs

§2.2.3 Second-order linear ODEs with constant coefficients

Consider the second-order linear ODE

$$x''(t) + b(t)x'(t) + c(t)x(t) = f(t)\,, \tag{11}$$

where b, c and f are given continuous functions.

We first consider the case $f \equiv 0$. In this case, the ODE is said to be ***homogeneous***, and the theory of differential equations shows that the solution space (that is, the collection of solutions) is two dimensional. In other words, there exist two linearly independent solutions φ_1 and φ_2 such that every solution x can be written as the linear combination of φ_1 and φ_2 or equivalently,

$$x(t) = C_1\varphi_1(t) + C_2\varphi_2(t) \quad \text{for some constantn } C_1 \text{ and } C_2.$$

§2.2 Some Basic Techniques of Solving ODEs

§2.2.3 Second-order linear ODEs with constant coefficients

Consider the second-order linear ODE

$$x''(t) + b(t)x'(t) + c(t)x(t) = f(t)\,, \tag{11}$$

where b, c and f are given continuous functions.

We first consider the case $f \equiv 0$. In this case, the ODE is said to be ***homogeneous***, and the theory of differential equations shows that the solution space (that is, the collection of solutions) is two dimensional. In other words, there exist two linearly independent solutions φ_1 and φ_2 such that every solution x can be written as the linear combination of φ_1 and φ_2 or equivalently,

$$x(t) = C_1\varphi_1(t) + C_2\varphi_2(t) \quad \text{for some constantn } C_1 \text{ and } C_2.$$

§2.2 Some Basic Techniques of Solving ODEs

In general, it is not easy to find linearly independent solution to homogeneous ODEs. Nevertheless, if $b(t) = b$ and $c(t) = c$ are constant functions, we can find linearly independent solution by looking at the **characteristic equation**

$$r^2 + br + c = 0 . \tag{12}$$

1. If (12) has two distinct real zeros r_1 and r_2, then
$$\varphi_1(t) = e^{r_1 t} \qquad \text{and} \qquad \varphi_2(t) = e^{r_2 t} .$$
2. If (12) has a repeated real zero r, then
$$\varphi_1(t) = e^{rt} \qquad \text{and} \qquad \varphi_2(t) = te^{rt} .$$
3. If (12) has complex zeros $\alpha \pm i\beta$, where $\alpha, \beta \in \mathbb{R}$, then
$$\varphi_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad \varphi_2(t) = e^{\alpha t}\sin(\beta t) .$$

When considering initial value problem, the constants C_1 and C_2 are determined by the initial conditions.

§2.2 Some Basic Techniques of Solving ODEs

In general, it is not easy to find linearly independent solution to homogeneous ODEs. Nevertheless, if $b(t) = b$ and $c(t) = c$ are constant functions, we can find linearly independent solution by looking at the **characteristic equation**

$$r^2 + br + c = 0\,. \tag{12}$$

1. If (12) has two distinct real zeros r_1 and r_2, then
$$\varphi_1(t) = e^{r_1 t} \qquad \text{and} \qquad \varphi_2(t) = e^{r_2 t}\,.$$
2. If (12) has a repeated real zero r, then
$$\varphi_1(t) = e^{rt} \qquad \text{and} \qquad \varphi_2(t) = te^{rt}\,.$$
3. If (12) has complex zeros $\alpha \pm i\beta$, where $\alpha, \beta \in \mathbb{R}$, then
$$\varphi_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad \varphi_2(t) = e^{\alpha t}\sin(\beta t)\,.$$

When considering initial value problem, the constants C_1 and C_2 are determined by the initial conditions.

§2.2 Some Basic Techniques of Solving ODEs

In general, it is not easy to find linearly independent solution to homogeneous ODEs. Nevertheless, if $b(t) = b$ and $c(t) = c$ are constant functions, we can find linearly independent solution by looking at the **characteristic equation**

$$r^2 + br + c = 0\,. \tag{12}$$

1. If (12) has two distinct real zeros r_1 and r_2, then
$$\varphi_1(t) = e^{r_1 t} \qquad \text{and} \qquad \varphi_2(t) = e^{r_2 t}\,.$$
2. If (12) has a repeated real zero r, then
$$\varphi_1(t) = e^{rt} \qquad \text{and} \qquad \varphi_2(t) = te^{rt}\,.$$
3. If (12) has complex zeros $\alpha \pm i\beta$, where $\alpha, \beta \in \mathbb{R}$, then
$$\varphi_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad \varphi_2(t) = e^{\alpha t}\sin(\beta t)\,.$$

When considering initial value problem, the constants C_1 and C_2 are determined by the initial conditions.

§2.2 Some Basic Techniques of Solving ODEs

In general, it is not easy to find linearly independent solution to homogeneous ODEs. Nevertheless, if $b(t) = b$ and $c(t) = c$ are constant functions, we can find linearly independent solution by looking at the **characteristic equation**

$$r^2 + br + c = 0\,. \tag{12}$$

1. If (12) has two distinct real zeros r_1 and r_2, then
$$\varphi_1(t) = e^{r_1 t} \qquad \text{and} \qquad \varphi_2(t) = e^{r_2 t}\,.$$
2. If (12) has a repeated real zero r, then
$$\varphi_1(t) = e^{rt} \qquad \text{and} \qquad \varphi_2(t) = te^{rt}\,.$$
3. If (12) has complex zeros $\alpha \pm i\beta$, where $\alpha, \beta \in \mathbb{R}$, then
$$\varphi_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad \varphi_2(t) = e^{\alpha t}\sin(\beta t)\,.$$

When considering initial value problem, the constants C_1 and C_2 are determined by the initial conditions.

§2.2 Some Basic Techniques of Solving ODEs

In general, it is not easy to find linearly independent solution to homogeneous ODEs. Nevertheless, if $b(t) = b$ and $c(t) = c$ are constant functions, we can find linearly independent solution by looking at the **characteristic equation**

$$r^2 + br + c = 0 . \tag{12}$$

1. If (12) has two distinct real zeros r_1 and r_2, then
$$\varphi_1(t) = e^{r_1 t} \qquad \text{and} \qquad \varphi_2(t) = e^{r_2 t} .$$
2. If (12) has a repeated real zero r, then
$$\varphi_1(t) = e^{rt} \qquad \text{and} \qquad \varphi_2(t) = te^{rt} .$$
3. If (12) has complex zeros $\alpha \pm i\beta$, where $\alpha, \beta \in \mathbb{R}$, then
$$\varphi_1(t) = e^{\alpha t}\cos(\beta t) \qquad \text{and} \qquad \varphi_2(t) = e^{\alpha t}\sin(\beta t) .$$

When considering initial value problem, the constants C_1 and C_2 are determined by the initial conditions.

§2.2 Some Basic Techniques of Solving ODEs

Example (Simple harmonic motion)

Consider the spring-mass system

$$m\ddot{x} + kx = 0\,, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,.$$

Rewrite the equation above as $\ddot{x} + \omega^2 x = 0$, where $\omega = \sqrt{k/m}$. Since the corresponding characteristic equation has two complex zeros $\pm\omega i$, we find that

$$x(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t)\,.$$

Using the initial data, we find that $C_1 = x_0$ and $C_2 = v_0/\omega$; thus the solution to the IVP above is given by

$$x(t) = x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t) = R\cos(\omega t - \phi)\,,$$

where $R = \sqrt{x_0^2 + \dfrac{v_0^2}{\omega^2}}$ and ϕ satisfies $\cos\phi = \dfrac{x_0}{R}$ and $\sin\phi = \dfrac{v_0}{R\omega}$.

§2.2 Some Basic Techniques of Solving ODEs

If b or c is not constant, there is a way to find a second solution which is linearly independent to **a known non-zero solution**. Suppose that $x = \varphi_1(t)$ satisfies

$$x''(t) + b(t)x'(t) + c(t)x(t) = 0\,. \tag{13}$$

We look for a solution φ_2 of the form $\varphi_2(t) = v(t)\varphi_1(t)$. If such a φ_2 is a solution to (13), then

$$\begin{aligned} 0 &= \varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t) \\ &= v''(t)\varphi_1(t) + 2v'(t)\varphi_1'(t) + b(t)v'(t)\varphi_1(t) \\ &\quad + v(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &= v''(t)\varphi_1(t) + v'(t)\big[2\varphi_1'(t) + b(t)\varphi_1(t)\big]\,. \end{aligned}$$

The equation above is a first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor:

§2.2 Some Basic Techniques of Solving ODEs

If b or c is not constant, there is a way to find a second solution which is linearly independent to **a known non-zero solution**. Suppose that $x = \varphi_1(t)$ satisfies

$$x''(t) + b(t)x'(t) + c(t)x(t) = 0\,. \tag{13}$$

We look for a solution φ_2 of the form $\varphi_2(t) = v(t)\varphi_1(t)$. If such a φ_2 is a solution to (13), then

$$\begin{aligned} 0 &= \varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t) \\ &= v''(t)\varphi_1(t) + 2v'(t)\varphi_1'(t) + b(t)v'(t)\varphi_1(t) \\ &\quad + v(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &= v''(t)\varphi_1(t) + v'(t)\big[2\varphi_1'(t) + b(t)\varphi_1(t)\big]\,. \end{aligned}$$

The equation above is a first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor:

§2.2 Some Basic Techniques of Solving ODEs

If b or c is not constant, there is a way to find a second solution which is linearly independent to **a known non-zero solution**. Suppose that $x = \varphi_1(t)$ satisfies

$$x''(t) + b(t)x'(t) + c(t)x(t) = 0\,. \tag{13}$$

We look for a solution φ_2 of the form $\varphi_2(t) = v(t)\varphi_1(t)$. If such a φ_2 is a solution to (13), then

$$\begin{aligned} 0 &= \varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t) \\ &= v''(t)\varphi_1(t) + 2v'(t)\varphi_1'(t) + b(t)v'(t)\varphi_1(t) \\ &\quad + v(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &= v''(t)\varphi_1(t) + v'(t)\big[2\varphi_1'(t) + b(t)\varphi_1(t)\big]\,. \end{aligned}$$

The equation above is a first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor:

§2.2 Some Basic Techniques of Solving ODEs

A second solution φ_2 is given by $\varphi_2(t) = v(t)\varphi_1(t)$, where v satisfies

$$v''(t)\varphi_1(t) + v'(t)\big[2\varphi_1'(t) + b(t)\varphi_1(t)\big] = 0\,.$$

The equation above is an first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor: since y satisfies

$$y' + \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)}y(t) = 0\,,$$

with B denoting an anti-derivative of b we have

$$\begin{aligned} y(t) &= C\exp\Big(-\int \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)}\,dt\Big) \\ &= C\exp\Big(-2\ln\big|\varphi(t)\big| - B(t)\Big) = \frac{C}{\varphi_1(t)^2}e^{-B(t)}\,. \end{aligned}$$

Therefore, another solution φ_2 is given by

$$\varphi_2(t) = v(t)\varphi_1(t) = \varphi_1(t)\int \frac{1}{\varphi_1(t)^2}e^{-B(t)}\,dt\,.$$

§2.2 Some Basic Techniques of Solving ODEs

A second solution φ_2 is given by $\varphi_2(t) = v(t)\varphi_1(t)$, where v satisfies

$$v''(t)\varphi_1(t) + v'(t)\left[2\varphi_1'(t) + b(t)\varphi_1(t)\right] = 0\,.$$

The equation above is an first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor: since y satisfies

$$y' + \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)} y(t) = 0\,,$$

with B denoting an anti-derivative of b we have

$$\begin{aligned} y(t) &= C\exp\left(-\int \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)}\,dt\right) \\ &= C\exp\left(-2\ln\left|\varphi(t)\right| - B(t)\right) = \frac{C}{\varphi_1(t)^2}e^{-B(t)}\,. \end{aligned}$$

Therefore, another solution φ_2 is given by

$$\varphi_2(t) = v(t)\varphi_1(t) = \varphi_1(t)\int \frac{1}{\varphi_1(t)^2}e^{-B(t)}\,dt\,.$$

§2.2 Some Basic Techniques of Solving ODEs

A second solution φ_2 is given by $\varphi_2(t) = v(t)\varphi_1(t)$, where v satisfies

$$v''(t)\varphi_1(t) + v'(t)\left[2\varphi_1'(t) + b(t)\varphi_1(t)\right] = 0\,.$$

The equation above is an first order ODE for $y(t) = v'(t)$ and can be solved using the method of integrating factor: since y satisfies

$$y' + \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)}y(t) = 0\,,$$

with B denoting an anti-derivative of b we have

$$\begin{aligned} y(t) &= C\exp\left(-\int \frac{2\varphi_1'(t) + b(t)\varphi_1(t)}{\varphi_1(t)}\,dt\right) \\ &= C\exp\left(-2\ln\left|\varphi(t)\right| - B(t)\right) = \frac{C}{\varphi_1(t)^2}e^{-B(t)}\,. \end{aligned}$$

Therefore, another solution φ_2 is given by

$$\varphi_2(t) = v(t)\varphi_1(t) = \varphi_1(t)\int \frac{1}{\varphi_1(t)^2}e^{-B(t)}\,dt\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Given that $y = \varphi_1(t) = \dfrac{1}{t}$ is a solution of

$$2t^2x'' + 3tx' - x = 0 \qquad \text{for } t > 0\,,$$

find a linearly independent solution of the equation.

Rewrite the ODE above as

$$x'' + \frac{3}{2t}x' - \frac{1}{2t^2}x = 0\,. \quad \Big(\text{so } b(t) = \frac{3}{2t}\Big)$$

Using the formula from previous page, we find that a linearly independent second solution is given by

$$\varphi_2(t) = \varphi_1(t)\int \frac{1}{\varphi_1(t)^2}e^{-B(t)}\,dt = \frac{1}{t}\int t^2 \exp\Big(-\frac{3}{2}\ln t\Big)\,dt = \frac{2}{3}\sqrt{t}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Given that $y = \varphi_1(t) = \dfrac{1}{t}$ is a solution of

$$2t^2x'' + 3tx' - x = 0 \qquad \text{for } t > 0\,,$$

find a linearly independent solution of the equation.

Rewrite the ODE above as

$$x'' + \frac{3}{2t}x' - \frac{1}{2t^2}x = 0\,. \quad \left(\text{so } b(t) = \frac{3}{2t}\right)$$

Using the formula from previous page, we find that a linearly independent second solution is given by

$$\varphi_2(t) = \varphi_1(t)\int \frac{1}{\varphi_1(t)^2} e^{-B(t)}\, dt = \frac{1}{t}\int t^2 \exp\left(-\frac{3}{2}\ln t\right) dt = \frac{2}{3}\sqrt{t}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Now we consider the general case that f **is not the zero function**. In this case, the theory of differential equations shows that the solution to (11) can be expressed as

$$x(t) = C_1\varphi_1(t) + C_2\,\varphi_2(t) + x_p(t)\,,$$

for some constants C_1 and C_2, where $\{\varphi_1, \varphi_2\}$ is a basis of the solution space of the corresponding homogeneous ODE, and x_p is a particular solution of (11). One such a particular solution can be found using **the method of variation of parameters/constants** as follows. Suppose that

$$x_p(t) = C_1(t)\varphi_1(t) + C_2(t)\varphi_2(t)$$

for some functions C_1, C_2 to be determined.

§2.2 Some Basic Techniques of Solving ODEs

Now we consider the general case that f **is not the zero function**. In this case, the theory of differential equations shows that the solution to (11) can be expressed as

$$x(t) = C_1\varphi_1(t) + C_2\,\varphi_2(t) + x_p(t)\,,$$

for some constants C_1 and C_2, where $\{\varphi_1, \varphi_2\}$ is a basis of the solution space of the corresponding homogeneous ODE, and x_p is a particular solution of (11). One such a particular solution can be found using **the method of variation of parameters/constants** as follows. Suppose that

$$x_p(t) = C_1(t)\varphi_1(t) + C_2(t)\varphi_2(t)$$

for some functions C_1, C_2 to be determined.

§2.2 Some Basic Techniques of Solving ODEs

First we assume that C_1, C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,.$$

Then $x_p'(t) = C_1(t)\varphi_1'(t) + C_2(t)\varphi_2'(t)$ which further implies that

$$\begin{aligned} f(t) &= x_p''(t) + b(t)x_p'(t) + c(t)x_p(t) \\ &= C_1'(t)\varphi_1'(t) + C_1(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &\quad + C_2'(t)\varphi_2'(t) + C_2(t)\big[\varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t)\big] \\ &= C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t)\,. \end{aligned}$$

Therefore, C_1 and C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$

$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

§2.2 Some Basic Techniques of Solving ODEs

First we assume that C_1, C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,.$$

Then $x_p'(t) = C_1(t)\varphi_1'(t) + C_2(t)\varphi_2'(t)$ which further implies that

$$\begin{aligned} f(t) &= x_p''(t) + b(t)x_p'(t) + c(t)x_p(t) \\ &= C_1'(t)\varphi_1'(t) + C_1(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &\quad + C_2'(t)\varphi_2'(t) + C_2(t)\big[\varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t)\big] \\ &= C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t)\,. \end{aligned}$$

Therefore, C_1 and C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$

$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

§2.2 Some Basic Techniques of Solving ODEs

First we assume that C_1, C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,.$$

Then $x_p'(t) = C_1(t)\varphi_1'(t) + C_2(t)\varphi_2'(t)$ which further implies that

$$\begin{aligned} f(t) &= x_p''(t) + b(t)x_p'(t) + c(t)x_p(t) \\ &= C_1'(t)\varphi_1'(t) + C_1(t)\big[\varphi_1''(t) + b(t)\varphi_1'(t) + c(t)\varphi_1(t)\big] \\ &\quad + C_2'(t)\varphi_2'(t) + C_2(t)\big[\varphi_2''(t) + b(t)\varphi_2'(t) + c(t)\varphi_2(t)\big] \\ &= C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t)\,. \end{aligned}$$

Therefore, C_1 and C_2 satisfy

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$
$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Let $W[\varphi_1, \varphi_2]$ denote the function $\varphi_1\varphi_2' - \varphi_1'\varphi_2$ (termed the **Wronskian** of φ_1 and φ_2). Then solving the system

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$

$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

we obtain that

$$C_1'(t) = -\frac{f(t)\varphi_2(t)}{W[\varphi_1, \varphi_2](t)} \qquad \text{and} \qquad C_2'(t) = \frac{f(t)\varphi_1(t)}{W[\varphi_1, \varphi_2](t)}\,.$$

As a consequence, a particular solution of (11) is given by

$$x_p(t) = -\varphi_1(t)\int \frac{f(t)\varphi_2(t)}{W[\varphi_1, \varphi_2](t)}\,dt + \varphi_2(t)\int \frac{f(t)\varphi_1(t)}{W[\varphi_1, \varphi_2](t)}\,dt\,.$$

We note that the indefinite integral has undetermined constants; thus the general solution to (11) is given by the formula above.

§2.2 Some Basic Techniques of Solving ODEs

Let $W[\varphi_1, \varphi_2]$ denote the function $\varphi_1\varphi_2' - \varphi_1'\varphi_2$ (termed the **Wronskian** of φ_1 and φ_2). Then solving the system

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$

$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

we obtain that

$$C_1'(t) = -\frac{f(t)\varphi_2(t)}{W[\varphi_1,\varphi_2](t)} \qquad \text{and} \qquad C_2'(t) = \frac{f(t)\varphi_1(t)}{W[\varphi_1,\varphi_2](t)}\,.$$

As a consequence, a particular solution of (11) is given by

$$x_p(t) = -\varphi_1(t)\int \frac{f(t)\varphi_2(t)}{W[\varphi_1,\varphi_2](t)}\,dt + \varphi_2(t)\int \frac{f(t)\varphi_1(t)}{W[\varphi_1,\varphi_2](t)}\,dt\,.$$

We note that the indefinite integral has undetermined constants; thus the general solution to (11) is given by the formula above.

§2.2 Some Basic Techniques of Solving ODEs

Let $W[\varphi_1, \varphi_2]$ denote the function $\varphi_1\varphi_2' - \varphi_1'\varphi_2$ (termed the **Wronskian** of φ_1 and φ_2). Then solving the system

$$C_1'(t)\varphi_1(t) + C_2'(t)\varphi_2(t) = 0\,,$$
$$C_1'(t)\varphi_1'(t) + C_2'(t)\varphi_2'(t) = f(t)\,.$$

we obtain that

$$C_1'(t) = -\frac{f(t)\varphi_2(t)}{W[\varphi_1, \varphi_2](t)} \qquad \text{and} \qquad C_2'(t) = \frac{f(t)\varphi_1(t)}{W[\varphi_1, \varphi_2](t)}\,.$$

As a consequence, a particular solution of (11) is given by

$$x_p(t) = -\varphi_1(t)\int \frac{f(t)\varphi_2(t)}{W[\varphi_1, \varphi_2](t)}\,dt + \varphi_2(t)\int \frac{f(t)\varphi_1(t)}{W[\varphi_1, \varphi_2](t)}\,dt\,.$$

We note that the indefinite integral has undetermined constants; thus the general solution to (11) is given by the formula above.

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the spring-mass system

$$m\ddot{x} + kx = F_0\,, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,,$$

where F_0 is a given constant force. Let $\varphi_1(t) = \cos(\omega t)$ and $\varphi_2(t) = \sin(\omega t)$, where $\omega = \sqrt{k/m}$. We note that previous example shows that $\{\varphi_1, \varphi_2\}$ is a basis of the solution space of the corresponding homogeneous ODE $m\ddot{x}+kx = 0$. To apply the formula of a particular solution, we first compute the Wronskain:

$$\begin{aligned} W[\varphi_1, \varphi_2](t) &= \varphi_1(t)\varphi_2'(t) - \varphi_1'(t)\varphi_2(t) \\ &= \omega\big[\cos^2(\omega t) + \sin^2(\omega t)\big] = \omega\,. \end{aligned}$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the spring-mass system

$$m\ddot{x} + kx = F_0 \,, \qquad x(0) = x_0 \,, \quad x'(0) = v_0 \,,$$

where F_0 is a given constant force. Let $\varphi_1(t) = \cos(\omega t)$ and $\varphi_2(t) = \sin(\omega t)$, where $\omega = \sqrt{k/m}$. We note that previous example shows that $\{\varphi_1, \varphi_2\}$ is a basis of the solution space of the corresponding homogeneous ODE $m\ddot{x} + kx = 0$. To apply the formula of a particular solution, we first compute the Wronskain:

$$\begin{aligned} W[\varphi_1, \varphi_2](t) &= \varphi_1(t)\varphi_2'(t) - \varphi_1'(t)\varphi_2(t) \\ &= \omega\big[\cos^2(\omega t) + \sin^2(\omega t)\big] = \omega \,. \end{aligned}$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the spring-mass system

$$m\ddot{x} + kx = F_0\,, \qquad x(0) = x_0\,, \quad x'(0) = v_0\,,$$

where F_0 is a given constant force. Let $\varphi_1(t) = \cos(\omega t)$ and $\varphi_2(t) = \sin(\omega t)$, where $\omega = \sqrt{k/m}$. We note that previous example shows that $\{\varphi_1, \varphi_2\}$ is a basis of the solution space of the corresponding homogeneous ODE $m\ddot{x}+kx = 0$. To apply the formula of a particular solution, we first compute the Wronskain:

$$\begin{aligned} W[\varphi_1, \varphi_2](t) &= \varphi_1(t)\varphi_2'(t) - \varphi_1'(t)\varphi_2(t) \\ &= \omega\big[\cos^2(\omega t) + \sin^2(\omega t)\big] = \omega\,. \end{aligned}$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

Having obtained the Wronskian, the formula for particular solutions to second-order linear ODEs implies that a particular solution to the ODE in the IVP is given by

$$x_p(t) = -\cos(\omega t)\int \frac{F_0/m \cdot \sin(\omega t)}{\omega}\,dt + \sin(\omega t)\int \frac{F_0/m \cdot \cos(\omega t)}{\omega}\,dt$$
$$= \frac{F_0}{m\omega^2}\big[\cos^2(\omega t) + \sin^2(\omega t)\big] = \frac{F_0}{k}\,.$$

Therefore, the general solution to the ODE in the IVP is

$$x(t) = C_1\cos(\omega t) + C_2\sin(\omega t) + \frac{F_0}{k}$$

and the initial conditions imply that $C_1 = x_0 - \dfrac{F_0}{k}$ and $C_2 = \dfrac{v_0}{\omega}$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion)

In this example we prove **Kepler's first and third laws of planetary motion**. Recall that in previous example we have shown that the polar coordinate (r, θ) of the location of a planet moving around a single sun satisfy a nonlinear second order ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2\,, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0\,. \tag{8b}$$

Since θ is one-to-one and continuously differentiable, the inverse function of θ exists and is also continuously differentiable (the Inverse Function Theorem for functions of one variable). Write $t = t(\theta)$, and every function of t can be viewed as a function of θ (via $f(t) \mapsto f(t(\theta))$).

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion)

In this example we prove **Kepler's first and third laws of planetary motion**. Recall that in previous example we have shown that the polar coordinate (r, θ) of the location of a planet moving around a single sun satisfy a nonlinear second order ODE

$$-\frac{GM}{r^2} = r'' - r(\theta')^2\,, \tag{8a}$$

$$2r'\theta' + r\theta'' = 0\,. \tag{8b}$$

Since θ is one-to-one and continuously differentiable, the inverse function of θ exists and is also continuously differentiable (the Inverse Function Theorem for functions of one variable). Write $t = t(\theta)$, and every function of t can be viewed as a function of θ (via $f(t) \mapsto f(t(\theta))$).

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\frac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\frac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \qquad \text{or equivalently,} \qquad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \frac{\ell}{m}\frac{\dot{r}}{r^2}$. Let $u = \frac{1}{r}$. Then $\dot{u} = -\frac{\dot{r}}{r^2}$ which implies that $r' = -\frac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\left(\theta'\right)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\dfrac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\dfrac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \quad \text{or equivalently,} \quad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot{r}}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot{u} = -\dfrac{\dot{r}}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\big(\theta'\big)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

For a function f of t, we let $\dot f(\theta)$ denote $\dfrac{d}{d\theta} f(t(\theta))$ and $\ddot f(\theta)$ denote $\dfrac{d^2}{d\theta^2} f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \qquad \text{or equivalently,} \qquad f' = \frac{\ell}{mr^2}\dot f;$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot r}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot u = -\dfrac{\dot r}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot u$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot u = -\frac{\ell^2}{m^2}\ddot u u^2;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\big(\theta'\big)^2 = -\frac{\ell^2}{m^2}\ddot u u^2 - \frac{\ell^2}{m^2}u^3.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\dfrac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\dfrac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \quad \text{or equivalently,} \quad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot{r}}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot{u} = -\dfrac{\dot{r}}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\big(\theta'\big)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

For a function f of t, we let $\dot f(\theta)$ denote $\dfrac{d}{d\theta} f(t(\theta))$ and $\ddot f(\theta)$ denote $\dfrac{d^2}{d\theta^2} f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \qquad \text{or equivalently,} \qquad f' = \frac{\ell}{mr^2}\dot f;$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot r}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot u = -\dfrac{\dot r}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot u$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot u = -\frac{\ell^2}{m^2}\ddot u u^2\,;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\big(\theta'\big)^2 = -\frac{\ell^2}{m^2}\ddot u u^2 - \frac{\ell^2}{m^2}u^3\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's $1^{\rm st}$ and $3^{\rm rd}$ laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\dfrac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\dfrac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \qquad \text{or equivalently,} \qquad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot{r}}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot{u} = -\dfrac{\dot{r}}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2\,;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GM\frac{1}{r^2} = r'' - r\big(\theta'\big)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's $1^{\rm st}$ and $3^{\rm rd}$ laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\dfrac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\dfrac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \quad \text{or equivalently,} \quad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot{r}}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot{u} = -\dfrac{\dot{r}}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2\,;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GMu^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - r\Big(\frac{\ell}{mr^2}\Big)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's $1^{\rm st}$ and $3^{\rm rd}$ laws of planetary motion - cont.)

For a function f of t, we let $\dot{f}(\theta)$ denote $\dfrac{d}{d\theta}f(t(\theta))$ and $\ddot{f}(\theta)$ denote $\dfrac{d^2}{d\theta^2}f(t(\theta))$. By the chain rule and the fact that $mr^2\theta' = \ell$,

$$\frac{d}{dt} = \frac{d\theta}{dt}\frac{d}{d\theta} = \theta'\frac{d}{d\theta} = \frac{\ell}{mr^2}\frac{d}{d\theta} \qquad \text{or equivalently,} \qquad f' = \frac{\ell}{mr^2}\dot{f};$$

thus $r' = \dfrac{\ell}{m}\dfrac{\dot{r}}{r^2}$. Let $u = \dfrac{1}{r}$. Then $\dot{u} = -\dfrac{\dot{r}}{r^2}$ which implies that $r' = -\dfrac{\ell}{m}\dot{u}$. Therefore,

$$r'' = -\frac{\ell}{m}\cdot\frac{\ell}{mr^2}\ddot{u} = -\frac{\ell^2}{m^2}\ddot{u}u^2\,;$$

thus (8a) and the fact that $mr^2\theta' = \ell$ show that

$$-GMu^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - r\Big(\frac{\ell}{mr^2}\Big)^2 = -\frac{\ell^2}{m^2}\ddot{u}u^2 - \frac{\ell^2}{m^2}u^3\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Therefore,

$$\ddot{u} + u = \frac{GMm^2}{\ell^2} = \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}.$$

A particular solution u_p to the ODE above is the constant function $u_p(\theta) = \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$; thus the general solution to the ODE above is

$$u(\theta) = C_1\cos\theta + C_2\sin\theta + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$$

for some constant $C \geqslant 0$ and angle ϕ. By the fact that $u = 1/r$, we find that the **polar equation for the orbit** of the planet is given by

$$r = \frac{1}{C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}} = \frac{A}{1 + \mathrm{e}\cos(\theta+\phi)},$$

where $A = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}$ and $\mathrm{e} = AC$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Therefore,

$$\ddot{u} + u = \frac{GMm^2}{\ell^2} = \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}\,.$$

A particular solution u_p to the ODE above is the constant function $u_p(\theta) = \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}$; thus the general solution to the ODE above is

$$u(\theta) = C_1 \cos\theta + C_2 \sin\theta + \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2} = C\cos(\theta + \phi) + \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}$$

for some constant $C \geqslant 0$ and angle ϕ. By the fact that $u = 1/r$, we find that the **polar equation for the orbit** of the planet is given by

$$r = \frac{1}{C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}} = \frac{A}{1 + \mathrm{e}\cos(\theta+\phi)}\,,$$

where $A = \frac{\mathrm{r}_0^2 \mathrm{v}_0^2}{GM}$ and $\mathrm{e} = AC$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Therefore,

$$\ddot{u} + u = \frac{GMm^2}{\ell^2} = \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}\,.$$

A particular solution u_p to the ODE above is the constant function $u_p(\theta) = \dfrac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}$; thus the general solution to the ODE above is

$$u(\theta) = C_1 \cos\theta + C_2 \sin\theta + \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2} = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}$$

for some constant $C \geqslant 0$ and angle ϕ. By the fact that $u = 1/r$, we find that the **polar equation for the orbit** of the planet is given by

$$r = \frac{1}{C\cos(\theta+\phi) + \dfrac{GM}{\mathrm{r}_0^2 \mathrm{v}_0^2}} = \frac{A}{1 + \mathrm{e}\cos(\theta+\phi)}\,,$$

where $A = \dfrac{\mathrm{r}_0^2 \mathrm{v}_0^2}{GM}$ and $\mathrm{e} = AC$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's $1^{\rm st}$ and $3^{\rm rd}$ laws of planetary motion - cont.)

Therefore,

$$\ddot{u} + u = \frac{GMm^2}{\ell^2} = \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}\,.$$

A particular solution u_p to the ODE above is the constant function $u_p(\theta) = \dfrac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$; thus the general solution to the ODE above is

$$u(\theta) = C_1\cos\theta + C_2\sin\theta + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$$

for some constant $C \geqslant 0$ and angle ϕ. By the fact that $u = 1/r$, we find that the **polar equation for the orbit** of the planet is given by

$$r = \frac{1}{C\cos(\theta+\phi) + \dfrac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}} = \frac{A}{1 + \mathrm{e}\cos(\theta+\phi)}\,,$$

where $A = \dfrac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}$ and $\mathrm{e} = AC$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's $1^{\rm st}$ and $3^{\rm rd}$ laws of planetary motion - cont.)

Therefore,

$$\ddot{u} + u = \frac{GMm^2}{\ell^2} = \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}\,.$$

A particular solution u_p to the ODE above is the constant function $u_p(\theta) = \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$; thus the general solution to the ODE above is

$$u(\theta) = C_1\cos\theta + C_2\sin\theta + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}$$

for some constant $C \geqslant 0$ and angle ϕ. By the fact that $u = 1/r$, we find that the **polar equation for the orbit** of the planet is given by

$$r = \frac{1}{C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}} = \frac{A}{1+\mathrm{e}\cos(\theta+\phi)}\,,$$

where $A = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}$ and $\mathrm{e} = AC$.

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

The polar equation of the orbit of a planet given by

$$r = \frac{A}{1 + \mathrm{e}\cos(\theta + \phi)}$$

represents a conic section (圓錐曲線) with eccentricity (離心率) e.

This proves **Kepler's first law of planetary motion:**

> The orbit of every planet is an ellipse with the Sun at one of the two foci.

Remark: The eccentricity e of a conic section C is a constant defined by

$$\mathrm{e} = \frac{\text{the distance from } P \text{ to the focus (焦點)}}{\text{the distance from } P \text{ to the directrix (準線)}} \quad \forall\, P \in C\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

The polar equation of the orbit of a planet given by

$$r = \frac{A}{1 + \mathrm{e}\cos(\theta + \phi)}$$

represents a conic section (圓錐曲線) with eccentricity (離心率) e. This proves **Kepler's first law of planetary motion:**

> The orbit of every planet is an ellipse with the Sun at one of the two foci.

Remark: The eccentricity e of a conic section C is a constant defined by

$$\mathrm{e} = \frac{\text{the distance from } P \text{ to the focus (焦點)}}{\text{the distance from } P \text{ to the directrix (準線)}} \qquad \forall\, P \in \mathrm{C}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

The polar equation of the orbit of a planet given by

$$r = \frac{A}{1 + \mathrm{e}\cos(\theta + \phi)}$$

represents a conic section (圓錐曲線) with eccentricity (離心率) e. This proves **Kepler's first law of planetary motion:**

> The orbit of every planet is an ellipse with the Sun at one of the two foci.

Remark: The eccentricity e of a conic section C is a constant defined by

$$\mathrm{e} = \frac{\text{the distance from } P \text{ to the focus (焦點)}}{\text{the distance from } P \text{ to the directrix (準線)}} \qquad \forall\, P \in \mathrm{C}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

$$\mathrm{e} = \frac{d(P,F)}{d(P,\mathrm{D})}$$

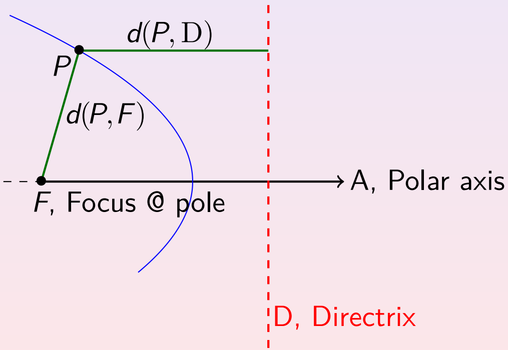


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

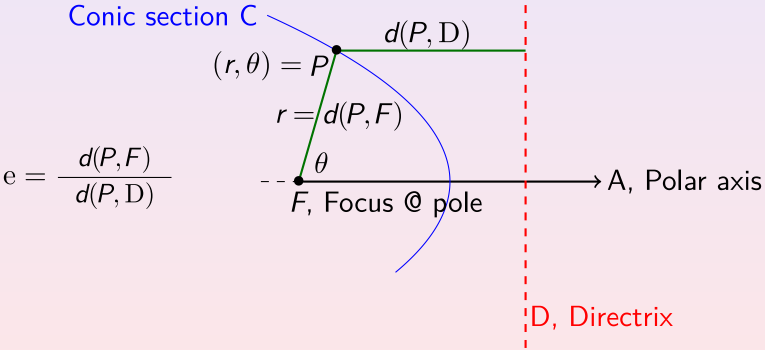


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

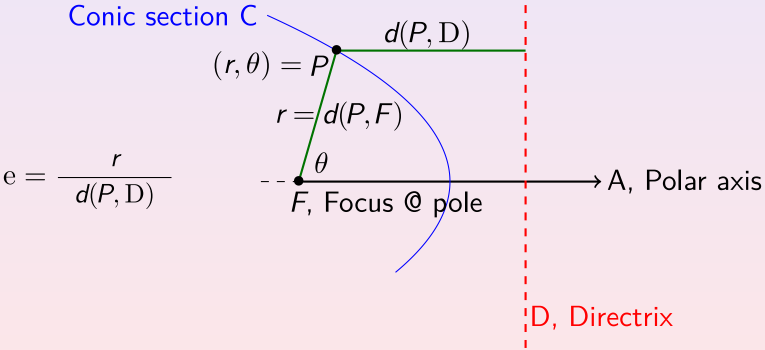


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

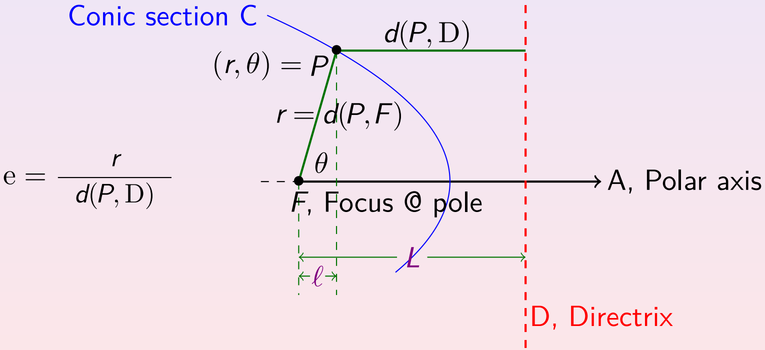


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

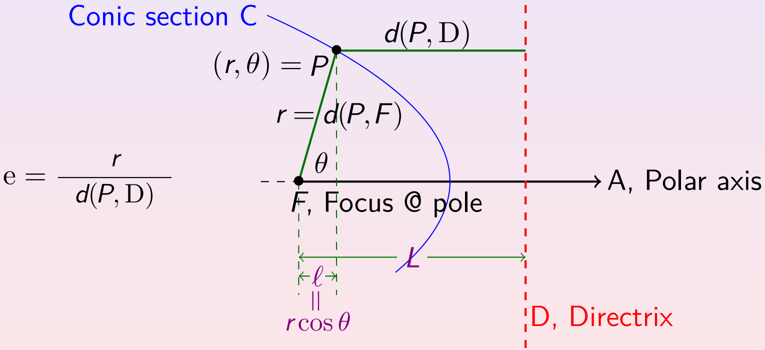


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

$$e = \frac{r}{L - r\cos\theta}$$

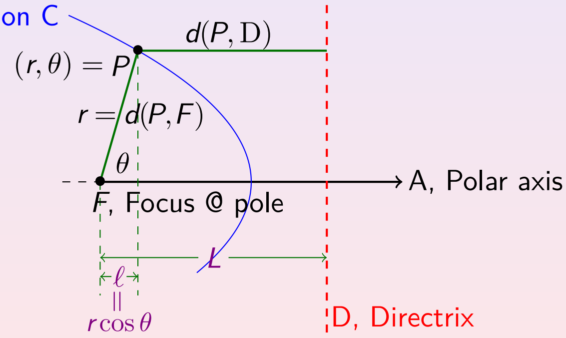


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

To derive the **polar equation for conic sections**, we introduce the polar coordinate in which the pole is (one of) the focus and the polar axis is perpendicular to and intersects the directrix.

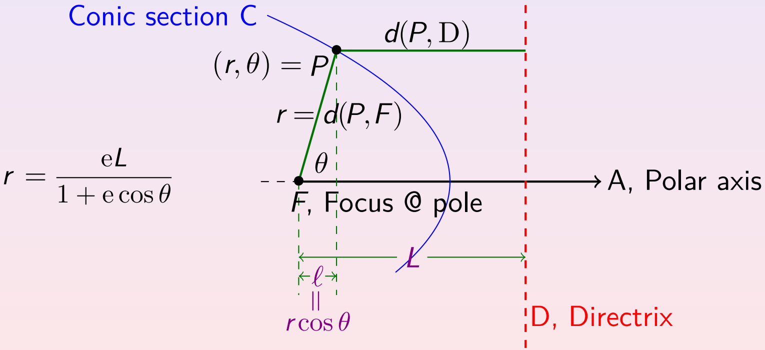


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

If a new polar axis A' is given by $\theta = \phi$ in the polar coordinate system with polar axis A, then the polar equation for a conic section with eccentricity e is given by $r = \dfrac{\mathrm{e}L}{1 + \mathrm{e}\cos(\theta + \phi)}$.

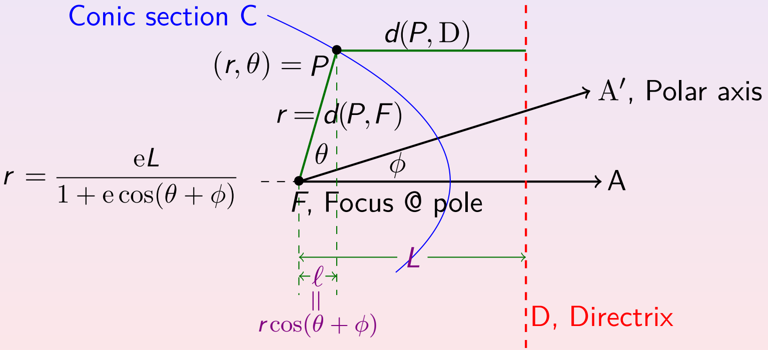


Figure 8: Polar representation of conic sections

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = \frac{1+\mathrm{e}\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$\frac{1}{\mathrm{r}_0} = C + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}.$$

This further implies that the eccentricity $\mathrm{e} \equiv AC = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}C$ is given by $\mathrm{e} = \frac{\mathrm{r}_0\mathrm{v}_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{A}{1+\mathrm{e}\cos(\theta+\phi)}.$$

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = \frac{1+\mathrm{e}\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$\frac{1}{\mathrm{r}_0} = C + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}.$$

This further implies that the eccentricity $\mathrm{e} \equiv AC = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}C$ is given by $\mathrm{e} = \frac{\mathrm{r}_0\mathrm{v}_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{A}{1+\mathrm{e}\cos(\theta+\phi)}.$$

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = \frac{1+\mathrm{e}\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$C = \frac{1}{\mathrm{r}_0} - \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}\,.$$

This further implies that the eccentricity $\mathrm{e} \equiv AC = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}C$ is given by $\mathrm{e} = \frac{\mathrm{r}_0\mathrm{v}_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{A}{1+\mathrm{e}\cos(\theta+\phi)}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{r_0^2 v_0^2} = \frac{1 + e\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$C = \frac{1}{r_0} - \frac{GM}{r_0^2 v_0^2}.$$

This further implies that the eccentricity $e \equiv AC = \frac{r_0^2 v_0^2}{GM} C$ is given by $e = \frac{r_0 v_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{A}{1 + e\cos(\theta+\phi)}.$$

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = \frac{1+\mathrm{e}\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$C = \frac{1}{\mathrm{r}_0} - \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}\,.$$

This further implies that the eccentricity $\mathrm{e} \equiv AC = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}C$ is given by $\mathrm{e} = \frac{\mathrm{r}_0\mathrm{v}_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{A}{1+\mathrm{e}\cos(\theta+\phi)}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Remark: Since we have proved that the orbit of a planet must be an ellipse, unlike the case of parabola or hyperbola the angular parameter θ in the equation

$$\frac{1}{r} = u = C\cos(\theta+\phi) + \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2} = \frac{1+\mathrm{e}\cos(\theta+\phi)}{A}$$

can be any real numbers. Therefore, the maximum of u is given by the reciprocal (倒數) of the perihelion and we have

$$C = \frac{1}{\mathrm{r}_0} - \frac{GM}{\mathrm{r}_0^2\mathrm{v}_0^2}\,.$$

This further implies that the eccentricity $\mathrm{e} \equiv AC = \frac{\mathrm{r}_0^2\mathrm{v}_0^2}{GM}C$ is given by $\mathrm{e} = \frac{\mathrm{r}_0\mathrm{v}_0^2}{GM} - 1$ and the polar equation of the ellipse is given by

$$r = \frac{(1+\mathrm{e})\mathrm{r}_0}{1+\mathrm{e}\cos(\theta+\phi)}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Recall that **Kepler's second law** of planetary motion shows that

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\,dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\,dt = \frac{\ell(t_2 - t_1)}{2m} = \frac{\mathrm{r}_0\mathrm{v}_0}{2}(t_2 - t_1)\,.$$

Let a, b be the semi-major axis (半長軸) and semi-minor axis (半短軸) of the orbit of a planet, and T be the orbital period (公轉週期). Then the identity above shows that

$$\pi ab = \int_0^{\mathrm{T}} \frac{1}{2} r^2\theta'\,dt = \frac{\mathrm{r}_0\mathrm{v}_0\mathrm{T}}{2}\,.$$

Therefore, by the fact that $b = a\sqrt{1 - \mathrm{e}^2}$,

$$\mathrm{T}^2 = \left(\frac{2\pi ab}{\mathrm{r}_0\mathrm{v}_0}\right)^2 = \frac{4\pi^2 a^4}{\mathrm{r}_0^2\mathrm{v}_0^2}(1 - \mathrm{e}^2) = \frac{4\pi^2 a^4}{GM} \cdot \frac{2GM - \mathrm{r}_0\mathrm{v}_0^2}{\mathrm{r}_0 GM}\,. \tag{14}$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Recall that **Kepler's second law** of planetary motion shows that

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\,dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\,dt = \frac{\ell(t_2-t_1)}{2m} = \frac{\mathrm{r}_0\mathrm{v}_0}{2}(t_2-t_1)\,.$$

Let a, b be the semi-major axis (半長軸) and semi-minor axis (半短軸) of the orbit of a planet, and T be the orbital period (公轉週期). Then the identity above shows that

$$\pi ab = \int_0^{\mathrm{T}} \frac{1}{2} r^2\theta'\,dt = \frac{\mathrm{r}_0\mathrm{v}_0\mathrm{T}}{2}\,.$$

Therefore, by the fact that $b = a\sqrt{1-\mathrm{e}^2}$,

$$\mathrm{T}^2 = \left(\frac{2\pi ab}{\mathrm{r}_0\mathrm{v}_0}\right)^2 = \frac{4\pi^2a^4}{\mathrm{r}_0^2\mathrm{v}_0^2}(1-\mathrm{e}^2) = \frac{4\pi^2a^4}{GM}\cdot\frac{2GM-\mathrm{r}_0\mathrm{v}_0^2}{\mathrm{r}_0GM}\,. \qquad (14)$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Recall that **Kepler's second law** of planetary motion shows that

$$\int_{t_1}^{t_2} \frac{1}{2} r^2(t)\theta'(t)\,dt = \int_{t_1}^{t_2} \frac{\ell}{2m}\,dt = \frac{\ell(t_2-t_1)}{2m} = \frac{\mathrm{r}_0\mathrm{v}_0}{2}(t_2-t_1)\,.$$

Let a, b be the semi-major axis (半長軸) and semi-minor axis (半短軸) of the orbit of a planet, and T be the orbital period (公轉週期). Then the identity above shows that

$$\pi ab = \int_0^{\mathrm{T}} \frac{1}{2} r^2\theta'\,dt = \frac{\mathrm{r}_0\mathrm{v}_0\mathrm{T}}{2}\,.$$

Therefore, by the fact that $b = a\sqrt{1-\mathrm{e}^2}$,

$$\mathrm{T}^2 = \Big(\frac{2\pi ab}{\mathrm{r}_0\mathrm{v}_0}\Big)^2 = \frac{4\pi^2 a^4}{\mathrm{r}_0^2\mathrm{v}_0^2}(1-\mathrm{e}^2) = \frac{4\pi^2 a^4}{GM}\cdot\frac{2GM-\mathrm{r}_0\mathrm{v}_0^2}{\mathrm{r}_0 GM}\,. \tag{14}$$

§2.2 Some Basic Techniques of Solving ODEs

Example (Kepler's 1^{st} and 3^{rd} laws of planetary motion - cont.)

