2011년 2학기 전산선박설계 강의자료
(Computer Aided Ship Design Lecture Note)

서울대학교 조선해양공학과

선박설계자동화연구실

이규열

CONTENTS

Part I. Optimization Method

Part II. Curve and Surface Modeling

Part III. Finite Element Method

2011년 2학기 전산선박설계 강의자료
(Computer Aided Ship Design Lecture Note)

Part I. Optimization Method

서울대학교 조선해양공학과

선박설계자동화연구실

이규열

CONTENTS

Computer Aided Ship Design

Part I. Optimization Method

- Ch. 1 Overview of Optimal Design

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.1 Overview of Optimal Design

Indeterminate Equation

Variable: x_1, x_2, x_3

Equation: $x_1 + x_2 + x_3 = 3$

✓ Number of variables: 3

✓ Number of equations: 1

Because the number of variables is larger than that of equations, these equation form an **indeterminate system**.

Solution for the indeterminate equation

:We assume
two unknown variables

Number of variables(3) - Number of equations(1)

Example) assume that $x_1 = 1, x_2 = 0$

→ $x_3 = 2$

Equation of straight line

$y = a_0 + a_1 x$ Where a_0, a_1: are given

✓ Number of variables: 2 x, y

✓ Number of equations: 1

☞ We can get the value of y by assuming x.

Find intersection point (x^*, y^*) of two straight lines

$$y = a_0 + a_1 x$$

Where a_0, a_1, b_0, b_1 are given

$$y = b_0 + b_1 x$$

✓ Number of variables: 2 x, y

✓ Number of equations: 2

Indeterminate Equation and Solution

Determinate equation

Variable: x_1, x_2, x_3

Equation: $f_1(x_1, x_2, x_3)=0$

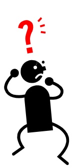
$f_2(x_1, x_2, x_3)=0$

$f_3(x_1, x_2, x_3)=0$

If f_1, f_2, and f_3 are linearly independent,

- ✓ Number of variables : 3
- ✓ Number of equations : 3

Since the number of equations is equal to that of variables, this problem can be solved.



What happens if $2 \times f_3 = f_2$?

f_2 and f_3 are linearly dependent.

Since the number of equations, which are linearly independent, is less than that of variables, these equations form an indeterminate systems.

Indeterminate equation

Variable: x_1, x_2, x_3

Equation: $f_1(x_1, x_2, x_3)=0$

$f_2(x_1, x_2, x_3)=0$

$f_3(x_1, x_2, x_3)=0$

If f_1 and f_2 are only linearly independent, ther

- ✓ Number of variables : 3
- ✓ Number of equations : 2

Since the number of equations is less than that of variables, one equation should be added to solve this problem.

Added Equation	Solution
$f_4^1 = 0$	(x_1^1, x_2^1, x_3^1)
$f_4^2 = 0$	(x_1^2, x_2^2, x_3^2)
:	

We can get many sets of solution by assuming different equations.
→ Indeterminate equation

We need a certain criteria to determine the proper solution. By adding the criteria, this problem can be formulated as an optimization problem.

Design

Esthetic* Design

Find(Design variables)

- Size, material, color, etc.

Constraints

- There are some constraints, but it is difficult to formulate them.
- By using the sense of designer, the constraints are satisfied.

Objective function(Criteria to determine the proper design variables)

- Preference, cost, etc.
- It is difficult to formulate the objective function.

Mathematical Model for Determination of the Main Dimensions(L,B,D,T,C_B) of a Ship(Summary)

- "Conceptual Ship Design Equation"

Find(Design variables) L, B, D, C_B
length breadth depth block coefficient

Given(Owner's requirement) DWT, V_{H_req}, $T_{max}(=T)$, V
deadweight, Required cargo hold capacity, maximum draft, ship speed

Physical constraint

→ Displacement - Weight equilibrium (Weight equation) – Equality constraint

$$L \cdot B \cdot T \cdot C_B \cdot \rho_{sw} \cdot C_\alpha = DWT_{given} + LWT(L,B,D,C_B)$$

$$= DWT_{given} + C_s \cdot L^{1.6}(B+D) + C_o \cdot L \cdot B$$

$$+ C_{power} \cdot (L \cdot B \cdot T \cdot C_B)^{2/3} \cdot V^3 \quad \cdots (2.3)$$

Economical constraints(Owner's requirements)

→ Required cargo hold capacity (Volume equation) - Equality constraint

$$V_{H\_req} = C_H \cdot L \cdot B \cdot D \cdots (3.1)$$

- DFOC(Daily Fuel Oil Consumption)
  : It is related with the resistance and propulsion.
- Delivery date
  : It is related with the shipbuilding process.

Regulatory constraint

→ Freeboard regulation(1966 ICLL) - Inequality constraint

$$D \geq T + C_{FB} \cdot D \cdots (4)$$

Objective Function(Criteria to determine the proper main dimensions)

$$Building\ Cost = C_{PS} \cdot C_s \cdot L^{1.6}(B+D) + C_{PO} \cdot C_o \cdot L \cdot B + C_{PM} \cdot C_{power} \cdot (L \cdot B \cdot T \cdot C_B)^{2/3} \cdot V^3$$

4 variables(L,B,D,C_B), 2 equality constraints,((2.3),(3.1)), 1 inequality constraint((4)) ➡ Optimization problem:

Determination of the Optimal Main Dimensions of a Ship

Optimization problem

➡ Minimize/maximize the objective function with the constraints on the design variables

Find(Design variables)

x_1, x_2, x_3, x_4

Equality constraint

$h(x_1, x_2, x_3, x_4) = 0$

Inequality constraint

$g(x_1, x_2, x_3, x_4) \le 0$

Objective function

$f\ (x_1, x_2, x_3, x_4)$

Engineering Design

Find(Design variables)

$L(=x_1), B(=x_2), D(=x_3), C_B(=x_4)$

length, breadth, depth, block coefficient

Weight equation

$$L \cdot B \cdot T \cdot C_B \cdot \rho_{sw} \cdot C_\alpha = DWT_{given} + LWT(L, B, D, C_B)$$

$$x_1 \cdot x_2 \cdot x_4 \cdot C_1 = C_2 + h'(x_1, x_2, x_3, x_4)$$

$$x_1 \cdot x_2 \cdot x_4 \cdot C_1 - C_2 - h'(x_1, x_2, x_3, x_4) = h(x_1, x_2, x_3, x_4) = 0$$

Objective Function(Criteria to determine the proper main dimensions)

$$Building\ Cost = C_{PS} \cdot C_s \cdot L^{1.6}(B + D) + C_{PO} \cdot C_o \cdot L \cdot B + C_{PM} \cdot C_{ma} \cdot NMCR$$

$$f(x_1, x_2, x_3, x_4) = C_3 \cdot x_1^{1.6} \cdot (x_2 + x_3) + C_4 \cdot x_1 \cdot x_2 + C_5$$

- $T, C_\alpha, \rho_{sw}, DWT_{given}, C_{PS}, C_s, C_{PO}, C_o, C_{PM}, C_{ma}, NMCR$ *are* Given

Characteristics of the constraint

✓ Physical constraints are usually formulated as the equality constraints.
(Example of ship design: Weight equation)

✓ Economical constraints, regulatory constraints, and constraints related with politics and culture are formulated as the inequality constraints.
(Example of ship design : Required cargo hold capacity(Volume equation), Freeboard regulation(1966 ICLL))

Classification of Optimization Problems and Optimization Methods

<table>
<tr><td rowspan="2"></td><td colspan="2">Unconstrained optimization problem</td><td colspan="3">Constrained optimization problem</td></tr>
<tr><td>Linear</td><td>Nonlinear</td><td>Linear</td><td colspan="2">Nonlinear</td></tr>
<tr><td>Objective function (example)</td><td>Minimize $f(\mathbf{x})$
$f(\mathbf{x}) = x_1 + 2x_2$</td><td>Minimize $f(\mathbf{x})$
$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$</td><td>Minimize $f(\mathbf{x})$
$f(\mathbf{x}) = x_1 + 2x_2$</td><td>Minimize $f(\mathbf{x})$
$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$</td><td>Minimize $f(\mathbf{x})$
$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$</td></tr>
<tr><td>Constraint (example)</td><td>None</td><td>None</td><td>$h(\mathbf{x}) = x_1 + 5x_2 = 0$
$g(\mathbf{x}) = -x_1 \le 0$</td><td>$h(\mathbf{x}) = x_1 + 5x_2 = 0$
$g(\mathbf{x}) = -x_1 \le 0$</td><td>$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$
$g_2(\mathbf{x}) = -x_1 \le 0$</td></tr>
<tr><td rowspan="3">Optimization methods -continuous value</td><td colspan="2" rowspan="3">① Direct search method
- Hooke & Jeeves method
- Nelder & Mead method

② Gradient method
- Steepest descent method
- Conjugate gradient method
- Newton method
- Davidon-Fletcher-Powell(DFP) method
- Broyden-Fletcher-Goldfarb-Shanno(BFGS) method</td><td>Linear programming (LP) method is usually used.</td><td colspan="2">Penalty Function Method: Converting the constrained optimization problem to the unconstrained optimization problem by using the penalty function, the problem can be solved using unconstrained optimization method.</td></tr>
<tr><td>Simplex Method
(Linear programming)</td><td rowspan="2">Quadratic programming(QP) method</td><td>SLP(Sequential Linear Programming)
First, linearize the nonlinear problem and then obtain the solution to this linear approximation problem using the linear programming method.
And ,then, repeat the linearization</td></tr>
<tr><td></td><td>Sequential Quadratic Programming(SQP) method
First, approximate a quadratic objective function and linear constraints, find the search direction and then obtain the solution to this quadratic programming problem in this direction. And ,then, repeat the approximation</td></tr>
<tr><td>Optimization methods - discrete value</td><td colspan="5">Integer programming: ① Cut algorithm ② Enumeration algorithm ③ Constructive algorithm</td></tr>
<tr><td>Heuristic optimization</td><td colspan="5">Genetic algorithm(GA), Ant algorithm, Simulated annealing, etc</td></tr>
</table>

Naval Architecture & Ocean Engineering

Computer Aided Ship Design

Part I. Optimization Method

- Ch. 2 Problem Statement of Optimal Design

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.2 Problem Statement of Optimal Design

2.1 Components of Optimal Design Problem

2.2 Formulation of Optimal Design Problem

2.3 Classification of Optimization Problems

2.4 Classification of Optimization Methods

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

2.1 Components of Optimal Design Problem(1)

☑ Design variable

- A set of variables that describes the system such as size and position, etc.
- It is also called 'Free variable' or 'Independent variable'.
- Dependent Variable