Moreover, the polar equation $r = \dfrac{\mathrm{r}_0(1+\mathrm{e})}{1+\mathrm{e}\cos(\theta+\phi)}$ implies that

$$r_{\max} = r\Big|_{\theta+\phi=\pi} = \mathrm{r}_0\frac{1+\mathrm{e}}{1-\mathrm{e}};$$

thus using the expression of e,

$$a = \frac{\mathrm{r}_0 + r_{\max}}{2} = \frac{\mathrm{r}_0}{1-\mathrm{e}} = \frac{\mathrm{r}_0 GM}{2GM - \mathrm{r}_0\mathrm{v}_0^2}.$$

Using the identity above in (14), we conclude that $\mathrm{T}^2 = \dfrac{4\pi^2}{GM}a^3$ which shows **the third law of Kepler:**

> The square of the orbital period of a planet is directly proportional to the cube of the semi-major axis of its orbit.

§2.2 Some Basic Techniques of Solving ODEs

§2.2.4 Linear systems with constant coefficients

A general linear system of ODEs takes the form

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + f_1(t)\,,\\
\frac{dx_2}{dt} &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + f_2(t)\,,\\
\vdots \quad &\qquad\qquad\qquad \vdots\\
\frac{dx_n}{dt} &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t)\,,
\end{aligned}$$

where the coefficients a_{ij}, where $1 \leqslant i,j \leqslant n$, and the forcing $f_1, \cdots, f_n$ are given functions. The linear system above is said to be **homogeneous** if $f_i(t) = 0$ for all $1 \leqslant i \leqslant n$; otherwise it is in-homogeneous. In this sub-section, we look for solutions of a linear system when all the a_{ij}'s are constant functions.

§2.2 Some Basic Techniques of Solving ODEs

§2.2.4 Linear systems with constant coefficients

A general linear system of ODEs takes the form

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + f_1(t)\,,\\
\frac{dx_2}{dt} &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + f_2(t)\,,\\
&\ \ \vdots \qquad\qquad\qquad\qquad \vdots\\
\frac{dx_n}{dt} &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t)\,,
\end{aligned}$$

where the coefficients a_{ij}, where $1 \leqslant i,j \leqslant n$, and the forcing $f_1, \cdots, f_n$ are given functions. The linear system above is said to be **homogeneous** if $f_i(t) = 0$ for all $1 \leqslant i \leqslant n$; otherwise it is in-homogeneous. In this sub-section, we look for solutions of a linear system when all the a_{ij}'s are constant functions.

§2.2 Some Basic Techniques of Solving ODEs

§2.2.4 Linear systems with constant coefficients

A general linear system of ODEs takes the form

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n + f_1(t)\,,\\
\frac{dx_2}{dt} &= a_{21}(t)x_1 + a_{22}(t)x_2 + \cdots + a_{2n}(t)x_n + f_2(t)\,,\\
&\vdots \qquad\qquad\qquad\qquad \vdots\\
\frac{dx_n}{dt} &= a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t)\,,
\end{aligned}$$

where the coefficients a_{ij}, where $1 \leqslant i,j \leqslant n$, and the forcing $f_1, \cdots, f_n$ are given functions. The linear system above is said to be **homogeneous** if $f_i(t) = 0$ for all $1 \leqslant i \leqslant n$; otherwise it is inhomogeneous. In this sub-section, we look for solutions of a linear system when all the a_{ij}'s are constant functions.

§2.2 Some Basic Techniques of Solving ODEs

In other words, we look for vector-valued function

$$\boldsymbol{x}(t) = \left[x_1(t), \cdots, x_n(t)\right]^{\mathrm{T}}$$

satisfying the ODE

$$\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t)\,,$$

where $\boldsymbol{A} = [a_{ij}]_{n\times n}$ is a constant matrix, $\boldsymbol{f}(t) = \left[f_1(t), \cdots, f_n(t)\right]^{\mathrm{T}}$.

We mimic the method of integrating factor and look for a matrix-valued function $\boldsymbol{M} = \boldsymbol{M}(t)$ such that

$$\frac{d}{dt}\left[\boldsymbol{M}(t)\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\left[\boldsymbol{x}'(t) - \boldsymbol{A}\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\boldsymbol{f}(t)\,.$$

This amounts to choose $\boldsymbol{M}$ satisfying

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

In other words, we look for vector-valued function

$$\boldsymbol{x}(t) = \left[x_1(t), \cdots, x_n(t)\right]^{\mathrm{T}}$$

satisfying the ODE

$$\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t),$$

where $\boldsymbol{A} = [a_{ij}]_{n\times n}$ is a constant matrix, $\boldsymbol{f}(t) = \left[f_1(t), \cdots, f_n(t)\right]^{\mathrm{T}}$.

We mimic the method of integrating factor and look for a matrix-valued function $\boldsymbol{M} = \boldsymbol{M}(t)$ such that

$$\frac{d}{dt}\left[\boldsymbol{M}(t)\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\left[\boldsymbol{x}'(t) - \boldsymbol{A}\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\boldsymbol{f}(t).$$

This amounts to choose $\boldsymbol{M}$ satisfying

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}.$$

§2.2 Some Basic Techniques of Solving ODEs

In other words, we look for vector-valued function

$$\boldsymbol{x}(t) = \left[x_1(t), \cdots, x_n(t)\right]^{\mathrm{T}}$$

satisfying the ODE

$$\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t)\,,$$

where $\boldsymbol{A} = [a_{ij}]_{n\times n}$ is a constant matrix, $\boldsymbol{f}(t) = \left[f_1(t), \cdots, f_n(t)\right]^{\mathrm{T}}$.

We mimic the method of integrating factor and look for a matrix-valued function $\boldsymbol{M} = \boldsymbol{M}(t)$ such that

$$\frac{d}{dt}\left[\boldsymbol{M}(t)\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\left[\boldsymbol{x}'(t) - \boldsymbol{A}\boldsymbol{x}(t)\right] = \boldsymbol{M}(t)\boldsymbol{f}(t)\,.$$

This amounts to choose $\boldsymbol{M}$ satisfying

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}\,.$$

§2.2 Some Basic Techniques of Solving ODEs

We note that the ODE $\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}$ shows that

$$\frac{d\boldsymbol{M}}{dt}(0) = -\boldsymbol{M}(0)\boldsymbol{A}, \quad \frac{d^2\boldsymbol{M}}{dt^2}(0) = -\frac{d\boldsymbol{M}}{dt}(0)\boldsymbol{A} = \boldsymbol{M}(0)\boldsymbol{A}^2,$$
$$\frac{d^3\boldsymbol{M}}{dt^3}(0) = -\frac{d^2\boldsymbol{M}}{dt^2}(0)\boldsymbol{A} = -\boldsymbol{M}(0)\boldsymbol{A}^3$$

and inductively we obtain $\frac{d^k\boldsymbol{M}}{dt^k}(0) = (-1)^k\boldsymbol{M}(0)\boldsymbol{A}^k$. Therefore, using the Taylor expansion we formally obtain that

$$\boldsymbol{M}(t) = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{d^k\boldsymbol{M}}{dt^k}(0)t^k = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}t^k\boldsymbol{M}(0)\boldsymbol{A}^k$$
$$= \boldsymbol{M}(0)\sum_{k=0}^{\infty}\frac{1}{k!}(-t\boldsymbol{A})^k.$$

This motivates the following

§2.2 Some Basic Techniques of Solving ODEs

We note that the ODE $\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}$ shows that

$$\frac{d\boldsymbol{M}}{dt}(0) = -\boldsymbol{M}(0)\boldsymbol{A}, \quad \frac{d^2\boldsymbol{M}}{dt^2}(0) = -\frac{d\boldsymbol{M}}{dt}(0)\boldsymbol{A} = \boldsymbol{M}(0)\boldsymbol{A}^2,$$
$$\frac{d^3\boldsymbol{M}}{dt^3}(0) = -\frac{d^2\boldsymbol{M}}{dt^2}(0)\boldsymbol{A} = -\boldsymbol{M}(0)\boldsymbol{A}^3$$

and inductively we obtain $\frac{d^k\boldsymbol{M}}{dt^k}(0) = (-1)^k\boldsymbol{M}(0)\boldsymbol{A}^k$. Therefore, using the Taylor expansion we formally obtain that

$$\boldsymbol{M}(t) = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{d^k\boldsymbol{M}}{dt^k}(0)t^k = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}t^k\boldsymbol{M}(0)\boldsymbol{A}^k$$
$$= \boldsymbol{M}(0)\sum_{k=0}^{\infty}\frac{1}{k!}(-t\boldsymbol{A})^k.$$

This motivates the following

§2.2 Some Basic Techniques of Solving ODEs

We note that the ODE $\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A}$ shows that

$$\frac{d\boldsymbol{M}}{dt}(0) = -\boldsymbol{M}(0)\boldsymbol{A}, \quad \frac{d^2\boldsymbol{M}}{dt^2}(0) = -\frac{d\boldsymbol{M}}{dt}(0)\boldsymbol{A} = \boldsymbol{M}(0)\boldsymbol{A}^2,$$
$$\frac{d^3\boldsymbol{M}}{dt^3}(0) = -\frac{d^2\boldsymbol{M}}{dt^2}(0)\boldsymbol{A} = -\boldsymbol{M}(0)\boldsymbol{A}^3$$

and inductively we obtain $\frac{d^k\boldsymbol{M}}{dt^k}(0) = (-1)^k\boldsymbol{M}(0)\boldsymbol{A}^k$. Therefore, using the Taylor expansion we formally obtain that

$$\boldsymbol{M}(t) = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{d^k\boldsymbol{M}}{dt^k}(0)t^k = \sum_{k=0}^{\infty}\frac{(-1)^k}{k!}t^k\boldsymbol{M}(0)\boldsymbol{A}^k$$
$$= \boldsymbol{M}(0)\sum_{k=0}^{\infty}\frac{1}{k!}(-t\boldsymbol{A})^k.$$

This motivates the following

§2.2 Some Basic Techniques of Solving ODEs

Definition

Let $\boldsymbol{B}$ be an $n \times n$ matrix. The exponential of $\boldsymbol{B}$, denoted by $e^{\boldsymbol{B}}$, is the series

$$e^{\boldsymbol{B}} = \mathbf{I} + \boldsymbol{B} + \frac{1}{2!}\boldsymbol{B}^2 + \frac{1}{3!}\boldsymbol{B}^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{B}^k .$$

Having defined the exponential of square matrices, we conclude that

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A} \quad \Leftrightarrow \quad \boldsymbol{M}(t) = \boldsymbol{M}(0)e^{-t\boldsymbol{A}} . \tag{15}$$

Remark: We note that the exponential of square matrices is given by an infinite series, so in principle we should check the convergence of the series before we can define it. Nevertheless, we will treat the convergence of the series as a fact for this requires some additional knowledge in analysis.

§2.2 Some Basic Techniques of Solving ODEs

Definition

Let $\boldsymbol{B}$ be an $n \times n$ matrix. The exponential of $\boldsymbol{B}$, denoted by $e^{\boldsymbol{B}}$, is the series

$$e^{\boldsymbol{B}} = \mathbf{I} + \boldsymbol{B} + \frac{1}{2!}\boldsymbol{B}^2 + \frac{1}{3!}\boldsymbol{B}^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{B}^k .$$

Having defined the exponential of square matrices, we conclude that

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A} \quad \Leftrightarrow \quad \boldsymbol{M}(t) = \boldsymbol{M}(0)e^{-t\boldsymbol{A}} . \tag{15}$$

Remark: We note that the exponential of square matrices is given by an infinite series, so in principle we should check the convergence of the series before we can define it. Nevertheless, we will treat the convergence of the series as a fact for this requires some additional knowledge in analysis.

§2.2 Some Basic Techniques of Solving ODEs

Definition

Let $\boldsymbol{B}$ be an $n \times n$ matrix. The exponential of $\boldsymbol{B}$, denoted by $e^{\boldsymbol{B}}$, is the series

$$e^{\boldsymbol{B}} = \mathbf{I} + \boldsymbol{B} + \frac{1}{2!}\boldsymbol{B}^2 + \frac{1}{3!}\boldsymbol{B}^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{B}^k .$$

Having defined the exponential of square matrices, we conclude that

$$\frac{d}{dt}\boldsymbol{M}(t) = -\boldsymbol{M}(t)\boldsymbol{A} \quad \Leftrightarrow \quad \boldsymbol{M}(t) = \boldsymbol{M}(0)e^{-t\boldsymbol{A}} . \tag{15}$$

Remark: We note that the exponential of square matrices is given by an infinite series, so in principle we should check the convergence of the series before we can define it. Nevertheless, we will treat the convergence of the series as a fact for this requires some additional knowledge in analysis.

§2.2 Some Basic Techniques of Solving ODEs

Before proceeding, let us establish a fundamental identity

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} = e^{(t+s)\boldsymbol{B}} \quad \text{for all square matrices } \boldsymbol{B} \text{ and } t, s \in \mathbb{R}\,.$$

To see this, we note that $e^{t\boldsymbol{B}}\boldsymbol{B} = \boldsymbol{B}e^{t\boldsymbol{B}}$ for all $t \in \mathbb{R}$ and

$$\frac{d}{dt}e^{t\boldsymbol{B}} = e^{t\boldsymbol{B}}\boldsymbol{B}\,.$$

Therefore, for each given $s \in \mathbb{R}$,

$$\frac{d}{dt}\Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big] = e^{t\boldsymbol{B}}\boldsymbol{B}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\boldsymbol{B} = \Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big]\boldsymbol{B}\,.$$

Using (15),

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}} = \Big[e^{0\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(0+s)\boldsymbol{B}}\Big]e^{t\boldsymbol{B}} = 0\,;$$

thus we establish the desired identity.

§2.2 Some Basic Techniques of Solving ODEs

Before proceeding, let us establish a fundamental identity

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} = e^{(t+s)\boldsymbol{B}} \quad \text{for all square matrices } \boldsymbol{B} \text{ and } t, s \in \mathbb{R}\,.$$

To see this, we note that $e^{t\boldsymbol{B}}\boldsymbol{B} = \boldsymbol{B}e^{t\boldsymbol{B}}$ for all $t \in \mathbb{R}$ and

$$\frac{d}{dt}e^{t\boldsymbol{B}} = e^{t\boldsymbol{B}}\boldsymbol{B}\,.$$

Therefore, for each given $s \in \mathbb{R}$,

$$\frac{d}{dt}\Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big] = e^{t\boldsymbol{B}}\boldsymbol{B}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\boldsymbol{B} = \Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big]\boldsymbol{B}\,.$$

Using (15),

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}} = \Big[e^{0\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(0+s)\boldsymbol{B}}\Big]e^{t\boldsymbol{B}} = 0\,;$$

thus we establish the desired identity.

§2.2 Some Basic Techniques of Solving ODEs

Before proceeding, let us establish a fundamental identity

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} = e^{(t+s)\boldsymbol{B}} \quad \text{for all square matrices } \boldsymbol{B} \text{ and } t, s \in \mathbb{R}\,.$$

To see this, we note that $e^{t\boldsymbol{B}}\boldsymbol{B} = \boldsymbol{B}e^{t\boldsymbol{B}}$ for all $t \in \mathbb{R}$ and

$$\frac{d}{dt}e^{t\boldsymbol{B}} = e^{t\boldsymbol{B}}\boldsymbol{B}\,.$$

Therefore, for each given $s \in \mathbb{R}$,

$$\frac{d}{dt}\Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big] = e^{t\boldsymbol{B}}\boldsymbol{B}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\boldsymbol{B} = \Big[e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}}\Big]\boldsymbol{B}\,.$$

Using (15),

$$e^{t\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(t+s)\boldsymbol{B}} = \Big[e^{0\boldsymbol{B}}e^{s\boldsymbol{B}} - e^{(0+s)\boldsymbol{B}}\Big]e^{t\boldsymbol{B}} = 0\,;$$

thus we establish the desired identity.

§2.2 Some Basic Techniques of Solving ODEs

Now we come back to solve for the ODE $\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t)$. We choose $\boldsymbol{M}(0) = \mathbf{I}$ so that an integrating factor $\boldsymbol{M}$ is given by $\boldsymbol{M}(t) = e^{-t\boldsymbol{A}}$. Therefore,

$$\frac{d}{dt}\big[e^{-t\boldsymbol{A}}\boldsymbol{x}(t)\big] = e^{-t\boldsymbol{A}}\big[\boldsymbol{x}'(t) - \boldsymbol{A}\boldsymbol{x}(t)\big] = e^{-t\boldsymbol{A}}\boldsymbol{f}(t)\,.$$

The equation above shows that

$$\boldsymbol{x}(t) = e^{t\boldsymbol{A}}\int e^{-t\boldsymbol{A}}\boldsymbol{f}(t)\,dt\,.$$

If an initial condition $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is imposed, the unique solution to the IVP is given by

$$\boldsymbol{x}(t) = e^{(t-t_0)\boldsymbol{A}}\boldsymbol{x}_0 + e^{t\boldsymbol{A}}\int_{t_0}^{t} e^{-s\boldsymbol{A}}\boldsymbol{f}(s)\,ds\,.$$

§2.2 Some Basic Techniques of Solving ODEs

Now we come back to solve for the ODE $\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t)$. We choose $\boldsymbol{M}(0) = \mathbf{I}$ so that an integrating factor $\boldsymbol{M}$ is given by $\boldsymbol{M}(t) = e^{-t\boldsymbol{A}}$. Therefore,

$$\frac{d}{dt}\big[e^{-t\boldsymbol{A}}\boldsymbol{x}(t)\big] = e^{-t\boldsymbol{A}}\big[\boldsymbol{x}'(t) - \boldsymbol{A}\boldsymbol{x}(t)\big] = e^{-t\boldsymbol{A}}\boldsymbol{f}(t)\,.$$

The equation above shows that

$$\boldsymbol{x}(t) = e^{t\boldsymbol{A}}\int e^{-t\boldsymbol{A}}\boldsymbol{f}(t)\,dt\,.$$

If an initial condition $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is imposed, the unique solution to the IVP is given by

$$\boldsymbol{x}(t) = e^{(t-t_0)\boldsymbol{A}}\boldsymbol{x}_0 + e^{t\boldsymbol{A}}\int_{t_0}^{t} e^{-s\boldsymbol{A}}\boldsymbol{f}(s)\,ds\,.$$

§2.2 Some Basic Techniques of Solving ODEs

- **The computation of e^{tB} for square matrix B**

By Jordan decomposition, every square matrix B can be written as

$$B = PJP^{-1},$$

where J takes the Jordan canonical form

$$J = \begin{bmatrix} J_1 & O & \cdots & O \\ O & J_2 & \ddots & O \\ \vdots & \ddots & \ddots & \vdots \\ O & \cdots & O & J_\ell \end{bmatrix}$$

in which each Jordan block J_r is a square matrix of the form $\lambda \mathbf{I}$ or

$$J_r = \lambda \mathbf{I} + \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

for some eigenvalue of B.

§2.2 Some Basic Techniques of Solving ODEs

Writing $\boldsymbol{B}$ in the form above, by the fact that $\boldsymbol{B}^k = \boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}$ we have

$$e^{t\boldsymbol{B}} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{B}^k = \boldsymbol{P}\Big(\sum_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{J}^k\Big)\boldsymbol{P}^{-1}.$$

Since $\boldsymbol{J}^k = \begin{bmatrix} \boldsymbol{J}_1^k & & & \\ & \boldsymbol{J}_2^k & & \\ & & \ddots & \\ & & & \boldsymbol{J}_\ell^k \end{bmatrix}$, we have

$$e^{t\boldsymbol{B}} = \boldsymbol{P}\begin{bmatrix} \sum\limits_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{J}_1^k & & & \\ & \sum\limits_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{J}_2^k & & \\ & & \ddots & \\ & & & \sum\limits_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{J}_\ell^k \end{bmatrix}\boldsymbol{P}^{-1}.$$

§2.2 Some Basic Techniques of Solving ODEs

Writing $\boldsymbol{B}$ in the form above, by the fact that $\boldsymbol{B}^k = \boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}$ we have

$$e^{t\boldsymbol{B}} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{B}^k = \boldsymbol{P}\Big(\sum_{k=0}^{\infty} \frac{1}{k!} t^k \boldsymbol{J}^k\Big)\boldsymbol{P}^{-1}.$$

Since $\boldsymbol{J}^k = \begin{bmatrix} \boldsymbol{J}_1^k & & & \\ & \boldsymbol{J}_2^k & & \\ & & \ddots & \\ & & & \boldsymbol{J}_\ell^k \end{bmatrix}$, we have

$$e^{t\boldsymbol{B}} = \boldsymbol{P} \begin{bmatrix} e^{t\boldsymbol{J}_1} & & & \\ & e^{t\boldsymbol{J}_2} & & \\ & & \ddots & \\ & & & e^{t\boldsymbol{J}_\ell} \end{bmatrix} \boldsymbol{P}^{-1}.$$

§2.2 Some Basic Techniques of Solving ODEs

For each r, since $\mathbf{I}$ commutes with the matrix $\mathbf{N} \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}$,

we have

$$\boldsymbol{J}_r^k = \left(\lambda\mathbf{I} + \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}\right)^k = \sum_{j=0}^{k} C_j^k \lambda^{k-j} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}^j .$$

Here the zeroth power of a square matrix is the identity matrix.

§2.2 Some Basic Techniques of Solving ODEs

By the fact that

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}^j = \begin{bmatrix} \overbrace{0 & \cdots & 0}^{j \text{ columns of } 0\text{'s here}} & 1 & 0 & \cdots & \cdots & 0 \\ 0 & & \ddots & 0 & 1 & 0 & \cdots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & 0 & 1 & 0 \\ \vdots & & & & & \ddots & 0 & 1 \\ \vdots & & & & & & \ddots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix},$$

§2.2 Some Basic Techniques of Solving ODEs

if $\boldsymbol{J}_r$ is an $m \times m$ matrix, we have

$$\boldsymbol{J}_r^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & C_2^k\lambda^{k-2} & \cdots & \cdots & C_{m-1}^k\lambda^{k-m+1} \\ 0 & \lambda^k & k\lambda^{k-1} & \ddots & \ddots & C_{m-2}^k\lambda^{k-m+2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & C_2^k\lambda^{k-2} \\ \vdots & \cdots & \cdots & 0 & \lambda^k & k\lambda^{k-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda^k \end{bmatrix}.$$

Here C_m^k is the number $\dfrac{k(k-1)\cdots(k-m+1)}{m!}$ so that $C_m^k = 0$ if $m > k$.

§2.2 Some Basic Techniques of Solving ODEs

Therefore, if J_r is an $m \times m$ matrix taking the form $J_r = \lambda \mathbf{I} + \mathbf{N}$,

$$\sum_{k=0}^{\infty} \frac{1}{k!} t^k J_r^k = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & \cdots & \cdots & \frac{t^{m-1}}{(m-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \ddots & \ddots & \frac{t^{m-2}}{(m-2)!}e^{\lambda t} \\ \vdots & 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} \\ \vdots & \vdots & \ddots & \ddots & 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & e^{\lambda t} \end{bmatrix}.$$

§2.2 Some Basic Techniques of Solving ODEs

Therefore, if J_r is an $m \times m$ matrix taking the form $J_r = \lambda\mathbf{I} + \mathbf{N}$,

$$e^{tJ_r} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & \cdots & \cdots & \frac{t^{m-1}}{(m-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \ddots & \ddots & \frac{t^{m-2}}{(m-2)!}e^{\lambda t} \\ \vdots & 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} \\ \vdots & \vdots & \ddots & \ddots & 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & e^{\lambda t} \end{bmatrix}.$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Let $\boldsymbol{J} = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$. Then J takes the form $\begin{bmatrix} \boldsymbol{J}_1 & & \\ & \boldsymbol{J}_2 & \\ & & \boldsymbol{J}_3 \end{bmatrix}$

so that

$$e^{t\boldsymbol{J}} = \begin{bmatrix} e^{2t} & te^{2t} & \frac{t^2}{2}e^{2t} & 0 & 0 & 0 \\ 0 & e^{2t} & te^{2t} & 0 & 0 & 0 \\ 0 & 0 & e^{2t} & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-3t} & te^{-3t} & 0 \\ 0 & 0 & 0 & 0 & e^{-3t} & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{5t} \end{bmatrix}.$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the ODE derived from studying the two masses three springs system:

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 + k_2(x_2 - x_1) + F_1 \,,$$
$$m_2 \frac{d^2 x_2}{dt^2} = -k_2(x_2 - x_1) - k_3 x_2 + F_2 \,.$$

Let $\boldsymbol{y} = [x_1, x_2, x_1', x_2']^{\mathrm{T}}$. Then

$$\boldsymbol{y}'(t) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{f} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\dfrac{k_1 + k_2}{m_1} & \dfrac{k_2}{m_1} & 0 & 0 \\ \dfrac{k_2}{m_2} & -\dfrac{k_2 + k_3}{m_2} & 0 & 0 \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{F_1}{m_1} \\ \dfrac{F_2}{m_2} \end{bmatrix} .$$

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the ODE derived from studying the two masses three springs system:

$$m_1 \frac{d^2 x_1}{dt^2} = -k_1 x_1 + k_2(x_2 - x_1) + F_1\,,$$
$$m_2 \frac{d^2 x_2}{dt^2} = -k_2(x_2 - x_1) - k_3 x_2 + F_2\,.$$

Let $\boldsymbol{y} = [x_1, x_2, x_1', x_2']^{\mathrm{T}}$. Then

$$\boldsymbol{y}'(t) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{f} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\dfrac{k_1 + k_2}{m_1} & \dfrac{k_2}{m_1} & 0 & 0 \\ \dfrac{k_2}{m_2} & -\dfrac{k_2 + k_3}{m_2} & 0 & 0 \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{F_1}{m_1} \\ \dfrac{F_2}{m_2} \end{bmatrix}.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

Suppose that $m_1 = m_2 = k_1 = k_2 = k_3 = 1$. If λ is an eigenvalue of $\boldsymbol{A}$, then

$$0 = \begin{vmatrix} -\lambda & 0 & 1 & 0 \\ 0 & -\lambda & 0 & 1 \\ -2 & 1 & -\lambda & 0 \\ 1 & -2 & 0 & -\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 1 \\ -2 & 0 & -\lambda \end{vmatrix} + \begin{vmatrix} 0 & -\lambda & 1 \\ -2 & 1 & 0 \\ 1 & -2 & -\lambda \end{vmatrix}$$

$$= -\lambda(-\lambda^3 - 2\lambda) + (4 - 1 + 2\lambda^2) = \lambda^4 + 4\lambda^2 + 3$$

$$= (\lambda^2 + 3)(\lambda^2 + 1)$$

which implies that the eigenvalues of $\boldsymbol{A}$ are $\pm\sqrt{3}i$ and $\pm i$. Corresponding eigenvectors are

$$\pm\sqrt{3}i \leftrightarrow \Big[\pm\frac{1}{\sqrt{3}}i, \mp\frac{1}{\sqrt{3}}i, -1, 1\Big]^{\mathrm{T}},$$

$$\pm i \leftrightarrow [\mp i, \mp i, 1, 1]^{\mathrm{T}};$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

Suppose that $m_1 = m_2 = k_1 = k_2 = k_3 = 1$. If λ is an eigenvalue of $\boldsymbol{A}$, then

$$0 = \begin{vmatrix} -\lambda & 0 & 1 & 0 \\ 0 & -\lambda & 0 & 1 \\ -2 & 1 & -\lambda & 0 \\ 1 & -2 & 0 & -\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 1 \\ -2 & 0 & -\lambda \end{vmatrix} + \begin{vmatrix} 0 & -\lambda & 1 \\ -2 & 1 & 0 \\ 1 & -2 & -\lambda \end{vmatrix}$$

$$= -\lambda(-\lambda^3 - 2\lambda) + (4 - 1 + 2\lambda^2) = \lambda^4 + 4\lambda^2 + 3$$

$$= (\lambda^2 + 3)(\lambda^2 + 1)$$

which implies that the eigenvalues of $\boldsymbol{A}$ are $\pm\sqrt{3}i$ and $\pm i$. Corresponding eigenvectors are

$$\pm\sqrt{3}i \leftrightarrow \Big[\pm\frac{1}{\sqrt{3}}i, \mp\frac{1}{\sqrt{3}}i, -1, 1\Big]^{\mathrm{T}},$$

$$\pm i \leftrightarrow \big[\mp i, \mp i, 1, 1\big]^{\mathrm{T}};$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

thus

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3}i & & & \\ & -\sqrt{3}i & & \\ & & i & \\ & & & -i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}.$$

Therefore,

$$e^{t\boldsymbol{A}} = \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^{i\sqrt{3}t} & & & \\ & e^{-i\sqrt{3}t} & & \\ & & e^{it} & \\ & & & e^{-it} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

thus

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3}i & & & \\ & -\sqrt{3}i & & \\ & & i & \\ & & & -i \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}.$$

Therefore,

$$e^{t\boldsymbol{A}} = \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^{i\sqrt{3}t} & & & \\ & e^{-i\sqrt{3}t} & & \\ & & e^{it} & \\ & & & e^{-it} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1}.$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

Using the formula for solutions of linear systems, we find that the general solution to the given ODE is given by

$$\begin{aligned}\boldsymbol{y}(t) &= \boldsymbol{P}\begin{bmatrix} e^{i\sqrt{3}t} & & & \\ & e^{-i\sqrt{3}t} & & \\ & & e^{it} & \\ & & & e^{-it}\end{bmatrix}\boldsymbol{P}^{-1}\int \boldsymbol{P}\begin{bmatrix} e^{-i\sqrt{3}t} & & & \\ & e^{i\sqrt{3}t} & & \\ & & e^{-it} & \\ & & & e^{it}\end{bmatrix}\boldsymbol{P}^{-1}\boldsymbol{f}(t)\,dt \\ &= \boldsymbol{P}\begin{bmatrix} e^{i\sqrt{3}t} & & & \\ & e^{-i\sqrt{3}t} & & \\ & & e^{it} & \\ & & & e^{-it}\end{bmatrix}\int \begin{bmatrix} e^{-i\sqrt{3}t} & & & \\ & e^{i\sqrt{3}t} & & \\ & & e^{-it} & \\ & & & e^{it}\end{bmatrix}\boldsymbol{P}^{-1}\boldsymbol{f}(t)\,dt\,,\end{aligned}$$

where $\boldsymbol{P} = \begin{bmatrix} \frac{1}{\sqrt{3}}i & -\frac{1}{\sqrt{3}}i & -i & i \\ -\frac{1}{\sqrt{3}}i & \frac{1}{\sqrt{3}}i & -i & i \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1\end{bmatrix}$ and $\boldsymbol{f}(t) = \begin{bmatrix} 0 \\ 0 \\ \frac{F_1}{m_1} \\ \frac{F_2}{m_2}\end{bmatrix}$.

§2.2 Some Basic Techniques of Solving ODEs

Example

Consider the linear system $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$, where

$$\boldsymbol{A} = \begin{bmatrix} 4 & -2 & 0 & 2 \\ 0 & 6 & -2 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & -2 & 0 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 2 & 1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 2 & 1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix}^{-1}$$

$$= \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}.$$

Using the formula

$$\boldsymbol{x}'(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t) \quad \Rightarrow \quad \boldsymbol{x}(t) = e^{t\boldsymbol{A}} \int e^{-t\boldsymbol{A}} \boldsymbol{f}(t)\, dt\,,$$

§2.2 Some Basic Techniques of Solving ODEs

Example (cont.)

we find that the general solution to the given ODE is

$$\boldsymbol{x}(t) = e^{t\boldsymbol{A}}\boldsymbol{C} = \boldsymbol{P}e^{t\boldsymbol{J}}\boldsymbol{P}^{-1}\boldsymbol{C}$$

$$= \begin{bmatrix} -2 & 0 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 2 & 1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \exp\left(t \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix} \right) \begin{bmatrix} -2 & 0 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 2 & 1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 0 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 2 & 1 & 0 \\ -2 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^{4t} & te^{4t} & 0 & 0 \\ 0 & e^{4t} & 0 & 0 \\ 0 & 0 & e^{4t} & 0 \\ 0 & 0 & 0 & e^{6t} \end{bmatrix} \begin{bmatrix} \bar{C}_1 \\ \bar{C}_2 \\ \bar{C}_3 \\ \bar{C}_4 \end{bmatrix}$$

for some constants $\bar{C}_1$, $\bar{C}_2$, $\bar{C}_3$ and $\bar{C}_4$.

§2.3 Solving IVP using matlab®

Step 1: Write the IVP in the vector form

$$\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0 .$$

Note that usually you need to write the IVP in a dimensionless form (under a proper choice of characteristic scale).

Step 2: Write (and save) the function $\boldsymbol{f}$ in matlab.

Step 3: Once the function $\boldsymbol{f}$ is saved, use the command "ode45" (based on the **adaptive Runge-Kutta** method) to solve the IVP:

```
[t,y] = ode45(@name of the function,[starting time, terminal time], initial data)
```

where the output of this command has two pieces t and y:

1. t is a column vector whose components are the samples of time at which the numerical solution evaluates.
2. y is a $m \times n$ matrix, where m is the total number of time samples, and n is the dimension of the vector y.

§2.3 Solving IVP using matlab®

Example

Consider solving the IVP (from the Lotka-Volterra model)

$$p' = -0.16p + 0.08pq\,, \qquad p(0) = 5\,,$$
$$q' = 4.5q - 0.9pq\,, \qquad q(0) = 3\,,$$

numerically using matlab. Let $\boldsymbol{y} = \begin{bmatrix} p \\ q \end{bmatrix}$, $\boldsymbol{f}(t, \boldsymbol{y}) = \begin{bmatrix} -0.16p + 0.08pq \\ 4.5q - 0.9pq \end{bmatrix}$.

First we write the function $\boldsymbol{f}$ (in the name "ODE_RHS"):

```
function yp = ODE_RHS(t,y)
p = y(1,1); q = y(2,1);
yp(1,1) = -0.16*p + 0.08*p*q;
yp(2,1) = 4.5*q - 0.9*p*q;
```

and then run

```
[t,y] = ode45(@ODE_RHS,[0,10],[5;3]);
```

§2.3 Solving IVP using matlab®

Example

Consider solving the IVP (from the Lotka-Volterra model)

$$p' = -0.16p + 0.08pq\,, \qquad p(0) = 5\,,$$
$$q' = 4.5q - 0.9pq\,, \qquad q(0) = 3\,,$$

numerically using matlab. Let $\boldsymbol{y} = \begin{bmatrix} p \\ q \end{bmatrix}$, $\boldsymbol{f}(t, \boldsymbol{y}) = \begin{bmatrix} -0.16p + 0.08pq \\ 4.5q - 0.9pq \end{bmatrix}$.

First we write the function $\boldsymbol{f}$ (in the name "ODE_RHS"):

```
function yp = ODE_RHS(t,y)
p = y(1,1); q = y(2,1);
yp(1,1) = -0.16*p + 0.08*p*q;
yp(2,1) = 4.5*q - 0.9*p*q;
```

and then run

```
[t,y] = ode45(@ODE_RHS,[0,10],[5;3]);
```

§2.3 Solving IVP using matlab®

Example

Consider solving the IVP (from the study of Kepler's laws of planetary motion)

$$-\frac{GMm}{r^2}\hat{r} = m\boldsymbol{r}'' , \qquad \boldsymbol{r}(0) = \boldsymbol{r}_0, \quad \boldsymbol{r}'(0) = \boldsymbol{r}_1 ,$$

under the settings: $GM = 1$, $\boldsymbol{r}_0 = [1;0]$ and $\boldsymbol{r}_1 = [0;0.6]$. We note that the IVP above can be written as

$$-\frac{GM}{(x^2+y^2)^{1.5}}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x'' \\ y'' \end{bmatrix} , \qquad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \boldsymbol{r}_0 , \quad \begin{bmatrix} x'(0) \\ y'(0) \end{bmatrix} = \boldsymbol{r}_1 .$$

One can follow the previous example and write the function on the right-hand side as a separate file; however, there is an easier way to do this if the right-hand side function is simple.

§2.3 Solving IVP using matlab®

Example

Consider solving the IVP (from the study of Kepler's laws of planetary motion)

$$-\frac{GMm}{r^2}\widehat{r} = m\boldsymbol{r}'' , \qquad \boldsymbol{r}(0) = \boldsymbol{r}_0, \quad \boldsymbol{r}'(0) = \boldsymbol{r}_1 ,$$

under the settings: $GM = 1$, $\boldsymbol{r}_0 = [1;0]$ and $\boldsymbol{r}_1 = [0;0.6]$. We note that the IVP above can be written as

$$-\frac{GM}{(x^2+y^2)^{1.5}}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x'' \\ y'' \end{bmatrix} , \qquad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \boldsymbol{r}_0 , \quad \begin{bmatrix} x'(0) \\ y'(0) \end{bmatrix} = \boldsymbol{r}_1 .$$

One can follow the previous example and write the function on the right-hand side as a separate file; however, there is an easier way to do this if the right-hand side function is simple.

§2.3 Solving IVP using matlab®

Example (Con't)

Let $\boldsymbol{z} = [z_1; z_2; z_3; z_4] \equiv [x; y; x'; y']$. Then $\boldsymbol{z}$ satisfies

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} z_3 \\ z_4 \\ -\dfrac{z_1}{(z_1^2+z_2^2)^{1.5}} \\ -\dfrac{z_2}{(z_1^2+z_2^2)^{1.5}} \end{bmatrix}, \qquad \boldsymbol{z}(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0.6 \end{bmatrix}.$$

Therefore, we execute the following codes

```
ODE_RHS = @(t,y) [y(3:4); -1/(norm(y(1:2))∧3)*y(1:2)];
[t,y] = ode45(@(t,y) ODE_RHS(t,y), [0,3], [1;0;0;0.6]);
```

to solve this problem numerically.

§2.3 Solving IVP using matlab®

Example (Con't)

If the right-hand side function has some parameters, one can write this function as a function of t, y, as well as these parameters (t and y have to be the first two variables). To use ode45, one runs

```
ODE_RHS = @(t,y,G,M) [y(3:4); -G*M/(norm(y(1:2))^3)*y(1:2)];
G = 1; M = 1;
[t,y] = ode45(@(t,y) ODE_RHS(t,y,G,M), [0,3], [1;0;0;0.6]);
plot(y(:,1),y(:,2),'b');
axis equal;
```

§2.3 Solving IVP using matlab®

Example

Consider finding the position where the function

$$f(x,y) = xe^{-x^2-y^2}$$

attains its global minimum or one of its local minimums. Using the idea of gradient flows, we compute

$$f_x(x,y) = (1-2x^2)e^{-x^2-y^2} \quad \text{and} \quad f_y(x,y) = -2xye^{-x^2-y^2}$$

and consider the IVP

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} f_x(x,y) \\ f_y(x,y) \end{bmatrix} = \begin{bmatrix} (2x^2-1)e^{-x^2-y^2} \\ 2xye^{-x^2-y^2} \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

where (x_0, y_0) is a point closed to the global minimum or a local minimum.

§2.3 Solving IVP using matlab®

Example

Consider finding the position where the function

$$f(x, y) = xe^{-x^2-y^2}$$

attains its global minimum or one of its local minimums. Using the idea of gradient flows, we compute

$$f_x(x, y) = (1 - 2x^2)e^{-x^2-y^2} \quad \text{and} \quad f_y(x, y) = -2xye^{-x^2-y^2}$$

and consider the IVP

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix} = \begin{bmatrix} (2x^2 - 1)e^{-x^2-y^2} \\ 2xye^{-x^2-y^2} \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

where (x_0, y_0) is a point closed to the global minimum or a local minimum.

§2.3 Solving IVP using matlab®

Example (Con't)

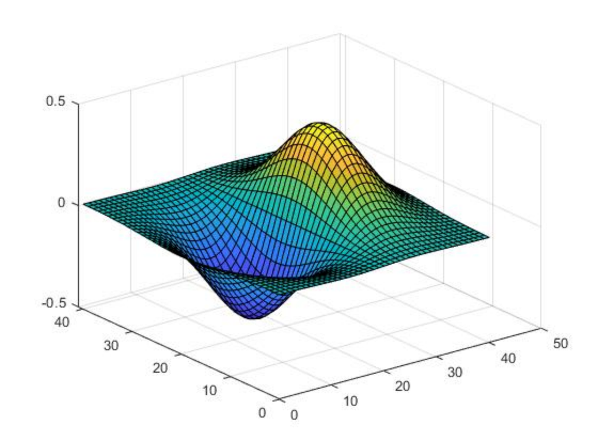


Figure 9: The graph of the function $f(x, y) = xe^{-x^2-y^2}$.

§2.3 Solving IVP using matlab®

Example (Con't)

We first write the function $-\nabla f$ by

```
function zp = ODE_RHS(t,z)
x = z(1,1); y = z(2,1);
zp(1,1) = (2*x∧2–1)*exp(–x∧2–y∧2);
zp(2,1) = 2*x*y*exp(–x∧2–y∧2);
```

and then (with a wild guess of a local minimum $(x_0, y_0) = (0.5, 0.5)$ in mind) run

```
[t,y] = ode45(@(t,y) ODE_RHS(t,y),[0,10],[0.5;0.5]);
```

The vector $y(\text{end}, :)$ may be very closed to $\lim\limits_{t\to\infty} y(t)$, a candidate of what we are after.

§2.3 Solving IVP using matlab®

Example (Con't)

We first write the function $-\nabla f$ by

```
function zp = ODE_RHS(t,z)
x = z(1,1); y = z(2,1);
zp(1,1) = (2*x∧2–1)*exp(–x∧2–y∧2);
zp(2,1) = 2*x*y*exp(–x∧2–y∧2);
```

and then (with a wild guess of a local minimum $(x_0, y_0) = (0.5, 0.5)$ in mind) run

```
[t,y] = ode45(@(t,y) ODE_RHS(t,y),[0,10],[0.5;0.5]);
```

The vector $y(\text{end}, :)$ may be very closed to $\lim_{t\to\infty} y(t)$, a candidate of what we are after.

§2.4 Boundary Value Problems

In this section we only consider ODE of the form

$$y'' + p(x)y' + q(x)y = g(x)\,,$$

where p, q and g are given functions, and $y = y(x)$ is the unknown function. Instead of imposing the initial condition $y(t_0) = y_0$ and $y'(t_0) = y_1$, sometimes the following four kinds of boundary condition can be imposed:

1. $y(\alpha) = y_0$, $y(\beta) = y_1$;
2. $y(\alpha) = y_0$, $y'(\beta) = y_1$;
3. $y'(\alpha) = y_0$, $y(\beta) = y_1$;
4. $y'(\alpha) = y_0$, $y'(\beta) = y_1$,

where α, β, y_0 and y_1 are given numbers. Such kind of combination of ODE and boundary condition is called a (two-point) ***boundary value problem*** (**BVP**), and a solution y to a BVP must be defined on the interval $I = [\alpha, \beta]$, as well as satisfy the ODE and the boundary condition.

§2.4 Boundary Value Problems

In this section we only consider ODE of the form

$$y'' + p(x)y' + q(x)y = g(x)\,,$$

where p, q and g are given functions, and $y = y(x)$ is the unknown function. Instead of imposing the initial condition $y(t_0) = y_0$ and $y'(t_0) = y_1$, sometimes the following four kinds of boundary condition can be imposed:

1. $y(\alpha) = y_0$, $y(\beta) = y_1$;
2. $y(\alpha) = y_0$, $y'(\beta) = y_1$;
3. $y'(\alpha) = y_0$, $y(\beta) = y_1$;
4. $y'(\alpha) = y_0$, $y'(\beta) = y_1$,

where α, β, y_0 and y_1 are given numbers. Such kind of combination of ODE and boundary condition is called a (two-point) ***boundary value problem*** (**BVP**), and a solution y to a BVP must be defined on the interval $I = [\alpha, \beta]$, as well as satisfy the ODE and the boundary condition.