  : A variable that is dependent on the design variable(independent variable)

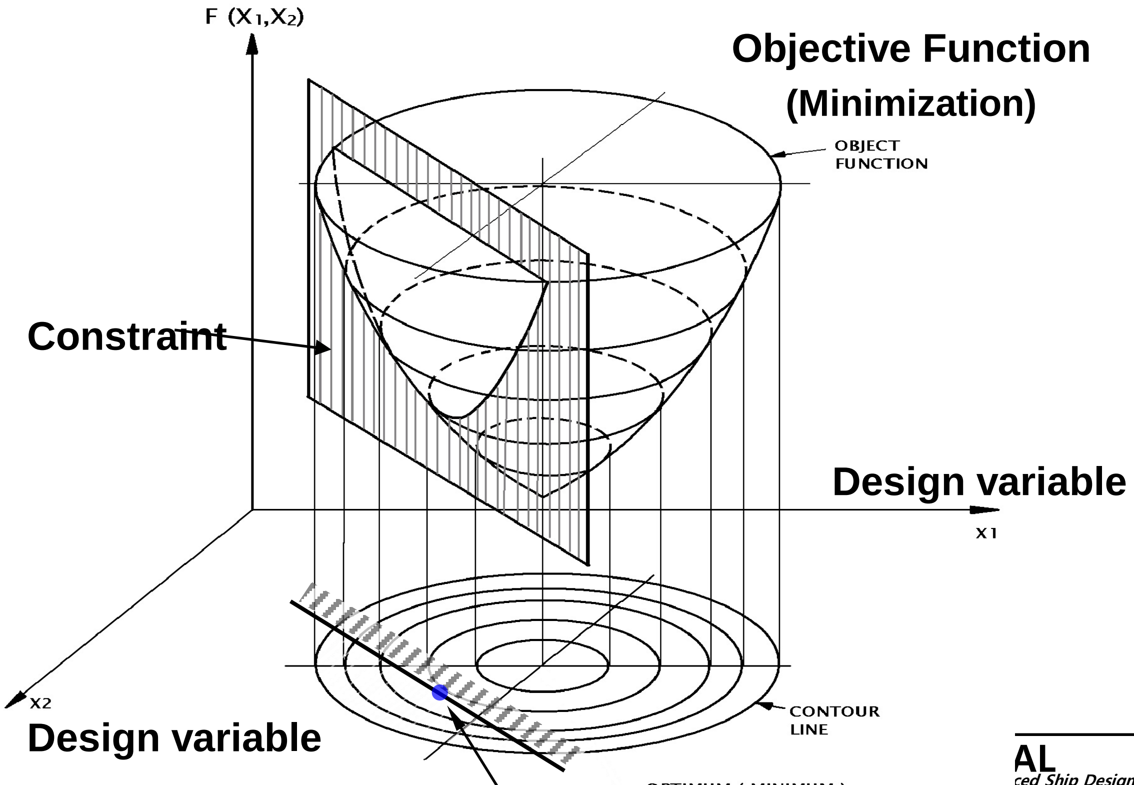
☑ Constraint

- A certain set of specified requirements and restrictions placed on a design
- Inequality Constraint, Equality Constraint

2.1 Components of Optimal Design Problem(2)

☑ Objective function

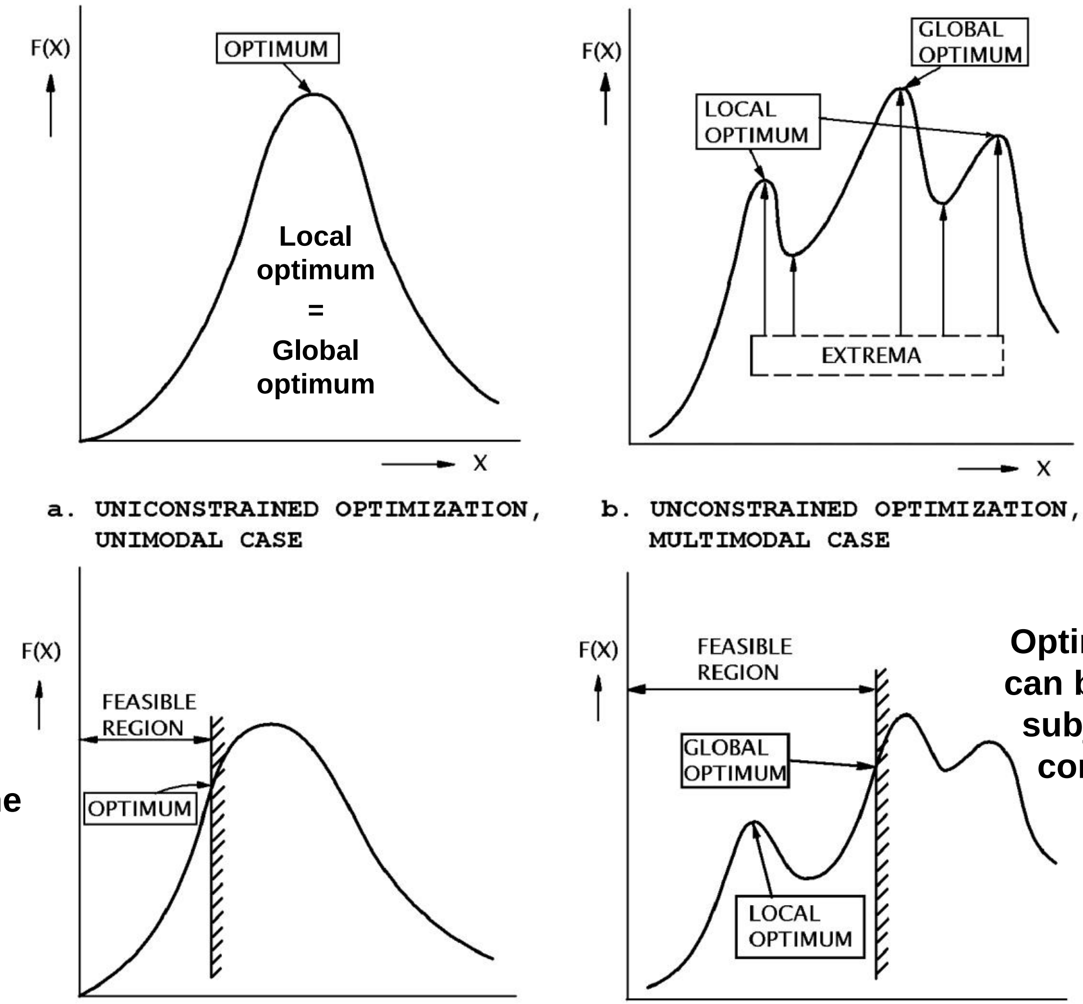
- A criteria to compare the different design and determine the proper design such as cost, profit, weight, etc.
- It is a function of the design variables.



2.1 Components of Optimal Design Problem(3)

Determination of the optimal design considering the objective function(maximization) and constraints

The region satisfying the constraint



a. UNICONSTRAINED OPTIMIZATION, UNIMODAL CASE

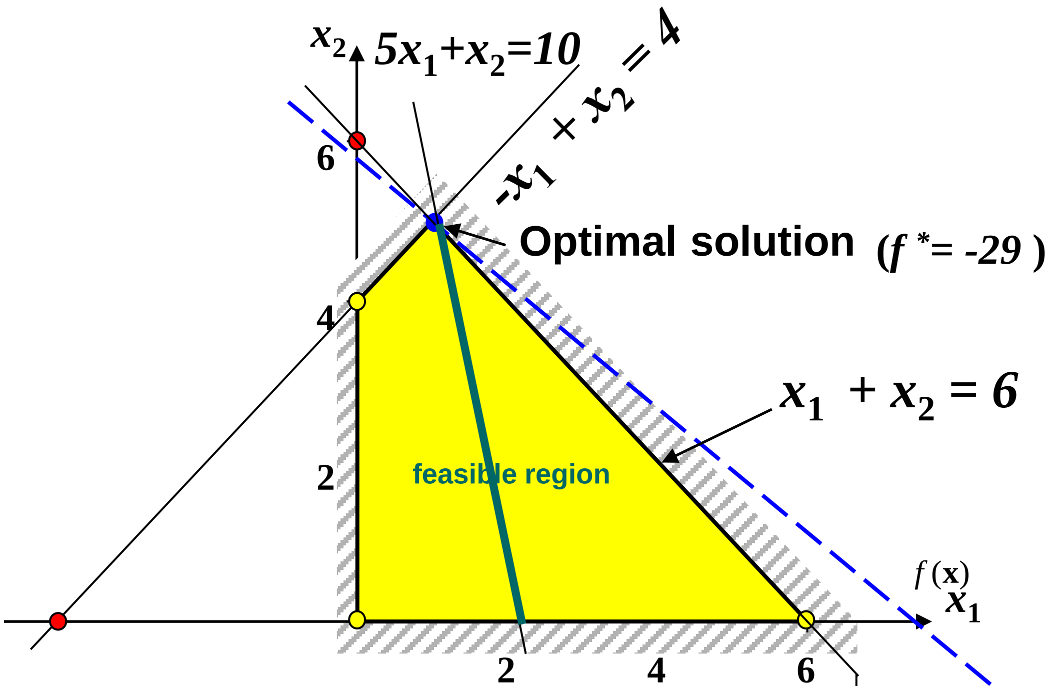
b. UNCONSTRAINED OPTIMIZATION, MULTIMODAL CASE

c. CONSTRAINED OPTIMIZATION, UNIMODAL CASE

d. CONSTRAINED OPTIMIZATION, MULTIMODAL CASE

Optimal design can be changed subject to the constraints.