§2.4 Boundary Value Problems

In this section we only consider ODE of the form

$$y'' + p(x)y' + q(x)y = g(x)\,,$$

where p, q and g are given functions, and $y = y(x)$ is the unknown function. Instead of imposing the initial condition $y(t_0) = y_0$ and $y'(t_0) = y_1$, sometimes the following four kinds of boundary condition can be imposed:

1. $y(\alpha) = y_0$, $y(\beta) = y_1$;
2. $y(\alpha) = y_0$, $y'(\beta) = y_1$;
3. $y'(\alpha) = y_0$, $y(\beta) = y_1$;
4. $y'(\alpha) = y_0$, $y'(\beta) = y_1$,

where α, β, y_0 and y_1 are given numbers. Such kind of combination of ODE and boundary condition is called a (two-point) ***boundary value problem*** (**BVP**), and a solution y to a BVP must be defined on the interval $I = [\alpha, \beta]$, as well as satisfy the ODE and the boundary condition.

§2.4 Boundary Value Problems

Example

In this example we reconsider the ODE in the spring-mass system

$$m\ddot{x} = -kx - r\dot{x} + f(t)\,.$$

We explain the meaning of the different boundary condition as follows:

1. $x(0) = x_0$ and $x(T) = x_1$: the initial and the terminal position of the mass are given.
2. $x(0) = x_0$ and $x'(T) = v_1$: the initial position and the terminal velocity of the mass are given.
3. $x'(0) = v_0$ and $x(T) = x_1$: the initial velocity and the terminal position of the mass are given.
4. $x'(0) = v_0$ and $x'(T) = v_1$: the initial and the terminal velocity of the mass are given.

§2.4 Boundary Value Problems

Example

In this example we reconsider the ODE in the spring-mass system

$$m\ddot{x} = -kx - r\dot{x} + f(t)\,.$$

We explain the meaning of the different boundary condition as follows:

1. $x(0) = x_0$ and $x(T) = x_1$: the initial and the terminal position of the mass are given.
2. $x(0) = x_0$ and $x'(T) = v_1$: the initial position and the terminal velocity of the mass are given.
3. $x'(0) = v_0$ and $x(T) = x_1$: the initial velocity and the terminal position of the mass are given.
4. $x'(0) = v_0$ and $x'(T) = v_1$: the initial and the terminal velocity of the mass are given.

§2.4 Boundary Value Problems

Example

In this example we reconsider the ODE in the spring-mass system

$$m\ddot{x} = -kx - r\dot{x} + f(t)\,.$$

We explain the meaning of the different boundary condition as follows:

1. $x(0) = x_0$ and $x(T) = x_1$: the initial and the terminal position of the mass are given.
2. $x(0) = x_0$ and $x'(T) = v_1$: the initial position and the terminal velocity of the mass are given.
3. $x'(0) = v_0$ and $x(T) = x_1$: the initial velocity and the terminal position of the mass are given.
4. $x'(0) = v_0$ and $x'(T) = v_1$: the initial and the terminal velocity of the mass are given.

§2.4 Boundary Value Problems

Example

In this example we reconsider the ODE in the spring-mass system

$$m\ddot{x} = -kx - r\dot{x} + f(t)\,.$$

We explain the meaning of the different boundary condition as follows:

1. $x(0) = x_0$ and $x(T) = x_1$: the initial and the terminal position of the mass are given.
2. $x(0) = x_0$ and $x'(T) = v_1$: the initial position and the terminal velocity of the mass are given.
3. $x'(0) = v_0$ and $x(T) = x_1$: the initial velocity and the terminal position of the mass are given.
4. $x'(0) = v_0$ and $x'(T) = v_1$: the initial and the terminal velocity of the mass are given.

§2.4 Boundary Value Problems

Example

Again we consider the ODE

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,.$$

This time we do not require that initial height $h(0)$ and the initial velocity $h'(0)$ are given but instead we want the object to reach certain height H at time $t = T$. Then the BVP is written as

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h(T) = H.$$

Similarly, if we want the object to reach certain velocity V at time $t = T$, then we have the BVP

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(T) = V.$$

§2.4 Boundary Value Problems

Example

Again we consider the ODE

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,.$$

This time we do not require that initial height $h(0)$ and the initial velocity $h'(0)$ are given but instead we want the object to reach certain height H at time $t = T$. Then the BVP is written as

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h(T) = H.$$

Similarly, if we want the object to reach certain velocity V at time $t = T$, then we have the BVP

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(T) = V.$$

§2.4 Boundary Value Problems

Example

Again we consider the ODE

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,.$$

This time we do not require that initial height $h(0)$ and the initial velocity $h'(0)$ are given but instead we want the object to reach certain height H at time $t = T$. Then the BVP is written as

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h(T) = H.$$

Similarly, if we want the object to reach certain velocity V at time $t = T$, then we have the BVP

$$m\frac{d^2h}{dt^2} = -\frac{GMm}{(R+h)^2}\,, \qquad h(0) = 0\,, \quad h'(T) = V.$$

§2.4 Boundary Value Problems

Consider the two-point boundary value problem

$$y'' + p(x)y' + q(x)y = g(x)\,, \qquad y(\alpha) = y_0, \quad y(\beta) = y_1\,. \qquad (16)$$

Let $z(x) = y(x) - \frac{x-\alpha}{\beta-\alpha}y_1 - \frac{x-\beta}{\alpha-\beta}y_0$. Then z satisfies

$$z'' + p(x)z' + q(x)z = G(x)\,, \qquad z(\alpha) = z(\beta) = 0\,,$$

where $G(x) = g(x) - p(x)\frac{y_0 - y_1}{\alpha - \beta} - q(x)\Big(\frac{x-\alpha}{\beta-\alpha}y_1 + \frac{x-\beta}{\alpha-\beta}y_0\Big)$. Therefore, in general we can assume the homogeneous boundary condition $y_0 = y_1 = 0$ in (16). Similarly, ODE $y'' + p(x)y' + q(x)y = g(x)$ with the other three kinds of boundary conditions can also be rewritten as a BVP with homogeneous boundary condition.

§2.4 Boundary Value Problems

Consider the two-point boundary value problem

$$y'' + p(x)y' + q(x)y = g(x)\,, \qquad y(\alpha) = y_0, \quad y(\beta) = y_1\,. \qquad (16)$$

Let $z(x) = y(x) - \dfrac{x-\alpha}{\beta-\alpha}y_1 - \dfrac{x-\beta}{\alpha-\beta}y_0$. Then z satisfies

$$z'' + p(x)z' + q(x)z = G(x)\,, \qquad z(\alpha) = z(\beta) = 0\,,$$

where $G(x) = g(x) - p(x)\dfrac{y_0 - y_1}{\alpha-\beta} - q(x)\Big(\dfrac{x-\alpha}{\beta-\alpha}y_1 + \dfrac{x-\beta}{\alpha-\beta}y_0\Big)$. Therefore, in general we can assume the homogeneous boundary condition $y_0 = y_1 = 0$ in (16). Similarly, ODE $y'' + p(x)y' + q(x)y = g(x)$ with the other three kinds of boundary conditions can also be rewritten as a BVP with homogeneous boundary condition.

§2.4 Boundary Value Problems

Consider the two-point boundary value problem

$$y'' + p(x)y' + q(x)y = g(x)\,, \qquad y(\alpha) = y_0, \quad y(\beta) = y_1\,. \qquad (16)$$

Let $z(x) = y(x) - \frac{x-\alpha}{\beta-\alpha}y_1 - \frac{x-\beta}{\alpha-\beta}y_0$. Then z satisfies

$$z'' + p(x)z' + q(x)z = G(x)\,, \qquad z(\alpha) = z(\beta) = 0\,,$$

where $G(x) = g(x) - p(x)\frac{y_0 - y_1}{\alpha - \beta} - q(x)\big(\frac{x-\alpha}{\beta-\alpha}y_1 + \frac{x-\beta}{\alpha-\beta}y_0\big)$. Therefore, in general we can assume the homogeneous boundary condition $y_0 = y_1 = 0$ in (16). Similarly, ODE $y'' + p(x)y' + q(x)y = g(x)$ with the other three kinds of boundary conditions can also be rewritten as a BVP with homogeneous boundary condition.

§2.4 Boundary Value Problems

Consider the two-point boundary value problem

$$y'' + p(x)y' + q(x)y = g(x)\,, \qquad y(\alpha) = y_0, \quad y(\beta) = y_1\,. \qquad (16)$$

Let $z(x) = y(x) - \dfrac{x-\alpha}{\beta-\alpha}y_1 - \dfrac{x-\beta}{\alpha-\beta}y_0$. Then z satisfies

$$z'' + p(x)z' + q(x)z = G(x)\,, \qquad z(\alpha) = z(\beta) = 0\,,$$

where $G(x) = g(x) - p(x)\dfrac{y_0 - y_1}{\alpha-\beta} - q(x)\Big(\dfrac{x-\alpha}{\beta-\alpha}y_1 + \dfrac{x-\beta}{\alpha-\beta}y_0\Big)$. Therefore, in general we can assume the homogeneous boundary condition $y_0 = y_1 = 0$ in (16). Similarly, ODE $y''+p(x)y'+q(x)y = g(x)$ with the other three kinds of boundary conditions can also be rewritten as a BVP with homogeneous boundary condition.

§2.4 Boundary Value Problems

Remark: Even though the initial value problem

$$y'' + p(t)y' + q(t)y = g(t)\,, \qquad y(t_0) = y_0\,, \quad y'(t_0) = y_1$$

looks quite similar to the boundary value problem (16), they actually differ in some very important ways. For example, if p, q, g are continuous, the initial value problem above always have a unique solution, while the boundary value problem (16) might have no solution or infinitely many solutions:

1. $y'' + y = 0$ with boundary condition $y(0) = y(\pi) = 0$ has infinite many solutions $y_c(x) = c\sin x$.
2. $y'' + y = \sin x$ with boundary condition $y(0) = y(\pi) = 0$ has no solution.

§2.4 Boundary Value Problems

Remark: Even though the initial value problem

$$y'' + p(t)y' + q(t)y = g(t)\,, \qquad y(t_0) = y_0\,, \quad y'(t_0) = y_1$$

looks quite similar to the boundary value problem (16), they actually differ in some very important ways. For example, if p, q, g are continuous, the initial value problem above always have a unique solution, while the boundary value problem (16) might have no solution or infinitely many solutions:

1. $y'' + y = 0$ with boundary condition $y(0) = y(\pi) = 0$ has infinite many solutions $y_c(x) = c\sin x$.
2. $y'' + y = \sin x$ with boundary condition $y(0) = y(\pi) = 0$ has no solution.

§2.4 Boundary Value Problems

Remark: Even though the initial value problem

$$y'' + p(t)y' + q(t)y = g(t)\,, \qquad y(t_0) = y_0\,, \quad y'(t_0) = y_1$$

looks quite similar to the boundary value problem (16), they actually differ in some very important ways. For example, if p, q, g are continuous, the initial value problem above always have a unique solution, while the boundary value problem (16) might have no solution or infinitely many solutions:

1. $y'' + y = 0$ with boundary condition $y(0) = y(\pi) = 0$ has infinite many solutions $y_c(x) = c\sin x$.
2. $y'' + y = \sin x$ with boundary condition $y(0) = y(\pi) = 0$ has no solution.

§2.4 Boundary Value Problems

Remark: Even though the initial value problem

$$y'' + p(t)y' + q(t)y = g(t)\,, \qquad y(t_0) = y_0\,, \quad y'(t_0) = y_1$$

looks quite similar to the boundary value problem (16), they actually differ in some very important ways. For example, if p, q, g are continuous, the initial value problem above always have a unique solution, while the boundary value problem (16) might have no solution or infinitely many solutions:

1. $y'' + y = 0$ with boundary condition $y(0) = y(\pi) = 0$ has infinite many solutions $y_c(x) = c\sin x$.
2. $y'' + y = \sin x$ with boundary condition $y(0) = y(\pi) = 0$ has no solution.

§2.4 Boundary Value Problems

On the other hand, there are cases that (16) has a unique solution. For example, the general solution to the boundary value problem

$$y'' + 2y = 0$$

is given by

$$y(x) = C_1 \cos \sqrt{2}x + C_2 \sin \sqrt{2}x;$$

thus to validate the boundary condition $y(0) = 1$ and $y(\pi) = 0$, we must have $C_1 = 1$ and $C_2 = -\cot \sqrt{2}\pi$. In other words, the solution $y(x) = \cos \sqrt{2}x - \cot \sqrt{2}\pi \sin \sqrt{2}x$.

§2.4 Boundary Value Problems

Theorem

Let α, β be real numbers and $\alpha < \beta$. Suppose that the function $f = f(x, y, p)$ is continuous on the set

$$D = \{(x, y, p) \,|\, x \in [\alpha, \beta], y, p \in \mathbb{R}\}$$

and the partial derivatives f_y and f_p are also continuous on D. If

1. *$f_y(x, y, p) > 0$ for all $(x, y, p) \in D$, and*
2. *there exists a constant $M > 0$ such that*

$$\big|f_p(x, y, p)\big| \leqslant M \qquad \forall\, (x, y, p) \in D\,,$$

then the boundary value problem

$$y'' = f(x, y, y') \qquad \forall\, x \in (\alpha, \beta),\ y(\alpha) = y(\beta) = 0$$

has a unique solution.

Chapter 3. Partial Differential Equations（偏微分方程）

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.1 The 1-dimensional conservation laws

Suppose that a substance of interest lives in a 1-dimensional space such as a tube. Let $u(x,t)$ be the **density** or **concentration** of the substance at position x and time t. Then

$$\int_x^{x+\Delta x} u(y,t)\,dy$$

is the total amount of the substance in the interval $I = [x, x+\Delta x]$ at time t; thus during the time period $[t, t+\Delta t]$, the change of the amount of the substance in the interval I in the time period $[t, t+\Delta t]$ is given by

$$\int_x^{x+\Delta x} u(y,t+\Delta t)\,dy - \int_x^{x+\Delta x} u(y,t)\,dy$$
$$= \int_x^{x+\Delta x} \big[u(y,t+\Delta t) - u(y,t)\big]\,dy\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.1 The 1-dimensional conservation laws

Suppose that a substance of interest lives in a 1-dimensional space such as a tube. Let $u(x,t)$ be the **density** or **concentration** of the substance at position x and time t. Then

$$\int_x^{x+\Delta x} u(y,t)\,dy$$

is the total amount of the substance in the interval $I=[x,x+\Delta x]$ at time t; thus during the time period $[t,t+\Delta t]$, the change of the amount of the substance in the interval I in the time period $[t,t+\Delta t]$ is given by

$$\int_x^{x+\Delta x} u(y,t+\Delta t)\,dy - \int_x^{x+\Delta x} u(y,t)\,dy = \int_x^{x+\Delta x} \big[u(y,t+\Delta t)-u(y,t)\big]\,dy\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.1 The 1-dimensional conservation laws

Suppose that a substance of interest lives in a 1-dimensional space such as a tube. Let $u(x,t)$ be the **density** or **concentration** of the substance at position x and time t. Then

$$\int_x^{x+\Delta x} u(y,t)\,dy$$

is the total amount of the substance in the interval $I=[x,x+\Delta x]$ at time t; thus during the time period $[t,t+\Delta t]$, the change of the amount of the substance in the interval I in the time period $[t,t+\Delta t]$ is given by

$$\int_x^{x+\Delta x} u(y,t+\Delta t)\,dy - \int_x^{x+\Delta x} u(y,t)\,dy$$
$$= \int_x^{x+\Delta x} \big[u(y,t+\Delta t)-u(y,t)\big]\,dy\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

On the other hand, there are two sources of changing the amount of the substance in the interval I:

1. a **flux** (通量，可先想成流率) that describes any effect that appears to **pass** or **travel** the substance through points.
2. a **source** that will **release** or **absorb** the substance in this interval.

Let f denote the flux and q denote the source. Then in the time interval $[t, t+\Delta t]$ the amount of the substance **flowing into / from the point** x is given by

$$\int_t^{t+\Delta t} f(x,t')\,dt'$$

while the amount of the substance **flowing out of / from the point** $x+\Delta x$ is given by

$$\int_t^{t+\Delta t} f(x+\Delta x,t')\,dt'.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

On the other hand, there are two sources of changing the amount of the substance in the interval I:

1. a **flux** (通量，可先想成流率) that describes any effect that appears to **pass** or **travel** the substance through points.
2. a **source** that will **release** or **absorb** the substance in this interval.

Let f denote the flux and q denote the source. Then in the time interval $[t, t+\Delta t]$ the amount of the substance **flowing into / from the point** x is given by

$$\int_t^{t+\Delta t} f(x, t')\, dt'$$

while the amount of the substance **flowing out of / from the point** $x + \Delta x$ is given by

$$\int_t^{t+\Delta t} f(x + \Delta x, t')\, dt' .$$

§3.1 Models with One Temporal Variable and One Spatial Variable

On the other hand, there are two sources of changing the amount of the substance in the interval I:

1. a **flux** (通量，可先想成流率) that describes any effect that appears to **pass** or **travel** the substance through points.
2. a **source** that will **release** or **absorb** the substance in this interval.

Let f denote the flux and q denote the source. Then in the time interval $[t, t+\Delta t]$ the amount of the substance **flowing into / from the point** x is given by

$$\int_t^{t+\Delta t} f(x, t')\, dt'$$

while the amount of the substance **flowing out of / from the point** $x + \Delta x$ is given by

$$\int_t^{t+\Delta t} f(x+\Delta x, t')\, dt' .$$

§3.1 Models with One Temporal Variable and One Spatial Variable

On the other hand, there are two sources of changing the amount of the substance in the interval I:

1. a **flux** (通量，可先想成流率) that describes any effect that appears to **pass** or **travel** the substance through points.
2. a **source** that will **release** or **absorb** the substance in this interval.

Let f denote the flux and q denote the source. Then in the time interval $[t, t+\Delta t]$ the amount of the substance **flowing into / from the point** x is given by

$$\int_t^{t+\Delta t} f(x, t')\, dt'$$

while the amount of the substance **flowing out of / from the point** $x+\Delta x$ is given by

$$\int_t^{t+\Delta t} f(x+\Delta x, t')\, dt'.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

On the other hand, there are two sources of changing the amount of the substance in the interval I:

1. a **flux** (通量，可先想成流率) that describes any effect that appears to **pass** or **travel** the substance through points.
2. a **source** that will **release** or **absorb** the substance in this interval.

Let f denote the flux and q denote the source. Then in the time interval $[t, t+\Delta t]$ the amount of the substance **flowing into / from the point** x is given by

$$\int_t^{t+\Delta t} f(x, t')\, dt'$$

while the amount of the substance **flowing out of / from the point** $x+\Delta x$ is given by

$$\int_t^{t+\Delta t} f(x+\Delta x, t')\, dt'.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Moreover, the **change** of the amount of the substance in the interval I in the time period $[t, t+\Delta t]$ **due to the source** is given by

$$\int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.$$

Therefore, the change of amount of the substance in the interval I in the time period $[t, t+\Delta t]$ is given by

$$\int_t^{t+\Delta t}\big[f(x,t')-f(x+\Delta x,t')\big]\,dt' + \int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.$$

As a consequence,

$$\begin{aligned}&\int_x^{x+\Delta x}\big[u(y,t+\Delta t)-u(y,t)\big]\,dy\\ &\quad=\int_t^{t+\Delta t}\big[f(x,t')-f(x+\Delta x,t')\big]\,dt' + \int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.\end{aligned}$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Moreover, the **change** of the amount of the substance in the interval I in the time period $[t, t+\Delta t]$ **due to the source** is given by

$$\int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.$$

Therefore, the change of amount of the substance in the interval I in the time period $[t, t+\Delta t]$ is given by

$$\int_t^{t+\Delta t}\big[f(x,t') - f(x+\Delta x,t')\big]\,dt' + \int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.$$

As a consequence,

$$\int_x^{x+\Delta x}\big[u(y,t+\Delta t) - u(y,t)\big]\,dy$$
$$= \int_t^{t+\Delta t}\big[f(x,t') - f(x+\Delta x,t')\big]\,dt' + \int_t^{t+\Delta t}\int_x^{x+\Delta x} q(y,t')\,dydt'\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Moreover, the **change** of the amount of the substance in the interval I in the time period $[t, t + \Delta t]$ **due to the source** is given by

$$\int_t^{t+\Delta t} \int_x^{x+\Delta x} q(y, t')\, dydt' .$$

Therefore, the change of amount of the substance in the interval I in the time period $[t, t + \Delta t]$ is given by

$$\int_t^{t+\Delta t} \big[f(x, t') - f(x + \Delta x, t')\big]\, dt' + \int_t^{t+\Delta t} \int_x^{x+\Delta x} q(y, t')\, dydt' .$$

As a consequence,

$$\begin{aligned} &\int_x^{x+\Delta x} \big[u(y, t + \Delta t) - u(y, t)\big]\, dy \\ &= \int_t^{t+\Delta t} \big[f(x, t') - f(x + \Delta x, t')\big]\, dt' + \int_t^{t+\Delta t} \int_x^{x+\Delta x} q(y, t')\, dydt' . \end{aligned}$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Dividing both sides of the resulting equation through by Δx and then passing to the limit as $\Delta x \to 0$, by the **fundamental theorem of Calculus** we find that (without any rigorous verification)

$$u(x, t+\Delta t) - u(x,t) = -\int_t^{t+\Delta t} \frac{\partial}{\partial x} f(x,t')\, dt' + \int_t^{t+\Delta t} q(x,t')\, dt' .$$

Similarly, dividing both sides of the equality above through Δt and then passing to the limit as $\Delta t \to 0$, the fundamental theorem of Calculus implies that

$$\frac{\partial}{\partial t} u(x,t) + \frac{\partial}{\partial x} f(x,t) = q(x,t) .$$

Fundamental Theorem of Calculus:

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} g(y)\, dy = g(x) \text{ if } g \text{ is continuous at } x.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Dividing both sides of the resulting equation through by Δx and then passing to the limit as $\Delta x \to 0$, by the **fundamental theorem of Calculus** we find that (without any rigorous verification)

$$u(x, t + \Delta t) - u(x, t) = -\int_t^{t+\Delta t} \frac{\partial}{\partial x} f(x, t')\, dt' + \int_t^{t+\Delta t} q(x, t')\, dt' \,.$$

Similarly, dividing both sides of the equality above through Δt and then passing to the limit as $\Delta t \to 0$, the fundamental theorem of Calculus implies that

$$\frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} f(x, t) = q(x, t) \,.$$

Fundamental Theorem of Calculus:

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} g(y)\, dy = g(x) \text{ if } g \text{ is continuous at } x.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Dividing both sides of the resulting equation through by Δx and then passing to the limit as $\Delta x \to 0$, by the **fundamental theorem of Calculus** we find that (without any rigorous verification)

$$u(x, t + \Delta t) - u(x, t) = -\int_t^{t+\Delta t} \frac{\partial}{\partial x} f(x, t')\, dt' + \int_t^{t+\Delta t} q(x, t')\, dt' \,.$$

Similarly, dividing both sides of the equality above through Δt and then passing to the limit as $\Delta t \to 0$, the fundamental theorem of Calculus implies that

$$\frac{\partial}{\partial t} u(x, t) + \frac{\partial}{\partial x} f(x, t) = q(x, t) \,.$$

Fundamental Theorem of Calculus:

$$\lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} g(y)\, dy = g(x) \text{ if } g \text{ is continuous at } x.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Example (Traffic flows (cont.))

Consider the traffic on the highway (parameterized by $\mathbb{R}$). Let u denote the car density (given in the number of vehicles per unit length). Then the flux f is a function of u with the property that

1. $f(u) = 0$ if $u = 0$ or $u > L$,
2. $f'(u) > 0$ if $u \in (0, u_{\max})$, and $f'(u) < 0$ if $u \in (u_{\max}, L)$.

Suppose that f is differentiable, and $f'(u) = c(u)$. Then

$$u_t(x,t) + c(u(x,t))u_x(x,t) = q(x,t) \qquad \forall\, x \in \mathbb{R}\,, t \in \mathbb{R}$$

which can be abbreviated as

$$u_t + c(u)u_x = q \qquad \text{in} \quad \mathbb{R} \times \mathbb{R}\,.$$

To complete the model, we also need to impose an initial condition

$$u(x,0) = u_0(x) \quad \forall\, x \in \mathbb{R} \quad (\text{or simply } u = u_0 \text{ on } \mathbb{R} \times \{t = 0\})\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Example (Traffic flows (cont.))

Consider the traffic on the highway (parameterized by $\mathbb{R}$). Let u denote the car density (given in the number of vehicles per unit length). Then the flux f is a function of u with the property that

1. $f(u) = 0$ if $u = 0$ or $u > L$,
2. $f'(u) > 0$ if $u \in (0, u_{\max})$, and $f'(u) < 0$ if $u \in (u_{\max}, L)$.

Suppose that f is differentiable, and $f'(u) = c(u)$. Then

$$u_t(x,t) + c(u(x,t))u_x(x,t) = q(x,t) \qquad \forall\, x \in \mathbb{R}\,,\, t \in \mathbb{R}$$

which can be abbreviated as

$$u_t + c(u)u_x = q \qquad \text{in} \quad \mathbb{R} \times \mathbb{R}\,.$$

To complete the model, we also need to impose an initial condition

$$u(x,0) = u_0(x) \quad \forall\, x \in \mathbb{R} \quad (\text{or simply } u = u_0 \text{ on } \mathbb{R} \times \{t = 0\})\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Example (Traffic flows (cont.))

Consider the traffic on the highway (parameterized by $\mathbb{R}$). Let u denote the car density (given in the number of vehicles per unit length). Then the flux f is a function of u with the property that

1. $f(u) = 0$ if $u = 0$ or $u > L$,
2. $f'(u) > 0$ if $u \in (0, u_{\max})$, and $f'(u) < 0$ if $u \in (u_{\max}, L)$.

Suppose that f is differentiable, and $f'(u) = c(u)$. Then

$$u_t(x,t) + c(u(x,t))u_x(x,t) = q(x,t) \qquad \forall\, x \in \mathbb{R}\,, t \in \mathbb{R}$$

which can be abbreviated as

$$u_t + c(u)u_x = q \qquad \text{in} \quad \mathbb{R} \times \mathbb{R}\,.$$

To complete the model, we also need to impose an initial condition

$$u(x,0) = u_0(x) \qquad \forall\, x \in \mathbb{R} \quad (\text{or simply } u = u_0 \text{ on } \mathbb{R} \times \{t = 0\})\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Example (Traffic flows (cont.))

Consider the traffic on the highway (parameterized by $\mathbb{R}$). Let u denote the car density (given in the number of vehicles per unit length). Then the flux f is a function of u with the property that

1. $f(u) = 0$ if $u = 0$ or $u > L$,
2. $f'(u) > 0$ if $u \in (0, u_{\max})$, and $f'(u) < 0$ if $u \in (u_{\max}, L)$.

Suppose that f is differentiable, and $f'(u) = c(u)$. Then

$$u_t(x,t) + c(u(x,t))u_x(x,t) = q(x,t) \qquad \forall\, x \in \mathbb{R}\,, t \in \mathbb{R}$$

which can be abbreviated as

$$u_t + c(u)u_x = q \qquad \text{in} \quad \mathbb{R} \times \mathbb{R}\,.$$

To complete the model, we also need to impose an initial condition

$$u(x,0) = u_0(x) \quad \forall\, x \in \mathbb{R} \quad (\text{or simply } u = u_0 \text{ on } \mathbb{R} \times \{t = 0\})\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.2 The 1-dimensional heat equations

Consider the heat distribution on a rod of length L: Parameterize the rod by $[0, L]$, and let t be the time variable. Let $\rho(x)$, $s(x)$, $\kappa(x)$ denote the **density**, **specific heat**, and the **thermal conductivity** of the rod at position $x \in (0, L)$, respectively, and $\vartheta(x, t)$ denote the **temperature** at position x and time t. For $0 < x < L$, and $\Delta x, \Delta t \ll 1$,

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[\vartheta(y, t+\Delta t) - \vartheta(y, t)\big]\,dy$$
$$= \int_t^{t+\Delta t} \big[\kappa(x+\Delta x)\vartheta_x(x+\Delta x, t') - \kappa(x)\vartheta_x(x, t')\big]\,dt',$$

where the left-hand side denotes the change of the total heat in the small section $(x, x+\Delta x)$, and the right-hand side denotes the heat flowing into the section from outside.

§3.1.2 The 1-dimensional heat equations

Consider the heat distribution on a rod of length L: Parameterize the rod by $[0, L]$, and let t be the time variable. Let $\rho(x)$, $s(x)$, $\kappa(x)$ denote the **density**, **specific heat**, and the **thermal conductivity** of the rod at position $x \in (0, L)$, respectively, and $\vartheta(x,t)$ denote the **temperature** at position x and time t. For $0 < x < L$, and $\Delta x, \Delta t \ll 1$,

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[\vartheta(y,t+\Delta t) - \vartheta(y,t)\big]\,dy = \int_t^{t+\Delta t} \big[\kappa(x+\Delta x)\vartheta_x(x+\Delta x,t') - \kappa(x)\vartheta_x(x,t')\big]\,dt'\,,$$

where the left-hand side denotes the change of the total heat in the small section $(x, x+\Delta x)$, and the right-hand side denotes the heat flowing into the section from outside.

§3.1.2 The 1-dimensional heat equations

Consider the heat distribution on a rod of length L: Parameterize the rod by $[0, L]$, and let t be the time variable. Let $\rho(x)$, $s(x)$, $\kappa(x)$ denote the **density**, **specific heat**, and the **thermal conductivity** of the rod at position $x \in (0, L)$, respectively, and $\vartheta(x, t)$ denote the **temperature** at position x and time t. For $0 < x < L$, and $\Delta x, \Delta t \ll 1$,

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[\vartheta(y, t+\Delta t) - \vartheta(y, t)\big]\, dy$$
$$= \int_t^{t+\Delta t} \big[\kappa(x+\Delta x)\vartheta_x(x+\Delta x, t') - \kappa(x)\vartheta_x(x, t')\big]\, dt' ,$$

where the left-hand side denotes the change of the total heat in the small section $(x, x+\Delta x)$, and the right-hand side denotes the heat flowing into the section from outside.

§3.1 Models with One Temporal Variable and One Spatial Variable

If there is a heat source Q, then the equation above has to be modified as

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[\vartheta(y,t+\Delta t)-\vartheta(y,t)\big]\,dy$$
$$=\int_t^{t+\Delta t}\big[\kappa(x+\Delta x)\vartheta_x(x+\Delta x,t')-\kappa(x)\vartheta_x(x,t')\big]\,dt'$$
$$+\int_t^{t+\Delta t}\int_x^{x+\Delta x} Q(y,t')\,dydt'\,.$$

Dividing both sides by Δx and passing to the limit as $\Delta x\to 0$, by the Fundamental Theorem of Calculus (assuming that all the functions appearing in the equation above are smooth enough) we obtain that

$$\rho(x)s(x)\big[\vartheta(x,t+\Delta t)-\vartheta(x,t)\big]$$
$$=\int_t^{t+\Delta t}\frac{\partial}{\partial x}\big[\kappa(x)\vartheta_x(x,t')\big]\,dt'+\int_t^{t+\Delta t}Q(x,t')\,dt'\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

If there is a heat source Q, then the equation above has to be modified as

$$\int_x^{x+\Delta x} \rho(y)s(y)\big[\vartheta(y,t+\Delta t)-\vartheta(y,t)\big]\,dy = \int_t^{t+\Delta t}\big[\kappa(x+\Delta x)\vartheta_x(x+\Delta x,t')-\kappa(x)\vartheta_x(x,t')\big]\,dt' + \int_t^{t+\Delta t}\int_x^{x+\Delta x} Q(y,t')\,dydt'\,.$$

Dividing both sides by Δx and passing to the limit as $\Delta x \to 0$, by the Fundamental Theorem of Calculus (assuming that all the functions appearing in the equation above are smooth enough) we obtain that

$$\rho(x)s(x)\big[\vartheta(x,t+\Delta t)-\vartheta(x,t)\big] = \int_t^{t+\Delta t}\frac{\partial}{\partial x}\big[\kappa(x)\vartheta_x(x,t')\big]\,dt' + \int_t^{t+\Delta t} Q(x,t')\,dt'\,.$$

Dividing both sides of the equation above by Δt and then passing to the limit to $\Delta t \to 0$, we obtain the **heat equation**

$$\rho(x)s(x)\frac{\partial}{\partial t}\vartheta(x,t) = \frac{\partial}{\partial x}\big[\kappa(x)\vartheta_x(x,t)\big] + Q(x,t) \quad 0 < x < L\,, t > 0\,.$$

Assuming uniform rod; that is, ρ, s, κ are constant functions, then the heat equation above reduces to that

$$\vartheta_t(x,t) = \alpha^2\vartheta_{xx}(x,t) + q(x,t)\,, \quad 0 < x < L\,,\ \ t > 0\,,$$

where $\alpha^2 = \dfrac{\kappa}{\rho s}$ is called the ***thermal diffusivity***.

§3.1 Models with One Temporal Variable and One Spatial Variable

To determine the state of the temperature, we need to impose an initial condition

$$\vartheta(x,0) = \vartheta_0(x) \qquad 0 < x < L$$

and a boundary condition (BC):

1. Temperature on the end-points of the rod is fixed: $\vartheta(0,t) = T_1$ and $\vartheta(L,t) = T_2$. This kind of boundary condition is called **Dirichlet** BC.
2. Insulation on the end-points of the rod: $\vartheta_x(0,t) = \vartheta_x(L,t) = 0$. This kind of boundary condition is called **Neumann** BC.
3. Mixed boundary conditions: $\vartheta(0,t) = T_1$ and $\vartheta_x(L,t) = 0$, or $\vartheta(L,t) = T_2$ and $\vartheta_x(0,t) = 0$.

§3.1 Models with One Temporal Variable and One Spatial Variable

To determine the state of the temperature, we need to impose an initial condition

$$\vartheta(x,0) = \vartheta_0(x) \qquad 0 < x < L$$

and a boundary condition (BC):

1. Temperature on the end-points of the rod is fixed: $\vartheta(0,t) = T_1$ and $\vartheta(L,t) = T_2$. This kind of boundary condition is called **Dirichlet** BC.
2. Insulation on the end-points of the rod: $\vartheta_x(0,t) = \vartheta_x(L,t) = 0$. This kind of boundary condition is called **Neumann** BC.
3. Mixed boundary conditions: $\vartheta(0,t) = T_1$ and $\vartheta_x(L,t) = 0$, or $\vartheta(L,t) = T_2$ and $\vartheta_x(0,t) = 0$.

§3.1 Models with One Temporal Variable and One Spatial Variable

To determine the state of the temperature, we need to impose an initial condition

$$\vartheta(x,0) = \vartheta_0(x) \qquad 0 < x < L$$

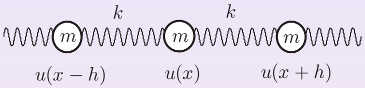
and a boundary condition (BC):

1. Temperature on the end-points of the rod is fixed: $\vartheta(0,t) = T_1$ and $\vartheta(L,t) = T_2$. This kind of boundary condition is called **Dirichlet** BC.
2. Insulation on the end-points of the rod: $\vartheta_x(0,t) = \vartheta_x(L,t) = 0$. This kind of boundary condition is called **Neumann** BC.
3. Mixed boundary conditions: $\vartheta(0,t) = T_1$ and $\vartheta_x(L,t) = 0$, or $\vartheta(L,t) = T_2$ and $\vartheta_x(0,t) = 0$.

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.3 The 1-dimensional wave equations

① From Hooke's law:



imagine an array of little weights of mass m interconnected with massless springs of length h, and the springs have a stiffness of k (see the figure). If $u(x, t)$ measures the distance from the equilibrium of the mass situated at position x and time t, then the forces exerted on the mass m at the location x are

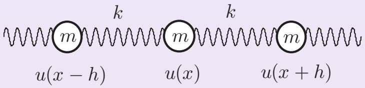
$$F_{\text{Newton}} = ma = m\frac{\partial^2 u}{\partial t^2}(x, t)\,,$$

$$\begin{aligned} F_{\text{Hooke}} &= k\big[u(x+h, t) - u(x, t)\big] - k\big[u(x, t) - u(x-h, t)\big] \\ &= k\big[u(x+h, t) - 2u(x, t) + u(x-h, t)\big]\,. \end{aligned}$$

§3.1 Models with One Temporal Variable and One Spatial Variable

§3.1.3 The 1-dimensional wave equations

① From Hooke's law:



imagine an array of little weights of mass m interconnected with massless springs of length h, and the springs have a stiffness of k (see the figure). If $u(x, t)$ measures the distance from the equilibrium of the mass situated at position x and time t, then the forces exerted on the mass m at the location x are

$$F_{\text{Newton}} = ma = m\frac{\partial^2 u}{\partial t^2}(x, t)\,,$$

$$\begin{aligned} F_{\text{Hooke}} &= k\big[u(x+h,t) - u(x,t)\big] - k\big[u(x,t) - u(x-h,t)\big] \\ &= k\big[u(x+h,t) - 2u(x,t) + u(x-h,t)\big]\,. \end{aligned}$$

§3.1 Models with One Temporal Variable and One Spatial Variable

The balance of force then implies that

$$m\frac{\partial^2 u}{\partial t^2}(x,t) = k\big[u(x+h,t) - 2u(x,t) + u(x-h,t)\big] .$$

If the array of weights consists of N weights spaced evenly over the length $L = (N+1)h$ of total mass $M = Nm$, and the total stiffness of the array $K = k/(N+1)$, then

$$\frac{\partial^2 u}{\partial t^2}(x,t) = \frac{N}{N+1}\frac{KL^2}{M}\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} .$$

Passing to the limit as $N \to \infty$ and $h \to 0$ (and assuming smoothness) we obtain the **wave equation**

$$u_{tt}(x,t) = c^2 u_{xx}(x,t) ,$$

where $c^2 = \dfrac{KL^2}{M}$.

§3.1 Models with One Temporal Variable and One Spatial Variable

The balance of force then implies that

$$m\frac{\partial^2 u}{\partial t^2}(x,t) = k\big[u(x+h,t) - 2u(x,t) + u(x-h,t)\big] .$$

If the array of weights consists of N weights spaced evenly over the length $L = (N+1)h$ of total mass $M = Nm$, and the total stiffness of the array $K = k/(N+1)$, then

$$\frac{\partial^2 u}{\partial t^2}(x,t) = \frac{N}{N+1}\frac{KL^2}{M}\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} .$$

Passing to the limit as $N \to \infty$ and $h \to 0$ (and assuming smoothness) we obtain the **wave equation**

$$u_{tt}(x,t) = c^2 u_{xx}(x,t) ,$$

where $c^2 = \dfrac{KL^2}{M}$.

§3.1 Models with One Temporal Variable and One Spatial Variable

The balance of force then implies that

$$m\frac{\partial^2 u}{\partial t^2}(x,t) = k\big[u(x+h,t) - 2u(x,t) + u(x-h,t)\big]\,.$$

If the array of weights consists of N weights spaced evenly over the length $L = (N+1)h$ of total mass $M = Nm$, and the total stiffness of the array $K = k/(N+1)$, then

$$\frac{\partial^2 u}{\partial t^2}(x,t) = \frac{N}{N+1}\,\frac{KL^2}{M}\,\frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2}\,.$$

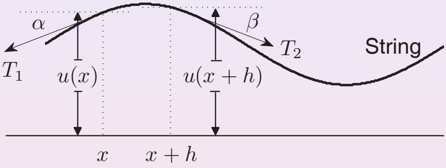
Passing to the limit as $N \to \infty$ and $h \to 0$ (and assuming smoothness) we obtain the **wave equation**

$$u_{tt}(x,t) = c^2 u_{xx}(x,t)\,,$$

where $c^2 = \dfrac{KL^2}{M}$.

§3.1 Models with One Temporal Variable and One Spatial Variable

② Equation of vibrating string: let $u(x,t)$ measure the distance of a string from its equilibrium, and $T(x,t)$ denote the tension of the string at position x and time t.



Assuming only motion in the vertical direction, the horizontal component of tensions $T_1 = T(x,t)$ and $T_2 = T(x+h,t)$ have to be the same; thus

$$T_1 \cos\alpha = T_2 \cos\beta\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

Noting that

$$\cos\alpha = \frac{1}{\sec\alpha} = \frac{1}{\sqrt{1+\tan^2\alpha}} = \frac{1}{\sqrt{1+\tan^2(\pi+\alpha)}} = \frac{1}{\sqrt{1+u_x(x,t)^2}},$$

$$\cos\beta = \frac{1}{\sec\beta} = \frac{1}{\sqrt{1+\tan^2\beta}} = \frac{1}{\sqrt{1+\tan^2(2\pi-\beta)}} = \frac{1}{\sqrt{1+u_x(x+h,t)^2}},$$

the identity $T_1\cos\alpha = T_2\cos\beta$ implies that the function

$$\frac{T(x,t)}{\sqrt{1+u_x(x,t)^2}}$$

is constant in x (but not necessary constant in t), and we denote this constant as $\tau(t)$.

§3.1 Models with One Temporal Variable and One Spatial Variable

By the fact that the vertical component of T_1 and T_2 induce the vertical motion, we obtain that

$$\begin{aligned}\int_x^{x+h} \mu(y) \frac{\partial^2 u(y,t)}{\partial t^2}\, dy &= -T_2 \sin\beta - T_1 \sin\alpha \\ &= -(T_2 \cos\beta) \tan\beta - (T_1 \cos\alpha) \tan\alpha \\ &= \tau(t) \tan(2\pi - \beta) - \tau(t) \tan(\pi + \alpha) \\ &= \tau(t) \big[u_x(x+h,t) - u_x(x,t)\big]\,,\end{aligned}$$

where μ denotes the density of the string, and the integral on the left-hand side is the total force due to the acceleration. Dividing both sides through by h and passing to the limit as $h \to 0$, we obtain

$$\mu(x) u_{tt}(x,t) = \tau(t) u_{xx}(x,t)\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

By the fact that the vertical component of T_1 and T_2 induce the vertical motion, we obtain that

$$\begin{aligned}\int_x^{x+h} \mu(y)\frac{\partial^2 u(y,t)}{\partial t^2}\,dy &= -T_2\sin\beta - T_1\sin\alpha \\ &= -(T_2\cos\beta)\tan\beta - (T_1\cos\alpha)\tan\alpha \\ &= \tau(t)\tan(2\pi-\beta) - \tau(t)\tan(\pi+\alpha) \\ &= \tau(t)\big[u_x(x+h,t) - u_x(x,t)\big]\,,\end{aligned}$$

where μ denotes the density of the string, and the integral on the left-hand side is the total force due to the acceleration. Dividing both sides through by h and passing to the limit as $h \to 0$, we obtain

$$\mu(x)u_{tt}(x,t) = \tau(t)u_{xx}(x,t)\,.$$

§3.1 Models with One Temporal Variable and One Spatial Variable

If there is an external forcing f acting on the string, then the derived wave equation becomes

$$\mu(x)u_{tt}(x,t) = \tau(t)u_{xx}(x,t) + f(x,t)\,.$$

If μ is constant in x and τ is constant in t (which is a reasonable assumption if the vibration of the string is very small and uniform), then the wave equation above reduces to

$$u_{tt}(x,t) = c^2 u_{xx}(x,t) + \frac{1}{\mu}f(x,t)\,,$$

where $c^2 = \tau/\mu$.

Initial conditions: Since the PDE is second order in t, to determined the state of the we need to impose two initial conditions

$$u(x,0) = \varphi(x)\,, \qquad u_t(x,0) = \psi(x) \qquad \forall\, x \in [0,L]\,,$$

where φ and ψ are given functions.

§3.1 Models with One Temporal Variable and One Spatial Variable

If there is an external forcing f acting on the string, then the derived wave equation becomes

$$\mu(x)u_{tt}(x,t) = \tau(t)u_{xx}(x,t) + f(x,t)\,.$$

If μ is constant in x and τ is constant in t (which is a reasonable assumption if the vibration of the string is very small and uniform), then the wave equation above reduces to

$$u_{tt}(x,t) = c^2 u_{xx}(x,t) + \frac{1}{\mu} f(x,t)\,,$$

where $c^2 = \tau/\mu$.

Initial conditions: Since the PDE is second order in t, to determined the state of the we need to impose two initial conditions

$$u(x,0) = \varphi(x)\,, \qquad u_t(x,0) = \psi(x) \qquad \forall\, x \in [0,L]\,,$$

where φ and ψ are given functions.

§3.1 Models with One Temporal Variable and One Spatial Variable

If there is an external forcing f acting on the string, then the derived wave equation becomes

$$\mu(x) u_{tt}(x,t) = \tau(t) u_{xx}(x,t) + f(x,t) \, .$$

If μ is constant in x and τ is constant in t (which is a reasonable assumption if the vibration of the string is very small and uniform), then the wave equation above reduces to

$$u_{tt}(x,t) = c^2 u_{xx}(x,t) + \frac{1}{\mu} f(x,t) \, ,$$

where $c^2 = \tau/\mu$.

Initial conditions: Since the PDE is second order in t, to determined the state of the we need to impose two initial conditions

$$u(x,0) = \varphi(x) \, , \qquad u_t(x,0) = \psi(x) \qquad \forall \, x \in [0,L] \, ,$$

where φ and ψ are given functions.

§3.1 Models with One Temporal Variable and One Spatial Variable

Wave equations: $\mu(x)u_{tt}(x,t) = \tau(t)u_{xx}(x,t) + f(x,t)$.

Initial conditions: Since the PDE is second order in t, to determined the state of the we need to impose two initial conditions

$$u(x,0) = \varphi(x)\,, \qquad u_t(x,0) = \psi(x) \qquad \forall\, x \in [0,L]\,,$$

where φ and ψ are given functions.

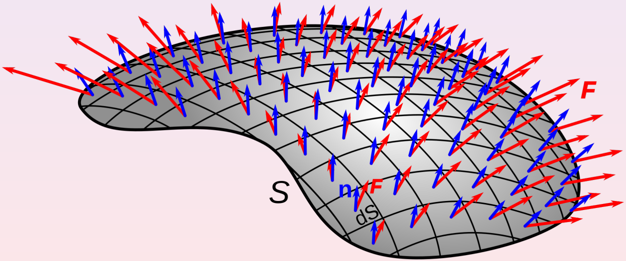
Boundary conditions:

1. Vibration string with fixed ends: $u(0,t) = u(L,t) = 0$ - this kind of boundary condition is also called **Dirichlet** BC.
2. Vibration string with free ends: $u_x(0,t) = u_x(L,t) = 0$ - this kind of boundary condition is also called **Neumann** BC.
3. Mixed boundary conditions: $u(0,t) = u_x(L,t) = 0$ or $u(L,t) = u_x(0,t) = 0$.

§3.3 Models with Several Spatial Variables

§3.3.1 Equation of continuity

Let u be the density of concentration of some physical quantity ($u = u(x,t)$) in a domain $\Omega \subseteq \mathbb{R}^n$, where $n = 2$ or $n = 3$, and let $\boldsymbol{F}$ be the flux of the quantity; that is, $\boldsymbol{F} \cdot \mathbf{n}\, dS$ is the **flow rate** of the quantity that passes through an area dS in the direction $\mathbf{n}$ normal to dS:



§3.3 Models with Several Spatial Variables

Then for a given open set $\mathcal{O} \subset\subset \Omega$ so that ∂O is piecewise smooth,

1. the change of the total amount of the quantity in $\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_{\mathcal{O}} \big[u(x, t + \Delta t) - u(x, t)\big]\, dx\,.$$
2. the total amount of the quantity flows out of $\mathcal{O}$ through $\partial\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\partial\mathcal{O}} (\boldsymbol{F} \cdot \mathbf{n})(x, s)\, dSds\,,$$
where $\mathbf{n}$ is the outward-pointing unit normal of $\partial\mathcal{O}$.
3. if there is a source of the quantity, the total amount of the quantity in $\mathcal{O}$ produced by the source from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\mathcal{O}} q(x, s) dxds\,,$$
where q is the strength of sources for the quantity.

§3.3 Models with Several Spatial Variables

Then for a given open set $\mathcal{O} \subset\subset \Omega$ so that ∂O is piecewise smooth,

1. the change of the total amount of the quantity in $\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_{\mathcal{O}} \big[u(x, t + \Delta t) - u(x, t)\big]\, dx\,.$$
2. the total amount of the quantity flows out of $\mathcal{O}$ through $\partial\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\partial\mathcal{O}} (\boldsymbol{F} \cdot \mathbf{n})(x, s)\, dSds\,,$$
where $\mathbf{n}$ is the outward-pointing unit normal of $\partial\mathcal{O}$.
3. if there is a source of the quantity, the total amount of the quantity in $\mathcal{O}$ produced by the source from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\mathcal{O}} q(x, s) dxds\,,$$
where q is the strength of sources for the quantity.

§3.3 Models with Several Spatial Variables

Then for a given open set $\mathcal{O} \subset\subset \Omega$ so that ∂O is piecewise smooth,

1. the change of the total amount of the quantity in $\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_{\mathcal{O}} \big[u(x, t + \Delta t) - u(x, t)\big]\, dx\,.$$
2. the total amount of the quantity flows out of $\mathcal{O}$ through $\partial\mathcal{O}$ from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\partial\mathcal{O}} (\boldsymbol{F} \cdot \mathbf{n})(x, s)\, dS ds\,,$$
where $\mathbf{n}$ is the outward-pointing unit normal of $\partial\mathcal{O}$.
3. if there is a source of the quantity, the total amount of the quantity in $\mathcal{O}$ produced by the source from time t to $t + \Delta t$ is
$$\int_t^{t+\Delta t} \int_{\mathcal{O}} q(x, s) dx ds\,,$$
where q is the strength of sources for the quantity.

§3.3 Models with Several Spatial Variables

Therefore, the balance of the amount of the quantity in $\mathcal{O}$ implies that

$$\int_{\mathcal{O}} \big[u(x, t+\Delta t) - u(x,t)\big]\, dx = -\int_t^{t+\Delta t}\int_{\partial\mathcal{O}} (\boldsymbol{F}\cdot\mathbf{n})(x,t')\, dS dt' + \int_t^{t+\Delta t}\int_{\mathcal{O}} q(x,t') dx dt'$$

for all "good" subset $\mathcal{O} \subseteq \Omega$, here a "good" set refers to a set with piecewise smooth boundary. Dividing both sides of the equation above by Δt and passing to the limit as $\Delta t \to 0$, we obtain that

$$\frac{d}{dt}\int_{\mathcal{O}} u(x,t)\, dx = -\int_{\partial\mathcal{O}} (\boldsymbol{F}\cdot\mathbf{n})(x,t)\, dS + \int_{\mathcal{O}} q(x,t)\, dx$$

for all "good" open subset $\mathcal{O} \subseteq \Omega$.

§3.3 Models with Several Spatial Variables

Therefore, the balance of the amount of the quantity in $\mathcal{O}$ implies that

$$\int_{\mathcal{O}} \big[u(x, t+\Delta t) - u(x,t)\big]\,dx = -\int_t^{t+\Delta t}\int_{\partial\mathcal{O}} (\boldsymbol{F}\cdot\mathbf{n})(x,t')\,dSdt' + \int_t^{t+\Delta t}\int_{\mathcal{O}} q(x,t')dxdt'$$

for all "good" subset $\mathcal{O} \subseteq \Omega$, here a "good" set refers to a set with piecewise smooth boundary. Dividing both sides of the equation above by Δt and passing to the limit as $\Delta t \to 0$, we obtain that

$$\frac{d}{dt}\int_{\mathcal{O}} u(x,t)\,dx = -\int_{\partial\mathcal{O}} (\boldsymbol{F}\cdot\mathbf{n})(x,t)\,dS + \int_{\mathcal{O}} q(x,t)\,dx$$

for all "good" open subset $\mathcal{O} \subseteq \Omega$.

§3.3 Models with Several Spatial Variables

If u is smooth, by the **divergence theorem** we find that

$$\int_{\mathcal{O}} u_t \, dx = \int_{\mathcal{O}} (-\mathrm{div}\boldsymbol{F} + q) \, dx \quad \text{for all "good" open subset } \mathcal{O} \subseteq \Omega,$$

or equivalently,

$$\int_{\mathcal{O}} \big(u_t + \mathrm{div}\boldsymbol{F} - q\big) \, dx = 0 \quad \text{for all "good" open subset } \mathcal{O} \subseteq \Omega.$$

Since $\mathcal{O}$ is given arbitrarily in Ω, we conclude that

$$u_t + \mathrm{div}\boldsymbol{F} = q \qquad \text{in} \quad \Omega \times (0, T)\,.$$

The equation above is called the **equation of continuity**.

The Divergence Theorem: Suppose that $\partial\mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If $\boldsymbol{F}$ is a smooth vector field, then $\displaystyle\int_{\partial\mathcal{O}} \boldsymbol{F}\cdot\mathbf{n}\, dS = \int_{\mathcal{O}} \mathrm{div}\boldsymbol{F}\, dx.$

§3.3 Models with Several Spatial Variables

If u is smooth, by the **divergence theorem** we find that

$$\int_{\mathcal{O}} u_t \, dx = \int_{\mathcal{O}} (-\operatorname{div} \boldsymbol{F} + q) \, dx \quad \text{for all "good" open subset } \mathcal{O} \subseteq \Omega,$$

or equivalently,

$$\int_{\mathcal{O}} \big(u_t + \operatorname{div} \boldsymbol{F} - q\big) \, dx = 0 \quad \text{for all "good" open subset } \mathcal{O} \subseteq \Omega.$$

Since $\mathcal{O}$ is given arbitrarily in Ω, we conclude that

$$u_t + \operatorname{div} \boldsymbol{F} = q \qquad \text{in} \quad \Omega \times (0, T)\,.$$

The equation above is called the **equation of continuity**.

The Divergence Theorem: Suppose that $\partial\mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If $\boldsymbol{F}$ is a smooth vector field, then $\displaystyle\int_{\partial\mathcal{O}} \boldsymbol{F} \cdot \mathbf{n} \, dS = \int_{\mathcal{O}} \operatorname{div} \boldsymbol{F} \, dx.$

§3.3 Models with Several Spatial Variables

- **The conservation of mass in fluid dynamics**

Let $\varrho(x,t)$ and $\boldsymbol{u}(x,t)$ denote the **density** and the **velocity** of a fluid at point x at time t. Then the **density flux** $\boldsymbol{F} = \rho \boldsymbol{u}$, and the equation of continuity reads

$$\varrho_t + \operatorname{div}(\varrho \boldsymbol{u}) = 0 \qquad \forall\, x \in \Omega\,, t \in \mathbb{R}\,.$$

In particular, if the density of a fluid is constant (incompressible fluid), then the velocity $\boldsymbol{u}$ of this fluid must satisfy

$$\operatorname{div}\boldsymbol{u} = 0 \qquad \text{in} \quad \Omega\,.$$

A vector field $\boldsymbol{u}$ satisfying $\operatorname{div}\boldsymbol{u} = 0$ everywhere in the domain is said to be **solenoidal** or **divergence-free**.

§3.3 Models with Several Spatial Variables

- **The conservation of mass in fluid dynamics**

Let $\varrho(x,t)$ and $\boldsymbol{u}(x,t)$ denote the **density** and the **velocity** of a fluid at point x at time t. Then the **density flux** $\boldsymbol{F} = \rho\boldsymbol{u}$, and the equation of continuity reads

$$\varrho_t + \operatorname{div}(\varrho\boldsymbol{u}) = 0 \qquad \forall\, x \in \Omega\,,\, t \in \mathbb{R}\,.$$

In particular, if the density of a fluid is constant (incompressible fluid), then the velocity $\boldsymbol{u}$ of this fluid must satisfy

$$\operatorname{div}\boldsymbol{u} = 0 \qquad \text{in} \quad \Omega\,.$$

A vector field $\boldsymbol{u}$ satisfying $\operatorname{div}\boldsymbol{u} = 0$ everywhere in the domain is said to be **solenoidal** or **divergence-free**.

§3.3 Models with Several Spatial Variables

§3.3.2 The heat equations

Let $\vartheta(x,t)$ defined on $\Omega \times (0,T]$ be the temperature of a material body at point $x \in \Omega$ at time $t \in (0,T]$, and $s(x)$, $\varrho(x)$, $\kappa(x)$ be the **specific heat**, **density**, and the inner **thermal conductivity** of the material body at x, and $Q(x,t)$ is the strength of the source of the heat energy. Then by the conservation of heat energy, similar to the derivation of the equation of continuity (with the **heat flux** $\boldsymbol{F} = -\kappa \nabla \vartheta$ in mind), we obtain that for any "good" open set $\mathcal{O} \subseteq \Omega$,

$$\frac{d}{dt}\int_{\mathcal{O}} s(x)\varrho(x)\vartheta(x,t)\,dx = \int_{\partial\mathcal{O}} \kappa(x)\nabla\vartheta(x,t)\cdot \mathbf{n}(x)\,dS + \int_{\mathcal{O}} Q(x,t)\,dx\,,$$

where $\mathbf{n}$ denotes the outward-pointing unit normal of $\mathcal{O}$.

§3.3 Models with Several Spatial Variables

§3.3.2 The heat equations

Let $\vartheta(x,t)$ defined on $\Omega \times (0,T]$ be the temperature of a material body at point $x \in \Omega$ at time $t \in (0,T]$, and $s(x)$, $\varrho(x)$, $\kappa(x)$ be the **specific heat**, **density**, and the inner **thermal conductivity** of the material body at x, and $Q(x,t)$ is the strength of the source of the heat energy. Then by the conservation of heat energy, similar to the derivation of the equation of continuity (with the **heat flux** $\boldsymbol{F} = -\kappa\nabla\vartheta$ in mind), we obtain that for any "good" open set $\mathcal{O} \subseteq \Omega$,

$$\frac{d}{dt}\int_{\mathcal{O}} s(x)\varrho(x)\vartheta(x,t)\,dx = \int_{\partial\mathcal{O}} \kappa(x)\nabla\vartheta(x,t)\cdot\mathbf{n}(x)\,dS + \int_{\mathcal{O}} Q(x,t)\,dx\,,$$

where $\mathbf{n}$ denotes the outward-pointing unit normal of $\mathcal{O}$.

§3.3 Models with Several Spatial Variables

Assume that ϑ is smooth, and $\mathcal{O}$ is a domain with piecewise smooth boundary. By the divergence theorem,

$$\int_{\mathcal{O}} s(x)\varrho(x)\vartheta_t(x,t)\,dx = \int_{\mathcal{O}} \mathrm{div}\big[\kappa(x)\nabla\vartheta(x,t)\big]\,dx + \int_{\mathcal{O}} Q(x,t)\,dx\,.$$

Since $\mathcal{O}$ is arbitrary, the equation above implies

$$s(x)\varrho(x)\vartheta_t(x,t) - \mathrm{div}\big[\kappa(x)\nabla\vartheta(x,t)\big] = Q(x,t) \quad \forall\, x\in\Omega\,,\, t\in(0,T]\,.$$

If s, ϱ and κ are constants (uniform material), then

$$\vartheta_t = \alpha^2\Delta\vartheta + q(x,t) \quad \forall\, x\in\Omega\,,\, t\in(0,T]\,,$$

where $\alpha^2 = \frac{\kappa}{s\varrho}$, $q = \frac{1}{s\varrho}Q$ and Δ is the Laplace operator (and $\Delta\vartheta$ reads laplacian theta) defined by

$$\Delta\vartheta \equiv \mathrm{div}(\nabla\vartheta) = \sum_{i=1}^{n} \frac{\partial^2\vartheta}{\partial x_i^2}\,.$$

This is the standard **heat equation**.

§3.3 Models with Several Spatial Variables

Assume that ϑ is smooth, and $\mathcal{O}$ is a domain with piecewise smooth boundary. By the divergence theorem,

$$\int_{\mathcal{O}} s(x)\varrho(x)\vartheta_t(x,t)\,dx = \int_{\mathcal{O}} \mathrm{div}\big[\kappa(x)\nabla\vartheta(x,t)\big]\,dx + \int_{\mathcal{O}} Q(x,t)\,dx\,.$$

Since $\mathcal{O}$ is arbitrary, the equation above implies

$$s(x)\varrho(x)\vartheta_t(x,t) - \mathrm{div}\big[\kappa(x)\nabla\vartheta(x,t)\big] = Q(x,t) \quad \forall\, x \in \Omega\,,\, t \in (0,T]\,.$$

If s, ϱ and κ are constants (uniform material), then

$$\vartheta_t = \alpha^2\Delta\vartheta + q(x,t) \quad \forall\, x \in \Omega\,,\, t \in (0,T]\,,$$

where $\alpha^2 = \dfrac{\kappa}{s\varrho}$, $q = \dfrac{1}{s\varrho}Q$ and Δ is the Laplace operator (and $\Delta\vartheta$ reads laplacian theta) defined by

$$\Delta\vartheta \equiv \mathrm{div}(\nabla\vartheta) = \sum_{i=1}^{n} \frac{\partial^2\vartheta}{\partial x_i^2}\,.$$

This is the standard **heat equation**.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

- **Initial condition**: $\vartheta(x,0) = \vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \varnothing$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x,t) = g(x,t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial \mathbf{N}} \equiv \mathbf{N} \cdot \nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

- **Initial condition**: $\vartheta(x, 0) = \vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \emptyset$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x, t) = g(x, t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial \mathbf{N}} \equiv \mathbf{N} \cdot \nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

- **Initial condition**: $\vartheta(x, 0) = \vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \emptyset$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x, t) = g(x, t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial \mathbf{N}} \equiv \mathbf{N} \cdot \nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

- **Initial condition**: $\vartheta(x,0) = \vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \varnothing$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x,t) = g(x,t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial \mathbf{N}} \equiv \mathbf{N} \cdot \nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

- **Initial condition**: $\vartheta(x,0)=\vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \varnothing$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x,t)=g(x,t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial \mathbf{N}} \equiv \mathbf{N}\cdot\nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial \mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

We need complementary conditions to specify a particular solution of the heat equation.

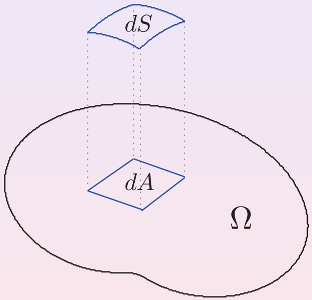
- **Initial condition**: $\vartheta(x,0) = \vartheta_0(x)$ for some given function $\vartheta_0(x)$.
- **Boundary condition**: if $\partial\Omega \neq \varnothing$, some boundary condition of u at $x \in \partial\Omega$ for all time have to be introduced by physical reason to specify a unique solution.

1. **Dirichlet boundary condition**: $\vartheta(x,t) = g(x,t)$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where g is a given function.
2. **Neumann boundary condition**: $\dfrac{\partial\vartheta}{\partial\mathbf{N}} = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where $\dfrac{\partial\vartheta}{\partial\mathbf{N}} \equiv \mathbf{N}\cdot\nabla\vartheta$ and g is a given function.
3. **Robin boundary condition**: $\dfrac{\partial\vartheta}{\partial\mathbf{N}} + h\vartheta = g$ for all $x \in \partial\Omega$ and $t \geqslant 0$, where h and g are given functions.

§3.3 Models with Several Spatial Variables

§3.3.3 The wave equations

Consider the membrane (of a drum) as a graph of a function $z = u(x_1, x_2)$ for $(x_1, x_2) \in \Omega$.



Question: If the deformation of the membrane is due to a small external force f, what is the relation between f and u?

§3.3 Models with Several Spatial Variables

Idea: The membrane stores certain energy $E(u)$ so that the deformation of the membrane changes the energy stored in the membrane which balances the work done by the external force f.

Suppose that an extra small external force $\triangle f = \triangle f(x_1, x_2)$ is suddenly added onto the membrane (so that the total force exerted on the membrane is $f + \triangle f$), and the membrane deforms to the surface $z = (u + \triangle u)(x_1, x_2)$ slowly (so the inertia does not have any effect). Then the extra energy needed to deform the membrane is $E(u + \triangle u) - E(u)$, while this extra work is done by the force $f + \triangle f$ given by

$$\int_\Omega (f + \triangle f) \triangle u \, dx \, .$$

Therefore,

$$E(u + \triangle u) - E(u) = \int_\Omega (f + \triangle f) \triangle u \, dx \, .$$

§3.3 Models with Several Spatial Variables

Idea: The membrane stores certain energy $E(u)$ so that the deformation of the membrane changes the energy stored in the membrane which balances the work done by the external force f.

Suppose that an extra small external force $\triangle f = \triangle f(x_1, x_2)$ is suddenly added onto the membrane (so that the total force exerted on the membrane is $f + \triangle f$), and the membrane deforms to the surface $z = (u + \triangle u)(x_1, x_2)$ slowly (so the inertia does not have any effect). Then the extra energy needed to deform the membrane is $E(u + \triangle u) - E(u)$, while this extra work is done by the force $f + \triangle f$ given by

$$\int_\Omega (f + \triangle f) \triangle u \, dx .$$

Therefore,

$$E(u + \triangle u) - E(u) = \int_\Omega (f + \triangle f) \triangle u \, dx .$$

§3.3 Models with Several Spatial Variables

Idea: The membrane stores certain energy $E(u)$ so that the deformation of the membrane changes the energy stored in the membrane which balances the work done by the external force f.

Suppose that an extra small external force $\triangle f = \triangle f(x_1, x_2)$ is suddenly added onto the membrane (so that the total force exerted on the membrane is $f + \triangle f$), and the membrane deforms to the surface $z = (u + \triangle u)(x_1, x_2)$ slowly (so the inertia does not have any effect). Then the extra energy needed to deform the membrane is $E(u + \triangle u) - E(u)$, while this extra work is done by the force $f + \triangle f$ given by

$$\int_\Omega (f + \triangle f)\triangle u\,dx.$$

Therefore,

$$E(u + \triangle u) - E(u) = \int_\Omega (f + \triangle f)\triangle u\,dx.$$

§3.3 Models with Several Spatial Variables

Even though we have assumed implicitly that $\triangle u$ is a function of $\triangle f$ (the deformation of the membrane is due to the change of external force), we can also assume that $\triangle f$ is a function of $\triangle u$ (so that we can modify $\triangle u$ independently). Let φ be an "**admissible**" function (which means that $t\varphi$ can be used as $\triangle u$ for each $t \ll 1$) and $\triangle u = t\varphi$. Then if $t \neq 0$,

$$\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega (f+\triangle f)\varphi\,dx\,.$$

Note that $\triangle f \to 0$ as $\triangle u \to 0$, so we have $\triangle f \to 0$ as $t \to 0$. Passing to the limit as $t \to 0$, we find that

$$\lim_{t\to 0}\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega f\varphi\,dx \quad \text{for all admissible } \varphi\,. \tag{17}$$

Equation (17) above always holds when considering time independent problems.

§3.3 Models with Several Spatial Variables

Even though we have assumed implicitly that $\triangle u$ is a function of $\triangle f$ (the deformation of the membrane is due to the change of external force), we can also assume that $\triangle f$ is a function of $\triangle u$ (so that we can modify $\triangle u$ independently). Let φ be an "**admissible**" function (which means that $t\varphi$ can be used as $\triangle u$ for each $t \ll 1$) and $\triangle u = t\varphi$. Then if $t \neq 0$,

$$\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega (f+\triangle f)\varphi\,dx\,.$$

Note that $\triangle f \to 0$ as $\triangle u \to 0$, so we have $\triangle f \to 0$ as $t \to 0$. Passing to the limit as $t \to 0$, we find that

$$\lim_{t\to 0}\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega f\varphi\,dx \quad \text{for all admissible } \varphi\,. \tag{17}$$

Equation (17) above always holds when considering time independent problems.

§3.3 Models with Several Spatial Variables

Even though we have assumed implicitly that $\triangle u$ is a function of $\triangle f$ (the deformation of the membrane is due to the change of external force), we can also assume that $\triangle f$ is a function of $\triangle u$ (so that we can modify $\triangle u$ independently). Let φ be an "**admissible**" function (which means that $t\varphi$ can be used as $\triangle u$ for each $t \ll 1$) and $\triangle u = t\varphi$. Then if $t \neq 0$,

$$\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega (f+\triangle f)\varphi\, dx\,.$$

Note that $\triangle f \to 0$ as $\triangle u \to 0$, so we have $\triangle f \to 0$ as $t \to 0$. Passing to the limit as $t \to 0$, we find that

$$\lim_{t\to 0} \frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega f\varphi\, dx \quad \text{for all admissible } \varphi\,. \tag{17}$$

Equation (17) above always holds when considering time independent problems.

§3.3 Models with Several Spatial Variables

Even though we have assumed implicitly that $\triangle u$ is a function of $\triangle f$ (the deformation of the membrane is due to the change of external force), we can also assume that $\triangle f$ is a function of $\triangle u$ (so that we can modify $\triangle u$ independently). Let φ be an "**admissible**" function (which means that $t\varphi$ can be used as $\triangle u$ for each $t \ll 1$) and $\triangle u = t\varphi$. Then if $t \neq 0$,

$$\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega (f+\triangle f)\varphi\,dx\,.$$

Note that $\triangle f \to 0$ as $\triangle u \to 0$, so we have $\triangle f \to 0$ as $t \to 0$. Passing to the limit as $t \to 0$, we find that

$$\lim_{t\to 0}\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega f\varphi\,dx \quad \text{for all admissible } \varphi\,. \tag{17}$$

Equation (17) above always holds when considering time independent problems.

§3.3 Models with Several Spatial Variables

Even though we have assumed implicitly that $\triangle u$ is a function of $\triangle f$ (the deformation of the membrane is due to the change of external force), we can also assume that $\triangle f$ is a function of $\triangle u$ (so that we can modify $\triangle u$ independently). Let φ be an "**admissible**" function (which means that $t\varphi$ can be used as $\triangle u$ for each $t \ll 1$) and $\triangle u = t\varphi$. Then if $t \neq 0$,

$$\frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega (f+\triangle f)\varphi\,dx\,.$$

Note that $\triangle f \to 0$ as $\triangle u \to 0$, so we have $\triangle f \to 0$ as $t \to 0$. Passing to the limit as $t \to 0$, we find that

$$\lim_{t\to 0} \frac{E(u+t\varphi)-E(u)}{t} = \int_\Omega f\varphi\,dx \quad \text{for all admissible } \varphi\,. \tag{17}$$

Equation (17) above always holds when considering time independent problems.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) \equiv \lim_{t\to 0} \frac{E(u+t\varphi) - E(u)}{t}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0} \sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A}$$

$$= \int_\Omega \operatorname{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi \operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,,$$

where identity $\operatorname{div}(\varphi \boldsymbol{F}) = \varphi \operatorname{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\begin{aligned}\delta E(u;\varphi) &\equiv \lim_{t\to 0} \frac{E(u+t\varphi) - E(u)}{t} \\ &= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0} \sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A} \\ &= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,,\end{aligned}$$

where identity $\mathrm{div}(\varphi \boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) \equiv \lim_{t\to 0} \frac{E(u+t\varphi) - E(u)}{t}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0} \sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,.$$

where identity $\mathrm{div}(\varphi \boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\,d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\,d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) = \lim_{t\to 0}\int_\Omega \mathrm{T}\frac{\sqrt{1+|\nabla u + t\nabla\varphi|^2} - \sqrt{1+|\nabla u|^2}}{t}\,d\mathbb{A}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0}\sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\,d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\,d\mathbb{A}$$

$$= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\,d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\,d\mathbb{A}\,.$$

where identity $\mathrm{div}(\varphi\boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) = \lim_{t\to 0} \int_\Omega \mathrm{T}\frac{\sqrt{1+|\nabla u + t\nabla\varphi|^2} - \sqrt{1+|\nabla u|^2}}{t}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0}\sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,.$$

where identity $\mathrm{div}(\varphi \boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) = \lim_{t\to 0} \int_\Omega \mathrm{T}\frac{\sqrt{1+|\nabla u + t\nabla\varphi|^2} - \sqrt{1+|\nabla u|^2}}{t}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0}\sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,.$$

where identity $\mathrm{div}(\varphi \boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

Suppose that the energy stored in the membrane is given by

$$E(u) = \int_\Omega \mathrm{T}\Big(\frac{dS}{d\mathbb{A}} - 1\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\big(\sqrt{1+|\nabla u|^2} - 1\big)\, d\mathbb{A}\,,$$

where T is called the tension of a membrane. In other words, to deform a membrane from its unforced equilibrium state to a surface S given by $z = u(x_1, x_2)$ requires the input of the energy $E(u)$.

Assuming that u is a smooth function, then

$$\delta E(u;\varphi) = \lim_{t\to 0} \int_\Omega \mathrm{T}\frac{\sqrt{1+|\nabla u + t\nabla\varphi|^2} - \sqrt{1+|\nabla u|^2}}{t}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{T}\Big(\frac{\partial}{\partial t}\Big|_{t=0}\sqrt{1+|\nabla u + t\nabla\varphi|^2}\Big)\, d\mathbb{A} = \int_\Omega \mathrm{T}\frac{\nabla u\cdot\nabla\varphi}{\sqrt{1+|\nabla u|^2}}\, d\mathbb{A}$$

$$= \int_\Omega \mathrm{div}\Big(\frac{\mathrm{T}\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A} - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,.$$

where identity $\mathrm{div}(\varphi \boldsymbol{F}) = \varphi\,\mathrm{div}\boldsymbol{F} + \boldsymbol{F}\cdot\nabla\varphi$ is used.

§3.3 Models with Several Spatial Variables

By the divergence theorem, with $\mathbf{N}$ denoting the outward-pointing unit normal on $\partial\Omega$,

$$\delta E(u;\varphi) = \int_{\partial\Omega} \frac{T\varphi\nabla u}{\sqrt{1+|\nabla u|^2}} \cdot \mathbf{N}\, ds - \int_{\Omega} \varphi\, \mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,;$$

thus (17) implies that

$$\int_{\Omega} \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}}\frac{\partial u}{\partial \mathbf{N}}\varphi\, ds = 0$$

for all admissible φ. In particular,

$$\int_{\Omega} \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} = 0$$

for all admissble φ that vanishes on $\partial\Omega$. Therefore,

$$\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f = 0 \qquad \text{in} \quad \Omega\,.$$

§3.3 Models with Several Spatial Variables

By the divergence theorem, with $\mathbf{N}$ denoting the outward-pointing unit normal on $\partial\Omega$,

$$\delta E(u;\varphi) = \int_{\partial\Omega} \frac{T\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\cdot\mathbf{N}\,ds - \int_\Omega \varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\,d\mathbb{A}\,;$$

thus (17) implies that

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\,d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}}\frac{\partial u}{\partial\mathbf{N}}\varphi\,ds = 0$$

for all admissible φ. In particular,

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\,d\mathbb{A} = 0$$

for all admissble φ that vanishes on $\partial\Omega$. Therefore,

$$\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f = 0 \qquad \text{in} \quad \Omega\,.$$

§3.3 Models with Several Spatial Variables

By the divergence theorem, with $\mathbf{N}$ denoting the outward-pointing unit normal on $\partial\Omega$,

$$\delta E(u;\varphi) = \int_{\partial\Omega} \frac{T\varphi\nabla u}{\sqrt{1+|\nabla u|^2}} \cdot \mathbf{N}\, ds - \int_\Omega \varphi\, \mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\, d\mathbb{A}\,;$$

thus (17) implies that

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}} \frac{\partial u}{\partial \mathbf{N}} \varphi\, ds = 0$$

for all admissible φ. In particular,

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} = 0$$

for all admissble φ that vanishes on $\partial\Omega$. Therefore,

$$\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f = 0 \qquad \text{in} \quad \Omega\,.$$

§3.3 Models with Several Spatial Variables

By the divergence theorem, with $\mathbf{N}$ denoting the outward-pointing unit normal on $\partial\Omega$,

$$\delta E(u;\varphi) = \int_{\partial\Omega} \frac{T\varphi\nabla u}{\sqrt{1+|\nabla u|^2}}\cdot\mathbf{N}\,ds - \int_{\Omega}\varphi\,\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\,d\mathbb{A}\,;$$

thus (17) implies that

$$\int_{\Omega}\Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)+f\Big]\varphi\,d\mathbb{A} - \int_{\partial\Omega}\frac{T}{\sqrt{1+|\nabla u|^2}}\frac{\partial u}{\partial\mathbf{N}}\varphi\,ds = 0$$

for all admissible φ. In particular,

$$\int_{\Omega}\Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)+f\Big]\varphi\,d\mathbb{A} = 0$$

for all admissble φ that vanishes on $\partial\Omega$. Therefore,

$$\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)+f=0 \qquad \text{in} \quad \Omega\,.$$

§3.3 Models with Several Spatial Variables

Therefore,

1. If the membrane is constrained on the boundary, then

$$-\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = f \quad \text{in} \quad \Omega\,, \qquad u = 0 \quad \text{on} \quad \partial\Omega\,.$$

2. If the membrane is not constrained on the boundary, then

$$\int_{\partial\Omega} \frac{\mathrm{T}}{\sqrt{1+|\nabla u|^2}} \frac{\partial u}{\partial \mathrm{N}} \varphi\, ds = 0 \quad \text{for all admissible } \varphi\,.$$

Therefore, by the assumption that $\mathrm{T} > 0$ everywhere we have $\frac{\partial u}{\partial \mathrm{N}} = 0$ on $\partial\Omega$. This shows that u satisfies

$$-\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = f \quad \text{in} \quad \Omega\,, \qquad \frac{\partial u}{\partial \mathrm{N}} = 0 \quad \text{on} \quad \partial\Omega\,.$$

§3.3 Models with Several Spatial Variables

Therefore,

1. If the membrane is constrained on the boundary, then

$$\begin{aligned} -\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= f \qquad \text{in} \quad \Omega\,, \\ u &= 0 \qquad \text{on} \quad \partial\Omega\,. \end{aligned}$$

2. If the membrane is not constrained on the boundary, then

$$\int_{\partial\Omega} \frac{\mathrm{T}}{\sqrt{1+|\nabla u|^2}} \frac{\partial u}{\partial \mathrm{N}} \varphi\, ds = 0 \quad \text{for all admissible } \varphi\,.$$

Therefore, by the assumption that $\mathrm{T} > 0$ everywhere we have $\frac{\partial u}{\partial \mathrm{N}} = 0$ on $\partial\Omega$. This shows that u satisfies

$$\begin{aligned} -\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= f \qquad \text{in} \quad \Omega\,, \\ \frac{\partial u}{\partial \mathrm{N}} &= 0 \qquad \text{on} \quad \partial\Omega\,. \end{aligned}$$

§3.3 Models with Several Spatial Variables

Therefore,

1. If the membrane is constrained on the boundary, then

$$
\begin{aligned}
-\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= f \qquad \text{in} \quad \Omega\,,\\
u &= 0 \qquad \text{on} \quad \partial\Omega\,.
\end{aligned}
$$

2. If the membrane is not constrained on the boundary, then

$$
\int_{\partial\Omega} \frac{\mathrm{T}}{\sqrt{1+|\nabla u|^2}} \frac{\partial u}{\partial \mathbf{N}} \varphi\, ds = 0 \quad \text{for all admissible } \varphi\,.
$$

Therefore, by the assumption that $\mathrm{T} > 0$ everywhere we have $\dfrac{\partial u}{\partial \mathbf{N}} = 0$ on $\partial\Omega$. This shows that u satisfies

$$
\begin{aligned}
-\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= f \qquad \text{in} \quad \Omega\,,\\
\frac{\partial u}{\partial \mathbf{N}} &= 0 \qquad \text{on} \quad \partial\Omega\,.
\end{aligned}
$$

§3.3 Models with Several Spatial Variables

Remark: If $u = 0$ on the boundary, we will **NOT** have an extra boundary condition $\dfrac{\partial u}{\partial \mathrm{N}} = 0$ on $\partial\Omega$ (even though at the first glance it seems the case). In fact, if $u = 0$ on $\partial\Omega$, then all possible displacement $\vartriangle u$ should also satisfy that $\vartriangle u = 0$ on $\partial\Omega$; thus φ also has to vanish on $\partial\Omega$ in the derivation of the equation (and this is what the term "admissible" refers to in this case). In other words, if the membrane is constrained, instead of obtaining

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}}\frac{\partial u}{\partial \mathrm{N}}\varphi\, ds = 0$$

for all admissible φ we should directly obtain

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Remark: If $u = 0$ on the boundary, we will **NOT** have an extra boundary condition $\dfrac{\partial u}{\partial \mathrm{N}} = 0$ on $\partial\Omega$ (even though at the first glance it seems the case). In fact, if $u = 0$ on $\partial\Omega$, then all possible displacement $\vartriangle u$ should also satisfy that $\vartriangle u = 0$ on $\partial\Omega$; thus φ also has to vanish on $\partial\Omega$ in the derivation of the equation (and this is what the term "admissible" refers to in this case). In other words, if the membrane is constrained, instead of obtaining

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}} \frac{\partial u}{\partial \mathrm{N}} \varphi\, ds = 0$$

for all admissible φ we should directly obtain

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Remark: If $u = 0$ on the boundary, we will **NOT** have an extra boundary condition $\dfrac{\partial u}{\partial \mathbf{N}} = 0$ on $\partial\Omega$ (even though at the first glance it seems the case). In fact, if $u = 0$ on $\partial\Omega$, then all possible displacement $\vartriangle u$ should also satisfy that $\vartriangle u = 0$ on $\partial\Omega$; thus φ also has to vanish on $\partial\Omega$ in the derivation of the equation (and this is what the term "admissible" refers to in this case). In other words, if the membrane is constrained, instead of obtaining

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} - \int_{\partial\Omega} \frac{T}{\sqrt{1+|\nabla u|^2}}\frac{\partial u}{\partial \mathbf{N}}\varphi\, ds = 0$$

for all admissible φ we should directly obtain

$$\int_\Omega \Big[\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) + f\Big]\varphi\, d\mathbb{A} = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Remark: By expanding the derivatives, we find that

$$\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} + \mathrm{T}\nabla u \cdot \nabla \frac{1}{\sqrt{1+|\nabla u|^2}}$$

$$= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} - \mathrm{T}\sum_{i,j=1}^{2} \frac{u_{x_i} u_{x_j} u_{x_i x_j}}{\sqrt{1+|\nabla u|^2}^3}\,.$$

Therefore, if $|\nabla u| \ll 1$, we find that

$$\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) \approx \operatorname{div}(\mathrm{T}\nabla u)\,;$$

thus if $|\nabla u| \ll 1$, the equations can be approximated by

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \quad & \text{in} \quad \Omega\,, \\ u = 0 \quad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{D}$$

and

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \quad & \text{in} \quad \Omega\,, \\ \frac{\partial u}{\partial \mathrm{N}} = 0 \quad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{N}$$

§3.3 Models with Several Spatial Variables

Remark: By expanding the derivatives, we find that

$$\begin{aligned}\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} + \mathrm{T}\nabla u\cdot\nabla\frac{1}{\sqrt{1+|\nabla u|^2}} \\ &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} - \mathrm{T}\sum_{i,j=1}^{2}\frac{u_{x_i}u_{x_j}u_{x_ix_j}}{\sqrt{1+|\nabla u|^2}^3}\,.\end{aligned}$$

Therefore, if $|\nabla u| \ll 1$, we find that

$$\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) \approx \operatorname{div}(\mathrm{T}\nabla u)\,;$$

thus if $|\nabla u| \ll 1$, the equations can be approximated by

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) &= f && \text{in } \Omega\,, \\ u &= 0 && \text{on } \partial\Omega\,. \end{aligned}\right. \tag{D}$$

and

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) &= f && \text{in } \Omega\,, \\ \frac{\partial u}{\partial \mathrm{N}} &= 0 && \text{on } \partial\Omega\,. \end{aligned}\right. \tag{N}$$

§3.3 Models with Several Spatial Variables

Remark: By expanding the derivatives, we find that

$$\begin{aligned}\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} + \mathrm{T}\nabla u\cdot\nabla\frac{1}{\sqrt{1+|\nabla u|^2}} \\ &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} - \mathrm{T}\sum_{i,j=1}^{2}\frac{u_{x_i}u_{x_j}u_{x_ix_j}}{\sqrt{1+|\nabla u|^2}^{3}}\,.\end{aligned}$$

Therefore, if $|\nabla u| \ll 1$, we find that

$$\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) \approx \operatorname{div}(\mathrm{T}\nabla u)\,;$$

thus if $|\nabla u| \ll 1$, the equations can be approximated by

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \quad & \text{in} \quad \Omega\,, \\ u = 0 \quad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{D}$$

and

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \quad & \text{in} \quad \Omega\,, \\ \frac{\partial u}{\partial \mathrm{N}} = 0 \quad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{N}$$

§3.3 Models with Several Spatial Variables

Remark: By expanding the derivatives, we find that

$$\begin{aligned}\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} + \mathrm{T}\nabla u\cdot\nabla\frac{1}{\sqrt{1+|\nabla u|^2}} \\ &= \frac{\operatorname{div}(\mathrm{T}\nabla u)}{\sqrt{1+|\nabla u|^2}} - \mathrm{T}\sum_{i,j=1}^{2}\frac{u_{x_i}u_{x_j}u_{x_ix_j}}{\sqrt{1+|\nabla u|^2}^3}\,.\end{aligned}$$

Therefore, if $|\nabla u| \ll 1$, we find that

$$\operatorname{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) \approx \operatorname{div}(\mathrm{T}\nabla u)\,;$$

thus if $|\nabla u| \ll 1$, the equations can be approximated by

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \qquad & \text{in} \quad \Omega\,, \\ u = 0 \qquad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{D}$$

and

$$\left\{\begin{aligned} -\operatorname{div}(\mathrm{T}\nabla u) = f \qquad & \text{in} \quad \Omega\,, \\ \frac{\partial u}{\partial \mathbf{N}} = 0 \qquad & \text{on} \quad \partial\Omega\,. \end{aligned}\right. \tag{N}$$

§3.3 Models with Several Spatial Variables

Let T be the tension, ϱ be the density, and f be the density of the external force which may depend on x and t. For the case of vibrating membranes, part of f induces the acceleration of the membrane which implies that

$$-\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega \times (0, T]$$

or under the assumption that $|\nabla u| \ll 1$, the PDE above is simplified as

$$-\mathrm{div}(\mathrm{T}\nabla u) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega \times (0, T].$$

This is in fact the **d'Alembert's principle** which states that the displacement u satisfies that

$$\int_\Omega \big[-\mathrm{T}\nabla u \cdot \nabla\varphi + (f - \varrho u_{tt})\varphi\big]\, dx = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Let T be the tension, ϱ be the density, and f be the density of the external force which may depend on x and t. For the case of vibrating membranes, part of f induces the acceleration of the membrane which implies that

$$-\operatorname{div}\left(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega \times (0, T]$$

or under the assumption that $|\nabla u| \ll 1$, the PDE above is simplified as

$$-\operatorname{div}(\mathrm{T}\nabla u) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega \times (0, T]\,.$$

This is in fact the **d'Alembert's principle** which states that the displacement u satisfies that

$$\int_\Omega \left[-\mathrm{T}\nabla u \cdot \nabla\varphi + (f - \varrho u_{tt})\varphi\right] dx = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Let T be the tension, ϱ be the density, and f be the density of the external force which may depend on x and t. For the case of vibrating membranes, part of f induces the acceleration of the membrane which implies that

$$-\mathrm{div}\Big(\frac{\mathrm{T}\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega\times(0,T]$$

or under the assumption that $|\nabla u| \ll 1$, the PDE above is simplified as

$$-\mathrm{div}(\mathrm{T}\nabla u) = f - \varrho u_{tt} \qquad \text{in} \quad \Omega\times(0,T]\,.$$

This is in fact the **d'Alembert's principle** which states that the displacement u satisfies that

$$\int_\Omega \big[-\mathrm{T}\nabla u\cdot\nabla\varphi + (f-\varrho u_{tt})\varphi\big]\,dx = 0$$

for all admissible φ.

§3.3 Models with Several Spatial Variables

Once the time derivative is involved in the PDEs, to fully determine the dynamics we need to impose initial conditions. For the wave equations, we need two initial conditions:

$$u(x,0) = \varphi(x)\,, \quad u_t(x,0) = \psi(x) \qquad \forall\, x \in \Omega\,,$$

where φ and ψ are given functions. Therefore, if $|\nabla u| \ll 1$,

1. Membrane fastened on the boundary:

$$\begin{cases} \varrho u_{tt} - \mathrm{div}(\mathrm{T}\nabla u) = f & \text{in} \quad \Omega \times (0,T]\,, \\ u = g & \text{on} \quad \partial\Omega \times (0,T]\,, \\ u(x,0) = \varphi(x)\,,\ u_t(x,0) = \psi(x) & \text{for all } x \in \Omega\,. \end{cases}$$

2. Membrane with free boundary:

$$\begin{cases} \varrho u_{tt} - \mathrm{div}(\mathrm{T}\nabla u) = f & \text{in} \quad \Omega \times (0,T]\,, \\ \dfrac{\partial u}{\partial \mathrm{N}} = 0 & \text{on} \quad \partial\Omega \times (0,T]\,, \\ u(x,0) = \varphi(x)\,,\ u_t(x,0) = \psi(x) & \text{for all } x \in \Omega\,. \end{cases}$$

§3.3 Models with Several Spatial Variables

Once the time derivative is involved in the PDEs, to fully determine the dynamics we need to impose initial conditions. For the wave equations, we need two initial conditions:

$$u(x,0) = \varphi(x)\,, \quad u_t(x,0) = \psi(x) \qquad \forall\, x \in \Omega\,,$$

where φ and ψ are given functions. Therefore, if $|\nabla u| \ll 1$,

1. Membrane fastened on the boundary:

$$\begin{cases} \varrho u_{tt} - \mathrm{div}(\mathrm{T}\nabla u) = f & \text{in} \quad \Omega \times (0,T]\,, \\ u = g & \text{on} \quad \partial\Omega \times (0,T]\,, \\ u(x,0) = \varphi(x)\,,\ u_t(x,0) = \psi(x) & \text{for all } x \in \Omega\,. \end{cases}$$

2. Membrane with free boundary:

$$\begin{cases} \varrho u_{tt} - \mathrm{div}(\mathrm{T}\nabla u) = f & \text{in} \quad \Omega \times (0,T]\,, \\ \dfrac{\partial u}{\partial \mathbf{N}} = 0 & \text{on} \quad \partial\Omega \times (0,T]\,, \\ u(x,0) = \varphi(x)\,,\ u_t(x,0) = \psi(x) & \text{for all } x \in \Omega\,. \end{cases}$$

§3.3 Models with Several Spatial Variables

§3.3.4 The Navier-Stokes equations

In this section we derive the governing equation for fluid velocity in a fluid system. Let Ω be the fluid domain in which the fluid flows, and ϱ and $\boldsymbol{u} = (u^1, u^2, u^3)$ be the density and the velocity of the fluid, respectively. Aside from the equation of continuity, at least an equation for the fluid velocity $\boldsymbol{u}$ is required to complete the system. Consider the conservation of momentum $\boldsymbol{m} = \varrho\boldsymbol{u}$.

§3.3 Models with Several Spatial Variables

By the fact that the rate of change of momentum of a body is equal to the resultant force acting on the body, the conservation of momentum states that for all $\mathcal{O} \subset\subset \Omega$ with (piecewise) smooth boundary,

$$\frac{d}{dt}\int_{\mathcal{O}} \boldsymbol{m}\,dx = -\int_{\partial\mathcal{O}} \boldsymbol{m}(\boldsymbol{u}\cdot\mathbf{n})\,dS + \int_{\partial\mathcal{O}} \boldsymbol{\sigma}\,dS + \int_{\mathcal{O}} \boldsymbol{f}\,dx\,,$$

where $\mathbf{n}$ is the **outward-pointing** unit normal of $\partial\mathcal{O}$ (so that the first integral on the right-hand side is due to the momentum flux), $\boldsymbol{f}$ is the external force (such as the gravity) on the fluid system (so that the third integral on the right-hand side is the source of momentum), and $\boldsymbol{\sigma}$ is the stress (應力) exerted by the fluid due to the friction (磨擦力)/viscosity (黏滯力) and the fluid pressure.