2.2 Formulation of Optimal Design Problem



Minimize: $f=-4x_1-5x_2$	**Minimize:** $f(\mathbf{x})$ — Objective Function
Subject to $-x_1+x_2\le 4$, $x_1+x_2\le 6$ ⇨ $-x_1+x_2-4\le 0$, $x_1+x_2-6\le 0$	**Subject to:** $g_j(\mathbf{x})\le 0,\ j=1,\cdots,m$ — Constraints : Inequality constraint
$5x_1+x_2=10$ ⇨ $5x_1+x_2-10=0$	$h_k(\mathbf{x})=0, k=1,\cdots,p$: Equality constraint
$0\le x_1, x_2$	$\mathbf{x}_l\le\mathbf{x}\le\mathbf{x}_u$: Constraint

2.3 Classification of Optimization Problems(1)

☑ Existence of the constraints

■ Unconstrained optimization problem(Unconstrained optimization problem)

- Minimize the objective function $f(\mathrm{x})$ without any constraints on the design variables x.

$$\text{Minimize } f(\mathrm{x})$$

■ Constrained optimization problem

- Minimize the objective function $f(\mathrm{x})$ with some constraints on the design variables x.

$$\begin{aligned} &\text{Minimize} && f(\mathrm{x}) \\ &\text{Subject to} && h(\mathrm{x})=0 \\ & && g(\mathrm{x})\leq 0 \end{aligned}$$

2.3 Classification of Optimization Problems(2)

☑ Number of the objective functions

■ Single-objective optimization problem

Minimize	$f(\mathrm{x})$
Subject to	$h(\mathrm{x})=0$
	$g(\mathrm{x})\leq 0$

■ Multi-objective optimization problem

- Weighting Method, Constraint Method

Minimize	$f_1(\mathrm{x}), f_2(\mathrm{x}), f_3(\mathrm{x})$
Subject to	$h(\mathrm{x})=0$
	$g(\mathrm{x})\leq 0$

2.3 Classification of Optimization Problems(3)

☑ Linearity of the objective function and constraints

■ Linear optimization problem

- The objective function($f(\mathbf{x})$) and constraints($h(\mathbf{x}),\ g(\mathbf{x})$) are linear functions of the design variables $\mathbf{x}$.

Minimize

$$f(\mathbf{x}) = x_1 + 2x_2$$

Subject to

$$h(\mathbf{x}) = x_1 + 5x_2 = 0$$

$$g(\mathbf{x}) = -x_1 \le 0$$

■ Nonlinear optimization problem

- The objective function($f(\mathbf{x})$) or constraints($h(\mathbf{x}),\ g(\mathbf{x})$) are nonlinear functions of the design variables $\mathbf{x}$.

Minimize

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

Subject to

$$h(\mathbf{x}) = x_1 + 5x_2 = 0$$

$$g(\mathbf{x}) = -x_1 \le 0$$

Minimize

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

Subject to

$$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$

$$g_2(\mathbf{x}) = -x_1 \le 0$$

2.3 Classification of Optimization Problems(4)

☑ Type of the design variables

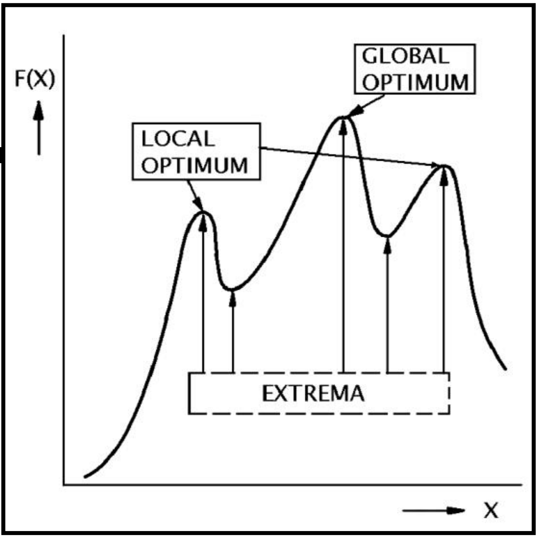
■ Continuous Problem

- The design variables in the optimization problem are continuous.

■ Discrete Problem

- The design variables in the optimization problem are discrete.
- It is also called the 'Combinatorial optimization problem'.
- Example) Integer programming problem

2.4 Classification of Optimization Method



☑ Global Optimization Methods

- **Advantage**
  - It is useful for solving the global optimization problem which has many local optimum solution.
- **Disadvantage**
  - It needs many iterations to obtain the optimum solution(much time).
- **Genetic Algorithms(GA), Simulated Annealing, etc.**

■ Local Optimization Methods

- **Advantage**
  - It needs relatively few iterations to obtain the optimum solution(less time).
- **Disadvantage**
  - It is only able to find the local optimum solution which is near to the starting point.
- **Sequential Quadratic Programming(SQP), Method of Feasible Directions(MFD), Multi-Start Optimization Method, etc.**

Computer Aided Ship Design

Part I. Optimization Method

- Ch.3 Unconstrained Optimization Method

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.3 Unconstrained Optimization Method
3.1 Gradient Method

1. Steepest Descent Method
2. Conjugate Gradient Method
3. Newton's Method
4. Davidon-Fletcher-Powell(DFP) Method
5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

3.1 Gradient Method

1. Steepest Descent Method(1)

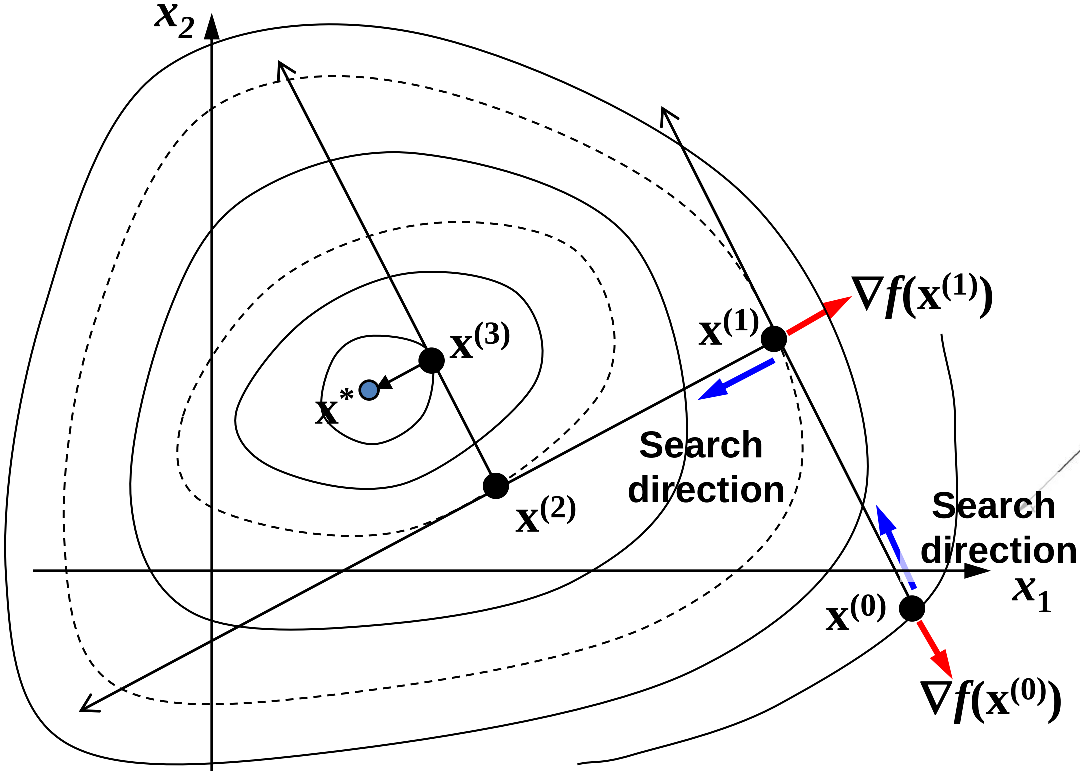
- Step 1: The search direction($\mathbf{d}$) is taken as the negative of the gradient of the objective function(f) at current iteration since the objective function f decrease mostly rapidly.
- The direction of gradient vector of f, $\nabla f(\mathbf{x})$, is the direction of maximum increase of f at $\mathbf{x}$

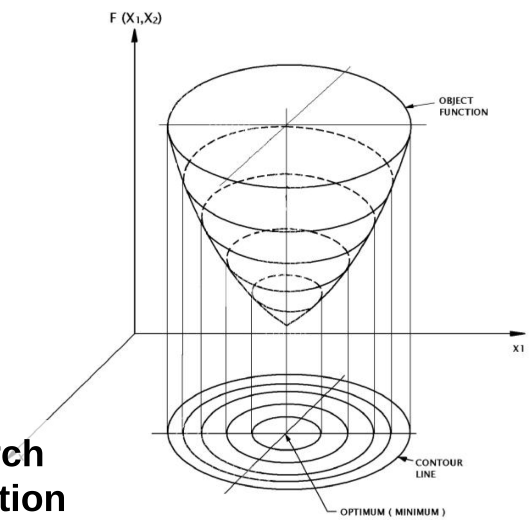
Search direction $$\mathbf{d} = -\mathbf{c} \equiv -\nabla f(\mathbf{x})$$

Ref) Appendix A.1:
Directional Derivative & Gradient Vector

- Step 2: Iterate successively to find the optimum design point.

ex) Minimize the objective function





3.1 Gradient Method
1. Steepest Descent Method(2): Example

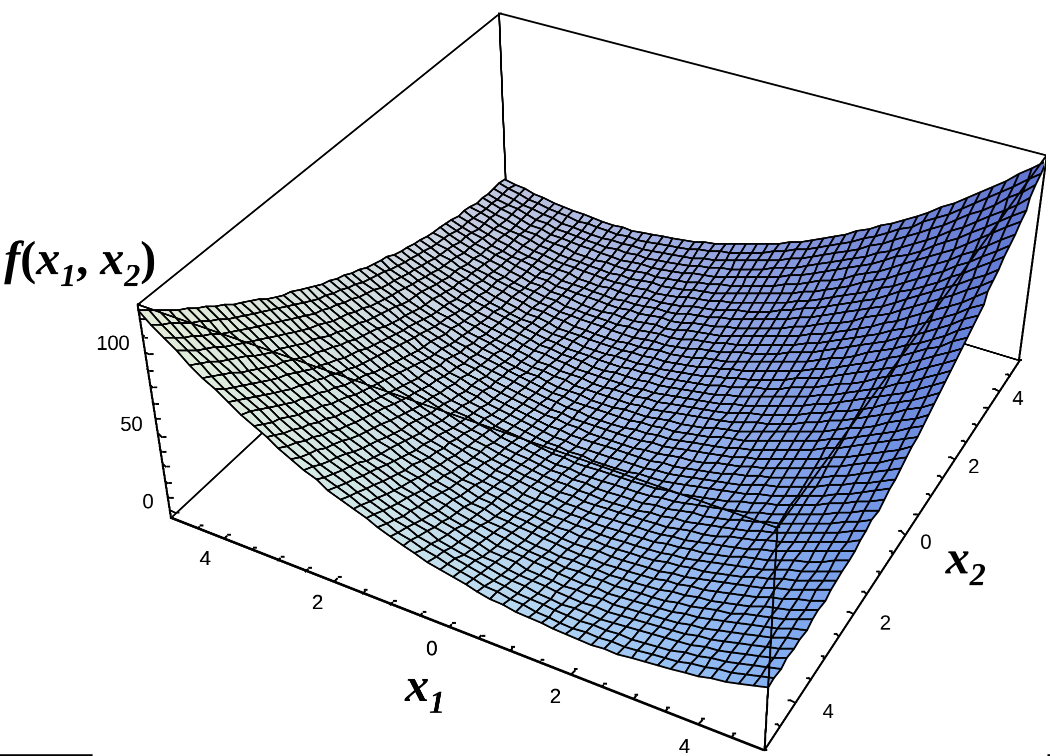
☑ **By using the steepest descent method, find the minimum design point in the following function of 2-variables.**