§3.3 Models with Several Spatial Variables

In the case of incompressible fluids, the stress is given by

$$\boldsymbol{\sigma} = 2\mu \mathrm{Def}\boldsymbol{u}\,\mathbf{n} - p\mathbf{n}\,,$$

where μ is called the **dynamical viscosity** (which may be a function of $\boldsymbol{u}$), p is the fluid **pressure**, and $\mathrm{Def}\boldsymbol{u}$, called the rate of strain tensor, is the symmetric part of the gradient of $\boldsymbol{u}$ given by

$$(\mathrm{Def}\boldsymbol{u})_{ij} = \frac{1}{2}\Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\Big)\,.$$

In other words, if $\mathbf{n} = (n_1, n_2, n_3)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$, then each component of $\boldsymbol{\sigma}$ is given by

$$\sigma_i = \mu \sum_{j=1}^{3} \Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\Big) n_j - p n_i\,.$$

The reason why the stress takes the form above will be explained later.

§3.3 Models with Several Spatial Variables

Therefore, writing $\boldsymbol{m} = (m^1, m^2, m^3)$ and $\boldsymbol{f} = (f^1, f^2, f^3)$, using the expression of $\boldsymbol{\sigma}$ in

$$\frac{d}{dt}\int_{\mathcal{O}} \boldsymbol{m}\, dx = -\int_{\partial\mathcal{O}} \boldsymbol{m}(\boldsymbol{u}\cdot\mathbf{n})\, dS + \int_{\partial\mathcal{O}} \boldsymbol{\sigma}\, dS + \int_{\mathcal{O}} \boldsymbol{f}\, dx\,,$$

we find that for each $1 \leqslant i \leqslant 3$ and all $\mathcal{O} \subset\subset \Omega$ with (piecewise) smooth boundary,

$$\begin{aligned}\frac{d}{dt}\int_{\mathcal{O}} m^i\, dx &+ \sum_{j=1}^{3}\int_{\partial\mathcal{O}} m^i u^j n_j\, dS \\ &= \sum_{j=1}^{3}\int_{\partial\mathcal{O}} \Big[\mu\Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\Big) n_j - p n_i\Big]\, dS + \int_{\mathcal{O}} f^i\, dx\,.\end{aligned}$$

§3.3 Models with Several Spatial Variables

Assuming the smoothness of the dependent variables, the divergence theorem imply that for each $1 \leqslant i \leqslant 3$,

$$\int_{\mathcal{O}} \Big\{ m^i_t + \sum_{j=1}^{3} \frac{\partial (m^i u^j)}{\partial x_j} + \frac{\partial p}{\partial x_i} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \Big[\mu \Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \Big) \Big] + f_i \Big\} \, dx = 0$$

for all regular domain $\mathcal{O} \subseteq \Omega$. As a consequence, we obtain the momentum equation

$$(\varrho \boldsymbol{u})_t + \operatorname{div}(\varrho \boldsymbol{u} \otimes \boldsymbol{u}) + \nabla p = \operatorname{div}(\mu \operatorname{Def} \boldsymbol{u}) + \boldsymbol{f} \quad \text{in} \quad \Omega \times (0, \infty),$$

where $\boldsymbol{u} \otimes \boldsymbol{u} = [u^i u^j]$ and for a matrix $a = [a_{ij}]$, $(\operatorname{div} a)_i \equiv \sum_{j=1}^{3} \frac{\partial a_{ij}}{\partial x_j}$.

The Divergence Theorem: Suppose that $\partial \mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If $\boldsymbol{F}$ is a smooth vector field, then $\displaystyle\int_{\partial \mathcal{O}} \boldsymbol{F} \cdot \mathbf{n} \, dS = \int_{\mathcal{O}} \operatorname{div} \boldsymbol{F} \, dx$.

§3.3 Models with Several Spatial Variables

Assuming the smoothness of the dependent variables, the divergence theorem imply that for each $1 \leqslant i \leqslant 3$,

$$\int_{\mathcal{O}} \Big\{ m^i_t + \sum_{j=1}^{3} \frac{\partial(m^i u^j)}{\partial x_j} + \frac{\partial p}{\partial x_i} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j}\Big[\mu\Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\Big)\Big] + f_i \Big\}\, dx = 0$$

for all regular domain $\mathcal{O} \subseteq \Omega$. As a consequence, we obtain the momentum equation

$$(\varrho \boldsymbol{u})_t + \operatorname{div}(\varrho \boldsymbol{u} \otimes \boldsymbol{u}) + \nabla p = \operatorname{div}(\mu \operatorname{Def} \boldsymbol{u}) + \boldsymbol{f} \quad \text{in} \quad \Omega \times (0, \infty)\,,$$

where $\boldsymbol{u} \otimes \boldsymbol{u} = [u^i u^j]$ and for a matrix $a = [a_{ij}]$, $(\operatorname{div} a)_i \equiv \sum_{j=1}^{3} \frac{\partial a_{ij}}{\partial x_j}$.

The Divergence Theorem: Suppose that $\partial\mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If g is a smooth function, then $\displaystyle\int_{\partial\mathcal{O}} g n_k \, dS = \int_{\mathcal{O}} \frac{\partial g}{\partial x_k}\, dx$ for each $1 \leqslant k \leqslant 3$.

§3.3 Models with Several Spatial Variables

Assuming the smoothness of the dependent variables, the divergence theorem imply that for each $1 \leqslant i \leqslant 3$,

$$\int_{\mathcal{O}} \left\{ m^i_t + \sum_{j=1}^{3} \frac{\partial(m^i u^j)}{\partial x_j} + \frac{\partial p}{\partial x_i} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j} \Big[\mu \Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \Big) \Big] + f_i \right\} dx = 0$$

for all regular domain $\mathcal{O} \subseteq \Omega$. As a consequence, we obtain the momentum equation

$$(\varrho \boldsymbol{u})_t + \operatorname{div}(\varrho \boldsymbol{u} \otimes \boldsymbol{u}) + \nabla p = \operatorname{div}(\mu \operatorname{Def} \boldsymbol{u}) + \boldsymbol{f} \quad \text{in} \quad \Omega \times (0, \infty),$$

where $\boldsymbol{u} \otimes \boldsymbol{u} = [u^i u^j]$ and for a matrix $a = [a_{ij}]$, $(\operatorname{div} a)_i \equiv \sum_{j=1}^{3} \frac{\partial a_{ij}}{\partial x_j}$.

The Divergence Theorem: Suppose that $\partial\mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If g is a smooth function, then $\int_{\partial\mathcal{O}} g n_k \, dS = \int_{\mathcal{O}} \frac{\partial g}{\partial x_k} \, dx$ for each $1 \leqslant k \leqslant 3$.

§3.3 Models with Several Spatial Variables

Assuming the smoothness of the dependent variables, the divergence theorem imply that for each $1 \leqslant i \leqslant 3$,

$$\int_{\mathcal{O}} \Big\{ m^i_t + \sum_{j=1}^{3} \frac{\partial(m^i u^j)}{\partial x_j} + \frac{\partial p}{\partial x_i} - \sum_{j=1}^{3} \frac{\partial}{\partial x_j}\Big[\mu\Big(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\Big)\Big] + f_i \Big\}\, dx = 0$$

for all regular domain $\mathcal{O} \subseteq \Omega$. As a consequence, we obtain the momentum equation

$$(\varrho \boldsymbol{u})_t + \mathrm{div}(\varrho \boldsymbol{u} \otimes \boldsymbol{u}) + \nabla p = \mathrm{div}(\mu \mathrm{Def} \boldsymbol{u}) + \boldsymbol{f} \quad \text{in} \quad \Omega \times (0, \infty),$$

where $\boldsymbol{u} \otimes \boldsymbol{u} = [u^i u^j]$ and for a matrix $a = [a_{ij}]$, $(\mathrm{div} a)_i \equiv \sum_{j=1}^{3} \frac{\partial a_{ij}}{\partial x_j}$.

The Divergence Theorem: Suppose that $\partial\mathcal{O}$ is smooth with outward-pointing unit normal $\mathbf{n}$. If g is a smooth function, then $\int_{\partial\mathcal{O}} g n_k \, dS = \int_{\mathcal{O}} \frac{\partial g}{\partial x_k}\, dx$ for each $1 \leqslant k \leqslant 3$.

§3.3 Models with Several Spatial Variables

Type of fluids:

1. **Newtonian fluids**: the viscosity μ is a constant.
2. **Non-Newtonian fluids**: the viscosity μ is a function of $\boldsymbol{u}$.

Consider the Newtonian case. If the fluids under consideration is incompressible, we let $\varrho = 1$ so that the equation of continuity and the momentum equation together imply the **Navier-Stokes equations**

$$\boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \quad \text{in} \quad \Omega\times(0,T)\,, \tag{18a}$$

$$\operatorname{div}\boldsymbol{u} = 0 \quad \text{in} \quad \Omega\times(0,T)\,, \tag{18b}$$

where the incompressibility condition (18b) is used to deduce that

$$\sum_{j=1}^{3}\frac{\partial}{\partial x_j}(u^i u^j) = \sum_{j=1}^{3}\Big(\frac{\partial u^i}{\partial x_j}u^j + u^i\frac{\partial u^j}{\partial x_j}\Big) = \sum_{j=1}^{3}\frac{\partial u^i}{\partial x_j}u^j \qquad \text{and}$$

§3.3 Models with Several Spatial Variables

Type of fluids:

1. **Newtonian fluids**: the viscosity μ is a constant.
2. **Non-Newtonian fluids**: the viscosity μ is a function of $\boldsymbol{u}$.

Consider the Newtonian case. If the fluids under consideration is incompressible, we let $\varrho = 1$ so that the equation of continuity and the momentum equation together imply the **Navier-Stokes equations**

$$\boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \quad \text{in} \quad \Omega\times(0,T)\,, \tag{18a}$$

$$\mathrm{div}\boldsymbol{u} = 0 \quad \text{in} \quad \Omega\times(0,T)\,, \tag{18b}$$

where the incompressibility condition (18b) is used to deduce that

$$\sum_{j=1}^{3}\frac{\partial}{\partial x_j}\Big[\mu\Big(\frac{\partial u^i}{\partial x_j}+\frac{\partial u^j}{\partial x_i}\Big)\Big] = \mu\sum_{j=1}^{3}\frac{\partial}{\partial x_j}\Big(\frac{\partial u^i}{\partial x_j}+\frac{\partial u^j}{\partial x_i}\Big) = \mu\sum_{j=1}^{3}\frac{\partial^2 u^i}{\partial x_j^2} = \mu\Delta u^i.$$

§3.3 Models with Several Spatial Variables

To fully determine the dynamics of fluids, in addition to the Navier-Stokes equations

$$\begin{aligned} \boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \qquad &\text{in} \quad \Omega\times(0,T)\,, \\ \mathrm{div}\boldsymbol{u} = 0 \qquad &\text{in} \quad \Omega\times(0,T)\,, \end{aligned}$$

we also need to impose initial and boundary conditions.

Initial conditions: $\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x)$ for all $x\in\Omega$.

Boundary condition:

1. **No-slip boundary condition**: $\boldsymbol{u} = \boldsymbol{0}$ on $\partial\Omega$.
2. **Navier boundary condition**: $\boldsymbol{u}\cdot\mathrm{N} = 0$ and $\mathrm{N}\times(\mu\mathrm{Def}\boldsymbol{u}\mathrm{N}) = \alpha(\mathrm{N}\times\boldsymbol{u})$ on $\partial\Omega$ for some constant $\alpha > 0$. This condition is based on the assumption that the traction force due to the viscous effect is proportional to the fluid velocity on the boundary.

§3.3 Models with Several Spatial Variables

To fully determine the dynamics of fluids, in addition to the Navier-Stokes equations

$$\begin{aligned} \boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \qquad &\text{in} \quad \Omega\times(0,T)\,, \\ \operatorname{div}\boldsymbol{u} = 0 \qquad &\text{in} \quad \Omega\times(0,T)\,, \end{aligned}$$

we also need to impose initial and boundary conditions.

Initial conditions: $\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x)$ for all $x\in\Omega$.

Boundary condition:

1. **No-slip boundary condition**: $\boldsymbol{u} = \boldsymbol{0}$ on $\partial\Omega$.
2. **Navier boundary condition**: $\boldsymbol{u}\cdot\mathrm{N} = 0$ and $\mathrm{N}\times(\mu\mathrm{Def}\boldsymbol{u}\mathrm{N}) = \alpha(\mathrm{N}\times\boldsymbol{u})$ on $\partial\Omega$ for some constant $\alpha > 0$. This condition is based on the assumption that the traction force due to the viscous effect is proportional to the fluid velocity on the boundary.

§3.3 Models with Several Spatial Variables

To fully determine the dynamics of fluids, in addition to the Navier-Stokes equations

$$\begin{aligned} \boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \qquad & \text{in} \quad \Omega\times(0,T)\,, \\ \operatorname{div}\boldsymbol{u} = 0 \qquad & \text{in} \quad \Omega\times(0,T)\,, \end{aligned}$$

we also need to impose initial and boundary conditions.

Initial conditions: $\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x)$ for all $x\in\Omega$.

Boundary condition:

1. **No-slip boundary condition**: $\boldsymbol{u} = \mathbf{0}$ on $\partial\Omega$.
2. **Navier boundary condition**: $\boldsymbol{u}\cdot \mathrm{N} = 0$ and $\mathrm{N}\times(\mu\mathrm{Def}\boldsymbol{u}\mathrm{N}) = \alpha(\mathrm{N}\times\boldsymbol{u})$ on $\partial\Omega$ for some constant $\alpha > 0$. This condition is based on the assumption that the traction force due to the viscous effect is proportional to the fluid velocity on the boundary.

§3.3 Models with Several Spatial Variables

To fully determine the dynamics of fluids, in addition to the Navier-Stokes equations

$$\begin{aligned} \boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \qquad &\text{in} \quad \Omega\times(0,T)\,, \\ \mathrm{div}\boldsymbol{u} = 0 \qquad &\text{in} \quad \Omega\times(0,T)\,, \end{aligned}$$

we also need to impose initial and boundary conditions.

Initial conditions: $\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x)$ for all $x \in \Omega$.

Boundary condition:

1. **No-slip boundary condition**: $\boldsymbol{u} = \mathbf{0}$ on $\partial\Omega$.
2. **Navier boundary condition**: $\boldsymbol{u}\cdot\mathrm{N} = 0$ and $\mathrm{N}\times(\mu\mathrm{Def}\boldsymbol{u}\mathrm{N}) = \alpha(\mathrm{N}\times\boldsymbol{u})$ on $\partial\Omega$ for some constant $\alpha > 0$. This condition is based on the assumption that the traction force due to the viscous effect is proportional to the fluid velocity on the boundary.

§3.3 Models with Several Spatial Variables

To fully determine the dynamics of fluids, in addition to the Navier-Stokes equations

$$\begin{aligned} \boldsymbol{u}_t + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \mu\Delta\boldsymbol{u} + \boldsymbol{f} \qquad & \text{in} \quad \Omega\times(0,T)\,, \\ \operatorname{div}\boldsymbol{u} = 0 \qquad & \text{in} \quad \Omega\times(0,T)\,, \end{aligned}$$

we also need to impose initial and boundary conditions.

Initial conditions: $\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x)$ for all $x\in\Omega$.

Boundary condition:

1. **No-slip boundary condition**: $\boldsymbol{u} = \mathbf{0}$ on $\partial\Omega$.
2. **Navier boundary condition**: $\boldsymbol{u}\cdot\mathrm{N} = 0$ and $\mathrm{N}\times(\mu\mathrm{Def}\boldsymbol{u}\mathrm{N}) = \alpha(\mathrm{N}\times\boldsymbol{u})$ on $\partial\Omega$ for some constant $\alpha > 0$. This condition is based on the assumption that the traction force due to the viscous effect is proportional to the fluid velocity on the boundary.

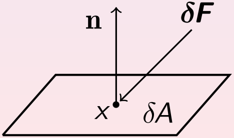
§3.3 Models with Several Spatial Variables

- **What is the stress/traction?**

Let Σ be a small piece of surface centered at x with area δA and $\mathbf{n}$ be a unit normal of Σ. The stress exerted by the fluid on the side toward which $\mathbf{n}$ points on the surface Σ ($\mathbf{n}$ 方向所指的這一側的流體對曲面 Σ 所施的應力) is defined as

$$\boldsymbol{\sigma}(x,\mathbf{n}) = \lim_{\delta A \to 0} \frac{\delta \boldsymbol{F}}{\delta A},$$

where $\delta\boldsymbol{F}$ is the force exerted on the surface by the fluid on that side (only one side is involved).



§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x,-\mathbf{n}) = -\boldsymbol{\sigma}(x,\mathbf{n})$.

Body force (that acts on every point of the body): $\triangle\ell\triangle w\triangle h\boldsymbol{f}$.

Surface force (due to the stress):

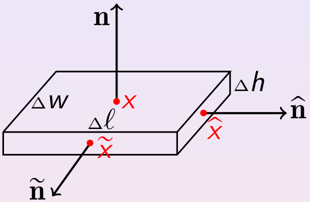
$$
\begin{aligned}
&\big[\boldsymbol{\sigma}(x,\mathbf{n}) + \boldsymbol{\sigma}(x-\triangle h\,\mathbf{n},-\mathbf{n})\big]\triangle\ell\triangle w\\
&+\big[\boldsymbol{\sigma}(\widehat{x},\widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x}-\triangle\ell\,\widehat{\mathbf{n}},-\widehat{\mathbf{n}})\big]\triangle w\triangle h\\
&+\big[\boldsymbol{\sigma}(\widetilde{x},\widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x}-\triangle w\,\widetilde{\mathbf{n}},-\widetilde{\mathbf{n}})\big]\triangle\ell\triangle h\,.
\end{aligned}
$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho\triangle\ell\triangle w\triangle h\boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.



Body force (that acts on every point of the body): $\triangle\ell\triangle w\triangle h \boldsymbol{f}$.

Surface force (due to the stress):

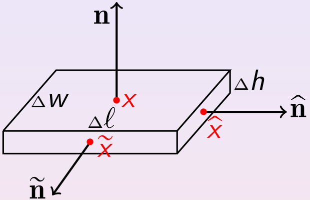
$$\begin{aligned}&\big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\,\mathbf{n}, -\mathbf{n})\big]\triangle\ell\triangle w\\ &+\big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\,\widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big]\triangle w\triangle h\\ &+\big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w\,\widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big]\triangle\ell\triangle h\,.\end{aligned}$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho\triangle\ell\triangle w\triangle h\,\boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.



Body force (that acts on every point of the body): $\triangle\ell\triangle w\triangle h\boldsymbol{f}$.

Surface force (due to the stress):

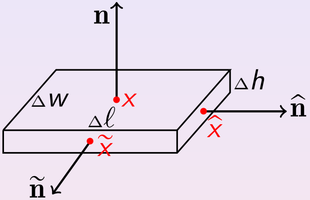
$$\begin{aligned}&\big[\boldsymbol{\sigma}(x,\mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\,\mathbf{n}, -\mathbf{n})\big]\triangle\ell\triangle w\\ &+\big[\boldsymbol{\sigma}(\widehat{x},\widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\,\widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big]\triangle w\triangle h\\ &+\big[\boldsymbol{\sigma}(\widetilde{x},\widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w\,\widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big]\triangle\ell\triangle h\,.\end{aligned}$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho\triangle\ell\triangle w\triangle h\boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.



Body force (that acts on every point of the body): $\triangle\ell\triangle w\triangle h\boldsymbol{f}$.

Surface force (due to the stress):

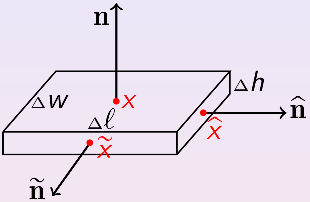
$$\begin{aligned}&\big[\boldsymbol{\sigma}(x,\mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\,\mathbf{n}, -\mathbf{n})\big]\triangle\ell\triangle w\\ &+\big[\boldsymbol{\sigma}(\widehat{x},\widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\,\widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big]\triangle w\triangle h\\ &+\big[\boldsymbol{\sigma}(\widetilde{x},\widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w\widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big]\triangle\ell\triangle h\,.\end{aligned}$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho\triangle\ell\triangle w\triangle h\boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.



Body force (that acts on every point of the body): $\triangle\ell\triangle w\triangle h\boldsymbol{f}$.

Surface force (due to the stress):

$$\begin{aligned}&\big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\,\mathbf{n}, -\mathbf{n})\big]\triangle\ell\triangle w\\ +&\big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\,\widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big]\triangle w\triangle h\\ +&\big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w\widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big]\triangle\ell\triangle h\,.\end{aligned}$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho\triangle\ell\triangle w\triangle h\boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:

  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.

Therefore,

$$\begin{aligned}\varrho \triangle\ell \triangle w \triangle h \boldsymbol{a} = {} & \triangle\ell \triangle w \triangle h \boldsymbol{f} \\ & + \big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\, \mathbf{n}, -\mathbf{n})\big] \triangle\ell \triangle w \\ & + \big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\, \widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big] \triangle w \triangle h \\ & + \big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w \widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big] \triangle\ell \triangle h\,.\end{aligned}$$

Body force (that acts on every point of the body): $\triangle\ell \triangle w \triangle h \boldsymbol{f}$.

Surface force (due to the stress):

$$\begin{aligned}& \big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\, \mathbf{n}, -\mathbf{n})\big] \triangle\ell \triangle w \\ & + \big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\, \widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big] \triangle w \triangle h \\ & + \big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w \widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big] \triangle\ell \triangle h\,.\end{aligned}$$

Balance of force: Let $\boldsymbol{a}$ denote the acceleration of the body. Then

$$\varrho \triangle\ell \triangle w \triangle h \boldsymbol{a} = \textbf{body force} + \textbf{surface force}\,.$$

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.

Therefore,

$$\begin{aligned} \varrho \triangle\ell \triangle w \triangle h\, \boldsymbol{a} = {} & \triangle\ell \triangle w \triangle h\, \boldsymbol{f} \\ & + \big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h\, \mathbf{n}, -\mathbf{n})\big] \triangle\ell \triangle w \\ & + \big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell\, \widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big] \triangle w \triangle h \\ & + \big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w\, \widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big] \triangle\ell \triangle h\,. \end{aligned}$$

Passing to the limit as $\triangle h \to 0$, we find that

$$\big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x, -\mathbf{n})\big] \triangle\ell \triangle w = 0$$

or to be more precise,

$$\int_{\square} \big[\boldsymbol{\sigma}(y, \mathbf{n}) + \boldsymbol{\sigma}(y, -\mathbf{n})\big]\, d\mathbb{A} = 0\,.$$

Dividing both sides by the area of $\square$ (that is, $\triangle\ell \triangle w$) and passing to the limit as $(\triangle\ell, \triangle w) \to (0, 0)$, we conclude the desired identity.

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:

  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.

Therefore,

$$\begin{aligned}\varrho\vartriangle\ell\vartriangle w\vartriangle h\,\boldsymbol{a} = {} & \vartriangle\ell\vartriangle w\vartriangle h\,\boldsymbol{f} \\ & + \big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \vartriangle h\,\mathbf{n}, -\mathbf{n})\big]\vartriangle\ell\vartriangle w \\ & + \big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \vartriangle\ell\,\widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big]\vartriangle w\vartriangle h \\ & + \big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \vartriangle w\,\widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big]\vartriangle\ell\vartriangle h\,.\end{aligned}$$

Passing to the limit as $\vartriangle h \to 0$, we find that

$$\big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x, -\mathbf{n})\big]\vartriangle\ell\vartriangle w = 0$$

or to be more precise,

$$\int_{\square}\big[\boldsymbol{\sigma}(y, \mathbf{n}) + \boldsymbol{\sigma}(y, -\mathbf{n})\big]\, d\mathbb{A} = 0\,.$$

Dividing both sides by the area of $\square$ (that is, $\vartriangle\ell\vartriangle w$) and passing to the limit as $(\vartriangle\ell, \vartriangle w) \to (0, 0)$, we conclude the desired identity.

§3.3 Models with Several Spatial Variables

- **General properties of the stress**:
  1. For a point x and a unit vector $\mathbf{n}$, $\boldsymbol{\sigma}(x, -\mathbf{n}) = -\boldsymbol{\sigma}(x, \mathbf{n})$.

Therefore,

$$\begin{aligned} \varrho \triangle\ell \triangle w \triangle h \boldsymbol{a} = {} & \triangle\ell \triangle w \triangle h \boldsymbol{f} \\ & + \big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x - \triangle h \mathbf{n}, -\mathbf{n})\big] \triangle\ell \triangle w \\ & + \big[\boldsymbol{\sigma}(\widehat{x}, \widehat{\mathbf{n}}) + \boldsymbol{\sigma}(\widehat{x} - \triangle\ell \widehat{\mathbf{n}}, -\widehat{\mathbf{n}})\big] \triangle w \triangle h \\ & + \big[\boldsymbol{\sigma}(\widetilde{x}, \widetilde{\mathbf{n}}) + \boldsymbol{\sigma}(\widetilde{x} - \triangle w \widetilde{\mathbf{n}}, -\widetilde{\mathbf{n}})\big] \triangle\ell \triangle h \,. \end{aligned}$$

Passing to the limit as $\triangle h \to 0$, we find that

$$\big[\boldsymbol{\sigma}(x, \mathbf{n}) + \boldsymbol{\sigma}(x, -\mathbf{n})\big] \triangle\ell \triangle w = 0$$

or to be more precise,

$$\int_{\square} \big[\boldsymbol{\sigma}(y, \mathbf{n}) + \boldsymbol{\sigma}(y, -\mathbf{n})\big] \, d\mathbb{A} = 0 \,.$$

Dividing both sides by the area of $\square$ (that is, $\triangle\ell \triangle w$) and passing to the limit as $(\triangle\ell, \triangle w) \to (0, 0)$, we conclude the desired identity.

§3.3 Models with Several Spatial Variables

2. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis of $\mathbb{R}^3$. For a point x and a unit vector $\mathbf{n} = (n_1, n_2, n_3)$,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \boldsymbol{\sigma}(x, \mathbf{e}_1)n_1 + \boldsymbol{\sigma}(x, \mathbf{e}_2)n_2 + \boldsymbol{\sigma}(x, \mathbf{e}_3)n_3 . \qquad (19)$$

In other words, the stress $\boldsymbol{\sigma}(x, \mathbf{n})$ can be expressed as a linear combination of $\boldsymbol{\sigma}(x, \mathbf{e}_1)$, $\boldsymbol{\sigma}(x, \mathbf{e}_2)$ and $\boldsymbol{\sigma}(x, \mathbf{e}_3)$.

Suppose that

$$\boldsymbol{\sigma}(x, \mathbf{e}_j) = \tau_{1j}\mathbf{e}_1 + \tau_{2j}\mathbf{e}_2 + \tau_{3j}\mathbf{e}_3 \qquad 1 \leqslant j \leqslant 3 ,$$

where $\tau_{ij} = \tau_{ij}(x)$. Then (19) implies that

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \Big(\sum_{i,j=1}^{3} \tau_{ij} n_j\Big)\mathbf{e}_i$$

or equivalently,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} .$$

§3.3 Models with Several Spatial Variables

2. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis of $\mathbb{R}^3$. For a point x and a unit vector $\mathbf{n} = (n_1, n_2, n_3)$,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \boldsymbol{\sigma}(x, \mathbf{e}_1)n_1 + \boldsymbol{\sigma}(x, \mathbf{e}_2)n_2 + \boldsymbol{\sigma}(x, \mathbf{e}_3)n_3 \,. \tag{19}$$

In other words, the stress $\boldsymbol{\sigma}(x, \mathbf{n})$ can be expressed as a linear combination of $\boldsymbol{\sigma}(x, \mathbf{e}_1)$, $\boldsymbol{\sigma}(x, \mathbf{e}_2)$ and $\boldsymbol{\sigma}(x, \mathbf{e}_3)$.

Suppose that

$$\boldsymbol{\sigma}(x, \mathbf{e}_j) = \tau_{1j}\mathbf{e}_1 + \tau_{2j}\mathbf{e}_2 + \tau_{3j}\mathbf{e}_3 \qquad 1 \leqslant j \leqslant 3\,,$$

where $\tau_{ij} = \tau_{ij}(x)$. Then (19) implies that

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \Big(\sum_{i,j=1}^{3} \tau_{ij} n_j\Big)\mathbf{e}_i$$

or equivalently,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} .$$

§3.3 Models with Several Spatial Variables

2. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis of $\mathbb{R}^3$. For a point x and a unit vector $\mathbf{n} = (n_1, n_2, n_3)$,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \boldsymbol{\sigma}(x, \mathbf{e}_1)n_1 + \boldsymbol{\sigma}(x, \mathbf{e}_2)n_2 + \boldsymbol{\sigma}(x, \mathbf{e}_3)n_3 \,. \qquad (19)$$

In other words, the stress $\boldsymbol{\sigma}(x, \mathbf{n})$ can be expressed as a linear combination of $\boldsymbol{\sigma}(x, \mathbf{e}_1)$, $\boldsymbol{\sigma}(x, \mathbf{e}_2)$ and $\boldsymbol{\sigma}(x, \mathbf{e}_3)$.

Suppose that

$$\boldsymbol{\sigma}(x, \mathbf{e}_j) = \tau_{1j}\mathbf{e}_1 + \tau_{2j}\mathbf{e}_2 + \tau_{3j}\mathbf{e}_3 \qquad 1 \leqslant j \leqslant 3\,,$$

where $\tau_{ij} = \tau_{ij}(x)$. Then (19) implies that

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \Big(\sum_{i,j=1}^{3} \tau_{ij} n_j\Big)\mathbf{e}_i$$

or equivalently,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} .$$

§3.3 Models with Several Spatial Variables

2. Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis of $\mathbb{R}^3$. For a point x and a unit vector $\mathbf{n} = (n_1, n_2, n_3)$,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \boldsymbol{\sigma}(x, \mathbf{e}_1)n_1 + \boldsymbol{\sigma}(x, \mathbf{e}_2)n_2 + \boldsymbol{\sigma}(x, \mathbf{e}_3)n_3 \,. \tag{19}$$

In other words, the stress $\boldsymbol{\sigma}(x, \mathbf{n})$ can be expressed as a linear combination of $\boldsymbol{\sigma}(x, \mathbf{e}_1)$, $\boldsymbol{\sigma}(x, \mathbf{e}_2)$ and $\boldsymbol{\sigma}(x, \mathbf{e}_3)$.

Suppose that

$$\boldsymbol{\sigma}(x, \mathbf{e}_j) = \tau_{1j}\mathbf{e}_1 + \tau_{2j}\mathbf{e}_2 + \tau_{3j}\mathbf{e}_3 \qquad 1 \leqslant j \leqslant 3 \,,$$

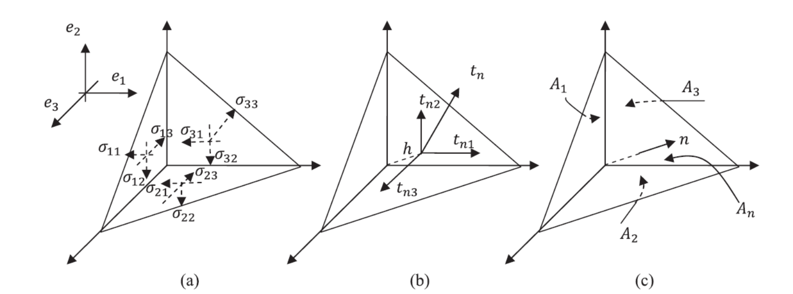
where $\tau_{ij} = \tau_{ij}(x)$. Then (19) implies that

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \Big(\sum_{i,j=1}^{3} \tau_{ij} n_j\Big)\mathbf{e}_i$$

or equivalently,

$$\boldsymbol{\sigma}(x, \mathbf{n}) = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} .$$

§3.3 Models with Several Spatial Variables

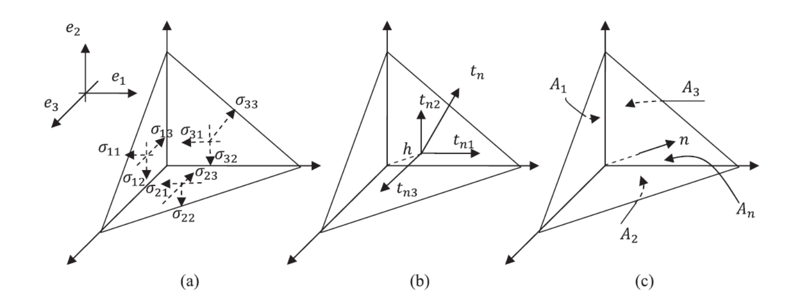


(a) On each side orthogonal to the coordinate axis, the stress is given by $\boldsymbol{\sigma}(-\mathbf{e}_j) = \sum_{k=1}^{3} \sigma_{jk}\mathbf{e}_k = -\sum_{j=1}^{3} \tau_{ij}\mathbf{e}_i$.

(b) On the "slant" side of the tetrahedron, the stress can be written as $\boldsymbol{\sigma}(\mathbf{n}) = t_n = t_{n1}\mathbf{e}_1 + t_{n2}\mathbf{e}_2 + t_{n3}\mathbf{e}_3$.

(c) By force balances, $\boldsymbol{\sigma}(\mathbf{n})A_n = \boldsymbol{\sigma}(\mathbf{e}_1)A_1 + \boldsymbol{\sigma}(\mathbf{e}_2)A_2 + \boldsymbol{\sigma}(\mathbf{e}_3)A_3$ which leads to (19).

§3.3 Models with Several Spatial Variables

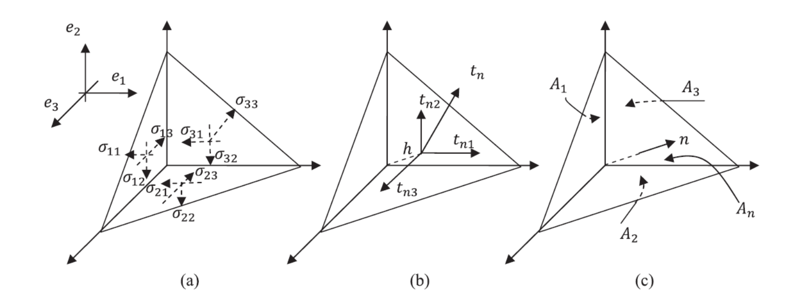


(a) On each side orthogonal to the coordinate axis, the stress is given by $\boldsymbol{\sigma}(-\mathbf{e}_j) = \sum_{k=1}^{3} \sigma_{jk}\mathbf{e}_k = -\sum_{j=1}^{3} \tau_{ij}\mathbf{e}_i$.

(b) On the "slant" side of the tetrahedron, the stress can be written as $\boldsymbol{\sigma}(\mathbf{n}) = t_n = t_{n1}\mathbf{e}_1 + t_{n2}\mathbf{e}_2 + t_{n3}\mathbf{e}_3$.

(c) By force balances, $\boldsymbol{\sigma}(\mathbf{n})A_n = \boldsymbol{\sigma}(\mathbf{e}_1)A_1 + \boldsymbol{\sigma}(\mathbf{e}_2)A_2 + \boldsymbol{\sigma}(\mathbf{e}_3)A_3$ which leads to (19).

§3.3 Models with Several Spatial Variables

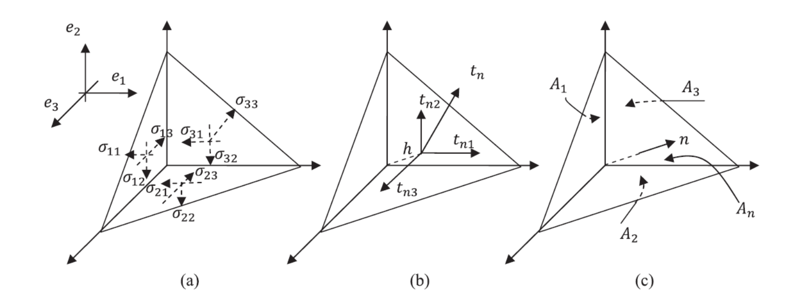


(a) On each side orthogonal to the coordinate axis, the stress is given by $\boldsymbol{\sigma}(-\mathbf{e}_j) = \sum_{k=1}^{3} \sigma_{jk}\mathbf{e}_k = -\sum_{j=1}^{3} \tau_{ij}\mathbf{e}_i$.

(b) On the "slant" side of the tetrahedron, the stress can be written as $\boldsymbol{\sigma}(\mathbf{n}) = t_n = t_{n1}\mathbf{e}_1 + t_{n2}\mathbf{e}_2 + t_{n3}\mathbf{e}_3$.

(c) By force balances, $\boldsymbol{\sigma}(\mathbf{n})A_n = \boldsymbol{\sigma}(\mathbf{e}_1)A_1 + \boldsymbol{\sigma}(\mathbf{e}_2)A_2 + \boldsymbol{\sigma}(\mathbf{e}_3)A_3$ which leads to (19).

§3.3 Models with Several Spatial Variables



(a) On each side orthogonal to the coordinate axis, the stress is given by $\boldsymbol{\sigma}(-\mathbf{e}_j) = \sum_{k=1}^{3} \sigma_{jk}\mathbf{e}_k = -\sum_{j=1}^{3} \tau_{ij}\mathbf{e}_i$.

(b) On the "slant" side of the tetrahedron, the stress can be written as $\boldsymbol{\sigma}(\mathbf{n}) = t_n = t_{n1}\mathbf{e}_1 + t_{n2}\mathbf{e}_2 + t_{n3}\mathbf{e}_3$.

(c) By force balances, $\boldsymbol{\sigma}(\mathbf{n})A_n = \boldsymbol{\sigma}(\mathbf{e}_1)A_1 + \boldsymbol{\sigma}(\mathbf{e}_2)A_2 + \boldsymbol{\sigma}(\mathbf{e}_3)A_3$ which leads to (19).

§3.3 Models with Several Spatial Variables

3. By the conservation of angular momentum, $\tau_{ij} = \tau_{ji}$ for all $1 \leqslant i, j \leqslant 3$. In other words, the matrix $\tau = [\tau_{ij}]$, called the stress tensor, is symmetric.

Reason: Relabel the point of interest as $\mathbf{0} = (0,0,0)$, and in the following we show that $\tau_{12}(\mathbf{0}) = \tau_{21}(\mathbf{0})$ (that $\tau_{13} = \tau_{31}$ and $\tau_{23} = \tau_{32}$ can be shown in a similar fashion). Consider the cube $[-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \times [-\frac{\Delta y}{2}, \frac{\Delta y}{2}] \times [-\frac{\Delta z}{2}, \frac{\Delta z}{2}]$ rotating about the z-axis.

§3.3 Models with Several Spatial Variables

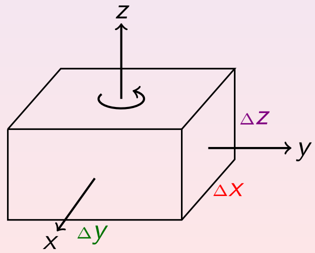
3. By the conservation of angular momentum, $\tau_{ij} = \tau_{ji}$ for all $1 \leqslant i, j \leqslant 3$. In other words, the matrix $\tau = [\tau_{ij}]$, called the stress tensor, is symmetric.

Reason: Relabel the point of interest as $\mathbf{0} = (0, 0, 0)$, and in the following we show that $\tau_{12}(\mathbf{0}) = \tau_{21}(\mathbf{0})$ (that $\tau_{13} = \tau_{31}$ and $\tau_{23} = \tau_{32}$ can be shown in a similar fashion). Consider the cube $[-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \times [-\frac{\Delta y}{2}, \frac{\Delta y}{2}] \times [-\frac{\Delta z}{2}, \frac{\Delta z}{2}]$ rotating about the z-axis.

§3.3 Models with Several Spatial Variables

3. By the conservation of angular momentum, $\tau_{ij} = \tau_{ji}$ for all $1 \leqslant i, j \leqslant 3$. In other words, the matrix $\tau = [\tau_{ij}]$, called the stress tensor, is symmetric.

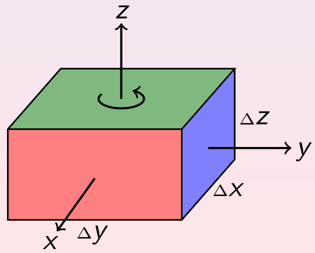
Reason: Relabel the point of interest as $\mathbf{0} = (0,0,0)$, and in the following we show that $\tau_{12}(\mathbf{0}) = \tau_{21}(\mathbf{0})$ (that $\tau_{13} = \tau_{31}$ and $\tau_{23} = \tau_{32}$ can be shown in a similar fashion). Consider the cube $[-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \times [-\frac{\Delta y}{2}, \frac{\Delta y}{2}] \times [-\frac{\Delta z}{2}, \frac{\Delta z}{2}]$ rotating about the z-axis.



§3.3 Models with Several Spatial Variables

3. By the conservation of angular momentum, $\tau_{ij} = \tau_{ji}$ for all $1 \leqslant i, j \leqslant 3$. In other words, the matrix $\tau = [\tau_{ij}]$, called the stress tensor, is symmetric.

Reason: Relabel the point of interest as $\mathbf{0} = (0,0,0)$, and in the following we show that $\tau_{12}(\mathbf{0}) = \tau_{21}(\mathbf{0})$ (that $\tau_{13} = \tau_{31}$ and $\tau_{23} = \tau_{32}$ can be shown in a similar fashion). Consider the cube $[-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \times [-\frac{\Delta y}{2}, \frac{\Delta y}{2}] \times [-\frac{\Delta z}{2}, \frac{\Delta z}{2}]$ rotating about the z-axis.



§3.3 Models with Several Spatial Variables

Torque about a point: Given a force $\mathbf{F}$ acting on a particle, the torque $\boldsymbol{\tau}$ on that particle about an fulcrum (支點) is defined as the cross product

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F},$$

where $\mathbf{r}$ is the particle's position vector relative to the fulcrum.

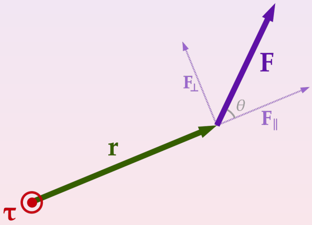


Figure 9: Torque in high school is given by $\mathrm{F}_\perp \mathrm{r}$ which is $\mathrm{F}\mathrm{r}\sin\theta$, where $\mathrm{F} = \|\mathbf{F}\|$ and $\mathrm{r} = \|\mathbf{r}\|$. Note that $\|\mathbf{r} \times \mathbf{F}\| = \|\mathbf{r}\|\|\mathbf{F}\|\sin\theta$.

§3.3 Models with Several Spatial Variables

Torque about an axis: Given a force $\mathbf{F}$ acting on a particle, the torque $\boldsymbol{\tau}$ on that particle about an axis is the **projection** of the cross product

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$$

onto the direction of the axis, where $\mathbf{r}$ is the particle's position vector relative to any point on the axis.

The net torque on a body determines the **rate of change** of the body's angular momentum $\boldsymbol{L} = \mathbf{r} \times \boldsymbol{p}$, where $\boldsymbol{p}$ is the linear momentum. Note that with m and $\boldsymbol{v}$ denoting the mass and the velocity of the point, respectively, we have $\boldsymbol{v} = \dfrac{d\mathbf{r}}{dt}$ and $\boldsymbol{p} = m\boldsymbol{v}$ so that

$$\frac{d\boldsymbol{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times m\boldsymbol{v}) = \boldsymbol{v} \times m\boldsymbol{v} + \mathbf{r} \times m\frac{d\boldsymbol{v}}{dt} = m\mathbf{r} \times \boldsymbol{a}\,.$$

§3.3 Models with Several Spatial Variables

Torque about an axis: Given a force $\mathbf{F}$ acting on a particle, the torque $\boldsymbol{\tau}$ on that particle about an axis is the **projection** of the cross product

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$$

onto the direction of the axis, where $\mathbf{r}$ is the particle's position vector relative to any point on the axis.

The net torque on a body determines the **rate of change** of the body's angular momentum $\boldsymbol{L} = \mathbf{r} \times \boldsymbol{p}$, where $\boldsymbol{p}$ is the linear momentum. Note that with m and $\boldsymbol{v}$ denoting the mass and the velocity of the point, respectively, we have $\boldsymbol{v} = \frac{d\mathbf{r}}{dt}$ and $\boldsymbol{p} = m\boldsymbol{v}$ so that

$$\frac{d\boldsymbol{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times m\boldsymbol{v}) = \boldsymbol{v} \times m\boldsymbol{v} + \mathbf{r} \times m\frac{d\boldsymbol{v}}{dt} = m\mathbf{r} \times \boldsymbol{a} .$$

§3.3 Models with Several Spatial Variables

Torque about an axis: Given a force $\mathbf{F}$ acting on a particle, the torque $\boldsymbol{\tau}$ on that particle about an axis is the **projection** of the cross product

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$$

onto the direction of the axis, where $\mathbf{r}$ is the particle's position vector relative to any point on the axis.

The net torque on a body determines the **rate of change** of the body's angular momentum $\boldsymbol{L} = \mathbf{r} \times \boldsymbol{p}$, where $\boldsymbol{p}$ is the linear momentum. Note that with m and $\boldsymbol{v}$ denoting the mass and the velocity of the point, respectively, we have $\boldsymbol{v} = \dfrac{d\boldsymbol{r}}{dt}$ and $\boldsymbol{p} = m\boldsymbol{v}$ so that

$$\frac{d\boldsymbol{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times m\boldsymbol{v}) = \boldsymbol{v} \times m\boldsymbol{v} + \mathbf{r} \times m\frac{d\boldsymbol{v}}{dt} = m\mathbf{r} \times \boldsymbol{a}\,.$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on the six faces of the cube is **the third component of**

$$\begin{aligned}
&\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(x,\frac{\Delta y}{2},0\Big)\times\sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A} \\
&\quad+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(x,-\frac{\Delta y}{2},0\Big)\times\sigma\Big(\big(x,-\frac{\Delta y}{2},z\big),-\mathbf{e}_2\Big)\,d\mathbb{A} \\
&\quad+\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(\frac{\Delta x}{2},y,0\Big)\times\sigma\Big(\big(\frac{\Delta x}{2},y,z\big),\mathbf{e}_1\Big)\,d\mathbb{A} \\
&\quad+\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(-\frac{\Delta x}{2},y,0\Big)\times\sigma\Big(\big(-\frac{\Delta x}{2},y,z\big),-\mathbf{e}_1\Big)\,d\mathbb{A} \\
&\quad+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} (x,y,0)\times\sigma\Big(\big(x,y,\frac{\Delta z}{2}\big),\mathbf{e}_3\Big)\,d\mathbb{A} \\
&\quad+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} (x,y,0)\times\sigma\Big(\big(x,y,-\frac{\Delta z}{2}\big),-\mathbf{e}_3\Big)\,d\mathbb{A}\,.
\end{aligned}$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on faces intersecting the y-axis is **the third component of**

$$\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(x,\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(x,-\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),-\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$=\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,0,0)\times\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,0,0)\times\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),-\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(0,\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \Big(0,-\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),-\mathbf{e}_2\Big)\,d\mathbb{A}\,.$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on faces intersecting the y-axis is **the third component of**

$$\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(x,\frac{\Delta y}{2},0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(x,-\frac{\Delta y}{2},0)\times\sigma\big((x,-\frac{\Delta y}{2},z),-\mathbf{e}_2\big)\,d\mathbb{A}$$
$$=\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(x,0,0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(x,0,0)\times\sigma\big((x,-\frac{\Delta y}{2},z),-\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(0,\frac{\Delta y}{2},0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}(0,-\frac{\Delta y}{2},0)\times\sigma\big((x,-\frac{\Delta y}{2},z),-\mathbf{e}_2\big)\,d\mathbb{A}\,.$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on faces intersecting the y-axis is **the third component of**

$$\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,\frac{\Delta y}{2},0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,-\frac{\Delta y}{2},0)\times\sigma\big((x,-\frac{\Delta y}{2},z),-\mathbf{e}_2\big)\,d\mathbb{A}$$
$$=\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,0,0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$-\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,0,0)\times\sigma\big((x,-\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (0,\frac{\Delta y}{2},0)\times\sigma\big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (0,\frac{\Delta y}{2},0)\times\sigma\big((x,-\frac{\Delta y}{2},z),\mathbf{e}_2\big)\,d\mathbb{A}\,.$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on faces intersecting the y-axis is **the third component of**

$$\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\Big(x,\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\Big(x,-\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),-\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$=\mathbf{e}_1\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}x\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$-\mathbf{e}_1\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}x\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\frac{\Delta y}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+\frac{\Delta y}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}\,.$$

§3.3 Models with Several Spatial Variables

The torque about the z-axis due to the stress on faces intersecting the y-axis is **the third component of**

$$\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\Big(x,\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$

$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\Big(x,-\frac{\Delta y}{2},0\Big)\times\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),-\mathbf{e}_2\Big)\,d\mathbb{A}$$

$$=\mathbf{e}_1\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}x\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$

$$-\mathbf{e}_1\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}x\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$

$$+\frac{\Delta y}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\sigma\Big(\Big(x,\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}$$

$$+\frac{\Delta y}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]}\sigma\Big(\Big(x,-\frac{\Delta y}{2},z\Big),\mathbf{e}_2\Big)\,d\mathbb{A}\,.$$

§3.3 Models with Several Spatial Variables

Dividing both sides by the volume of the cube and passing to the limit as $(\Delta x, \Delta y, \Delta z) \to (0,0,0)$, by the fact that

$$\lim_{\Delta y \to 0} \frac{1}{\Delta y}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$- \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big(\big(x,-\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\Big]$$
$$= \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\frac{\partial \sigma}{\partial y}\Big(\big(x,0,z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\,,$$

and

$$\lim_{(\Delta x,\Delta z)\to(0,0)} \frac{1}{\Delta x \Delta z}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}$$
$$+ \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\Big] = 2\sigma(\mathbf{0},\mathbf{e}_2)\,,$$

§3.3 Models with Several Spatial Variables

Dividing both sides by the volume of the cube and passing to the limit as $(\Delta x, \Delta y, \Delta z) \to (0,0,0)$, by the fact that

$$\begin{aligned}\lim_{\Delta y \to 0} \frac{1}{\Delta y}\Big[&\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A} \\ &- \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big(\big(x,-\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\Big] \\ &= \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\frac{\partial\sigma}{\partial y}\Big(\big(x,0,z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\,,\end{aligned}$$

and

$$\begin{aligned}\lim_{(\Delta x,\Delta z)\to(0,0)} \frac{1}{\Delta x\Delta z}\Big[&\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A} \\ &+ \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(x,\frac{\Delta y}{2},z\big),\mathbf{e}_2\Big)\,d\mathbb{A}\Big] = 2\sigma(\mathbf{0},\mathbf{e}_2)\,,\end{aligned}$$

§3.3 Models with Several Spatial Variables

Dividing both sides by the volume of the cube and passing to the limit as $(\Delta x, \Delta y, \Delta z) \to (0,0,0)$, by the fact that

$$\lim_{\Delta y \to 0} \frac{1}{\Delta y}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A} - \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\sigma\Big((x,-\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A}\Big]$$
$$= \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\frac{\partial\sigma}{\partial y}\Big((x,0,z),\mathbf{e}_2\Big)\,d\mathbb{A}\,,$$

and

$$\lim_{(\Delta x,\Delta z)\to(0,0)} \frac{1}{\Delta x\Delta z}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A} + \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A}\Big] = 2\sigma(\mathbf{0},\mathbf{e}_2)\,,$$

§3.3 Models with Several Spatial Variables

we find that

$$
\begin{aligned}
&\lim_{(\Delta x,\Delta y,\Delta z)\to(0,0,0)} \frac{1}{\Delta x\Delta y\Delta z} \\
&\quad\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,\frac{\Delta y}{2},0)\times\sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A} \\
&\qquad+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,-\frac{\Delta y}{2},0)\times\sigma\Big((x,-\frac{\Delta y}{2},z),-\mathbf{e}_2\Big)\,d\mathbb{A}\Big] \\
&=\lim_{(\Delta x,\Delta z)\to(0,0)} \frac{1}{\Delta x\Delta z}\Big[\mathbf{e}_1\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} x\frac{\partial\sigma}{\partial y}\Big((x,0,z),\mathbf{e}_2\Big)\,d\mathbb{A} \\
&\qquad+\frac{1}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A} \\
&\qquad+\frac{1}{2}\mathbf{e}_2\times\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big((x,\frac{\Delta y}{2},z),\mathbf{e}_2\Big)\,d\mathbb{A}\Big] \\
&=\mathbf{e}_2\times\sigma(\mathbf{0},\mathbf{e}_2)\,.
\end{aligned}
$$

§3.3 Models with Several Spatial Variables

Similarly,

$$\lim_{(\Delta x,\Delta y,\Delta z)\to(0,0,0)} \frac{1}{\Delta x\Delta y\Delta z}\Big[\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(\frac{\Delta x}{2},y,z\big),\mathbf{e}_1\Big)\times\big(\frac{\Delta x}{2},y,0\big)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(-\frac{\Delta x}{2},y,z\big),-\mathbf{e}_1\Big)\times\big(-\frac{\Delta x}{2},y,0\big)\,d\mathbb{A}\Big]$$
$$=\mathbf{e}_1\times\sigma\big(\mathbf{0},\mathbf{e}_1\big)\,.$$

Moreover,

$$\frac{1}{\Delta x\Delta y\Delta z}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} \sigma\Big(\big(x,y,\frac{\Delta z}{2}\big),\mathbf{e}_3\Big)\times(x,y,0)\,d\mathbb{A}$$
$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} \sigma\Big(\big(x,y,-\frac{\Delta z}{2}\big),-\mathbf{e}_3\Big)\times(x,y,0)\,d\mathbb{A}\Big]$$
$$\to 0 \quad \text{as} \quad (\Delta x,\Delta y,\Delta z)\to(0,0,0)\,.$$

§3.3 Models with Several Spatial Variables

Similarly,

$$\lim_{(\Delta x,\Delta y,\Delta z)\to(0,0,0)} \frac{1}{\Delta x\Delta y\Delta z}$$

$$\Big[\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(\frac{\Delta x}{2},y,z\big),\mathbf{e}_1\Big)\times\big(\frac{\Delta x}{2},y,0\big)\,d\mathbb{A}$$

$$+\int_{[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} \sigma\Big(\big(-\frac{\Delta x}{2},y,z\big),-\mathbf{e}_1\Big)\times\big(-\frac{\Delta x}{2},y,0\big)\,d\mathbb{A}\Big]$$

$$= \mathbf{e}_1\times\sigma\big(\mathbf{0},\mathbf{e}_1\big)\,.$$

Moreover,

$$\frac{1}{\Delta x\Delta y\Delta z}\Big[\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} \sigma\Big(\big(x,y,\frac{\Delta z}{2}\big),\mathbf{e}_3\Big)\times(x,y,0)\,d\mathbb{A}$$

$$+\int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]} \sigma\Big(\big(x,y,-\frac{\Delta z}{2}\big),-\mathbf{e}_3\Big)\times(x,y,0)\,d\mathbb{A}\Big]$$

$$\to 0 \qquad \text{as} \quad (\Delta x,\Delta y,\Delta z)\to(0,0,0)\,.$$

§3.3 Models with Several Spatial Variables

Therefore, by the fact that

$$\sigma(\boldsymbol{x},\mathbf{n}) = \begin{bmatrix} \tau_{11}(\boldsymbol{x}) & \tau_{12}(\boldsymbol{x}) & \tau_{13}(\boldsymbol{x}) \\ \tau_{21}(\boldsymbol{x}) & \tau_{22}(\boldsymbol{x}) & \tau_{23}(\boldsymbol{x}) \\ \tau_{31}(\boldsymbol{x}) & \tau_{32}(\boldsymbol{x}) & \tau_{33}(\boldsymbol{x}) \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}, \quad \mathbf{n} = (n_1, n_2, n_3)^{\mathrm{T}},$$

we find that

$$\begin{aligned}
&\lim_{(\Delta x,\Delta y,\Delta z)\to(0,0,0)} \frac{\text{the torque about the } z\text{-axis due to the stress}}{\Delta x \Delta y \Delta z} \\
&= \big[\mathbf{e}_2 \times \sigma(\mathbf{0},\mathbf{e}_2)\big] \cdot \mathbf{e}_3 + \big[\mathbf{e}_1 \times \sigma(\mathbf{0},\mathbf{e}_1)\big] \cdot \mathbf{e}_3 \\
&= \big[(0,1,0) \times \big(\tau_{12}(\mathbf{0}), \tau_{22}(\mathbf{0}), \tau_{32}(\mathbf{0})\big)\big] \cdot (0,0,1) \\
&\quad + \big[(1,0,0) \times \big(\tau_{11}(\mathbf{0}), \tau_{21}(\mathbf{0}), \tau_{31}(\mathbf{0})\big)\big] \cdot (0,0,1) \\
&= \tau_{21}(\mathbf{0}) - \tau_{12}(\mathbf{0}).
\end{aligned}$$

§3.3 Models with Several Spatial Variables

Therefore, by the fact that

$$\sigma(\boldsymbol{x},\mathbf{n}) = \begin{bmatrix} \tau_{11}(\boldsymbol{x}) & \tau_{12}(\boldsymbol{x}) & \tau_{13}(\boldsymbol{x}) \\ \tau_{21}(\boldsymbol{x}) & \tau_{22}(\boldsymbol{x}) & \tau_{23}(\boldsymbol{x}) \\ \tau_{31}(\boldsymbol{x}) & \tau_{32}(\boldsymbol{x}) & \tau_{33}(\boldsymbol{x}) \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}, \quad \mathbf{n} = (n_1, n_2, n_3)^{\mathrm{T}},$$

we find that

$$\begin{aligned} &\lim_{(\Delta x,\Delta y,\Delta z)\to(0,0,0)} \frac{\text{the torque about the } z\text{-axis due to the stress}}{\Delta x \Delta y \Delta z} \\ &= \big[\mathbf{e}_2 \times \sigma(\mathbf{0},\mathbf{e}_2)\big] \cdot \mathbf{e}_3 + \big[\mathbf{e}_1 \times \sigma(\mathbf{0},\mathbf{e}_1)\big] \cdot \mathbf{e}_3 \\ &= \big[(0,1,0) \times \big(\tau_{12}(\mathbf{0}), \tau_{22}(\mathbf{0}), \tau_{32}(\mathbf{0})\big)\big] \cdot (0,0,1) \\ &\quad + \big[(1,0,0) \times \big(\tau_{11}(\mathbf{0}), \tau_{21}(\mathbf{0}), \tau_{31}(\mathbf{0})\big)\big] \cdot (0,0,1) \\ &= \tau_{21}(\mathbf{0}) - \tau_{12}(\mathbf{0}). \end{aligned}$$

§3.3 Models with Several Spatial Variables

Now, the torque about the z-axis due to the body force is

$$\mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,y,0)\times \boldsymbol{f}(x,y,z)\, d\mathbb{V}\,,$$

and the total torque contributes to the rate of change of the third component of the angular momentum so that

$$\begin{aligned}\mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} &(x,y,0)\times \rho(x,y,z)\boldsymbol{a}(x,y,z)\, d\mathbb{V} \\ = \mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} &(x,y,0)\times \boldsymbol{f}(x,y,z)\, d\mathbb{V} \\ &+ \text{the torque about the } z\text{-axis due to the stress}\,.\end{aligned}$$

Dividing both sides by the volume of the cube ans passing to the limit as $(\Delta x, \Delta y, \Delta z) \to (0,0,0)$, the limit involving volume integrals are zero; thus we conclude that

$$\tau_{21}(\mathbf{0}) - \tau_{12}(\mathbf{0}) = 0\,.$$

§3.3 Models with Several Spatial Variables

Now, the torque about the z-axis due to the body force is

$$\mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,y,0)\times \boldsymbol{f}(x,y,z)\, d\mathbb{V},$$

and the total torque contributes to the rate of change of the third component of the angular momentum so that

$$\begin{aligned}
&\mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,y,0)\times \rho(x,y,z)\boldsymbol{a}(x,y,z)\, d\mathbb{V} \\
&= \mathbf{e}_3 \cdot \int_{[-\frac{\Delta x}{2},\frac{\Delta x}{2}]\times[-\frac{\Delta y}{2},\frac{\Delta y}{2}]\times[-\frac{\Delta z}{2},\frac{\Delta z}{2}]} (x,y,0)\times \boldsymbol{f}(x,y,z)\, d\mathbb{V} \\
&\quad + \text{the torque about the } z\text{-axis due to the stress}.
\end{aligned}$$

Dividing both sides by the volume of the cube ans passing to the limit as $(\Delta x,\Delta y,\Delta z)\to(0,0,0)$, the limit involving volume integrals are zero; thus we conclude that

$$\tau_{21}(\mathbf{0}) - \tau_{12}(\mathbf{0}) = 0.$$

§3.3 Models with Several Spatial Variables

4. What is the form of $\sigma(\boldsymbol{x}, \mathbf{n})$?

• First note that the force due to the pressure is always perpendicular to the surface under consideration; thus

$$\sigma(\boldsymbol{x}, \mathbf{n}) = -p(\boldsymbol{x})\,\mathbf{n} + \Sigma(\boldsymbol{x})\mathbf{n} = (\Sigma - p\mathbf{I})(\boldsymbol{x})\mathbf{n}$$

for some symmetric matrix $\Sigma = [\bar{\sigma}_{ij}]$.

• The presence of Σ is due to the internal friction of fliuds and is called the ***viscous stress tensor***.

• The friction of fluids occurs only when different fluid particles move with different velocities. Therefore, Σ must depend on $\nabla \boldsymbol{u} = \left[\frac{\partial u^i}{\partial x_j}\right]$, where $\boldsymbol{u} = (u^1, u^2, u^3)$.

• If the velocity gradient is small, we can assume that Σ is linear in $\nabla \boldsymbol{u}$. Therefore,

$$\bar{\sigma}_{ij} = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell}\,.$$

§3.3 Models with Several Spatial Variables

4. What is the form of $\sigma(\boldsymbol{x}, \mathbf{n})$?

   • First note that the force due to the pressure is always perpendicular to the surface under consideration; thus

$$\sigma(\boldsymbol{x}, \mathbf{n}) = -p(\boldsymbol{x})\mathbf{n} + \Sigma(\boldsymbol{x})\mathbf{n} = (\Sigma - p\mathbf{I})(\boldsymbol{x})\mathbf{n}$$

   for some symmetric matrix $\Sigma = [\bar{\sigma}_{ij}]$.

   • The presence of Σ is due to the internal friction of fliuds and is called the ***viscous stress tensor***.

   • The friction of fluids occurs only when different fluid particles move with different velocities. Therefore, Σ must depend on $\nabla \boldsymbol{u} = \left[\frac{\partial u^i}{\partial x_j}\right]$, where $\boldsymbol{u} = (u^1, u^2, u^3)$.

   • If the velocity gradient is small, we can assume that Σ is linear in $\nabla \boldsymbol{u}$. Therefore,

$$\bar{\sigma}_{ij} = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell}.$$

§3.3 Models with Several Spatial Variables

4. What is the form of $\sigma(\boldsymbol{x}, \mathbf{n})$?

• First note that the force due to the pressure is always perpendicular to the surface under consideration; thus

$$\sigma(\boldsymbol{x}, \mathbf{n}) = -p(\boldsymbol{x})\mathbf{n} + \Sigma(\boldsymbol{x})\mathbf{n} = (\Sigma - p\mathbf{I})(\boldsymbol{x})\mathbf{n}$$

for some symmetric matrix $\Sigma = [\bar{\sigma}_{ij}]$.

• The presence of Σ is due to the internal friction of fliuds and is called the ***viscous stress tensor***.

• The friction of fluids occurs only when different fluid particles move with different velocities. Therefore, Σ must depend on $\nabla \boldsymbol{u} = \left[\frac{\partial u^i}{\partial x_j}\right]$, where $\boldsymbol{u} = (u^1, u^2, u^3)$.

• If the velocity gradient is small, we can assume that Σ is linear in $\nabla \boldsymbol{u}$. Therefore,

$$\bar{\sigma}_{ij} = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell} .$$

§3.3 Models with Several Spatial Variables

4. What is the form of $\sigma(\boldsymbol{x}, \mathbf{n})$?

• First note that the force due to the pressure is always perpendicular to the surface under consideration; thus

$$\sigma(\boldsymbol{x}, \mathbf{n}) = -p(\boldsymbol{x})\mathbf{n} + \Sigma(\boldsymbol{x})\mathbf{n} = (\Sigma - p\mathbf{I})(\boldsymbol{x})\mathbf{n}$$

for some symmetric matrix $\Sigma = [\bar{\sigma}_{ij}]$.

• The presence of Σ is due to the internal friction of fliuds and is called the ***viscous stress tensor***.

• The friction of fluids occurs only when different fluid particles move with different velocities. Therefore, Σ must depend on $\nabla \boldsymbol{u} = \left[\frac{\partial u^i}{\partial x_j}\right]$, where $\boldsymbol{u} = (u^1, u^2, u^3)$.

• If the velocity gradient is small, we can assume that Σ is linear in $\nabla \boldsymbol{u}$. Therefore,

$$\bar{\sigma}_{ij} = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell}.$$

§3.3 Models with Several Spatial Variables

Since $\bar{\sigma}_{ij} = \bar{\sigma}_{ji}$,

$$\bar{\sigma}_{ij} = \frac{1}{2}\left(\bar{\sigma}_{ij} + \bar{\sigma}_{ji}\right) = \sum_{k,\ell=1}^{3} \frac{a^{ijk\ell} + a^{jik\ell}}{2} \frac{\partial u^k}{\partial x_\ell}.$$

Therefore, W.L.O.G. we can assume that $a^{ijk\ell} = a^{jik\ell}$.

- Write

$$\bar{\sigma}_{ij} = \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\left(\frac{\partial u^k}{\partial x_\ell} + \frac{\partial u^\ell}{\partial x_k}\right) + \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\left(\frac{\partial u^k}{\partial x_\ell} - \frac{\partial u^\ell}{\partial x_k}\right).$$

Since we do not expect any viscous effect (internal friction) to be present if the fluid is in a state of pure rotation, we find that $\bar{\sigma}$ is independent of $\frac{\partial u^k}{\partial x_\ell} - \frac{\partial u^\ell}{\partial x_k}$; thus

$$\bar{\sigma}_{ij} = \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\left(\frac{\partial u^k}{\partial x_\ell} + \frac{\partial u^\ell}{\partial x_k}\right) = \sum_{k,\ell=1}^{3} \frac{a^{ijk\ell} + a^{ij\ell k}}{2} \frac{\partial u^k}{\partial x_\ell}.$$

Therefore, W.L.O.G. we can also assume that $a^{ijk\ell} = a^{ij\ell k}$.

§3.3 Models with Several Spatial Variables

Since $\bar{\sigma}_{ij} = \bar{\sigma}_{ji}$,

$$\bar{\sigma}_{ij} = \frac{1}{2}\big(\bar{\sigma}_{ij} + \bar{\sigma}_{ji}\big) = \sum_{k,\ell=1}^{3} \frac{a^{ijk\ell} + a^{jik\ell}}{2} \frac{\partial u^k}{\partial x_\ell}.$$

Therefore, W.L.O.G. we can assume that $a^{ijk\ell} = a^{jik\ell}$.

- Write

$$\bar{\sigma}_{ij} = \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\Big(\frac{\partial u^k}{\partial x_\ell} + \frac{\partial u^\ell}{\partial x_k}\Big) + \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\Big(\frac{\partial u^k}{\partial x_\ell} - \frac{\partial u^\ell}{\partial x_k}\Big).$$

Since we do not expect any viscous effect (internal friction) to be present if the fluid is in a state of pure rotation, we find that $\bar{\sigma}$ is independent of $\dfrac{\partial u^k}{\partial x_\ell} - \dfrac{\partial u^\ell}{\partial x_k}$; thus

$$\bar{\sigma}_{ij} = \frac{1}{2}\sum_{k,\ell=1}^{3} a^{ijk\ell}\Big(\frac{\partial u^k}{\partial x_\ell} + \frac{\partial u^\ell}{\partial x_k}\Big) = \sum_{k,\ell=1}^{3} \frac{a^{ijk\ell} + a^{ij\ell k}}{2} \frac{\partial u^k}{\partial x_\ell}.$$

Therefore, W.L.O.G. we can also assume that $a^{ijk\ell} = a^{ij\ell k}$.

§3.3 Models with Several Spatial Variables

- Consider the case of laminar flow (層流) that $\mathbf{n} = \mathbf{e}_s$ and $\boldsymbol{u} = u(x_s)\mathbf{e}_r$ for some $s \neq r$. Then

$$\big[\boldsymbol{\sigma}(\boldsymbol{x}, \mathbf{n})\big]^i = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell}(\boldsymbol{x})\delta_{sj} = a^{isrs}\frac{\partial u}{\partial x_s}(\boldsymbol{x}) .$$

Since in this case the drag force due to the friction is in direction $\mathbf{e}_r$, we find that $a^{isrs} = 0$ if $i \neq r$ and $r \neq s$.

On the other hand, if $i = r$ (and $r \neq s$), we let $a^{rsrs} = \mu$ for all $r \neq s$ so that

$$\boldsymbol{\sigma}(\boldsymbol{x}, \mathbf{e}_s) = \mu \frac{\partial u}{\partial x_s}(\boldsymbol{x})\mathbf{e}_r .$$

- Therefore, the only possible non-zero $a^{ijk\ell}$ terms are:

$$a^{iiii} \text{ for all } 1 \leqslant i \leqslant 3 ,$$
$$a^{ikkk}, a^{kikk}, a^{iikk}, a^{ikik}, a^{ikki} \text{ with } i \neq k ,$$
$$a^{iik\ell}, a^{k\ell ii} \text{ with distinct } i, k, \ell .$$

§3.3 Models with Several Spatial Variables

- Consider the case of laminar flow (層流) that $\mathbf{n} = \mathbf{e}_s$ and $\boldsymbol{u} = u(x_s)\mathbf{e}_r$ for some $s \neq r$. Then

$$\big[\boldsymbol{\sigma}(\boldsymbol{x},\mathbf{n})\big]^i = \sum_{k,\ell=1}^{3} a^{ijk\ell}\frac{\partial u^k}{\partial x_\ell}(\boldsymbol{x})\delta_{sj} = a^{isrs}\frac{\partial u}{\partial x_s}(\boldsymbol{x})\,.$$

Since in this case the drag force due to the friction is in direction $\mathbf{e}_r$, we find that $a^{isrs} = 0$ if $i \neq r$ and $r \neq s$.

On the other hand, if $i = r$ (and $r \neq s$), we let $a^{rsrs} = \mu$ for all $r \neq s$ so that

$$\boldsymbol{\sigma}(\boldsymbol{x},\mathbf{e}_s) = \mu\frac{\partial u}{\partial x_s}(\boldsymbol{x})\mathbf{e}_r\,.$$

- Therefore, the only possible non-zero $a^{ijk\ell}$ terms are:

$$a^{iiii} \text{ for all } 1 \leqslant i \leqslant 3\,,$$
$$a^{ikkk}, a^{kikk}, a^{iikk}, a^{ikik}, a^{ikki} \text{ with } i \neq k\,,$$
$$a^{iik\ell}, a^{k\ell ii} \text{ with distinct } i,k,\ell\,.$$

§3.3 Models with Several Spatial Variables

- Consider the case of laminar flow (層流) that $\mathbf{n} = \mathbf{e}_s$ and $\boldsymbol{u} = u(x_s)\mathbf{e}_r$ for some $s \neq r$. Then

$$\big[\boldsymbol{\sigma}(\boldsymbol{x}, \mathbf{n})\big]^i = \sum_{k,\ell=1}^{3} a^{ijk\ell} \frac{\partial u^k}{\partial x_\ell}(\boldsymbol{x})\delta_{sj} = a^{isrs} \frac{\partial u}{\partial x_s}(\boldsymbol{x})\,.$$

Since in this case the drag force due to the friction is in direction $\mathbf{e}_r$, we find that $a^{isrs} = 0$ if $i \neq r$ and $r \neq s$.

On the other hand, if $i = r$ (and $r \neq s$), we let $a^{rsrs} = \mu$ for all $r \neq s$ so that

$$\boldsymbol{\sigma}(\boldsymbol{x}, \mathbf{e}_s) = \mu \frac{\partial u}{\partial x_s}(\boldsymbol{x})\mathbf{e}_r\,.$$

- Therefore, the only possible non-zero $a^{ijk\ell}$ terms are:

$$a^{iiii} \text{ for all } 1 \leqslant i \leqslant 3\,,$$
$$a^{ikkk}\,, a^{kikk}\,, a^{iikk}\,, a^{ikik}\,, a^{ikki} \text{ with } i \neq k\,,$$
$$a^{iik\ell}\,, a^{k\ell ii} \text{ with distinct } i, k, \ell\,.$$

§3.3 Models with Several Spatial Variables

Due to the isotropy (that is, properties are the same in all directions) and the symmetry of $a^{ijk\ell}$, we have

ⓐ $a^{1111} = a^{2222} = a^{3333} = A$.

ⓑ $a^{1222} = a^{2122} = a^{1333} = a^{3133} = a^{2111} = a^{1211} = a^{2333} = a^{3233} = a^{3111} = a^{1311} = a^{3222} = a^{2322} = B$.

ⓒ $a^{1122} = a^{2211} = a^{1133} = a^{3311} = a^{2233} = a^{3322} = C$.

ⓓ $a^{1212} = a^{2112} = a^{2121} = a^{1221} = a^{1313} = a^{3113} = a^{3131} = a^{1331} = a^{2323} = a^{3223} = a^{3232} = a^{3223} = \mu$.

ⓔ $a^{1123} = a^{1132} = a^{2213} = a^{2231} = a^{3312} = a^{3321} = D$.

ⓕ $a^{2311} = a^{3211} = a^{1322} = a^{3122} = a^{1233} = a^{2133} = E$.

The simplest case is $A = B = C = \mu = D = E$ and μ is a constant. In such a case, $\sigma(\cdot, \mathbf{n}) = \mu \mathrm{Def} u\, \mathbf{n}$ or more precise,

$$\sigma(\boldsymbol{x}, \mathbf{n})^i = \frac{\mu}{2} \sum_{j=1}^{3} \left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \right)(\boldsymbol{x}) n_j(\boldsymbol{x}) .$$

§3.3 Models with Several Spatial Variables

Due to the isotropy (that is, properties are the same in all directions) and the symmetry of $a^{ijk\ell}$, we have

ⓐ $a^{1111} = a^{2222} = a^{3333} = A$.

ⓑ $a^{1222} = a^{2122} = a^{1333} = a^{3133} = a^{2111} = a^{1211} = a^{2333} = a^{3233} = a^{3111} = a^{1311} = a^{3222} = a^{2322} = B$.

ⓒ $a^{1122} = a^{2211} = a^{1133} = a^{3311} = a^{2233} = a^{3322} = C$.

ⓓ $a^{1212} = a^{2112} = a^{2121} = a^{1221} = a^{1313} = a^{3113} = a^{3131} = a^{1331} = a^{2323} = a^{3223} = a^{3232} = a^{3223} = \mu$.

ⓔ $a^{1123} = a^{1132} = a^{2213} = a^{2231} = a^{3312} = a^{3321} = D$.

ⓕ $a^{2311} = a^{3211} = a^{1322} = a^{3122} = a^{1233} = a^{2133} = E$.

The simplest case is $A = B = C = \mu = D = E$ and μ is a constant. In such a case, $\sigma(\cdot, \mathbf{n}) = \mu \mathrm{Def} u\, \mathbf{n}$ or more precise,

$$\sigma(\boldsymbol{x}, \mathbf{n})^i = \frac{\mu}{2} \sum_{j=1}^{3} \left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \right)(\boldsymbol{x}) n_j(\boldsymbol{x}) \,.$$

Chapter 4. Optimization Problems and Calculus of Variations（最佳化問題與變分）

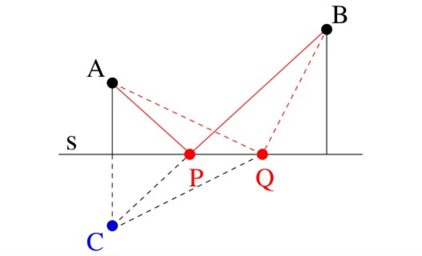
§4.1 Examples of Optimization Problems

§4.1.1 Heron's principle

Given a straight line L and two points a, b on a plane P, find a point x on L such that $|\overline{ax}| + |\overline{bx}|$ is minimal.

Theorem

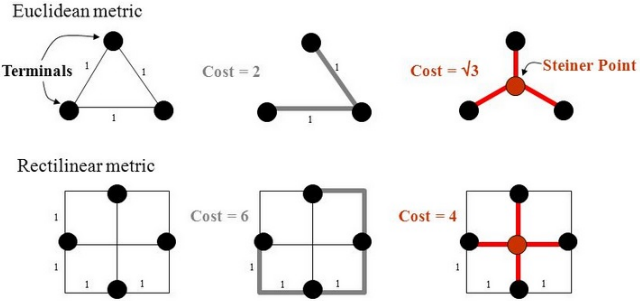
If x is a point of L such that the sum $|\overline{ax}| + |\overline{bx}|$ is the least possible, then the lines $\overline{ax}$ and $\overline{bx}$ form equal angles with the line L.



§4.1 Examples of Optimization Problems

§4.1.2 Steiner's tree problem

The minimum spanning tree problem: given a set V of points (vertices), interconnect them by a network (graph) of shortest length, where the length is the sum of the lengths of all edges. In the Steiner tree problem, extra intermediate vertices and edges may be added to the graph in order to reduce the length of the spanning tree.



§4.1 Examples of Optimization Problems

§4.1.3 Separation problem (分群問題)

Suppose that we are given two types of points in $\mathbb{R}^n$: points of type A $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ and points of type B $\boldsymbol{x}_{m+1}, \boldsymbol{x}_{m+2}, \cdots, \boldsymbol{x}_{m+p}$. The goal of the separation problem is to find a ***linear separator***, a hyperplane of the form

$$H(\boldsymbol{\omega}, \beta) \equiv \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \boldsymbol{\omega} \cdot \boldsymbol{x} + \beta = 0 \right\}$$

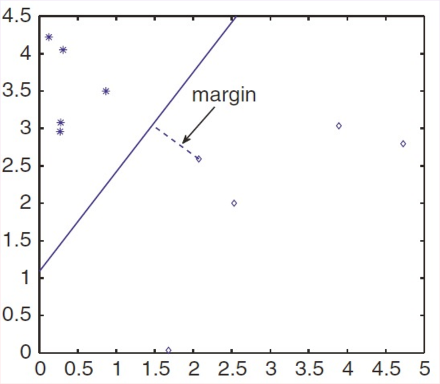
for which

1. points of type A and points of type B are on opposite sides of the hyperplane, and
2. the hyperplane is the "farthest" as possible from all points.

§4.1 Examples of Optimization Problems

The margin of the separator is the distance of the separator from the closest point, as illustrated in the figure below. In mathematics,

$$\text{margin} = \min_{1 \leqslant i \leqslant m+p} \frac{|\boldsymbol{\omega} \cdot \boldsymbol{x}_i + \beta|}{\|\boldsymbol{\omega}\|_2}.$$



§4.1 Examples of Optimization Problems

The separation problem will thus consist of finding the linear separator with the largest margin:

$$\max_{(\boldsymbol{\omega},\beta)\in\mathbb{R}^{n+1}} \left\{ \min_{1\leqslant i\leqslant m+p} \frac{|\boldsymbol{\omega}\cdot \mathbf{x}_i+\beta|}{\|\boldsymbol{\omega}\|_2} \right\}$$

subject to the following constraints:

$$\begin{cases} \boldsymbol{\omega}\cdot \mathbf{x}_i+\beta<0 & \text{for } 1\leqslant i\leqslant m\,, \\ \boldsymbol{\omega}\cdot \mathbf{x}_i+\beta>0 & \text{for } m+1\leqslant i\leqslant m+p\,. \end{cases}$$

§4.1 Examples of Optimization Problems

§4.1.4 Dido's problem (Isoperimetric problem)

For a simple closed curve C in the plane, let $\ell(C)$ denote the length of the curve. **The isoperimetric problem** is to find a closed curve C satisfying $\ell(C) = L$ which encloses the largest area.

Theorem

If $A(C)$ denotes the area enclosed by the curve C, then

$$\ell(C)^2 \geqslant 4\pi A(C) \qquad \text{for every simple closed curve } C, \tag{20}$$

and "$=$" holds if and only if C is a circle.

Inequality (20) is called the ***isoperimetric inequality***（等周不等式）.

§4.1 Examples of Optimization Problems

§4.1.4 Dido's problem (Isoperimetric problem)

For a simple closed curve C in the plane, let $\ell(C)$ denote the length of the curve. **The isoperimetric problem** is to find a closed curve C satisfying $\ell(C) = L$ which encloses the largest area.

Theorem

If $A(C)$ denotes the area enclosed by the curve C, then

$$\ell(C)^2 \geqslant 4\pi A(C) \qquad \textit{for every simple closed curve } C, \tag{20}$$

and "=" holds if and only if C is a circle.

Inequality (20) is called the ***isoperimetric inequality***（等周不等式）.

§4.1 Examples of Optimization Problems

§4.1.4 Dido's problem (Isoperimetric problem)

For a simple closed curve C in the plane, let $\ell(C)$ denote the length of the curve. **The isoperimetric problem** is to find a closed curve C satisfying $\ell(C) = L$ which encloses the largest area.

Theorem

If $A(C)$ denotes the area enclosed by the curve C, then

$$\ell(C)^2 \geqslant 4\pi A(C) \qquad \textit{for every simple closed curve } C, \tag{20}$$

and "$=$" holds if and only if C is a circle.

Inequality (20) is called the ***isoperimetric inequality***（等周不等式）.

§4.1 Examples of Optimization Problems

Sketch of the proof.

Let $\mathcal{P}_n$ denote the collection of simple closed polygon with $2n$ sides and with perimeter L. We look for P in $\mathcal{P}_n$ which encloses the largest area. For given points $B_1, \cdots, B_m$, let $[B_1, B_2, \cdots, B_m, B_1]$ denote the polygon with edges $\overline{B_1B_2}$, $\overline{B_2B_3}$, $\cdots$, $\overline{B_{m-1}B_m}$ and $\overline{B_mB_1}$. Suppose that

$$P_n = [A_1, A_2, \cdots, A_n, A_{n+1}, \cdots, A_{2n}, A_1]$$

is a polygon in $\mathcal{P}_n$ which encloses the largest area. We use the notion $A_j = A_k$ if $j = k \pmod{2n}$.

Claim I: P_n is convex.

Claim II: For all $j \in \mathbb{N}$, $|\overline{A_jA_{j+1}}| = |\overline{A_{j+1}A_{j+2}}|$.

Claim III: For all $j \in \mathbb{N}$, the two polygons $[A_j, A_{j+1}, \cdots, A_{j+n}, A_j]$ and $[A_{j+n}, A_{j+n+1}, \cdots, A_{j+2n}, A_{j+n}]$ enclose the same area. □

§4.1 Examples of Optimization Problems

Sketch of the proof.

Let $\mathcal{P}_n$ denote the collection of simple closed polygon with $2n$ sides and with perimeter L. We look for P in $\mathcal{P}_n$ which encloses the largest area. For given points $B_1, \cdots, B_m$, let $[B_1, B_2, \cdots, B_m, B_1]$ denote the polygon with edges $\overline{B_1B_2}$, $\overline{B_2B_3}$, $\cdots$, $\overline{B_{m-1}B_m}$ and $\overline{B_mB_1}$. Suppose that

$$P_n = [A_1, A_2, \cdots, A_n, A_{n+1}, \cdots, A_{2n}, A_1]$$

is a polygon in $\mathcal{P}_n$ which encloses the largest area. We use the notion $A_j = A_k$ if $j = k \pmod{2n}$.

Claim I: P_n is convex.

Claim II: For all $j \in \mathbb{N}$, $|\overline{A_jA_{j+1}}| = |\overline{A_{j+1}A_{j+2}}|$.

Claim III: For all $j \in \mathbb{N}$, the two polygons $[A_j, A_{j+1}, \cdots, A_{j+n}, A_j]$ and $[A_{j+n}, A_{j+n+1}, \cdots, A_{j+2n}, A_{j+n}]$ enclose the same area. □

§4.1 Examples of Optimization Problems

Sketch of the proof.

Let $\mathcal{P}_n$ denote the collection of simple closed polygon with $2n$ sides and with perimeter L. We look for P in $\mathcal{P}_n$ which encloses the largest area. For given points $B_1, \cdots, B_m$, let $[B_1, B_2, \cdots, B_m, B_1]$ denote the polygon with edges $\overline{B_1B_2}$, $\overline{B_2B_3}$, $\cdots$, $\overline{B_{m-1}B_m}$ and $\overline{B_mB_1}$. Suppose that

$$P_n = [A_1, A_2, \cdots, A_n, A_{n+1}, \cdots, A_{2n}, A_1]$$

is a polygon in $\mathcal{P}_n$ which encloses the largest area. We use the notion $A_j = A_k$ if $j = k \pmod{2n}$.

Claim I: P_n is convex.

Claim II: For all $j \in \mathbb{N}$, $|\overline{A_jA_{j+1}}| = |\overline{A_{j+1}A_{j+2}}|$.

Claim III: For all $j \in \mathbb{N}$, the two polygons $[A_j, A_{j+1}, \cdots, A_{j+n}, A_j]$ and $[A_{j+n}, A_{j+n+1}, \cdots, A_{j+2n}, A_{j+n}]$ enclose the same area. □

§4.1 Examples of Optimization Problems

Sketch of the proof.

Let $\mathcal{P}_n$ denote the collection of simple closed polygon with $2n$ sides and with perimeter L. We look for P in $\mathcal{P}_n$ which encloses the largest area. For given points $B_1, \cdots, B_m$, let $[B_1, B_2, \cdots, B_m, B_1]$ denote the polygon with edges $\overline{B_1B_2}$, $\overline{B_2B_3}$, $\cdots$, $\overline{B_{m-1}B_m}$ and $\overline{B_mB_1}$. Suppose that

$$P_n = [A_1, A_2, \cdots, A_n, A_{n+1}, \cdots, A_{2n}, A_1]$$

is a polygon in $\mathcal{P}_n$ which encloses the largest area. We use the notion $A_j = A_k$ if $j = k \pmod{2n}$.

Claim I: P_n is convex.

Claim II: For all $j \in \mathbb{N}$, $|\overline{A_jA_{j+1}}| = |\overline{A_{j+1}A_{j+2}}|$.

Claim III: For all $j \in \mathbb{N}$, the two polygons $[A_j, A_{j+1}, \cdots, A_{j+n}, A_j]$ and $[A_{j+n}, A_{j+n+1}, \cdots, A_{j+2n}, A_{j+n}]$ enclose the same area. □

§4.1 Examples of Optimization Problems

Sketch of the proof.

Let $\mathcal{P}_n$ denote the collection of simple closed polygon with $2n$ sides and with perimeter L. We look for P in $\mathcal{P}_n$ which encloses the largest area. For given points $B_1, \cdots, B_m$, let $[B_1, B_2, \cdots, B_m, B_1]$ denote the polygon with edges $\overline{B_1 B_2}$, $\overline{B_2 B_3}$, $\cdots$, $\overline{B_{m-1} B_m}$ and $\overline{B_m B_1}$. Suppose that

$$P_n = [A_1, A_2, \cdots, A_n, A_{n+1}, \cdots, A_{2n}, A_1]$$

is a polygon in $\mathcal{P}_n$ which encloses the largest area. We use the notion $A_j = A_k$ if $j = k \pmod{2n}$.

Claim I: P_n is convex.

Claim II: For all $j \in \mathbb{N}$, $|\overline{A_j A_{j+1}}| = |\overline{A_{j+1} A_{j+2}}|$.

Claim III: For all $j \in \mathbb{N}$, the two polygons $[A_j, A_{j+1}, \cdots, A_{j+n}, A_j]$ and $[A_{j+n}, A_{j+n+1}, \cdots, A_{j+2n}, A_{j+n}]$ enclose the same area. □

§4.1 Examples of Optimization Problems

Proof (cont.)

Claim IV: For $1 < j < n+1$, $\overline{A_1A_j} \perp \overline{A_jA_{n+1}}$ at A_j.

Proof of Claim IV: If $\overline{A_1A_j}$ is not perpendicular to $\overline{A_jA_{n+1}}$ at A_j, we can adjust the position of A_1 to A_1', and adjust accordingly the positions of $A_2, \cdots, A_{j-1}$ to $A_2', \cdots, A_{j-1}'$ so that the polygon $[A_1, A_2, \cdots, A_j, A_1]$ is the identical (in shape) to $[A_1', A_2', \cdots, A_{j-1}', A_j, A_1']$. We note that the area enclosed by the polygon $[A_1', \cdots, A_{j-1}', A_j, A_{j+1}, \cdots, A_{n+1}, A_1']$ is larger than the area enclosed by the polygon $[A_1, \cdots, A_{n+1}, A_1]$. **(End of proof of Claim IV)** □

§4.1 Examples of Optimization Problems

Proof (cont.)

By **Claim IV**, $A_j's$ locates on a circle (with diameter $|A_1A_{n+1}|$). Let r_n be the radius of the circle in which P_n is inscribed. Then $4nr_n \sin\frac{\pi}{2n} = L$ and the area A_n enclosed by P_n is

$$A_n = nr_n^2 \sin\frac{\pi}{n} = \frac{L^2}{8n}\cot\frac{\pi}{2n};$$

thus $A_{n+1} \geqslant A_n$ for all $n \in \mathbb{N}$. The circle C with radius r has length L and encloses the largest area among all simple closed curves with length L and $L^2 = 4\pi A$. □

§4.1 Examples of Optimization Problems

On the other hand, the optimization problem can be reformulated by looking for "minimizer" of a certain functional in the space of piece-wise continuously differentiable closed curve. To be more precise, we look for curves C that can be parameterized, using the arc-length, by vector-valued function $\boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j}$ in the set

$$\mathcal{A} = \Big\{ \boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} \,\Big|\, x, y \in \mathcal{D}^1([0, L]; \mathbb{R}), \boldsymbol{r}(0) = \boldsymbol{r}(L), \\ |\dot{\boldsymbol{r}}(s)|^2 = 1 \text{ for all } s \in [0, L] \Big\},$$

where $\mathcal{D}^1([a, b]; \mathbb{R})$ denotes the collection of continuous, piecewise continuously differentiable real-valued functions defined on $[a, b]$ so that the functional

$$-\int_0^L \big[x(s)\dot{y}(s) - \dot{x}(s)y(s)\big]\, ds$$

is minimized.

§4.1 Examples of Optimization Problems

On the other hand, the optimization problem can be reformulated by looking for "minimizer" of a certain functional in the space of piece-wise continuously differentiable closed curve. To be more precise, we look for curves C that can be parameterized, using the arc-length, by vector-valued function $\boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j}$ in the set

$$\mathcal{A} = \Big\{ \boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} \,\Big|\, x, y \in \mathcal{D}^1([0, L]; \mathbb{R}), \boldsymbol{r}(0) = \boldsymbol{r}(L), \\ |\dot{\boldsymbol{r}}(s)|^2 = 1 \text{ for all } s \in [0, L] \Big\},$$

where $\mathcal{D}^1([a, b]; \mathbb{R})$ denotes the collection of continuous, piecewise continuously differentiable real-valued functions defined on $[a, b]$ so that the functional

$$-\int_0^L \big[x(s)\dot{y}(s) - \dot{x}(s)y(s)\big]\, ds$$

is minimized.