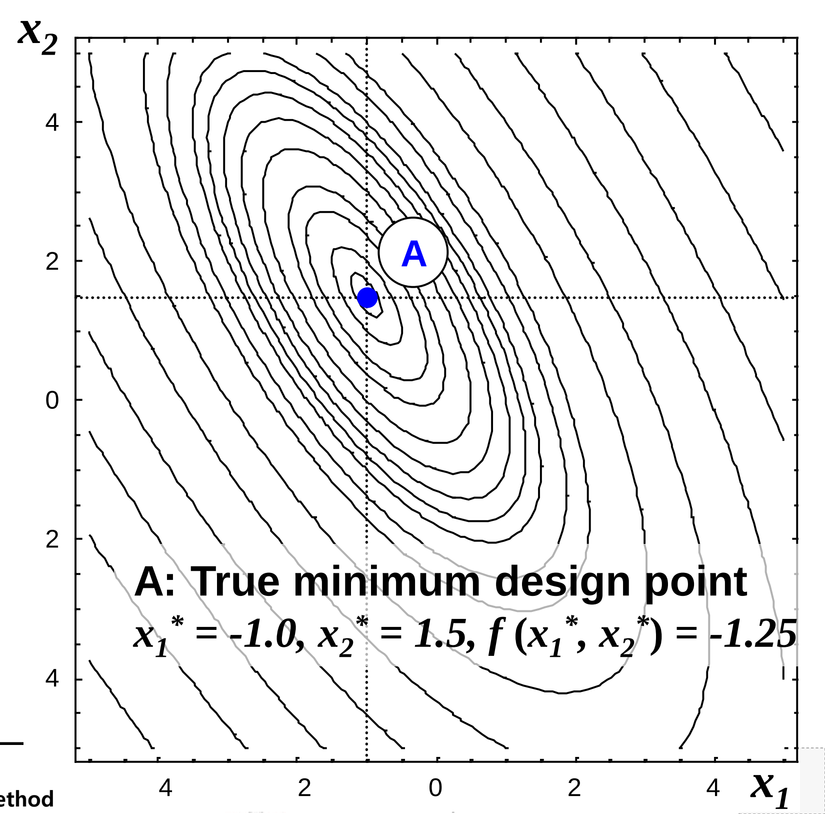
Given: Starting design point $x^{(0)} = (0, 0)$, convergence tolerance $\varepsilon = 0.001$

Find: $x^{(1)}$, $x^{(2)}$

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

➡ **Optimization problem with two unknown variables**





3.1 Gradient Method
1. Steepest Descent Method(3): Example

Minimize $f(x_1,x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1,x_2) = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix}$$

☑ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\nabla f(\mathbf{x}^{(0)}) = \nabla f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha^{(0)}\nabla f(\mathbf{x}^{(0)})$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \alpha\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

Replacing $\alpha^{(0)}$ to α for convenience

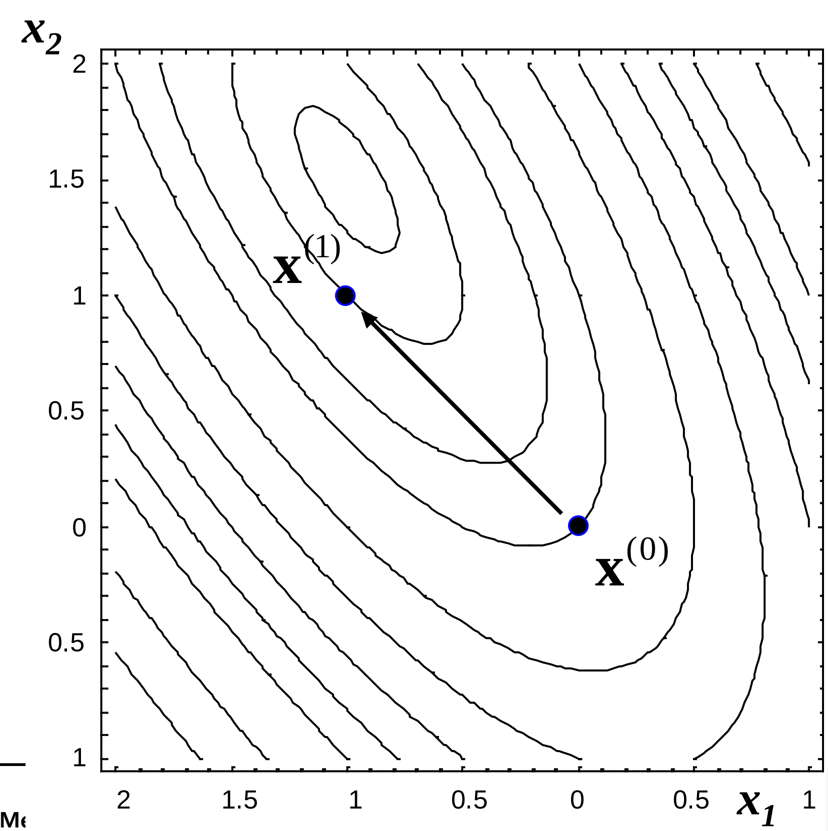
Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ into the objective function

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\boxed{\frac{df(\mathbf{x}^{(1)})}{d\alpha}} = 2\alpha - 2 = 0 \;\rightarrow\; \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

How can we differentiate f with respect to α?



1. Steepest Descent Method(4): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $x^{(0)} = (0, 0)$

☑ 2nd Iteration: Find $\mathbf{x}^{(2)}$

$$\nabla f(\mathbf{x}^{(1)}) = \nabla f\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}1+4x_1+2x_2\\-1+2x_1+2x_2\end{pmatrix} = \begin{pmatrix}-1\\-1\end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - \alpha^{(1)}\nabla f(\mathbf{x}^{(1)})$$

$$= \begin{pmatrix}-1\\1\end{pmatrix} - \alpha\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}-1+\alpha\\1+\alpha\end{pmatrix}$$

Replacing $\alpha^{(1)}$ to α for convenience

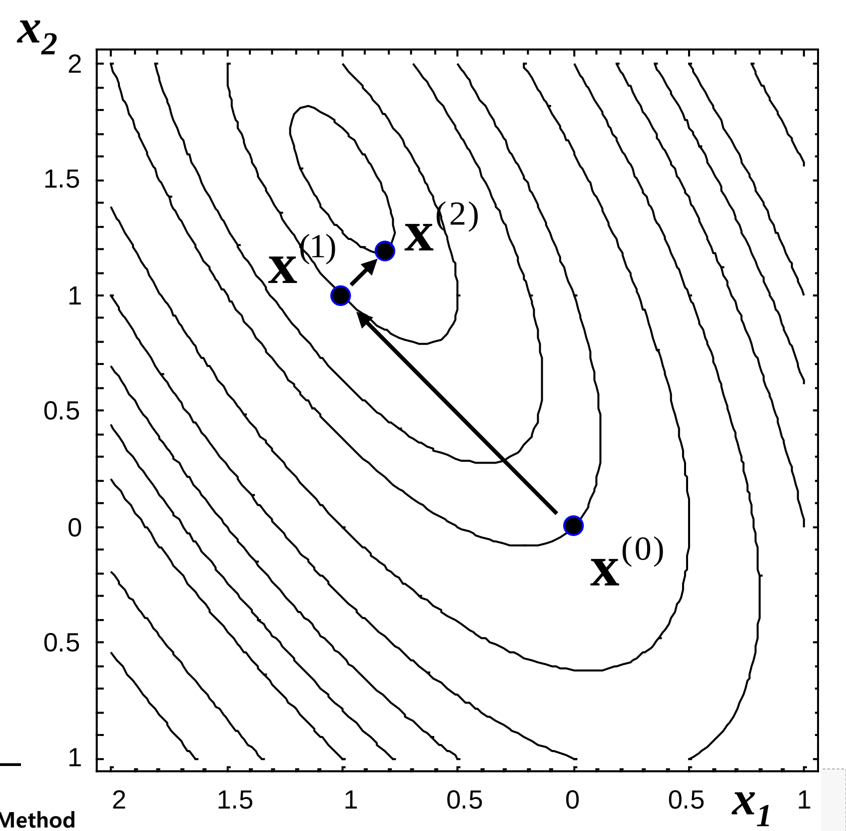
Substituting $\mathbf{x}^{(2)} = (-1+\alpha, 1+\alpha)$ into the objective function

$$f(\mathbf{x}^{(2)}) = 5\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$,

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 10\alpha - 2 = 0 \rightarrow \alpha = 0.2$$

$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix}-0.8\\1.2\end{pmatrix}$$