§4.1 Examples of Optimization Problems

On the other hand, the optimization problem can be reformulated by looking for "minimizer" of a certain functional in the space of piece-wise continuously differentiable closed curve. To be more precise, we look for curves C that can be parameterized, using the arc-length, by vector-valued function $\boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j}$ in the set

$$\mathcal{A} = \Big\{ \boldsymbol{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} \,\Big|\, x, y \in \mathcal{D}^1([0, L]; \mathbb{R}), \boldsymbol{r}(0) = \boldsymbol{r}(L), |\dot{\boldsymbol{r}}(s)|^2 = 1 \text{ for all } s \in [0, L] \Big\},$$

where $\mathcal{D}^1([a, b]; \mathbb{R})$ denotes the collection of continuous, piecewise continuously differentiable real-valued functions defined on $[a, b]$ so that the functional

$$-\int_0^L \big[x(s)\dot{y}(s) - \dot{x}(s)y(s) \big]\, ds$$

is minimized.

§4.1 Examples of Optimization Problems

§4.1.5 Minimal surface of revolution

This is a problem of finding a curve C connecting two given points (x_0, y_0) and (x_1, y_1), where $x_0 < x_1$, such that its surface of revolution has the least surface area. Given a function $y = y(x)$ satisfying $y(x_0) = y_0$ and $y(x_1) = y_1$, the surface of revolution of the curve $C = \{(x, y(x)) \,|\, y \in \mathcal{D}^1([x_0, x_1]; \mathbb{R}), y(x_0) = y_0, y(x_1) = y_1\}$ is

$$2\pi \int_{x_0}^{x_1} y(x)\sqrt{1 + y'(x)^2}\, dx\,.$$

Therefore, the problem of minimal surface of revolution is to find a function $y \in \mathcal{A} \equiv \{y \in \mathcal{D}^1([x_0, x_1]; \mathbb{R}) \,|\, y(x_0) = y_0, y(x_1) = y_1\}$ which minimizes the functional

$$I(y) = 2\pi \int_{x_0}^{x_1} y(x)\sqrt{1 + y'(x)^2}\, dx\,.$$

§4.1 Examples of Optimization Problems

§4.1.5 Minimal surface of revolution

This is a problem of finding a curve C connecting two given points (x_0, y_0) and (x_1, y_1), where $x_0 < x_1$, such that its surface of revolution has the least surface area. Given a function $y = y(x)$ satisfying $y(x_0) = y_0$ and $y(x_1) = y_1$, the surface of revolution of the curve $C = \{(x, y(x)) \,|\, y \in \mathcal{D}^1([x_0, x_1]; \mathbb{R}), y(x_0) = y_0, y(x_1) = y_1\}$ is

$$2\pi \int_{x_0}^{x_1} y(x)\sqrt{1 + y'(x)^2}\, dx .$$

Therefore, the problem of minimal surface of revolution is to find a function $y \in \mathcal{A} \equiv \{y \in \mathcal{D}^1([x_0, x_1]; \mathbb{R}) \,|\, y(x_0) = y_0, y(x_1) = y_1\}$ which minimizes the functional

$$I(y) = 2\pi \int_{x_0}^{x_1} y(x)\sqrt{1 + y'(x)^2}\, dx .$$

§4.1 Examples of Optimization Problems

§4.1.6 Newton's problem

The Newton problem is to find a curve C connecting two given points (x_0, y_0) and (x_1, y_1), where $x_0 < x_1$, such that its surface of revolution has the least resistance from the air when it moves along x-axis with speed v (or velocity $v\,\mathbf{i}$).

Let u be the normal component of the velocity (given some surface of revolution) $\left(\text{thus } u = \frac{dy}{ds}v = \frac{y'v}{\sqrt{1+y'^2}}\right)$. Suppose that for each surface element dS (at point (x, y, z)), the resistance force is

$$\left[\varphi(u)dS\right]\mathbf{N}$$

for some function φ, where $\mathbf{N}$ is the unit normal of the surface with negative first component (which means the resistance force points to the left).

§4.1 Examples of Optimization Problems

§4.1.6 Newton's problem

The Newton problem is to find a curve C connecting two given points (x_0, y_0) and (x_1, y_1), where $x_0 < x_1$, such that its surface of revolution has the least resistance from the air when it moves along x-axis with speed v (or velocity $v\,\mathbf{i}$).

Let u be the normal component of the velocity (given some surface of revolution) $\left(\text{thus } u = \dfrac{dy}{ds}v = \dfrac{y'v}{\sqrt{1+y'^2}}\right)$. Suppose that for each surface element dS (at point (x, y, z)), the resistance force is

$$\big[\varphi(u)dS\big]\mathbf{N}$$

for some function φ, where $\mathbf{N}$ is the unit normal of the surface with negative first component (which means the resistance force points to the left).

§4.1 Examples of Optimization Problems

If the surface of revolution is given by the curve $y = y(x)$, then with ds denoting the infinitesimal arc-length, for each slice of the surface the total force acting on this slice is $2\pi y \varphi(u) ds (\mathbf{N} \cdot \mathbf{e}_1)$ (the $\mathbf{e}_2$ and $\mathbf{e}_3$ components all cancel out); thus by the fact that $\frac{dy}{ds} = (\mathbf{N} \cdot \mathbf{e}_1)$, the total resistance force (in magnitude) is

$$I(y) = 2\pi \int_{x_0}^{x_1} y\varphi(u) ds \frac{dy}{ds} = 2\pi \int_{x_0}^{x_1} yy' \varphi\Big(\frac{y'v}{\sqrt{1+y'^2}}\Big) dx .$$

Therefore, the Newton problem can be formulated as "finding a function $y \in \mathcal{A} \equiv \{y \in \mathscr{D}^1([x_0, x_1]; \mathbb{R}) \,|\, y(x_0) = y_0, y(x_1) = y_1\}$ which minimizes $I(y)$".

Newton's model: $\varphi(u) = u^2$.

§4.1 Examples of Optimization Problems

If the surface of revolution is given by the curve $y = y(x)$, then with ds denoting the infinitesimal arc-length, for each slice of the surface the total force acting on this slice is $2\pi y\varphi(u)ds(\mathbf{N}\cdot\mathbf{e}_1)$ (the $\mathbf{e}_2$ and $\mathbf{e}_3$ components all cancel out); thus by the fact that $\dfrac{dy}{ds} = (\mathbf{N}\cdot\mathbf{e}_1)$, the total resistance force (in magnitude) is

$$I(y) = 2\pi\int_{x_0}^{x_1} y\varphi(u)ds\frac{dy}{ds} = 2\pi\int_{x_0}^{x_1} yy'\varphi\Big(\frac{y'v}{\sqrt{1+y'^2}}\Big)dx\,.$$

Therefore, the Newton problem can be formulated as "finding a function $y \in \mathcal{A} \equiv \{y \in \mathscr{D}^1([x_0,x_1];\mathbb{R}) \,|\, y(x_0) = y_0, y(x_1) = y_1\}$ which minimizes $I(y)$".

Newton's model: $\varphi(u) = u^2$.

§4.1 Examples of Optimization Problems

§4.1.7 Brachistochrone problem (最速下降曲線問題)

A brachistochrone curve, meaning "shortest time" or curve of fastest descent, is the curve that would carry an idealized point-like body, starting at rest and moving along the curve, without friction, under constant gravity, to a given end point in the shortest time. For given two points $(0,0)$ and (a,b), where $a>0$ and $b<0$, what is the brachistochrone curve connecting $(0,0)$ and (a,b)?

Given a curve parameterized by $\{(x,y(x))\,|\,x\in[0,a]\}$ for some function $y\in\mathscr{D}^1([0,a];\mathbb{R})$, the total time required to travel from $(0,0)$ to (a,b) is given by

$$T(y)=\int_0^a \frac{\sqrt{1+y'(x)^2}}{\sqrt{-2gy(x)}}\,dx\,.$$

§4.1 Examples of Optimization Problems

§4.1.7 Brachistochrone problem (最速下降曲線問題)

A brachistochrone curve, meaning "shortest time" or curve of fastest descent, is the curve that would carry an idealized point-like body, starting at rest and moving along the curve, without friction, under constant gravity, to a given end point in the shortest time. For given two points $(0,0)$ and (a,b), where $a>0$ and $b<0$, what is the brachistochrone curve connecting $(0,0)$ and (a,b)?

Given a curve parameterized by $\{(x,y(x))\,|\,x\in[0,a]\}$ for some function $y\in\mathscr{D}^1([0,a];\mathbb{R})$, the total time required to travel from $(0,0)$ to (a,b) is given by

$$T(y)=\int_0^a \frac{\sqrt{1+y'(x)^2}}{\sqrt{-2gy(x)}}\,dx\,.$$

§4.1 Examples of Optimization Problems

Therefore, the brachistochrone problem can be formulated as finding $y \in \mathcal{A} = \{y \in \mathcal{D}^1([0,a];\mathbb{R}) \,|\, y(0) = 0, y(a) = b\}$ such that $T(y)$ is minimized. In other words, the minimizer $\hat{y}$ satisfies that

$$T(\hat{y}) = \inf_{y \in \mathcal{A}} \int_0^a \frac{\sqrt{1 + y'(x)^2}}{\sqrt{-2gy(x)}}\, dx\,.$$

§4.1 Examples of Optimization Problems

§4.1.8 Plateau's problem - minimal surface problem

The minimal surface problem is to find a (smooth) surface Σ whose boundary is a given curve C but has the minimal surface area. Consider the simplest case that the orthogonal projection from space onto the xy-plane is a bijection between the curve C and the boundary of a simply connected region Ω on the xy-plane. In this case, there exists a continuous function $f: \partial\Omega \to \mathbb{R}$ so that

$$C = \left\{ x\mathbf{i} + y\mathbf{j} + f(x,y)\mathbf{k} \,\middle|\, (x,y) \in \partial\Omega \right\}.$$

The goal is then to find a (smooth) function $z = u(x,y)$ defined on Ω such that $u = f$ on $\partial\Omega$ and

$$\int_\Omega \sqrt{1 + u_x(x,y)^2 + u_y(x,y)^2}\, dA = \min_{v \in \mathcal{A}} \int_\Omega \sqrt{1 + v_x(x,y)^2 + v_y(x,y)^2}\, dA,$$

§4.1 Examples of Optimization Problems

§4.1.8 Plateau's problem - minimal surface problem

The minimal surface problem is to find a (smooth) surface Σ whose boundary is a given curve C but has the minimal surface area. Consider the simplest case that the orthogonal projection from space onto the xy-plane is a bijection between the curve C and the boundary of a simply connected region Ω on the xy-plane. In this case, there exists a continuous function $f\colon \partial\Omega \to \mathbb{R}$ so that

$$C = \left\{ x\mathbf{i} + y\mathbf{j} + f(x,y)\mathbf{k} \,\middle|\, (x,y) \in \partial\Omega \right\}.$$

The goal is then to find a (smooth) function $z = u(x,y)$ defined on Ω such that $u = f$ on $\partial\Omega$ and

$$\int_\Omega \sqrt{1 + u_x(x,y)^2 + u_y(x,y)^2}\, dA = \min_{v \in \mathcal{A}} \int_\Omega \sqrt{1 + v_x(x,y)^2 + v_y(x,y)^2}\, dA \,,$$

§4.1 Examples of Optimization Problems

where $\mathcal{A}$ is the admissible set

$$\mathcal{A} = \left\{ v : \overline{\Omega} \to \mathbb{R} \,\middle|\, v \text{ is (piecewise) differentiable on } \Omega \text{ and } v = f \text{ on } \partial\Omega \right\}.$$

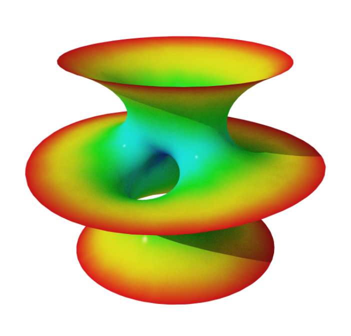


Figure 1: Costa's Minimal Surface - the minimal surface with three circles as prescribed boundaries.

§4.1 Examples of Optimization Problems

§4.1.9 Image processing

An image can often be viewed as a function defined on a square domain. In many problems in image processing, the goal is to recover an ideal image u from an observation f, where u is a perfect original image describing a real scene, f is an observed image, which is a degraded version of u. The degradation can be due to:

1. Signal transmission: there can be some noise (random perturbation).
2. Defects of the imaging system: there can be some blur (deterministic perturbation).

The simplest modelization is the following:

$$f = Ku + n,$$

where n is the noise, and K is the blur, a linear operator.

§4.1 Examples of Optimization Problems

§4.1.9 Image processing

An image can often be viewed as a function defined on a square domain. In many problems in image processing, the goal is to recover an ideal image u from an observation f, where u is a perfect original image describing a real scene, f is an observed image, which is a degraded version of u. The degradation can be due to:

1. Signal transmission: there can be some noise (random perturbation).
2. Defects of the imaging system: there can be some blur (deterministic perturbation).

The simplest modelization is the following:

$$f = Ku + n\,,$$

where n is the noise, and K is the blur, a linear operator.

§4.1 Examples of Optimization Problems

§4.1.9 Image processing

An image can often be viewed as a function defined on a square domain. In many problems in image processing, the goal is to recover an ideal image u from an observation f, where u is a perfect original image describing a real scene, f is an observed image, which is a degraded version of u. The degradation can be due to:

1. Signal transmission: there can be some noise (random perturbation).
2. Defects of the imaging system: there can be some blur (deterministic perturbation).

The simplest modelization is the following:

$$f = Ku + n,$$

where n is the noise, and K is the blur, a linear operator.

§4.1 Examples of Optimization Problems

The following assumptions are classical:

1. K is known (but often not invertible);
2. Only some statistics (mean, variance, $\cdots$) are known of n.

A classical approach in the image processing problems consists in introducing a regularization term L which admits a unique solution of the optimization problem

$$\inf_{u \in \mathcal{A}} \left(\int_\Omega |f - Ku|^2 \, dx + \lambda L(u) \right),$$

where $\mathcal{A}$ is an admissible set which describes the requirement for the real images, and L is a non-negative function (with certain requirements that we will not explore here).

§4.1 Examples of Optimization Problems

The following assumptions are classical:

1. K is known (but often not invertible);
2. Only some statistics (mean, variance, $\cdots$) are known of n.

A classical approach in the image processing problems consists in introducing a regularization term L which admits a unique solution of the optimization problem

$$\inf_{u\in\mathcal{A}}\left(\int_\Omega |f-Ku|^2\,dx+\lambda L(u)\right),$$

where $\mathcal{A}$ is an admissible set which describes the requirement for the real images, and L is a non-negative function (with certain requirements that we will not explore here).

§4.1 Examples of Optimization Problems

Example

Suppose that the polluted image f is solely due to noise (so $K = \mathrm{Id}$, the identity map). The ROF model is a model for denoise which requires the minimization of the functional

$$\int_\Omega |f-u|^2\, dA + \lambda \int_\Omega |\nabla u|\, dA\,,$$

where u should picked up in the admissible set

$$\mathcal{A} = \Big\{ u : \Omega \to \mathbb{R} \,\Big|\, u \text{ is continuous and piecewise differentiable with } \int_\Omega |\nabla u|\, dA < \infty \Big\}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Let $[a, b] \subseteq \mathbb{R}$, $L : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be continuous. We consider the problem of minimizing the functional

$$I(y) = \int_a^b L(x, y(x), y'(x))\, dx$$

for $y \in \mathcal{C}^1([a, b]; \mathbb{R})$ or $\mathcal{D}^1([a, b]; \mathbb{R})$, and y satisfies the boundary condition $y(a) = A_0, y(b) = B_0$, where $\mathcal{C}^1([a, b]; \mathbb{R})$ denotes the space of continuously differentiable real-valued functions defined on $[a, b]$, and $\mathcal{D}^1([a, b]; \mathbb{R})$ denotes the space of continuous, piecewise continuously differentiable real-valued functions defined on $[a, b]$.

§4.2 Simplest Problem in Calculus of Variations

In other words, with $\mathcal{A}$ denoting either the set

$$\left\{y \in \mathcal{C}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0\right\}$$

or

$$\left\{y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0\right\},$$

we consider the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

The function L is called the ***Lagrangian***.

In the following discussion, we write $L = L(x, y, p)$ and let $\underset{z \in \mathcal{A}}{\arg\min}\, I(z)$ denote the minimizer, if exists, of the minimization problem $\min_{z \in \mathcal{A}} I(z)$. In other word, if $y = \underset{z \in \mathcal{A}}{\arg\min}\, I(z)$, then $y \in \mathcal{A}$ and

$$I(y) \leqslant I(z) \qquad \forall\, z \in \mathcal{A}\,.$$

§4.2 Simplest Problem in Calculus of Variations

In other words, with $\mathcal{A}$ denoting either the set

$$\left\{y \in \mathcal{C}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0\right\}$$

or

$$\left\{y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0\right\},$$

we consider the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx.$$

The function L is called the ***Lagrangian***.

In the following discussion, we write $L = L(x,y,p)$ and let $\underset{z\in\mathcal{A}}{\arg\min}\, I(z)$ denote the minimizer, if exists, of the minimization problem $\min\limits_{z\in\mathcal{A}} I(z)$. In other word, if $y = \underset{z\in\mathcal{A}}{\arg\min}\, I(z)$, then $y \in \mathcal{A}$ and

$$I(y) \leqslant I(z) \qquad \forall\, z \in \mathcal{A}.$$

§4.2 Simplest Problem in Calculus of Variations

Remark: Let

$$\mathcal{X} = \{y \in \mathcal{C}^1([a,b];\mathbb{R}) \mid y(a) = A_0, y(b) = B_0\}$$
$$\mathcal{Y} = \{y \in \mathcal{D}^1([a,b];\mathbb{R}) \mid y(a) = A_0, y(b) = B_0\}.$$

Then $\arg\min\limits_{z\in\mathcal{X}} I(z)$, if exists, equals $\arg\min\limits_{z\in\mathcal{Y}} I(z)$. To see this, we first note that $\min\limits_{z\in\mathcal{X}} I(z) \geqslant \min\limits_{z\in\mathcal{Y}} I(z)$; thus for $\arg\min\limits_{z\in\mathcal{X}} I(z) \neq \arg\min\limits_{z\in\mathcal{Y}} I(z)$ to hold, we must have $\widehat{y} \in \mathcal{Y}\backslash\mathcal{X}$ such that $I(\widehat{y}) < \min\limits_{z\in\mathcal{X}} I(z)$. By smooth $\widehat{y}$ at corners, we obtain $\bar{y} \in \mathcal{X}$ such that $I(\bar{y}) < \min\limits_{z\in\mathcal{X}} I(z)$, a contradiction.

However, it is possible that there are only minimizers in $\mathcal{D}^1([a,b];\mathbb{R})$.

§4.2 Simplest Problem in Calculus of Variations

Remark: Let

$$\mathcal{X} = \left\{ y \in \mathcal{C}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}$$
$$\mathcal{Y} = \left\{ y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}.$$

Then $\arg\min\limits_{z \in \mathcal{X}} I(z)$, if exists, equals $\arg\min\limits_{z \in \mathcal{Y}} I(z)$. To see this, we first note that $\min\limits_{z \in \mathcal{X}} I(z) \geqslant \min\limits_{z \in \mathcal{Y}} I(z)$; thus for $\arg\min\limits_{z \in \mathcal{X}} I(z) \neq \arg\min\limits_{z \in \mathcal{Y}} I(z)$ to hold, we must have $\widehat{y} \in \mathcal{Y} \backslash \mathcal{X}$ such that $I(\widehat{y}) < \min\limits_{z \in \mathcal{X}} I(z)$. By smooth $\widehat{y}$ at corners, we obtain $\bar{y} \in \mathcal{X}$ such that $I(\bar{y}) < \min\limits_{z \in \mathcal{X}} I(z)$, a contradiction.

However, it is possible that there are only minimizers in $\mathcal{D}^1([a,b];\mathbb{R})$.

§4.2 Simplest Problem in Calculus of Variations

Remark: Let

$$\mathcal{X} = \left\{ y \in \mathcal{C}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}$$
$$\mathcal{Y} = \left\{ y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}.$$

Then $\underset{z\in\mathcal{X}}{\arg\min}\, I(z)$, if exists, equals $\underset{z\in\mathcal{Y}}{\arg\min}\, I(z)$. To see this, we first note that $\min\limits_{z\in\mathcal{X}} I(z) \geqslant \min\limits_{z\in\mathcal{Y}} I(z)$; thus for $\underset{z\in\mathcal{X}}{\arg\min}\, I(z) \neq \underset{z\in\mathcal{Y}}{\arg\min}\, I(z)$ to hold, we must have $\widehat{y} \in \mathcal{Y}\backslash\mathcal{X}$ such that $I(\widehat{y}) < \min\limits_{z\in\mathcal{X}} I(z)$. By smooth $\widehat{y}$ at corners, we obtain $\bar{y} \in \mathcal{X}$ such that $I(\bar{y}) < \min\limits_{z\in\mathcal{X}} I(z)$, a contradiction.

However, it is possible that there are only minimizers in $\mathcal{D}^1([a,b];\mathbb{R})$.

§4.2 Simplest Problem in Calculus of Variations

§4.2.1 First variation of I

Let

$$\mathcal{A} = \left\{ y \in \mathcal{D}^1([a, b]; \mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}$$

and

$$\mathcal{N} = \left\{ \eta \in \mathcal{D}^1([a, b]; \mathbb{R}) \,\middle|\, \eta(a) = \eta(b) = 0 \right\},$$

called the ***admissible set*** and the ***test function space***, respectively.

For $y \in \mathcal{A}$, $\eta \in \mathcal{N}$ and $\epsilon \in \mathbb{R}$, let $J(\epsilon) = I(y + \epsilon\eta)$ and consider the following quotient

$$\frac{J(\epsilon) - J(0)}{\epsilon} = \frac{1}{\epsilon} \int_a^b \left[L(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x)) - L(x, y(x), y'(x)) \right] dx$$

for all $\epsilon \neq 0$.

§4.2 Simplest Problem in Calculus of Variations

§4.2.1 First variation of I

Let

$$\mathcal{A} = \left\{ y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, y(a) = A_0, y(b) = B_0 \right\}$$

and

$$\mathcal{N} = \left\{ \eta \in \mathcal{D}^1([a,b];\mathbb{R}) \,\middle|\, \eta(a) = \eta(b) = 0 \right\},$$

called the ***admissible set*** and the ***test function space***, respectively.

For $y \in \mathcal{A}$, $\eta \in \mathcal{N}$ and $\epsilon \in \mathbb{R}$, let $J(\epsilon) = I(y + \epsilon\eta)$ and consider the following quotient

$$\frac{J(\epsilon) - J(0)}{\epsilon} = \frac{1}{\epsilon}\int_a^b \big[L(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x)) - L(x, y(x), y'(x))\big]\,dx$$

for all $\epsilon \neq 0$.

§4.2 Simplest Problem in Calculus of Variations

Assume that L_y and L_p are continuous, then

$$\lim_{\epsilon\to 0}\frac{J(\epsilon)-J(0)}{\epsilon}$$
$$=\int_a^b \big[L_y(x,y(x),y'(x))\eta(x)+L_p(x,y(x),y'(x))\eta'(x)\big]\,dx\,.$$

This limit, denoted by $\delta I(y;\eta)$ or $\frac{\delta I}{\delta\eta}(y)$, is called the ***first variation*** of I at y along η.

Theorem

If $y=\arg\min_{z\in\mathcal{A}} I(z)$ is a minimizer of I, then $\delta I(y;\eta)=0$ for all $\eta\in\mathcal{N}$.

Sketch of proof.

If y is a minimizer of I, then $I(y)\leqslant I(y+\epsilon\eta)$ for all $\epsilon\in\mathbb{R}$ since $y+\epsilon\eta\in\mathcal{A}$. □

§4.2 Simplest Problem in Calculus of Variations

Assume that L_y and L_p are continuous, then

$$\lim_{\epsilon\to 0}\frac{J(\epsilon)-J(0)}{\epsilon} = \int_a^b \left[L_y(x,y(x),y'(x))\eta(x) + L_p(x,y(x),y'(x))\eta'(x)\right] dx\,.$$

This limit, denoted by $\delta I(y;\eta)$ or $\dfrac{\delta I}{\delta\eta}(y)$, is called the ***first variation*** of I at y along η.

Theorem

If $y = \arg\min_{z\in\mathcal{A}} I(z)$ is a minimizer of I, then $\delta I(y;\eta) = 0$ for all $\eta\in\mathcal{N}$.

Sketch of proof.

If y is a minimizer of I, then $I(y) \leqslant I(y+\epsilon\eta)$ for all $\epsilon\in\mathbb{R}$ since $y+\epsilon\eta\in\mathcal{A}$. □

§4.2 Simplest Problem in Calculus of Variations

Assume that L_y and L_p are continuous, then

$$\lim_{\epsilon\to 0}\frac{J(\epsilon)-J(0)}{\epsilon} = \int_a^b \big[L_y(x,y(x),y'(x))\eta(x) + L_p(x,y(x),y'(x))\eta'(x)\big]\,dx\,.$$

This limit, denoted by $\delta I(y;\eta)$ or $\dfrac{\delta I}{\delta\eta}(y)$, is called the ***first variation*** of I at y along η.

Theorem

If $y = \operatorname*{arg\,min}_{z\in\mathcal{A}} I(z)$ is a minimizer of I, then $\delta I(y;\eta) = 0$ for all $\eta\in\mathcal{N}$.

Sketch of proof.

If y is a minimizer of I, then $I(y) \leqslant I(y+\epsilon\eta)$ for all $\epsilon\in\mathbb{R}$ since $y+\epsilon\eta\in\mathcal{A}$. □

§4.2 Simplest Problem in Calculus of Variations

Assume that L_y and L_p are continuous, then

$$\lim_{\epsilon \to 0} \frac{J(\epsilon) - J(0)}{\epsilon} = \int_a^b \big[L_y(x, y(x), y'(x))\eta(x) + L_p(x, y(x), y'(x))\eta'(x)\big]\, dx\,.$$

This limit, denoted by $\delta I(y;\eta)$ or $\dfrac{\delta I}{\delta \eta}(y)$, is called the ***first variation*** of I at y along η.

Theorem

If $y = \arg\min\limits_{z \in \mathcal{A}} I(z)$ is a minimizer of I, then $\delta I(y;\eta) = 0$ for all $\eta \in \mathcal{N}$.

Sketch of proof.

If y is a minimizer of I, then $I(y) \leqslant I(y + \epsilon\eta)$ for all $\epsilon \in \mathbb{R}$ since $y + \epsilon\eta \in \mathcal{A}$. □

§4.2 Simplest Problem in Calculus of Variations

Assume that L_y and L_p are continuous, then

$$\lim_{\epsilon\to 0}\frac{J(\epsilon)-J(0)}{\epsilon} = \int_a^b \big[L_y(x,y(x),y'(x))\eta(x)+L_p(x,y(x),y'(x))\eta'(x)\big]\,dx\,.$$

This limit, denoted by $\delta I(y;\eta)$ or $\dfrac{\delta I}{\delta\eta}(y)$, is called the ***first variation*** of I at y along η.

Theorem

If $y=\displaystyle\arg\min_{z\in\mathcal{A}} I(z)$ is a minimizer of I, then $\delta I(y;\eta)=0$ for all $\eta\in\mathcal{N}$.

Sketch of proof.

If y is a minimizer of I, then $J(0)\leqslant J(\epsilon)$ for all $\epsilon\in\mathbb{R}$; thus J attains it minimum at 0 so that $J'(0)=0$. □

§4.2 Simplest Problem in Calculus of Variations

Definition

The integral equation

$$\int_a^b \left[L_y(x, y(x), y'(x))\eta(x) + L_p(x, y(x), y'(x))\eta'(x)\right] dx = 0$$

for all $\eta \in \mathcal{N}$ is called the ***weak form*** of the ***Euler-Lagrange equation*** associated with the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

§4.2 Simplest Problem in Calculus of Variations

The weak form of the Euler-Lagrange equation does not seem to tell us too much about how y should look like, and we prefer to see if the minimizer satisfies a differential equation. In order to see what differential equation the minimizer satisfies, we need some basic lemmas.

Lemma

If $y \in \mathcal{C}([a,b];\mathbb{R})$ and $\displaystyle\int_a^b y(x)\eta(x)\,dx = 0$ for all $\eta \in \mathcal{C}([a,b];\mathbb{R})$, then $y \equiv 0$.

Proof.

By assumption,

$$\int_a^b y(x)^2\,dx = 0\,;$$

thus by the fact that y is continuous, $y \equiv 0$. □

§4.2 Simplest Problem in Calculus of Variations

The weak form of the Euler-Lagrange equation does not seem to tell us too much about how y should look like, and we prefer to see if the minimizer satisfies a differential equation. In order to see what differential equation the minimizer satisfies, we need some basic lemmas.

Lemma

If $y \in \mathcal{C}([a,b];\mathbb{R})$ *and* $\displaystyle\int_a^b y(x)\eta(x)\,dx = 0$ *for all* $\eta \in \mathcal{C}([a,b];\mathbb{R})$, *then* $y \equiv 0$.

Proof.

By assumption,

$$\int_a^b y(x)^2\,dx = 0\,;$$

thus by the fact that y is continuous, $y \equiv 0$. □

§4.2 Simplest Problem in Calculus of Variations

The weak form of the Euler-Lagrange equation does not seem to tell us too much about how y should look like, and we prefer to see if the minimizer satisfies a differential equation. In order to see what differential equation the minimizer satisfies, we need some basic lemmas.

Lemma

If $y \in \mathcal{C}([a,b];\mathbb{R})$ and $\displaystyle\int_a^b y(x)\eta(x)\,dx = 0$ for all $\eta \in \mathcal{C}([a,b];\mathbb{R})$, then $y \equiv 0$.

Proof.

By assumption,

$$\int_a^b y(x)^2\,dx = 0\,;$$

thus by the fact that y is continuous, $y \equiv 0$. □

§4.2 Simplest Problem in Calculus of Variations

The weak form of the Euler-Lagrange equation does not seem to tell us too much about how y should look like, and we prefer to see if the minimizer satisfies a differential equation. In order to see what differential equation the minimizer satisfies, we need some basic lemmas.

Lemma

If $y \in \mathcal{C}([a,b];\mathbb{R})$ and $\displaystyle\int_a^b y(x)\eta(x)\,dx = 0$ for all $\eta \in \mathcal{C}([a,b];\mathbb{R})$, then $y \equiv 0$.

Remark: It requires more analysis to show the following conclusion:

If $y \in \mathcal{C}([a,b];\mathbb{R})$ and $\displaystyle\int_a^b y(x)\eta(x)\,dx = 0$ for all $\eta \in \mathcal{D}^1([a,b];\mathbb{R})$, then $y \equiv 0$.

§4.2 Simplest Problem in Calculus of Variations

Lemma

If $y \in \mathcal{C}([a, b]; \mathbb{R})$ *and* $\displaystyle\int_a^b y(x)\eta'(x)\,dx = 0$ *for all* $\eta \in \mathcal{N}$, *then* $y \equiv c$ *for some constant* c.

Proof.

Let $\eta(x) = \displaystyle\int_a^x \big(y(t) - c\big)\,dt$, where the constant c is chosen so that $\displaystyle\int_a^b \big(y(t) - c\big)\,dt = 0$. Then $\eta \in \mathcal{N}$ and

$$\int_a^b |y(x) - c|^2\,dx = \int_a^b \big(y(x) - c\big)\eta'(x)\,dx = -c\int_a^b \eta'(x)\,dx$$
$$= c\big(\eta(a) - \eta(b)\big) = 0\,.$$

Therefore, $y(x) = c$ for all $x \in [a, b]$. □

§4.2 Simplest Problem in Calculus of Variations

Lemma

If $y \in \mathscr{C}([a,b];\mathbb{R})$ *and* $\displaystyle\int_a^b y(x)\eta'(x)\,dx = 0$ *for all* $\eta \in \mathcal{N}$, *then* $y \equiv c$ *for some constant* c.

Proof.

Let $\eta(x) = \displaystyle\int_a^x \big(y(t) - c\big)\,dt$, where the constant c is chosen so that $\displaystyle\int_a^b \big(y(t) - c\big)\,dt = 0$. Then $\eta \in \mathcal{N}$ and

$$\int_a^b \big|y(x) - c\big|^2\,dx = \int_a^b \big(y(x) - c\big)\eta'(x)\,dx = -c\int_a^b \eta'(x)\,dx$$
$$= c\big(\eta(a) - \eta(b)\big) = 0\,.$$

Therefore, $y(x) = c$ for all $x \in [a,b]$. □

§4.2 Simplest Problem in Calculus of Variations

Lemma

If $y \in \mathcal{C}([a,b];\mathbb{R})$ *and* $\displaystyle\int_a^b y(x)\eta'(x)\,dx = 0$ *for all* $\eta \in \mathcal{N}$, *then* $y \equiv c$ *for some constant* c.

Proof.

Let $\eta(x) = \displaystyle\int_a^x \big(y(t) - c\big)\,dt$, where the constant c is chosen so that $\displaystyle\int_a^b \big(y(t) - c\big)\,dt = 0$. Then $\eta \in \mathcal{N}$ and

$$\begin{aligned}\int_a^b \big|y(x) - c\big|^2\,dx &= \int_a^b \big(y(x) - c\big)\eta'(x)\,dx = -c\int_a^b \eta'(x)\,dx \\ &= c\big(\eta(a) - \eta(b)\big) = 0\,.\end{aligned}$$

Therefore, $y(x) = c$ for all $x \in [a,b]$. □

§4.2 Simplest Problem in Calculus of Variations

Lemma

If $y, z \in \mathscr{C}([a,b];\mathbb{R})$ *satisfy*

$$\int_a^b \big[y(x)\eta(x) + z(x)\eta'(x)\big]\,dx = 0 \qquad \forall\, \eta \in \mathcal{N}\,, \tag{21}$$

then $z \in \mathscr{C}^1([a,b];\mathbb{R})$ *and* $z'(x) = y(x)$ *for all* $x \in [a,b]$.

Proof.

Let $z_1(x) = \int_a^x y(t)\,dt$. Integration-by-parts provides that

$$\int_a^b y(x)\eta(x)\,dx = z_1(x)\eta(x)\Big|_{x=a}^{x=b} - \int_a^b z_1(x)\eta'(x)\,dx\,;$$

thus (21) implies that

$$\int_a^b \big[z(x) - z_1(x)\big]\eta'(x)\,dx = 0 \qquad \forall\, \eta \in \mathcal{N}\,. \qquad \square$$

§4.2 Simplest Problem in Calculus of Variations

Lemma

If $y, z \in \mathscr{C}([a,b];\mathbb{R})$ *satisfy*

$$\int_a^b \big[y(x)\eta(x) + z(x)\eta'(x)\big]\,dx = 0 \qquad \forall\, \eta \in \mathcal{N}\,, \tag{21}$$

then $z \in \mathscr{C}^1([a,b];\mathbb{R})$ *and* $z'(x) = y(x)$ *for all* $x \in [a,b]$.

Proof.

Let $z_1(x) = \int_a^x y(t)\,dt$. Integration-by-parts provides that

$$\int_a^b y(x)\eta(x)\,dx = z_1(x)\eta(x)\Big|_{x=a}^{x=b} - \int_a^b z_1(x)\eta'(x)\,dx\,;$$

thus (21) implies that

$$\int_a^b \big[z(x) - z_1(x)\big]\eta'(x)\,dx = 0 \qquad \forall\, \eta \in \mathcal{N}\,.$$

□

§4.2 Simplest Problem in Calculus of Variations

Proof (cont.)

By the previous lemma, $z(x) - z_1(x) = C$ for some constant C. Therefore, $z(x) = C + \int_a^x y(t)\,dt$ which implies that $z \in \mathscr{C}^1([a,b];\mathbb{R})$ and $z'(x) = y(x)$. □

§4.2 Simplest Problem in Calculus of Variations

Lemma

Suppose that $y, z \in \mathscr{C}([a,b];\mathbb{R})$ *and* z *is not a constant function. If*

$$\int_a^b y(x)\eta'(x)\,dx = 0 \quad \forall\, \eta \in \mathcal{N} \text{ and } \eta \text{ satisfies } \int_a^b z(x)\eta'(x)\,dx = 0\,,$$

then there are constants $\lambda, \mu \in \mathbb{R}$ *such that* $y(x) = \lambda z(x) + \mu$.

Proof.

Let $\eta(x) = \int_a^x \big(y(t) - \lambda z(t) - \mu\big)\,dt$, where λ, μ are chosen so that $\eta(b) = 0$ and $\int_a^b z(x)\eta'(x)\,dx = 0$; that is,

$$\lambda \int_a^b z(x)\,dx + \mu \int_a^b dx = \int_a^b y(x)\,dx\,,$$

$$\lambda \int_a^b z^2(x)\,dx + \mu \int_a^b z(x)\,dx = \int_a^b y(x)z(x)\,dx\,. \qquad \square$$

§4.2 Simplest Problem in Calculus of Variations

Lemma

Suppose that $y, z \in \mathcal{C}([a,b];\mathbb{R})$ and z is not a constant function. If $\int_a^b y(x)\eta'(x)\,dx = 0 \quad \forall\, \eta \in \mathcal{N}$ and η satisfies $\int_a^b z(x)\eta'(x)\,dx = 0$, then there are constants $\lambda, \mu \in \mathbb{R}$ such that $y(x) = \lambda z(x) + \mu$.

Proof.

Let $\eta(x) = \int_a^x \big(y(t) - \lambda z(t) - \mu\big)\,dt$, where λ, μ are chosen so that $\eta(b) = 0$ and $\int_a^b z(x)\eta'(x)\,dx = 0$; that is,

$$\lambda \int_a^b z(x)\,dx + \mu \int_a^b dx = \int_a^b y(x)\,dx\,,$$

$$\lambda \int_a^b z^2(x)\,dx + \mu \int_a^b z(x)\,dx = \int_a^b y(x)z(x)\,dx\,. \qquad \square$$

§4.2 Simplest Problem in Calculus of Variations

Proof (cont.)

Since z is not a constant, the Cauchy-Schwarz inequality implies that the system above has a unique solution (λ, μ). Since $\eta \in \mathcal{N}$ and satisfies $\int_a^b z(x)\eta'(x)\,dx = 0$, we have

$$\int_a^b \big|y(x) - \lambda z(x) - \mu\big|^2\,dx = \int_a^b \big(y(x) - \lambda z(x) - \mu\big)\eta'(x)\,dx$$
$$= -\mu \int_a^b \eta'(x)\,dx = 0\,;$$

thus $y(x) = \lambda z(x) + \mu$ for all $x \in [a, b]$. □

§4.2 Simplest Problem in Calculus of Variations

Proof (cont.)

Since z is not a constant, the Cauchy-Schwarz inequality implies that the system above has a unique solution (λ, μ). Since $\eta \in \mathcal{N}$ and satisfies $\displaystyle\int_a^b z(x)\eta'(x)\,dx = 0$, we have

$$\begin{aligned}\int_a^b \big|y(x) - \lambda z(x) - \mu\big|^2\,dx &= \int_a^b \big(y(x) - \lambda z(x) - \mu\big)\eta'(x)\,dx \\ &= -\mu\int_a^b \eta'(x)\,dx = 0\,;\end{aligned}$$

thus $y(x) = \lambda z(x) + \mu$ for all $x \in [a, b]$. □

§4.2 Simplest Problem in Calculus of Variations

§4.2.2 The Euler-Lagrange equation

Recall that the weak form of the Euler-Lagrange equation associated with the minimization problem $\inf\limits_{y \in \mathcal{A}} I(y)$ is

$$\int_a^b \big[L_y(x, y(x), y'(x))\eta(x) + L_p(x, y(x), y'(x))\eta'(x)\big]\, dx = 0 \quad \forall\, \eta \in \mathcal{N}.$$

Theorem

Suppose that L, L_y, L_p are continuous. If $\widehat{y} \in \mathcal{A}$ is a minimizer of the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx\,,$$

then

$$\frac{d}{dx} L_p(x, \widehat{y}(x), \widehat{y}'(x)) = L_y(x, \widehat{y}(x), \widehat{y}'(x))$$

for point x at which $\widehat{y}'$ is continuous.

§4.2 Simplest Problem in Calculus of Variations

Definition

The differential equation

$$\frac{d}{dx}L_p(x, y(x), y'(x)) = L_y(x, y(x), y'(x))$$

is called (the ***strong form*** of) the Euler-Lagrange equation associated with the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

Remark: The theorem above is essentially due to Du Bois-Reymond, so the Euler-Lagrange equation is also called the Du Bois-Reymond equation.

§4.2 Simplest Problem in Calculus of Variations

Definition

The differential equation

$$\frac{d}{dx}L_p(x, y(x), y'(x)) = L_y(x, y(x), y'(x))$$

is called (the ***strong form*** of) the Euler-Lagrange equation associated with the minimization problem

$$\inf_{y\in\mathcal{A}} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

Remark: The theorem above is essentially due to Du Bois-Reymond, so the Euler-Lagrange equation is also called the Du Bois-Reymond equation.

§4.2 Simplest Problem in Calculus of Variations

Example

The Lagrangian for the minimal surface of revolution problem is $L(x, y, p) = y\sqrt{1+p^2}$, so the Euler-Lagrange equation for the minimal surface of revolution problem is

$$\frac{d}{dx}\frac{yy'}{\sqrt{1+y'^2}} = \sqrt{1+y'^2}\,.$$

Example

The Lagrangian for Newton's problem is

$$L(x, y, p) = yp\varphi\Big(\frac{pv}{\sqrt{1+p^2}}\Big)\,,$$

so the Euler-Lagrange equation for Newton's problem (with $\varphi(u) = u^2$) is

$$\frac{d}{dx}\frac{yy'^2(y'^2+3)}{(1+y'^2)^2} = \frac{y'^3}{1+y'^2}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example

The Lagrangian for the minimal surface of revolution problem is $L(x, y, p) = y\sqrt{1+p^2}$, so the Euler-Lagrange equation for the minimal surface of revolution problem is

$$\frac{d}{dx}\frac{yy'}{\sqrt{1+y'^2}} = \sqrt{1+y'^2}\,.$$

Example

The Lagrangian for Newton's problem is

$$L(x, y, p) = yp\varphi\Big(\frac{pv}{\sqrt{1+p^2}}\Big)\,,$$

so the Euler-Lagrange equation for Newton's problem (with $\varphi(u) = u^2$) is

$$\frac{d}{dx}\frac{yy'^2(y'^2+3)}{(1+y'^2)^2} = \frac{y'^3}{1+y'^2}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example

Now we consider the brachistochrone problem. Making the change of variable $y \mapsto -y$ (and ignoring $\sqrt{2g}$ in the denominator), we rewritten the minimization problem as

$$\inf_{y \in \mathcal{A}} \int_0^a \frac{\sqrt{1 + y'(x)^2}}{\sqrt{y(x)}}\, dx$$

where $\mathcal{A} = \big\{ y \in \mathscr{D}^1([0,a];\mathbb{R}) \,\big|\, y(0) = 0, y(a) = -b \big\}$. Therefore, $L(x,y,p) = \dfrac{\sqrt{1+p^2}}{\sqrt{y}}$ which implies that the Euler-Lagrange equation for the brachistochrone problem is

$$\frac{d}{dx} \frac{y'}{\sqrt{y}\sqrt{1+y'^2}} = -\frac{\sqrt{1+y'^2}}{2y^{\frac{3}{2}}}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Theorem

Suppose that $\widehat{y} \in \mathcal{D}^1([a,b];\mathbb{R})$ *satisfies the Euler-Lagrange equation*

$$\frac{d}{dx} L_p(x, \widehat{y}(x), \widehat{y}'(x)) = L_y(x, \widehat{y}(x), \widehat{y}'(x)) .$$

If for some $x \in (a,b)$, L_{px}, L_{py} *are continuous at* $(x, \widehat{y}(x), \widehat{y}'(x))$, $L_{pp}(x, \widehat{y}(x), \widehat{y}'(x)) \neq 0$, *and* $\widehat{y}'$ *is continuous at* x, *then* $\widehat{y}''(x)$ *exists.*

Remark: Let $\widehat{y} = \arg\min\limits_{z \in \mathcal{A}} I(z)$. If L_{px}, L_{py}, L_{pp} are continuous at $(x, \widehat{y}(x), \widehat{y}'(x))$, $L_{pp}(x, \widehat{y}(x), \widehat{y}'(x)) \neq 0$, and $\widehat{y}'$ is continuous in a neighborhood of x, then $\widehat{y}''$ exists in a neighborhood of x and is continuous there.

§4.2 Simplest Problem in Calculus of Variations

Theorem

Suppose that $\widehat{y} \in \mathcal{D}^1([a,b];\mathbb{R})$ *satisfies the Euler-Lagrange equation*

$$\frac{d}{dx}L_p(x,\widehat{y}(x),\widehat{y}'(x)) = L_y(x,\widehat{y}(x),\widehat{y}'(x))\,.$$

If for some $x \in (a,b)$, L_{px}, L_{py} *are continuous at* $(x,\widehat{y}(x),\widehat{y}'(x))$, $L_{pp}(x,\widehat{y}(x),\widehat{y}'(x)) \neq 0$, *and* $\widehat{y}'$ *is continuous at* x, *then* $\widehat{y}''(x)$ *exists.*

Remark: Let $\widehat{y} = \mathop{\arg\min}\limits_{z\in\mathcal{A}} I(z)$. If L_{px}, L_{py}, L_{pp} are continuous at $(x,\widehat{y}(x),\widehat{y}'(x))$, $L_{pp}(x,\widehat{y}(x),\widehat{y}'(x)) \neq 0$, and $\widehat{y}'$ is continuous in a neighborhood of x, then $\widehat{y}''$ exists in a neighborhood of x and is continuous there.