3.1 Gradient Method
1. Steepest Descent Method(5): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

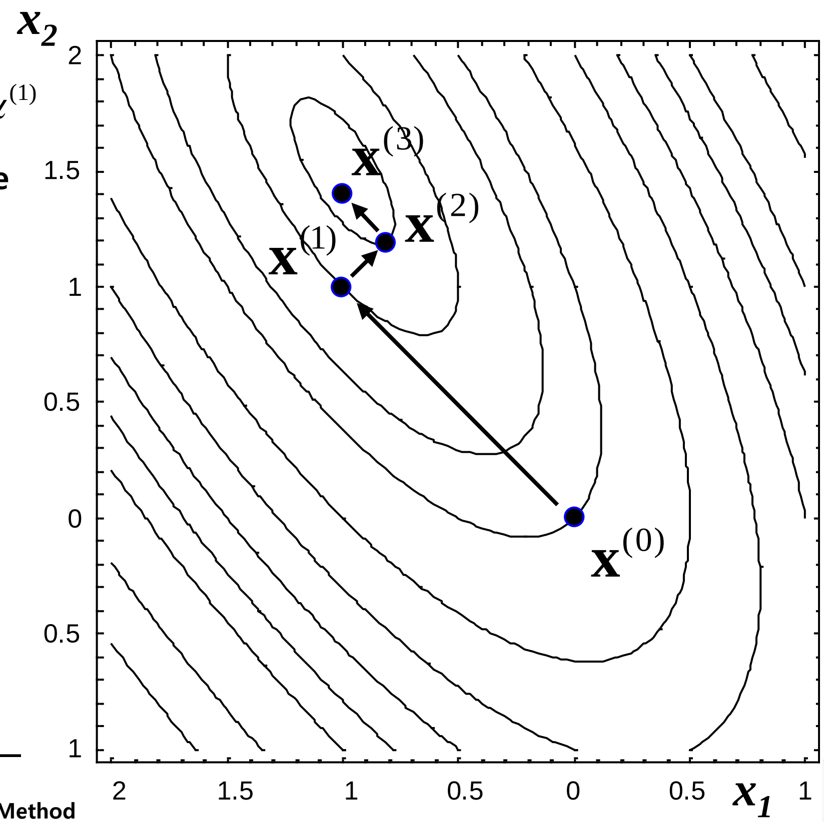
■ **3rd Iteration: Find** $\mathbf{x}^{(3)}$

$$\nabla f(\mathbf{x}^{(2)}) = \nabla f\begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix} = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix} = \begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix}$$

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} - \alpha^{(2)}\nabla f(\mathbf{x}^{(2)})$$

$$= \begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix} - \alpha\begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix} = \begin{pmatrix} -0.8-0.2\alpha \\ 1.2+0.2\alpha \end{pmatrix}$$

Replacing $\alpha^{(1)}$ **to** α **for convenience**



Substituting $\mathbf{x}^{(3)} = (-0.8-0.2\alpha, 1.2+0.2\alpha)$ **into the objective function**

$$f(\mathbf{x}^{(3)}) = 0.04\alpha^2 - 0.08\alpha - 1.2$$

To minimize $f(\mathbf{x}^{(3)})$**,**

$$\frac{df(\mathbf{x}^{(3)})}{d\alpha} = 0.08\alpha - 0.08 = 0 \quad \rightarrow \quad \alpha = 1.0$$

$$\therefore \mathbf{x}^{(3)} = \begin{pmatrix} -1 \\ 1.4 \end{pmatrix}$$

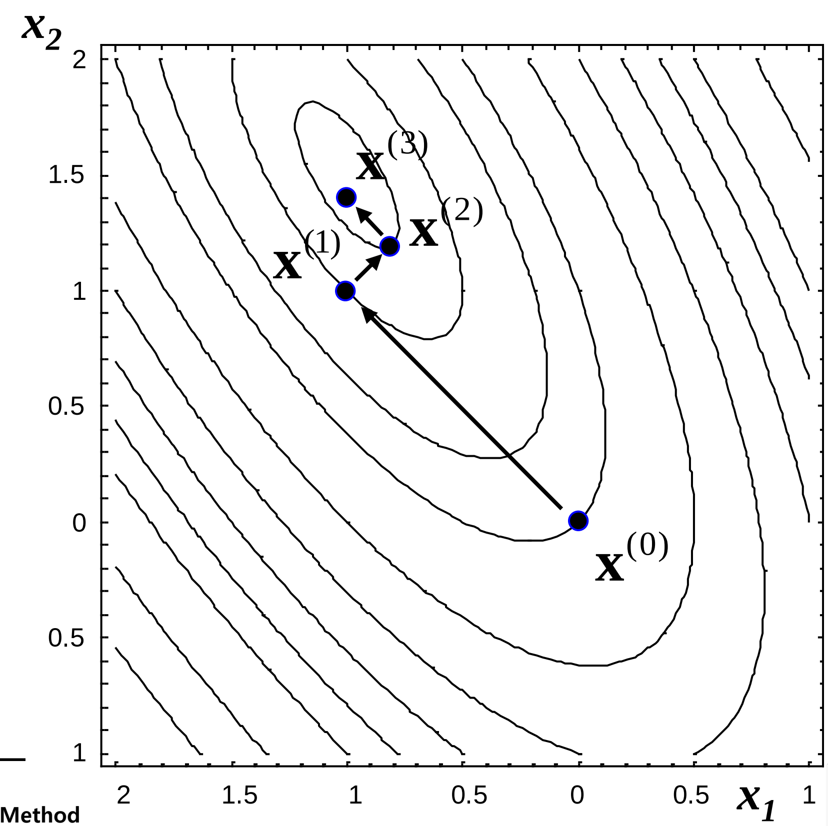
3.1 Gradient Method
1. Steepest Descent Method(6): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ **4th Iteration: Find the minimum design point.**

To obtain the minimum design point, we have to iterate.

If $\left|\mathrm{x}^{(k+1)} - \mathrm{x}^{(k)}\right| \le \varepsilon$**, then stop the iterative process because** $\mathbf{x}^{(k+1)}$ **can be assumed as the minimum design point.**



[Reference] Differentiation of Function of x with Respect to the Another Variable

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$f(x_1, x_2) = f(\mathbf{x})$: f **is the function of** $\mathbf{x}$.

$\mathbf{x}^{(1)} = (-\alpha, \alpha)$: $\mathbf{x}^{(1)}$ **is the function of** α

→Substituting $\mathbf{x}^{(1)}$ **into** f, f **is ,then, function of** α **and can be differentiated** f **with respect to** α **.**

In the similar way, we can consider the followings:

To minimize $f(\mathbf{x}^* + \Delta\mathbf{x})$,

The second-order Taylor series expansion of $f(\mathbf{x}^* + \Delta\mathbf{x})$

$$f(\mathbf{x}^* + \Delta\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$

$$f(\mathbf{x}^* + \Delta\mathbf{x}) - f(\mathbf{x}^*) = \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$

In the above equation, we assume that x* is constant and $\Delta\mathbf{x}$ **is a variable.**

$$f(\Delta\mathbf{x}) = \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$

To minimize f,

$$\frac{df(\Delta\mathbf{x})}{d\Delta\mathbf{x}} = \mathbf{c} + \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = 0$$

$$\Rightarrow \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = -\mathbf{c}$$

$$\Rightarrow \boxed{\Delta\mathbf{x} = -\mathbf{H}(\mathbf{x}^*)^{-1}\mathbf{c}}$$ **'Newton's method'**

Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\boxed{\frac{df(\mathbf{x}^{(1)})}{d\alpha}} = 2\alpha - 2 = 0 \;\rightarrow\; \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

How can we differentiate f **with respect to** α**?**

2. Conjugate Gradient Method(1)

☑ This method requires only a simple modification to the steepest descent method and dramatically improves the convergence rate of the optimization process.

☑ The current steepest descent direction is modified by adding a scaled direction used in the previous iteration.

- Step 1 : Estimate a starting design point as $\mathbf{x}^{(0)}$. Set the iteration counter $k=0$. Also, specify a tolerance ε for stopping criterion. Calculate

$$\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} \equiv -\nabla f(\mathbf{x}^{(0)})$$

Check stopping criterion. If $\left\|\mathbf{c}^{(0)}\right\| < \varepsilon$, then stop. Otherwise, go to Step 4(note that Step 1 of the conjugate gradient and steepest descent method is the same).

3.1 Gradient Method
2. Conjugate Gradient Method(2)

■ **Step 2 :Compute the gradient of the objective function as** $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$.
If $\|\mathbf{c}^{(k)}\| < \varepsilon$, **then stop; otherwise continue.**

■ **Step 3 : Calculate the new search direction as**

$$\boxed{\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta_k \mathbf{d}^{(k-1)}} \rightarrow \text{Previous search direction}$$

$$\beta_k = (\|\mathbf{c}^{(k)}\| / \|\mathbf{c}^{(k-1)}\|)^2$$

The current search direction is calculated by adding a scaled direction used in the previous iteration.

■ **Step 4 : Compute a step size** $\alpha = \alpha_k$ **to minimize** $f(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})$.

■ **Step 5 : Change the design point as follows, set** $k = k+1$ **and go to Step 2.**

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

3.1 Gradient Method
2. Conjugate Gradient Method(3) : Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

☑ **Ist Iteration: Find $\mathbf{x}^{(1)}$**

$$\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} = -\nabla f\left(\mathbf{x}^{(0)}\right) = -\nabla f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$= -\begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

Replacing α_0 to α for convenience

Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ into the objective function

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$

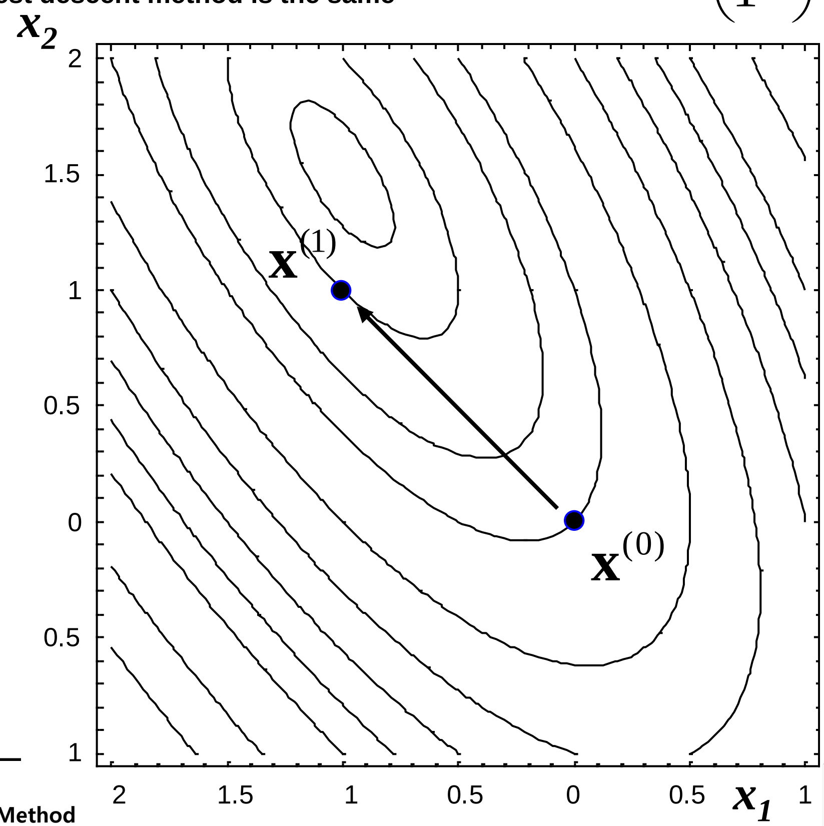
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \quad \rightarrow \quad \alpha = 1.0$$

$$\therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Note: Step 1 of the conjugate gradient and steepest descent method is the same



Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

■ 2nd Iteration-Find $\mathbf{x}^{(2)}$

$$\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{d}^{(0)} = -\nabla f\left(\mathbf{x}^{(0)}\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta_k \mathbf{d}^{(k-1)}$$

$$\beta_k = (\|\mathbf{c}^{(k)}\| / \|\mathbf{c}^{(k-1)}\|)^2$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

Compute the gradient of the objective function as

$$\mathbf{c}^{(1)} = \nabla f\left(\mathbf{x}^{(1)}\right)$$

$$= \nabla f \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Calculate the new search direction as

$$\mathbf{d}^{(1)} = -\mathbf{c}^{(1)} + \beta_1 \mathbf{d}^{(0)} = -\mathbf{c}^{(1)} + \frac{\left\|\nabla f\left(\mathbf{x}^{(1)}\right)\right\|^2}{\left\|\nabla f\left(\mathbf{x}^{(0)}\right)\right\|^2} \mathbf{d}^{(0)}$$

$$= -\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \frac{2}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

2. Conjugate Gradient Method(5): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 + 2\alpha \end{pmatrix}$$ **Replacing** α_1 **to** α **for convenience**

Substituting $\mathbf{x}^{(2)} = (-1, 1 + 2\alpha)$ **into the objective function**

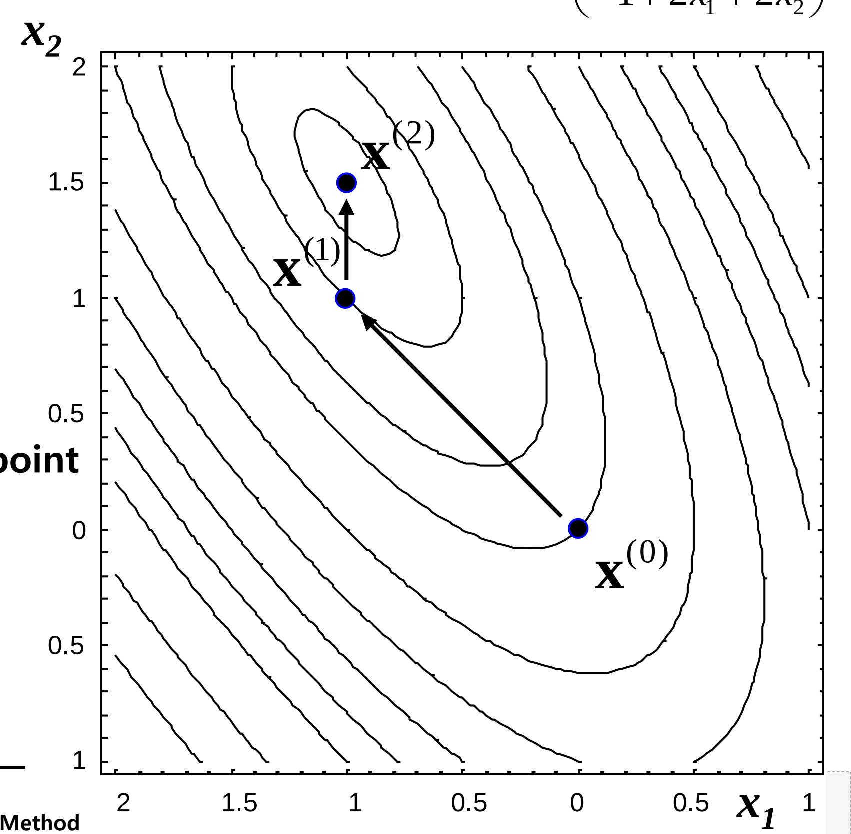
$$f(\mathbf{x}^{(2)}) = 4\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 8\alpha - 2 = 0 \quad \rightarrow \quad \alpha = 0.25$$

$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

→Minimum design point



Check stopping criterion.

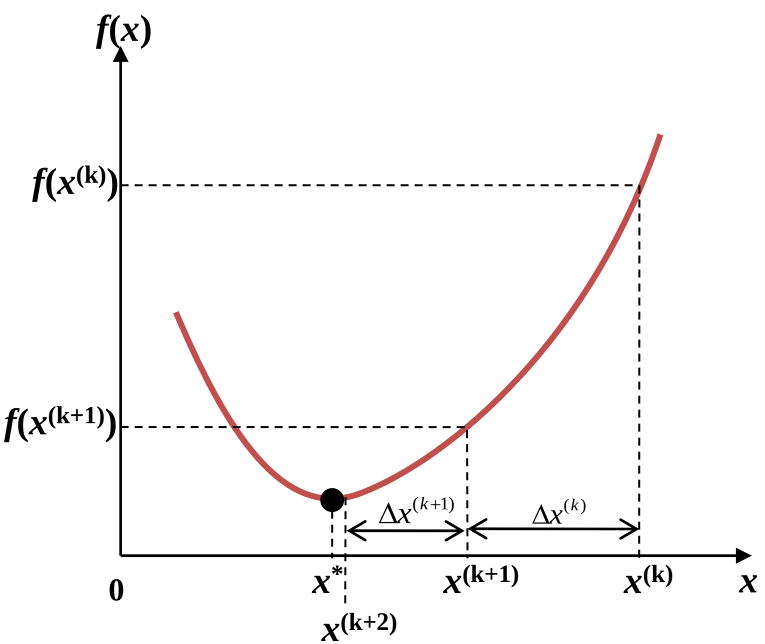
$$\mathbf{c}^{(2)} = \nabla f\left(\mathbf{x}^{(2)}\right) = \nabla f \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\| \mathbf{c}^{(2)} \right\| = 0 < \varepsilon \quad$$ **→Stop!**

3. Newton's Method(1)

Given: $f(x)$

Find: x^* **which minimizes** $f(x)$



Assume that $f(x)$ has minimum at $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(k)}$ using the second-order Taylor expansion.

$$f(x^{(k)} + \Delta x^{(k)}) = f(x^{(k)}) + \frac{df(x^{(k)})}{dx}\Delta x^{(k)} + \frac{1}{2}\frac{d^2 f(x^{(k)})}{dx^2}\left(\Delta x^{(k)}\right)^2 + \cancel{O\left(\left(\Delta x^{(k)}\right)^3\right)}$$

In this equation, $x^{(k)}$ is a constant and $\Delta x^{(k)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(k)}$.

$$f(x^{(k)} + \Delta x^{(k)}) = f(x^{(k)}) + \frac{df(x^{(k)})}{dx}\Delta x^{(k)} + \frac{1}{2}\frac{d^2 f(x^{(k)})}{dx^2}\left(\Delta x^{(k)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(k)}$.

$$\frac{df(x^{(k)} + \Delta x^{(k)})}{d\Delta x^{(k)}} = \frac{df(x^{(k)})}{dx} + \frac{d^2 f(x^{(k)})}{dx^2}\Delta x^{(k)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(k)}$ in design.

$$\Delta x^{(k)} = \left(-\frac{df(x^{(k)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(k)})}{dx^2}\right)$$

Is | $\Delta x^{(k)}$ | < ε ?

NO → $k = k + 1$

YES → **Set $x^* = x^{(k+1)}$ and stop the iteration**

3.1 Gradient Method

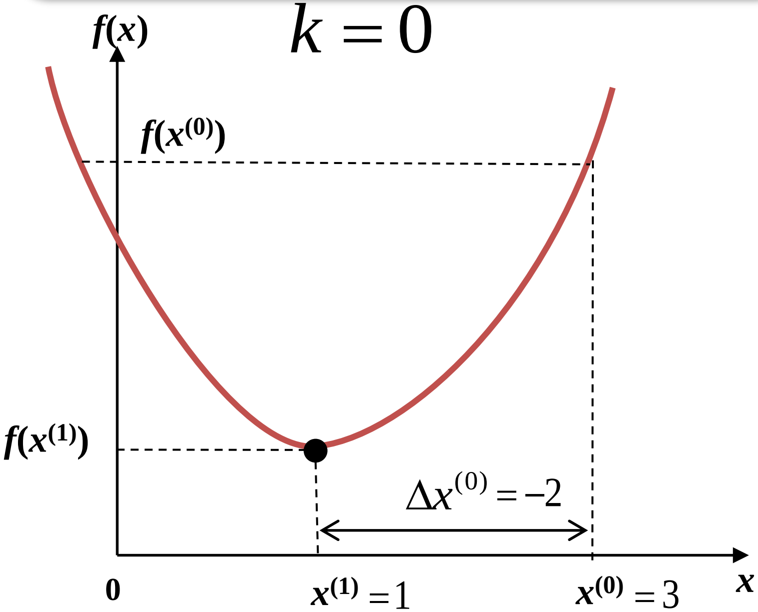
3. Newton's Method(2): Example

Given: $f(x) = x^2 - 2x + 2$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$k = 0$



Assume that $f(x)$ has minimum at $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(0)}$ using the second-order Taylor expansion.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}(\Delta x^{(0)})^2 + \cancel{O((\Delta x^{(0)})^3)}$$

In this equation, $x^{(0)}$ is a constant and $\Delta x^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(0)}$.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}(\Delta x^{(0)})^2$$

Differentiate this equation with respect to $\Delta x^{(0)}$.

$$\frac{df(x^{(0)} + \Delta x^{(0)})}{d\Delta x^{(0)}} = \frac{df(x^{(0)})}{dx} + \frac{d^2 f(x^{(0)})}{dx^2}\Delta x^{(0)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(0)}$ in design.

$$\Delta x^{(0)} = \left(-\frac{df(x^{(0)})}{dx}\right) / \left(\frac{d^2 f(x^{(0)})}{dx^2}\right)$$

$$= (-2x+2)_{x=3} / (2)_{x=3} = -2$$

Is | $\Delta x^{(0)}$ |< ε ?