§4.2 Simplest Problem in Calculus of Variations

Example (A minimization problem whose minimizer is not in $\mathscr{C}^1$)

Let $\mathcal{A} = \{y \in \mathscr{D}^1([0,1];\mathbb{R}) \,|\, y(0) = y(1) = 0\}$. Consider the minimization problem

$$\inf_{y\in\mathcal{A}} \int_0^1 \big(y'(x)^2 - 1\big)^2 \, dx\,;$$

that is, we assume $L(x,y,p) = (p^2-1)^2$. The Euler-Lagrange equation associated with this minimization problem is

$$\frac{d}{dx}\frac{d}{dp}\Big|_{p=y'(x)} (p^2-1)^2 = 0$$

which, together with the fact that $L_{pp}(x,y,p) = 12p^2 - 4$, implies that if $p^2 \neq \frac{1}{3}$ the minimizer $\widehat{y}$ satisfies

$$2\widehat{y}'^2\widehat{y}'' + (\widehat{y}'^2 - 1)\widehat{y}'' = 0$$

for points at which $\widehat{y}'$ is continuous.

§4.2 Simplest Problem in Calculus of Variations

Example (cont.)

Therefore, $\widehat{y}''(3\widehat{y}'^2 - 1) = 0$ for points at which $\widehat{y}'$ is continuous if $\widehat{y}'^2 \neq \dfrac{1}{3}$. Therefore, $\widehat{y}'' = 0$ if $\widehat{y}'^2 \neq \dfrac{1}{3}$ which implies that $\widehat{y}'$ is piecewise constant. The minimizer is then saw-tooth like function with slope ± 1, and there are only $\mathcal{D}^1$-minimizers.

§4.2 Simplest Problem in Calculus of Variations

Remark on the extensions of the simplest problem of Calculus of Variations:

1. **Higher derivatives**: The Lagrangian might involves higher order derivatives of y. For example, we can consider the minimization problem

$$\inf_{y \in \mathcal{A}} \int_a^b L(x, y(x), y'(x), y''(x))\, dx\,,$$

where

$$\mathcal{A} = \Big\{ y \in \mathscr{D}^2([a,b];\mathbb{R}) \,\Big|\, y(a) = A_0, y(b) = B_0, \\ y'(a) = A_1, y'(b) = B_1 \Big\}\,.$$

We note that the corresponding test function space is

$$\mathcal{N} = \big\{ y \in \mathscr{D}^2([a,b];\mathbb{R}) \,\big|\, y(a) = y(b) = y'(a) = y'(b) = 0 \big\}\,.$$

§4.2 Simplest Problem in Calculus of Variations

If $\widehat{y}$ is a minimizer, then $J(\epsilon) = I(\widehat{y} + \epsilon\eta)$ attains its minimum at $\epsilon = 0$ for all $\eta \in \mathcal{N}$. This implies $J'(0) = 0$ for all $\eta \in \mathcal{N}$, and this condition gives the ***weak form*** of the Euler-Lagrange equation associated with this minimization problem: write $L = L(x, y, p, q)$,

$$\int_a^b \Big[L_y(x, \widehat{y}(x), \widehat{y}'(x), \widehat{y}''(x))\eta(x) + L_p(x, \widehat{y}(x), \widehat{y}'(x), \widehat{y}''(x))\eta'(x) \\ + L_q(x, \widehat{y}(x), \widehat{y}'(x), \widehat{y}''(x))\eta''(x) \Big] dx = 0$$

for all $\eta \in \mathcal{N}$.

§4.2 Simplest Problem in Calculus of Variations

2. **Free ends**: This is to consider the minimization problem

$$\inf_{y \in \mathcal{D}^1([a,b];\mathbb{R})} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

In this case, the test function space is then $\mathcal{N} = \mathcal{D}^1([a,b];\mathbb{R})$. The same argument implies that

$$\int_a^b \big[L_y(x, \widehat{y}(x), \widehat{y}'(x))\eta(x) + L_p(x, \widehat{y}(x), \widehat{y}'(x))\eta'(x)\big]\, dx = 0 \quad (22)$$

for all $\eta \in \mathcal{N}$ if $\widehat{y}$ is a minimizer. In particular, (22) holds for all $\eta \in \big\{y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\big|\, y(a) = y(b) = 0\big\}$; thus the 3rd lemma shows that if L_y and L_p are continuous, then

$$\frac{d}{dx} L_p(x, \widehat{y}(x), \widehat{y}'(x)) = L_y(x, \widehat{y}(x), \widehat{y}'(x))$$

for point x at which $\widehat{y}'$ is continuous.

§4.2 Simplest Problem in Calculus of Variations

2. **Free ends**: This is to consider the minimization problem

$$\inf_{y \in \mathcal{D}^1([a,b];\mathbb{R})} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

In this case, the test function space is then $\mathcal{N} = \mathcal{D}^1([a,b];\mathbb{R})$. The same argument implies that

$$\int_a^b \big[L_y(x,\widehat{y}(x),\widehat{y}'(x))\eta(x) + L_p(x,\widehat{y}(x),\widehat{y}'(x))\eta'(x)\big]\, dx = 0 \quad (22)$$

for all $\eta \in \mathcal{N}$ if $\widehat{y}$ is a minimizer. In particular, (22) holds for all $\eta \in \big\{ y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\big|\, y(a) = y(b) = 0 \big\}$; thus the 3rd lemma shows that if L_y and L_p are continuous, then

$$\frac{d}{dx} L_p(x,\widehat{y}(x),\widehat{y}'(x)) = L_y(x,\widehat{y}(x),\widehat{y}'(x))$$

for point x at which $\widehat{y}'$ is continuous.

§4.2 Simplest Problem in Calculus of Variations

2. **Free ends**: This is to consider the minimization problem

$$\inf_{y \in \mathcal{D}^1([a,b];\mathbb{R})} \int_a^b L(x, y(x), y'(x))\, dx\,.$$

In this case, the test function space is then $\mathcal{N} = \mathcal{D}^1([a,b];\mathbb{R})$. The same argument implies that

$$\int_a^b \big[L_y(x, \widehat{y}(x), \widehat{y}'(x))\eta(x) + L_p(x, \widehat{y}(x), \widehat{y}'(x))\eta'(x)\big]\, dx = 0 \quad (22)$$

for all $\eta \in \mathcal{N}$ if $\widehat{y}$ is a minimizer. In particular, (22) holds for all $\eta \in \{y \in \mathcal{D}^1([a,b];\mathbb{R}) \,\big|\, y(a) = y(b) = 0\}$; thus the 3rd lemma shows that if L_y and L_p are continuous, then

$$\frac{d}{dx} L_p(x, \widehat{y}(x), \widehat{y}'(x)) = L_y(x, \widehat{y}(x), \widehat{y}'(x))$$

for point x at which $\widehat{y}'$ is continuous.

§4.2 Simplest Problem in Calculus of Variations

Integrating-by-parts of (22) further implies that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b))\eta(b) - L_p(a, \widehat{y}(a), \widehat{y}'(a))\eta(a) = 0 \quad \forall\, \eta \in \mathcal{N}\,.$$

Choosing $\eta \in \mathcal{N}$ so that $\eta(a) = 1$ and $\eta(b) = 0$ (such η always exists), we find that

$$L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0\,.$$

Similarly, the choice of $\eta \in \mathcal{N}$ satisfying $\eta(a) = 0$ and $\eta(b) = 1$ shows that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b)) = 0\,.$$

Therefore,

(i) The Euler-Lagrange/Du Bois-Reymond equation holds.

(ii) $L_p(b, \widehat{y}(b), \widehat{y}'(b)) = L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0$ - this is called the ***natural boundary condition***.

§4.2 Simplest Problem in Calculus of Variations

Integrating-by-parts of (22) further implies that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b))\eta(b) - L_p(a, \widehat{y}(a), \widehat{y}'(a))\eta(a) = 0 \quad \forall\, \eta \in \mathcal{N}\,.$$

Choosing $\eta \in \mathcal{N}$ so that $\eta(a) = 1$ and $\eta(b) = 0$ (such η always exists), we find that

$$L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0\,.$$

Similarly, the choice of $\eta \in \mathcal{N}$ satisfying $\eta(a) = 0$ and $\eta(b) = 1$ shows that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b)) = 0\,.$$

Therefore,

(i) The Euler-Lagrange/Du Bois-Reymond equation holds.

(ii) $L_p(b, \widehat{y}(b), \widehat{y}'(b)) = L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0$ - this is called the ***natural boundary condition***.

§4.2 Simplest Problem in Calculus of Variations

Integrating-by-parts of (22) further implies that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b))\eta(b) - L_p(a, \widehat{y}(a), \widehat{y}'(a))\eta(a) = 0 \quad \forall \eta \in \mathcal{N} .$$

Choosing $\eta \in \mathcal{N}$ so that $\eta(a) = 1$ and $\eta(b) = 0$ (such η always exists), we find that

$$L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0 .$$

Similarly, the choice of $\eta \in \mathcal{N}$ satisfying $\eta(a) = 0$ and $\eta(b) = 1$ shows that

$$L_p(b, \widehat{y}(b), \widehat{y}'(b)) = 0 .$$

Therefore,

(i) The Euler-Lagrange/Du Bois-Reymond equation holds.

(ii) $L_p(b, \widehat{y}(b), \widehat{y}'(b)) = L_p(a, \widehat{y}(a), \widehat{y}'(a)) = 0$ - this is called the ***natural boundary condition***.

§4.2 Simplest Problem in Calculus of Variations

3. **Several dependent variables**: Let

$$\mathcal{A} = \Big\{ \boldsymbol{y} = (y_1, \cdots, y_n) : [a, b] \to \mathbb{R}^n \,\Big|\, y_j \in \mathcal{D}^1([a, b]; \mathbb{R}) \text{ for } 1 \leqslant j \leqslant n, \boldsymbol{y}(a) = \boldsymbol{A}_0, \boldsymbol{y}(b) = \boldsymbol{B}_0 \Big\}$$

or (when considering minimization problems with free ends)

$$\mathcal{A} = \Big\{ \boldsymbol{y} = (y_1, \cdots, y_n) : [a, b] \to \mathbb{R}^n \,\Big|\, y_j \in \mathcal{D}^1([a, b]; \mathbb{R}) \text{ for } 1 \leqslant j \leqslant n \Big\} \equiv \mathcal{D}^1([a, b]; \mathbb{R}^n),$$

and $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Consider the minimization problem

$$\inf_{\boldsymbol{y} \in \mathcal{A}} \int_a^b L(x, \boldsymbol{y}(x), \boldsymbol{y}'(x))\, dx.$$

§4.2 Simplest Problem in Calculus of Variations

Write $L = L(x, y_1, \cdots, y_n, p_1, \cdots, p_n)$. Then the Du Bois-Reymond equation is

$$\frac{d}{dx} L_{p_i}(x, \boldsymbol{y}(x), \boldsymbol{y}'(x)) = L_{y_i}(x, \boldsymbol{y}(x), \boldsymbol{y}'(x)) \quad \text{for} \quad 1 \leqslant i \leqslant n\,.$$

When considering free ends problem, natural boundary conditions

$$L_{p_i}(b, \widehat{\boldsymbol{y}}(b), \widehat{\boldsymbol{y}}'(b)) = L_{p_i}(a, \widehat{\boldsymbol{y}}(a), \widehat{\boldsymbol{y}}'(a)) = 0 \quad \text{for} \quad 1 \leqslant i \leqslant n$$

have to be imposed for the minimizer $\boldsymbol{y}$.

§4.2 Simplest Problem in Calculus of Variations

4. **Several independent variables**: Let $\Omega \subseteq \mathbb{R}^n$ be bounded open set, and $L : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ (here we write $L = L(x, y, p_1, \cdots, p_n)$) be continuous. Consider the minimization problem

$$\inf_{y \in \mathcal{A}} \int_\Omega L(x, y(x), \nabla y(x))\, dx\,,$$

where $\mathcal{A}$ could be

ⓘ $\mathcal{A} = \{y \in \mathcal{D}^1(\overline{\Omega}; \mathbb{R}) \,|\, y = f \text{ on } \partial\Omega\}$ (with corresponding $\mathcal{N} = \{\eta \in \mathcal{D}^1(\overline{\Omega}; \mathbb{R}) \,|\, \eta = 0 \text{ on } \partial\Omega\}$) when considering the fixed-end problem, or

ⓘⓘ $\mathcal{A} = \mathcal{D}^1(\overline{\Omega}; \mathbb{R})$ (with corresponding $\mathcal{N} = \mathcal{D}^1(\overline{\Omega}; \mathbb{R})$) when considering the free-end problem.

§4.2 Simplest Problem in Calculus of Variations

Define $J(\epsilon) = I(\hat{y} + \epsilon\eta)$, where $\hat{y} \in \mathcal{A}$ is a possible minimizer, $\eta \in \mathcal{N}$ and $\epsilon \in \mathbb{R}$. The ***weak form*** of the Euler-Lagrange equation is $J'(0) = 0$:

$$\int_\Omega \Big[L_y(x, \hat{y}(x), \nabla\hat{y}(x))\eta(x) + (\nabla_p L)(x, \hat{y}(x), \nabla\hat{y}(x)) \cdot \nabla_x \eta(x)\Big]\, dx = 0$$

for all $\eta \in \mathcal{N}$, where $\nabla_p L = \Big(\frac{\partial L}{\partial p_1}, \frac{\partial L}{\partial p_2}, \cdots, \frac{\partial L}{\partial p_n}\Big)$ is the gradient of L in p-variable. By the divergence theorem, the ***strong form*** of the Euler-Lagrange equation is

$$\mathrm{div}_x\big[(\nabla_p L)(x, \hat{y}(x), \nabla\hat{y}(x))\big] = L_y(x, \hat{y}(x), \nabla\hat{y}(x))\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (The minimal surface)

In this example we revisit Plateau's problem. Suppose that $\Omega \subseteq \mathbb{R}^2$ is a bounded set with boundary parameterized by $(x(t), y(t))$ for $t \in I$, and $C \subseteq \mathbb{R}^3$ is a closed curve parameterized by $(x(t), y(t), f(x(t), y(t)))$ for some given function f. We want to find a surface having C as its boundary with minimal surface area. Then the goal is to find a function u with the property that $u = f$ on $\partial\Omega$ that minimizes the functional

$$A(w) = \int_\Omega \sqrt{1 + |\nabla w|^2}\, dA\,.$$

Let $\varphi \in \mathscr{D}^1(\overline{\Omega}; \mathbb{R})$, and define

$$\delta A(u; \varphi) = \lim_{t \to 0} \frac{A(u + \epsilon\varphi) - A(u)}{\epsilon} = \int_\Omega \frac{\nabla u \cdot \nabla \varphi}{\sqrt{1 + |\nabla u|^2}}\, dA\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (The minimal surface)

In this example we revisit Plateau's problem. Suppose that $\Omega \subseteq \mathbb{R}^2$ is a bounded set with boundary parameterized by $(x(t), y(t))$ for $t \in I$, and $C \subseteq \mathbb{R}^3$ is a closed curve parameterized by $(x(t), y(t), f(x(t), y(t)))$ for some given function f. We want to find a surface having C as its boundary with minimal surface area. Then the goal is to find a function u with the property that $u = f$ on $\partial\Omega$ that minimizes the functional

$$A(w) = \int_\Omega \sqrt{1 + |\nabla w|^2}\, dA\,.$$

Let $\varphi \in \mathcal{D}^1(\overline{\Omega}; \mathbb{R})$, and define

$$\delta A(u; \varphi) = \lim_{t \to 0} \frac{A(u + \epsilon\varphi) - A(u)}{\epsilon} = \int_\Omega \frac{\nabla u \cdot \nabla \varphi}{\sqrt{1 + |\nabla u|^2}}\, dA\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (The minimal surface (cont.))

If u minimizes A, then $\delta A(u;\varphi) = 0$ for all $\varphi \in \mathcal{D}^1(\Omega;\mathbb{R})$ satisfying $\varphi = 0$ on $\partial\Omega$. Assuming that $u \in \mathcal{C}^2(\overline{\Omega};\mathbb{R})$, by the divergence theorem (or Green's Theorem in divergence form) we find that u satisfies

$$\operatorname{div}\Big(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\Big) = 0\,,$$

or expanding the bracket using the Leibnitz rule, we obtain the ***minimal surface equation***

$$(1+u_y^2)u_{xx} - 2u_xu_yu_{xy} + (1+u_x^2)u_{yy} = 0 \qquad \text{in} \quad \Omega\,.$$

§4.2 Simplest Problem in Calculus of Variations

5. **Non-affine admissible set**: We note that in Dido's problem the admissible set $\mathcal{A}$ is not an affine space (a translation of a vector space). In a minimization problem, the admissible set $\mathcal{A}$ in general is **not** an affine space so there is no obvious test function spaces $\mathcal{N}$ to work on. See the following two examples for deriving the weak form of the Euler-Lagrange equation for minimizers.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit)

We rephrase Dido's problem as finding a simple closed curve C enclosing a fixed number A of area with shortest perimeter. Let

$$\mathcal{A} = \Big\{ \boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} \in \mathcal{D}^1([0,1];\mathbb{R}^2) \,\Big|\, \boldsymbol{r}(0) = \boldsymbol{r}(1)\,,\ \int_0^1 \big[x(t)\dot{y}(t) - y(t)\dot{x}(t)\big]\,dt = 2\mathrm{A} \Big\}$$

and $I(\boldsymbol{r}) = \int_0^1 |\boldsymbol{r}'(t)|\,dt$. We would like to study the minimization problem $\inf\limits_{\boldsymbol{r}\in\mathcal{A}} I(\boldsymbol{r})$.

The difficulty of this particular formulation is that $\mathcal{A}$ is not an affine space so there is "no" corresponding test functions space to compute the first variation as before.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit)

We rephrase Dido's problem as finding a simple closed curve C enclosing a fixed number A of area with shortest perimeter. Let

$$\mathcal{A} = \Big\{ \boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} \in \mathscr{D}^1([0,1];\mathbb{R}^2) \,\Big|\, \boldsymbol{r}(0) = \boldsymbol{r}(1)\,,\ \int_0^1 \big[x(t)\dot{y}(t) - y(t)\dot{x}(t)\big]\,dt = 2\mathrm{A} \Big\}$$

and $I(\boldsymbol{r}) = \int_0^1 |\boldsymbol{r}'(t)|\,dt$. We would like to study the minimization problem $\inf\limits_{\boldsymbol{r}\in\mathcal{A}} I(\boldsymbol{r})$.

The difficulty of this particular formulation is that $\mathcal{A}$ is not an affine space so there is "no" corresponding test functions space to compute the first variation as before.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

To see how we derive the Euler-Lagrange equation for this minimization problem for a minimizer $\widehat{\boldsymbol{r}} = \widehat{x}\mathbf{i} + \widehat{y}\mathbf{j}$, we introduce a family of curves $\boldsymbol{r}(t;\epsilon) = x(t;\epsilon)\mathbf{i} + y(t;\epsilon)\mathbf{j} \in \mathcal{A}$, where $\epsilon \in \mathbb{R}$ is a parameter that will be passed to the limit, such that

1. $\boldsymbol{r}(t;0) = \widehat{\boldsymbol{r}}(t)$;
2. $\boldsymbol{r}(0;\epsilon) = \boldsymbol{r}(1;\epsilon)$;
3. $\boldsymbol{r}$ is also differentiable in ϵ.

By the fact that $\boldsymbol{r} \in \mathcal{A}$,

$$\int_0^1 \big[x(t;\epsilon)\dot{y}(t;\epsilon) - y(t;\epsilon)\dot{x}(t;\epsilon)\big]\,dt = 2\mathrm{A}\,;$$

thus

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_0^1 \big[x(t;\epsilon)\dot{y}(t;\epsilon) - y(t;\epsilon)\dot{x}(t;\epsilon)\big]\,dt = 0\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

To see how we derive the Euler-Lagrange equation for this minimization problem for a minimizer $\widehat{\boldsymbol{r}} = \widehat{x}\mathbf{i} + \widehat{y}\mathbf{j}$, we introduce a family of curves $\boldsymbol{r}(t;\epsilon) = x(t;\epsilon)\mathbf{i} + y(t;\epsilon)\mathbf{j} \in \mathcal{A}$, where $\epsilon \in \mathbb{R}$ is a parameter that will be passed to the limit, such that

1. $\boldsymbol{r}(t;0) = \widehat{\boldsymbol{r}}(t)$;
2. $\boldsymbol{r}(0;\epsilon) = \boldsymbol{r}(1;\epsilon)$;
3. $\boldsymbol{r}$ is also differentiable in ϵ.

By the fact that $\boldsymbol{r} \in \mathcal{A}$,

$$\int_0^1 \big[x(t;\epsilon)\dot{y}(t;\epsilon) - y(t;\epsilon)\dot{x}(t;\epsilon)\big]\,dt = 2\mathrm{A}\,;$$

thus

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_0^1 \big[x(t;\epsilon)\dot{y}(t;\epsilon) - y(t;\epsilon)\dot{x}(t;\epsilon)\big]\,dt = 0\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Denote $\delta\boldsymbol{r}(t) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\boldsymbol{r}(t;\epsilon) = \delta x(t)\mathbf{i} + \delta y(t)\mathbf{j}$. Then

$$\int_0^1 \big[(\delta x)\dot{\hat{y}} + \hat{x}(\dot{\delta y}) - (\delta y)\dot{\hat{x}} - \hat{y}(\dot{\delta x})\big]\,dt = 0\,.$$

For each possible minimizer $\hat{\boldsymbol{r}}$, the relation above induces a linear vector space

$$\mathcal{N}_{\hat{\boldsymbol{r}}} = \Big\{\delta\boldsymbol{r} = \delta x\mathbf{i} + \delta y\mathbf{j} \in \mathcal{C}^1([0,1];\mathbb{R}^2)\,\Big|\,\int_0^1 \big[\hat{x}(\dot{\delta y}) - \hat{y}(\dot{\delta x})\big]\,dt = 0\Big\}.$$

Now we look for a minimizer $\hat{\boldsymbol{r}} \in \mathcal{C}^2([0,1];\mathbb{R}^2)$. We note that if we are able to find a minimizer in $\mathcal{C}^2([0,1];\mathbb{R}^2)$ (thus a $\mathcal{C}^1$-minimizer), it must also be a minimizer in $\mathcal{D}^1([0,1];\mathbb{R}^2)$. Since $\hat{\boldsymbol{r}} \in \mathcal{C}^2([0,1];\mathbb{R}^2)$ is a minimizer, the function $J(\epsilon) \equiv I(\boldsymbol{r}(t;\epsilon))$ attains its minimum at $\epsilon = 0$.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Denote $\delta \boldsymbol{r}(t) = \dfrac{d}{d\epsilon}\Big|_{\epsilon=0} \boldsymbol{r}(t;\epsilon) = \delta x(t)\mathbf{i} + \delta y(t)\mathbf{j}$. Then

$$\int_0^1 \big[(\delta x)\dot{\widehat{y}} + \widehat{x}(\dot{\delta y}) - (\delta y)\dot{\widehat{x}} - \widehat{y}(\dot{\delta x})\big]\, dt = 0\,.$$

For each possible minimizer $\widehat{\boldsymbol{r}}$, the relation above induces a linear vector space

$$\mathcal{N}_{\widehat{\boldsymbol{r}}} = \Big\{ \delta \boldsymbol{r} = \delta x \mathbf{i} + \delta y \mathbf{j} \in \mathcal{C}^1([0,1];\mathbb{R}^2) \,\Big|\, \int_0^1 \big[\widehat{x}(\dot{\delta y}) - \widehat{y}(\dot{\delta x})\big]\, dt = 0 \Big\}.$$

Now we look for a minimizer $\widehat{\boldsymbol{r}} \in \mathcal{C}^2([0,1];\mathbb{R}^2)$. We note that if we are able to find a minimizer in $\mathcal{C}^2([0,1];\mathbb{R}^2)$ (thus a $\mathcal{C}^1$-minimizer), it must also be a minimizer in $\mathcal{D}^1([0,1];\mathbb{R}^2)$. Since $\widehat{\boldsymbol{r}} \in \mathcal{C}^2([0,1];\mathbb{R}^2)$ is a minimizer, the function $J(\epsilon) \equiv I(\boldsymbol{r}(t;\epsilon))$ attains its minimum at $\epsilon = 0$.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

This yields that $J'(0) = 0$ or more precisely,

$$\int_0^1 \frac{\hat{\boldsymbol{r}}'(t) \cdot (\delta \boldsymbol{r})'(t)}{|\hat{\boldsymbol{r}}'(t)|}\, dt = 0\,,$$

where we note that $\delta \boldsymbol{r} \in \mathcal{N}_{\hat{\boldsymbol{r}}}$. In other words, $\hat{\boldsymbol{r}}$ satisfies

$$\int_0^1 \frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|} \cdot (\delta \boldsymbol{r})'(t)\, dt = 0 \qquad \forall\, \delta \boldsymbol{r} \in \mathcal{N}_{\hat{\boldsymbol{r}}}\,,$$

and by the 4th lemma there exists $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ such that

$$\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|} = \big(\lambda_1 \hat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda_2 \hat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Since $\hat{\boldsymbol{r}} = (\hat{x}, \hat{y}) \in \mathcal{C}^2([0,1];\mathbb{R}^2)$, we differentiate the equation above and obtain that

$$\Big(\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}\Big)' = \lambda_1 \hat{y}'(t)\mathbf{i} + \lambda_2 \hat{x}'(t)\mathbf{j}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

This yields that $J'(0) = 0$ or more precisely,

$$\int_0^1 \frac{\widehat{\boldsymbol{r}}'(t) \cdot (\delta \boldsymbol{r})'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\, dt = 0\,,$$

where we note that $\delta \boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}}$. In other words, $\widehat{\boldsymbol{r}}$ satisfies

$$\int_0^1 \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} \cdot (\delta \boldsymbol{r})'(t)\, dt = 0 \qquad \forall\, \delta \boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}}\,,$$

and by the 4th lemma there exists $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ such that

$$\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} = \big(\lambda_1 \widehat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda_2 \widehat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Since $\widehat{\boldsymbol{r}} = (\widehat{x}, \widehat{y}) \in \mathcal{C}^2([0,1]; \mathbb{R}^2)$, we differentiate the equation above and obtain that

$$\Big(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\Big)' = \lambda_1 \widehat{y}'(t)\mathbf{i} + \lambda_2 \widehat{x}'(t)\mathbf{j}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

This yields that $J'(0) = 0$ or more precisely,

$$\int_0^1 \frac{\widehat{\boldsymbol{r}}'(t) \cdot (\delta \boldsymbol{r})'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\, dt = 0\,,$$

where we note that $\delta \boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}}$. In other words, $\widehat{\boldsymbol{r}}$ satisfies

$$\int_0^1 \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} \cdot (\delta \boldsymbol{r})'(t)\, dt = 0 \qquad \forall\, \delta \boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}}\,,$$

and by the 4th lemma there exists $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ such that

$$\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} = \big(\lambda_1 \widehat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda_2 \widehat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Since $\widehat{\boldsymbol{r}} = (\widehat{x}, \widehat{y}) \in \mathscr{C}^2([0,1];\mathbb{R}^2)$, we differentiate the equation above and obtain that

$$\left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' = \lambda_1 \widehat{y}'(t)\mathbf{i} + \lambda_2 \widehat{x}'(t)\mathbf{j}\,.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, taking the inner product of the equation above with the unit tangent vector $\dfrac{\widehat{\boldsymbol{r}}'}{|\widehat{\boldsymbol{r}}'|}$, we find that for all $t \in [0,1]$,

$$0 = \left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right) \cdot \left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' = \big(\lambda_1 \widehat{y}'(t)\mathbf{i} + \lambda_2 \widehat{x}'(t)\mathbf{j}\big) \cdot \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}$$

$$= (\lambda_2 + \lambda_1) \frac{\widehat{x}'(t)\,\widehat{y}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}$$

which implies that $\lambda_2 = -\lambda_1 = \lambda$ (for otherwise $\widehat{x}'\widehat{y}' = 0$ which shows that the trajectory is a straight line); thus

$$\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} = \big(-\lambda \widehat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda \widehat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Note that $\lambda \neq 0$ for otherwise the unit tangent vector is constant which implies that $\widehat{\boldsymbol{r}}$ is a parametrization of a straight line.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, taking the inner product of the equation above with the unit tangent vector $\frac{\widehat{\boldsymbol{r}}'}{|\widehat{\boldsymbol{r}}'|}$, we find that for all $t \in [0,1]$,

$$0 = \Big(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\Big) \cdot \Big(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\Big)' = \big(\lambda_1 \widehat{y}'(t)\mathbf{i} + \lambda_2 \widehat{x}'(t)\mathbf{j}\big) \cdot \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}$$

$$= (\lambda_2 + \lambda_1)\frac{\widehat{x}'(t)\,\widehat{y}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}$$

which implies that $\lambda_2 = -\lambda_1 = \lambda$ (for otherwise $\widehat{x}'\widehat{y}' = 0$ which shows that the trajectory is a straight line); thus

$$\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} = \big(-\lambda\widehat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda\widehat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Note that $\lambda \neq 0$ for otherwise the unit tangent vector is constant which implies that $\widehat{\boldsymbol{r}}$ is a parametrization of a straight line.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, taking the inner product of the equation above with the unit tangent vector $\dfrac{\hat{\boldsymbol{r}}'}{|\hat{\boldsymbol{r}}'|}$, we find that for all $t \in [0,1]$,

$$0 = \Big(\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}\Big) \cdot \Big(\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}\Big)' = \big(\lambda_1 \hat{y}'(t)\mathbf{i} + \lambda_2 \hat{x}'(t)\mathbf{j}\big) \cdot \frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}$$
$$= (\lambda_2 + \lambda_1)\frac{\hat{x}'(t)\,\hat{y}'(t)}{|\hat{\boldsymbol{r}}'(t)|}$$

which implies that $\lambda_2 = -\lambda_1 = \lambda$ (for otherwise $\hat{x}'\hat{y}' = 0$ which shows that the trajectory is a straight line); thus

$$\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|} = \big(-\lambda\hat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda\hat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Note that $\lambda \neq 0$ for otherwise the unit tangent vector is constant which implies that $\hat{\boldsymbol{r}}$ is a parametrization of a straight line.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, taking the inner product of the equation above with the unit tangent vector $\dfrac{\hat{\boldsymbol{r}}'}{|\hat{\boldsymbol{r}}'|}$, we find that for all $t \in [0,1]$,

$$0 = \Big(\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}\Big) \cdot \Big(\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}\Big)' = \big(\lambda_1 \hat{y}'(t)\mathbf{i} + \lambda_2 \hat{x}'(t)\mathbf{j}\big) \cdot \frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|}$$

$$= (\lambda_2 + \lambda_1)\frac{\hat{x}'(t)\,\hat{y}'(t)}{|\hat{\boldsymbol{r}}'(t)|}$$

which implies that $\lambda_2 = -\lambda_1 = \lambda$ (for otherwise $\hat{x}'\hat{y}' = 0$ which shows that the trajectory is a straight line); thus

$$\frac{\hat{\boldsymbol{r}}'(t)}{|\hat{\boldsymbol{r}}'(t)|} = \big(-\lambda\hat{y}(t) + \mu_1\big)\mathbf{i} + \big(\lambda\hat{x}(t) + \mu_2\big)\mathbf{j}\,.$$

Note that $\lambda \neq 0$ for otherwise the unit tangent vector is constant which implies that $\hat{\boldsymbol{r}}$ is a parametrization of a straight line.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, with $\widetilde{\boldsymbol{r}}$ denoting the vector

$$\widetilde{\boldsymbol{r}}(t) = \widetilde{x}(t)\mathbf{i} + \widetilde{y}(t)\mathbf{j} \equiv \Big(\widehat{x}(t) + \frac{\mu_2}{\lambda}\Big)\mathbf{i} + \Big(\widehat{y}(t) - \frac{\mu_1}{\lambda}\Big)\mathbf{j},$$

we have

$$\frac{\widetilde{\boldsymbol{r}}'(t)}{|\widetilde{\boldsymbol{r}}'(t)|} = -\lambda\widetilde{y}(t)\mathbf{i} + \lambda\widetilde{x}(t)\mathbf{j}.$$

Finally, taking the inner product of the equation above with the (position) vector $\widetilde{\boldsymbol{r}}$, we conclude that

$$\frac{d}{dt}\big|\widetilde{\boldsymbol{r}}(t)\big|^2 = 0.$$

Therefore, the closed curve having fixed length and enclosing the largest area must be a circle.

§4.2 Simplest Problem in Calculus of Variations

Example (Isoperimetric Inequality - revisit (cont.))

Therefore, with $\widetilde{\boldsymbol{r}}$ denoting the vector

$$\widetilde{\boldsymbol{r}}(t) = \widetilde{x}(t)\,\mathbf{i} + \widetilde{y}(t)\,\mathbf{j} \equiv \Big(\widehat{x}(t) + \frac{\mu_2}{\lambda}\Big)\mathbf{i} + \Big(\widehat{y}(t) - \frac{\mu_1}{\lambda}\Big)\mathbf{j}\,,$$

we have

$$\frac{\widetilde{\boldsymbol{r}}'(t)}{|\widetilde{\boldsymbol{r}}'(t)|} = -\lambda\widetilde{y}(t)\,\mathbf{i} + \lambda\widetilde{x}(t)\,\mathbf{j}\,.$$

Finally, taking the inner product of the equation above with the (position) vector $\widetilde{\boldsymbol{r}}$, we conclude that

$$\frac{d}{dt}\big|\widetilde{\boldsymbol{r}}(t)\big|^2 = 0\,.$$

Therefore, the closed curve having fixed length and enclosing the largest area must be a circle.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere)

Consider finding the shortest path on the unit sphere connecting two points A_0 and B_0 (on the same sphere). In other words, we are interested in the minimization problem

$$\inf_{\boldsymbol{r}\in\mathcal{A}} \int_0^1 \left|\boldsymbol{r}'(t)\right| dt\,,$$

where $\mathcal{A} = \big\{\boldsymbol{r} \in \mathcal{D}^1([0,1];\mathbb{R}^3) \,\big|\, \boldsymbol{r}(0) = A_0, \boldsymbol{r}(1) = B_0, |\boldsymbol{r}(t)| = 1 \;\forall\, t\big\}$.

Similar to the previous example, we introduce a family of curves $\boldsymbol{r}(t;\epsilon)$, where $\epsilon \in \mathbb{R}$ is a parameter that will be passed to the limit, such that

① $\boldsymbol{r}(t;0) = \widehat{\boldsymbol{r}}(t)$; ② $\boldsymbol{r}(0;\epsilon) = A_0$; ③ $\boldsymbol{r}(1;\epsilon) = B_0$; ④ $\boldsymbol{r}$ is also differentiable in ϵ,

where $\widehat{\boldsymbol{r}}$ gives the shortest path connecting A_0 and B_0.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere)

Consider finding the shortest path on the unit sphere connecting two points A_0 and B_0 (on the same sphere). In other words, we are interested in the minimization problem

$$\inf_{\boldsymbol{r}\in\mathcal{A}} \int_0^1 \left|\boldsymbol{r}'(t)\right| dt\,,$$

where $\mathcal{A} = \left\{\boldsymbol{r}\in \mathscr{D}^1([0,1];\mathbb{R}^3) \,\middle|\, \boldsymbol{r}(0) = A_0, \boldsymbol{r}(1) = B_0, |\boldsymbol{r}(t)| = 1\ \forall\, t\right\}$.

Similar to the previous example, we introduce a family of curves $\boldsymbol{r}(t;\epsilon)$, where $\epsilon\in\mathbb{R}$ is a parameter that will be passed to the limit, such that

① $\boldsymbol{r}(t;0) = \widehat{\boldsymbol{r}}(t)$; ② $\boldsymbol{r}(0;\epsilon) = A_0$; ③ $\boldsymbol{r}(1;\epsilon) = B_0$; ④ $\boldsymbol{r}$ is also differentiable in ϵ,

where $\widehat{\boldsymbol{r}}$ gives the shortest path connecting A_0 and B_0.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Since the minimizer $\widehat{\boldsymbol{r}}$ satisfies that $\widehat{\boldsymbol{r}} \in \mathcal{A}$ (that is, $|\widehat{\boldsymbol{r}}| = 1$), we find that $\widehat{\boldsymbol{r}}'(t) \cdot \widehat{\boldsymbol{r}}(t) = 0$ whenever $\widehat{\boldsymbol{r}}'(t)$ exists. Therefore, we can assume that

$\widehat{\boldsymbol{r}}(t), \widehat{\boldsymbol{r}}'(t), (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t)$ are linearly independent if $\widehat{\boldsymbol{r}}'(t) \neq \mathbf{0}$.

Denote $\delta \boldsymbol{r}(t) = \dfrac{d}{d\epsilon}\Big|_{\epsilon=0} \boldsymbol{r}(t;\epsilon)$. Then the fact that $\boldsymbol{r} \in \mathcal{A}$ again implies that $\delta\boldsymbol{r} \cdot \widehat{\boldsymbol{r}} = 0$; thus we shall introduce $\mathcal{N}_{\widehat{\boldsymbol{r}}}$ as

$$\mathcal{N}_{\widehat{\boldsymbol{r}}} = \Big\{ \delta\boldsymbol{r} \in \mathscr{C}^1([0,1];\mathbb{R}^2) \,\Big|\, \widehat{\boldsymbol{r}}(t) \cdot \delta\boldsymbol{r}(t) = 0 \text{ for all } t \in [0,1] \Big\};$$

thus we find that

$$\mathcal{N}_{\widehat{\boldsymbol{r}}} = \operatorname{span}\big(\widehat{\boldsymbol{r}}', \widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}\big) = \big\{ a\widehat{\boldsymbol{r}}' + b(\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}) \,\big|\, a, b \in \mathbb{R} \big\}.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Since the minimizer $\widehat{\boldsymbol{r}}$ satisfies that $\widehat{\boldsymbol{r}} \in \mathcal{A}$ (that is, $|\widehat{\boldsymbol{r}}| = 1$), we find that $\widehat{\boldsymbol{r}}'(t) \cdot \widehat{\boldsymbol{r}}(t) = 0$ whenever $\widehat{\boldsymbol{r}}'(t)$ exists. Therefore, we can assume that

$$\widehat{\boldsymbol{r}}(t), \widehat{\boldsymbol{r}}'(t), (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t) \text{ are linearly independent if } \widehat{\boldsymbol{r}}'(t) \neq \mathbf{0}.$$

Denote $\delta\boldsymbol{r}(t) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \boldsymbol{r}(t;\epsilon)$. Then the fact that $\boldsymbol{r} \in \mathcal{A}$ again implies that $\delta\boldsymbol{r} \cdot \widehat{\boldsymbol{r}} = 0$; thus we shall introduce $\mathcal{N}_{\widehat{\boldsymbol{r}}}$ as

$$\mathcal{N}_{\widehat{\boldsymbol{r}}} = \Big\{\delta\boldsymbol{r} \in \mathcal{C}^1([0,1];\mathbb{R}^2) \,\Big|\, \widehat{\boldsymbol{r}}(t) \cdot \delta\boldsymbol{r}(t) = 0 \text{ for all } t \in [0,1]\Big\};$$

thus we find that

$$\mathcal{N}_{\widehat{\boldsymbol{r}}} = \operatorname{span}\big(\widehat{\boldsymbol{r}}', \widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}\big) = \big\{a\widehat{\boldsymbol{r}}' + b(\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}) \,\big|\, a, b \in \mathbb{R}\big\}.$$

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Now suppose that $\widehat{\boldsymbol{r}} \in \mathcal{C}^2([0,1];\mathbb{R}^3)$. Similar to the previous example, we obtain that

$$0 = \frac{d}{d\epsilon}\Big|_{\epsilon=0} \int_0^1 \big|\boldsymbol{r}'(t;\epsilon)\big|\,dt = \int_0^1 \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} \cdot (\delta\boldsymbol{r})'(t)\,dt \qquad \forall\, \delta\boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}},$$

and integrating by parts further shows that for $\delta\boldsymbol{r} \in \mathcal{N}_{\widehat{\boldsymbol{r}}}$,

$$\begin{aligned} 0 &= \frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|} \cdot (\delta\boldsymbol{r})(t)\Big|_{t=0}^{t=1} - \int_0^1 \Big(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\Big)' \cdot (\delta\boldsymbol{r})(t)\,dt \\ &= -\int_0^1 \Big(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\Big)' \cdot (\delta\boldsymbol{r})(t)\,dt\,, \end{aligned}$$

where we have use the fact that $(\delta\boldsymbol{r})(0) = (\delta\boldsymbol{r})(1) = \mathbf{0}$ to eliminate the boundary contributions.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Since $\left(\dfrac{\widehat{\boldsymbol{r}}'}{|\widehat{\boldsymbol{r}}'|}\right)' \cdot \widehat{\boldsymbol{r}}' = 0$, we conclude from the structure of $\mathcal{N}_{\widehat{\boldsymbol{r}}}$ that

$$\int_0^1 b(t)\left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t)\, dt = 0 \qquad \forall\, b \in \mathcal{C}([0,1];\mathbb{R})$$

which (by the first lemma) shows that

$$\left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t) = 0 \qquad \forall\, t \in [0,1]\,.$$

By the fact that $\widehat{\boldsymbol{r}}' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}) = 0$, the identity above further shows that

$$\widehat{\boldsymbol{r}}''(t) \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t) = 0 \qquad \forall\, t \in [0,1]\,. \tag{23}$$

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Since $\left(\dfrac{\widehat{\boldsymbol{r}}'}{|\widehat{\boldsymbol{r}}'|}\right)' \cdot \widehat{\boldsymbol{r}}' = 0$, we conclude from the structure of $\mathcal{N}_{\widehat{\boldsymbol{r}}}$ that

$$\int_0^1 b(t)\left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t)\, dt = 0 \qquad \forall\, b \in \mathcal{C}([0,1];\mathbb{R})$$

which (by the first lemma) shows that

$$\left(\frac{\widehat{\boldsymbol{r}}'(t)}{|\widehat{\boldsymbol{r}}'(t)|}\right)' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t) = 0 \qquad \forall\, t \in [0,1]\,.$$

By the fact that $\widehat{\boldsymbol{r}}' \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}) = 0$, the identity above further shows that

$$\widehat{\boldsymbol{r}}''(t) \cdot (\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})(t) = 0 \qquad \forall\, t \in [0,1]\,. \tag{23}$$

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Now suppose that the parametrization of the shortest path satisfies that $|\hat{\boldsymbol{r}}'(t)| = $ constant; that is, the motion along the shortest path has constant speed. Then $\hat{\boldsymbol{r}}'(t)\cdot\hat{\boldsymbol{r}}''(t) = 0$ for all $t \in [0,1]$; thus

$$\hat{\boldsymbol{r}}'' = c\hat{\boldsymbol{r}} + d(\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}}) \quad \text{for some functions } c \text{ and } d \text{ of } t.$$

Identity (??) further shows that $d = 0$; thus $\hat{\boldsymbol{r}}'' = c\hat{\boldsymbol{r}}$ so that

$$(\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}})' = \hat{\boldsymbol{r}}'' \times \hat{\boldsymbol{r}} = c\hat{\boldsymbol{r}} \times \hat{\boldsymbol{r}} = \boldsymbol{0}.$$

As a consequence, $\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}}$ is a constant vector $\boldsymbol{c}$ which further implies that $\hat{\boldsymbol{r}} \cdot \boldsymbol{c} = 0$. Therefore, the trajectory lies on a plane passing through the origin which shows that the shortest path connecting two points on the sphere must be part of a great circle.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Now suppose that the parametrization of the shortest path satisfies that $|\hat{\boldsymbol{r}}'(t)| = $ constant; that is, the motion along the shortest path has constant speed. Then $\hat{\boldsymbol{r}}'(t) \cdot \hat{\boldsymbol{r}}''(t) = 0$ for all $t \in [0,1]$; thus

$$\hat{\boldsymbol{r}}'' = c\hat{\boldsymbol{r}} + d(\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}}) \quad \text{for some functions } c \text{ and } d \text{ of } t\,.$$

Identity (**??**) further shows that $d = 0$; thus $\hat{\boldsymbol{r}}'' = c\hat{\boldsymbol{r}}$ so that

$$(\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}})' = \hat{\boldsymbol{r}}'' \times \hat{\boldsymbol{r}} = c\hat{\boldsymbol{r}} \times \hat{\boldsymbol{r}} = \mathbf{0}\,.$$

As a consequence, $\hat{\boldsymbol{r}}' \times \hat{\boldsymbol{r}}$ is a constant vector $\boldsymbol{c}$ which further implies that $\hat{\boldsymbol{r}} \cdot \boldsymbol{c} = 0$. Therefore, the trajectory lies on a plane passing through the origin which shows that the shortest path connecting two points on the sphere must be part of a great circle.

§4.2 Simplest Problem in Calculus of Variations

Example (Geodesic on unit sphere (cont.))

Now suppose that the parametrization of the shortest path satisfies that $|\widehat{\boldsymbol{r}}'(t)| = $ constant; that is, the motion along the shortest path has constant speed. Then $\widehat{\boldsymbol{r}}'(t) \cdot \widehat{\boldsymbol{r}}''(t) = 0$ for all $t \in [0,1]$; thus

$$\widehat{\boldsymbol{r}}'' = c\widehat{\boldsymbol{r}} + d(\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}) \quad \text{for some functions } c \text{ and } d \text{ of } t\,.$$

Identity (**??**) further shows that $d = 0$; thus $\widehat{\boldsymbol{r}}'' = c\widehat{\boldsymbol{r}}$ so that

$$(\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}})' = \widehat{\boldsymbol{r}}'' \times \widehat{\boldsymbol{r}} = c\widehat{\boldsymbol{r}} \times \widehat{\boldsymbol{r}} = \mathbf{0}\,.$$

As a consequence, $\widehat{\boldsymbol{r}}' \times \widehat{\boldsymbol{r}}$ is a constant vector $\boldsymbol{c}$ which further implies that $\widehat{\boldsymbol{r}} \cdot \boldsymbol{c} = 0$. Therefore, the trajectory lies on a plane passing through the origin which shows that the shortest path connecting two points on the sphere must be part of a great circle.

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.
If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.