NO

$k = k + 1$

$= 0 + 1 = 1$

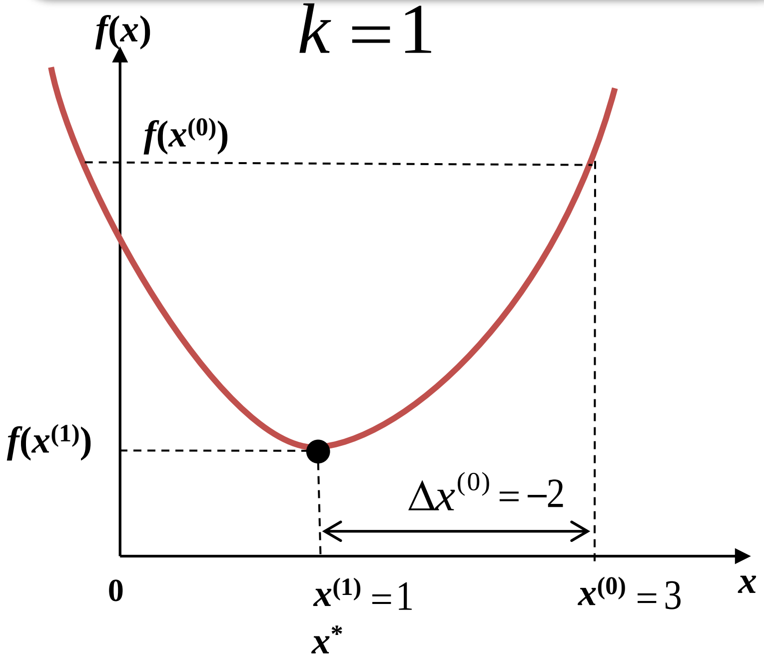
3.1 Gradient Method
3. Newton's Method(3): Example

Given: $f(x) = x^2 - 2x + 2$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$k = 1$



Assume that $f(x)$ has minimum at $x^{(2)} = x^{(1)} + \Delta x^{(1)}$.

Consider the quadratic approximation of the function $f(x)$ at $x = x^{(1)}$ using the second-order Taylor expansion.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2 + \cancel{O\left(\left(\Delta x^{(1)}\right)^3\right)}$$

In this equation, $x^{(1)}$ is a constant and $\Delta x^{(1)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(1)}$.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{df(x^{(1)} + \Delta x^{(1)})}{d\Delta x^{(1)}} = \frac{df(x^{(1)})}{dx} + \frac{d^2 f(x^{(1)})}{dx^2}\Delta x^{(1)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(1)}$ in design.

$$\Delta x^{(1)} = \left(-\frac{df(x^{(1)})}{dx}\right) / \left(\frac{d^2 f(x^{(1)})}{dx^2}\right)$$

$$= \left(-2x + 2\right)_{x=1} / \left(2\right)_{x=1} = 0$$

Is $|\Delta x^{(1)}| < \varepsilon$?

YES

Set $x^* = x^{(1)}$ and stop the iteration

Is it possible to find the x* which minimizes a cubic function at once?

3.1 Gradient Method
3. Newton's Method(4): Example

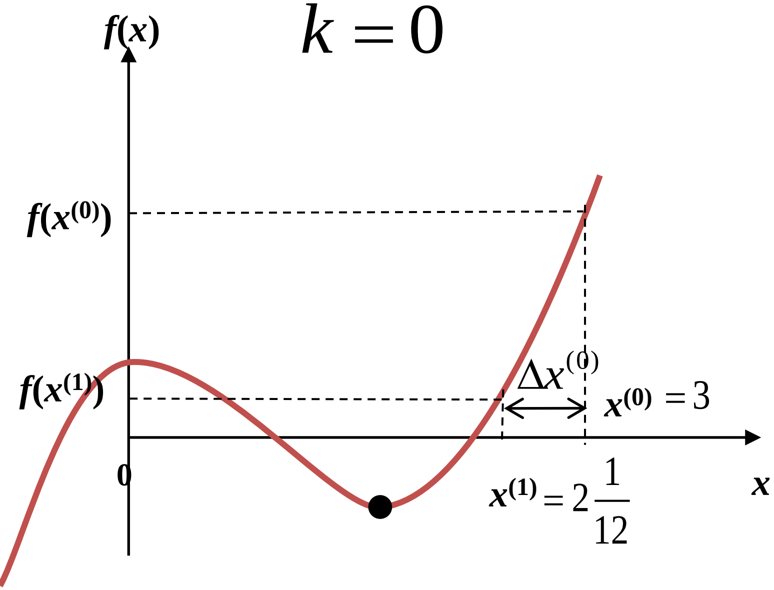
Is it possible to find the x* which minimizes a cubic function at once?

Given: $f(x) = x^3 - 3x^2 + 2x$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$$k = 0$$



Assume that $f(x)$ has minimum at $x^{(1)} = x^{(0)} + \Delta x^{(0)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(0)}$ using the second-order Taylor expansion.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2 + \cancel{O\left(\left(\Delta x^{(0)}\right)^3\right)}$$

In this equation, $x^{(0)}$ is a constant and $\Delta x^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(0)}$.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(0)}$.

$$\frac{df(x^{(0)} + \Delta x^{(0)})}{d\Delta x^{(0)}} = \frac{df(x^{(0)})}{dx} + \frac{d^2 f(x^{(0)})}{dx^2}\Delta x^{(0)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(0)}$ in design.

$$\Delta x^{(0)} = \left(-\frac{df(x^{(0)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(0)})}{dx^2}\right)$$

$$= \left(-3x^2 + 6x - 2\right)_{x=3} / \left(6x - 6\right)_{x=3} = -\frac{11}{12}$$

Is $|\ \Delta x^{(0)}\ | < \varepsilon$?

NO

$$k = k + 1$$
$$= 0 + 1 = 1$$

3.1 Gradient Method

3. Newton's Method(5): Example

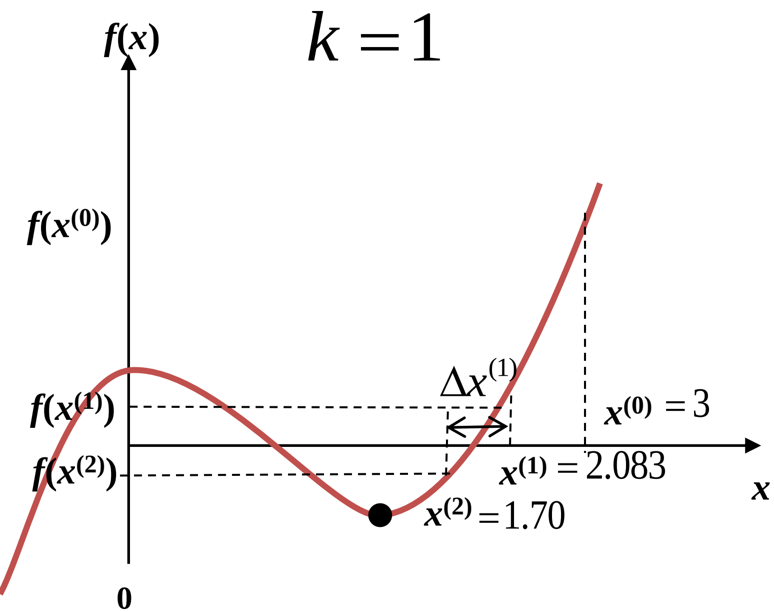
Is it possible to find the x* which minimizes a cubic function at once?

Given: $f(x) = x^3 - 3x^2 + 2x$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$$k = 1$$



Why is it not possible to find the x* which minimizes a cubic function at once?

Assume that $f(x)$ has minimum at $x^{(2)} = x^{(1)} + \Delta x^{(1)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(1)}$ using the second-order Taylor expansion.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2 + \cancel{O\left(\left(\Delta x^{(1)}\right)^3\right)}$$

In this equation, $x^{(1)}$ is a constant and $\Delta x^{(1)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(1)}$.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{df(x^{(1)} + \Delta x^{(1)})}{d\Delta x^{(1)}} = \frac{df(x^{(1)})}{dx} + \frac{d^2 f(x^{(1)})}{dx^2}\Delta x^{(1)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(1)}$ in design.

$$\Delta x^{(1)} = \left(-\frac{df(x^{(1)})}{dx}\right) / \left(\frac{d^2 f(x^{(1)})}{dx^2}\right)$$

$$= \left(-3x^2 + 6x - 2\right)_{x=\frac{25}{12}} / \left(6x - 6\right)_{x=\frac{25}{12}} = -0.388$$

Is $|\ \Delta x^{(1)}\ | < \varepsilon$?

NO

$$k = k + 1$$
$$= 1 + 1 = 2$$

Since the second-order Taylor expansion is just an approximation for $f(x)$ at the point $x^{(0)}$, $x^{(1)}$ will probably not be the precise minimum design point of $f(x)$.

3.1 Gradient Method

3. Newton's Method(6): Example of Function of Two Variables

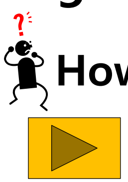
Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}, \quad \mathbf{H}(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

☑ 1st Iteration: Find $\mathbf{x}^{(1)}$

Assume that $f(x)$ has minimum at $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}$.

Consider the quadratic approximation of the function $f(\mathbf{x})$ at $\mathbf{x}=\mathbf{x}^{(0)}$ using the second-order Taylor expansion.



How?

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}^{(0)}$$

In this equation, $\mathbf{x}^{(0)}$ is a constant and $\Delta\mathbf{x}^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta\mathbf{x}^{(0)}$.

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}^{(0)}$$

3. Newton's Method(7): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ **1st Iteration: Find** $\mathbf{x}^{(1)}$

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}$$

How?

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

Differentiate this equation with respect to $\Delta x^{(0)}$.

$$\frac{\partial f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)})}{\partial(\Delta\mathbf{x}^{(0)})} = \nabla f(\mathbf{x}^{(0)}) + \mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = 0 \longrightarrow$$

The necessary condition for minimization of function $f(x_1, x_2)$

Calculate the small change $\Delta x^{(0)}$ **in design.**

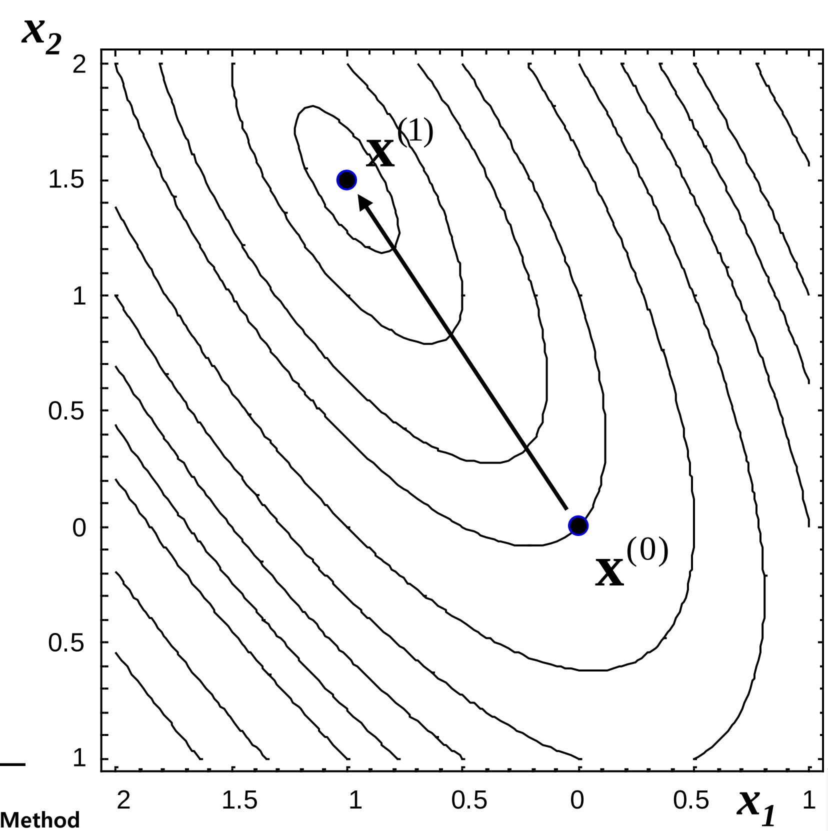
$$\mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = -\nabla f(\mathbf{x}^{(0)})$$

$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\downarrow \quad \left(-\nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \ \mathbf{H}(\mathbf{x}^{(0)}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \right)$$

$$\begin{pmatrix} \Delta\mathbf{x}_1^{(0)} \\ \Delta\mathbf{x}_2^{(0)} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} \Delta\mathbf{x}_1^{(0)} \\ \Delta\mathbf{x}_2^{(0)} \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

$$\therefore x^{(1)} = x^{(0)} + \Delta\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$



3. Newton's Method(8): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

☑ 2nd Iteration-Find $\mathbf{x}^{(2)}$

$$x^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

In the same way as 1st Iteration,

Assume that $f(x)$ **has minimum at** $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}$.

Consider the quadratic approximation of the function $f(\mathbf{x})$ **at** $\mathbf{x}=\mathbf{x}^{(1)}$ **using the second-order Taylor expansion.**

$$f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}) = f(\mathbf{x}^{(1)}) + \nabla f(\mathbf{x}^{(1)})^T \Delta\mathbf{x}^{(1)} + \frac{1}{2}(\Delta\mathbf{x}^{(1)})^T \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)}$$

In this equation, $\mathbf{x}^{(1)}$ **is a constant and** $\Delta x^{(1)}$ **is a variable. So, the following equation is a quadratic function in terms of** $\Delta x^{(1)}$.

$$f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}) = f(\mathbf{x}^{(1)}) + \nabla f(\mathbf{x}^{(1)})^T \Delta\mathbf{x}^{(1)} + \frac{1}{2}(\Delta\mathbf{x}^{(1)})^T \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)}$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{\partial f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)})}{\partial(\Delta\mathbf{x}^{(1)})} = \nabla f(\mathbf{x}^{(1)}) + \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)} = 0 \longrightarrow$$

The necessary condition for minimization of function $f(x_1, x_2)$

3. Newton's Method(9): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

$$x^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

☑ 2nd Iteration-Find $\mathbf{x}^{(2)}$

Calculate the small change $\Delta x^{(1)}$ in design.

$$\mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)} = -\nabla f(\mathbf{x}^{(1)})$$

$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\left(-\nabla f(\mathbf{x}^{(1)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{H}(\mathbf{x}^{(1)}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \right)$$

$$\begin{pmatrix} \Delta\mathbf{x}_1^{(1)} \\ \Delta\mathbf{x}_2^{(1)} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} \Delta\mathbf{x}_1^{(1)} \\ \Delta\mathbf{x}_2^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

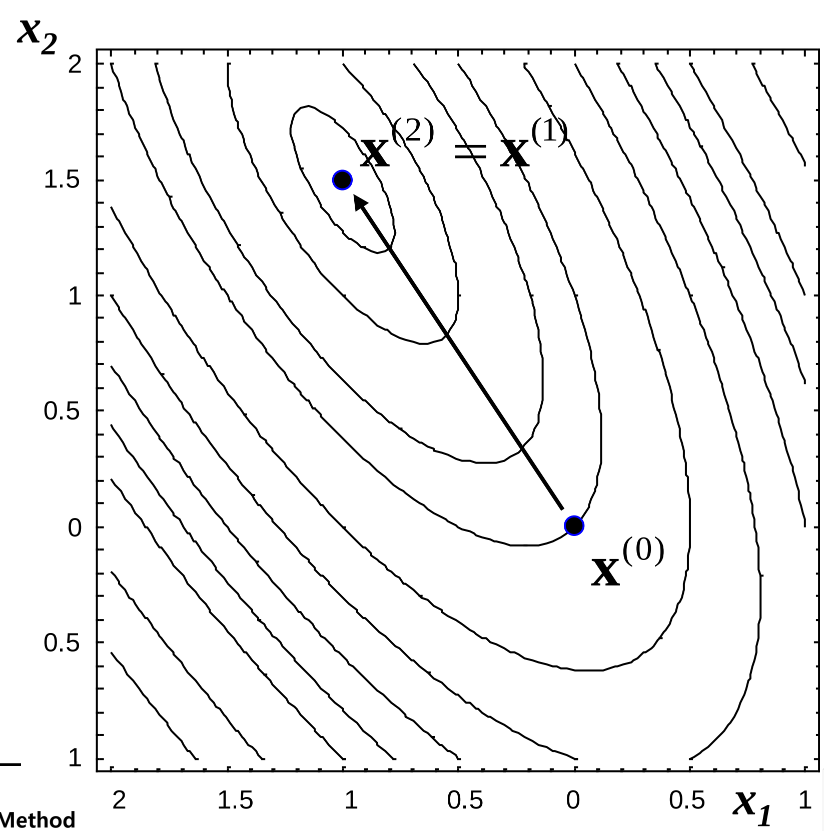
$$\therefore x^{(2)} = x^{(1)} + \Delta\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

→Optimal design point

Check stopping criterion.

$$\left|\Delta\mathbf{x}^{(1)}\right| = 0 < \varepsilon$$

→Stop!



3. Modified Newton's Method(1)

☑ In this method, we treat $\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$ of the Newton's method as **the search direction** and use any of the one-dimensional search methods to calculate the step size in the search direction.

- **Step 1 :** Estimate a starting design point $\mathbf{x}^{(0)}$.
  Set iteration counter $k = 0$. Specify a tolerance ε for the stopping criterion.

- **Step 2 :** Calculate $c_i^{(k)} = \partial f(\mathbf{x}^{(k)}) / \partial x_i$ for $i = 1$ to n. If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$, stop the iterative process. Otherwise, continue.

- **Step 3 :** Calculate the Hessian matrix $\mathbf{H}^{(k)}$ at current design point $\mathbf{x}^{(k)}$.

$$\mathbf{H}(\mathbf{x}^{(k)}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right], \quad i = 1, \cdots, n; \ \ j = 1, \cdots, n$$

3. Modified Newton's Method(2)

■ **Step 4 : Calculate the search direction as follows:**

$$\boxed{\mathbf{d}^{(k)} = \Delta\mathbf{x}^{(k)} = -\mathbf{H}^{-1}\mathbf{c}^{(k)}}$$

When $f(\mathbf{x}^* + \Delta\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$,

the necessary condition for minimization of this function is as follows:

$$df(\Delta\mathbf{x})/d\Delta\mathbf{x} = \mathbf{c} + \mathbf{H}(\mathbf{x}^*)\Delta\mathbf{x} = 0$$

$$\Rightarrow \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = -\mathbf{c} \Rightarrow \Delta\mathbf{x} = -\mathbf{H}(\mathbf{x}^*)^{-1}\mathbf{c}$$

■ **Step 5 : Update the design point as** $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)}$**, where** α **is calculated to minimize** $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$**. Any one-dimensional search method may be used to calculate** α**.**

■ **Step 6 : Set** $k = k+1$ **and go to Step 2.**

3. Disadvantages of the Newton's Method

The Newton's method is not very useful in practice, due to following features of the method:

1. It requires the storing of the $n \times n$ matrix $\mathbf{H}(\mathbf{x}^{(k)})$.

2. It becomes very difficult and sometimes, impossible to compute the elements of the matrix $\mathbf{H}(\mathbf{x}^{(k)})$.

3. It requires the inversion of the matrix $\mathbf{H}(\mathbf{x}^{(k)})$ at each iteration.

4. It requires the evaluation of the quantity $\mathbf{H}(\mathbf{x}^{(k)})^{-1}\nabla f(\mathbf{x}^{(k)})$ at each iteration.

4. Davidon-Fletcher-Powell(DFP) Method(1)

☑ **This method builds an approximation for the inverse of the Hessian matrix of $f(\mathbf{x})$ using only the first derivatives.**

- **Step 1 : Estimate a starting design point $\mathbf{x}^{(0)}$.**

  Choose a symmetric positive definite nxn matrix $\mathbf{A}^{(0)}$ as an approximation for the inverse of the Hessian matrix of the objective function. In the absence of more information, $\mathbf{A}^{(0)} = \mathbf{I}$ may be chosen. Also, specify a tolerance ε for the stopping criterion. Set $k = 0$. Compute the gradient vector as $\mathbf{d}^{(0)} = \mathbf{c}^{(0)} \equiv \nabla f(\mathbf{x}^{(0)})$.

- **Step 2 : Calculate the norm of the gradient vector as $\left\|\mathbf{c}^{(k)}\right\|$.**

  If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$, then stop the iterative process. Otherwise, continue (note that Step 1 and 2 of this method and the steepest descent method are the same).

4. Davidon-Fletcher-Powell(DFP) Method(2)

- **Step 3 : Calculate the search direction as follows:**

$$\boxed{\mathbf{d}^{(k)} = -\mathbf{A}^{(k)}\mathbf{c}^{(k)}}$$

That is, $\mathbf{A}$ matrix is used as an estimate for the inverse of the Hessian matrix $\mathbf{H}^{-1}$ of the objective function.

- **Step 4 : Compute optimum step size $\alpha_k = \alpha$ to minimize $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$.**

- **Step 5 : Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k\mathbf{d}^{(k)}$.**

$\mathbf{d}^{(k)}$: search direction

$\alpha^{(k)}$: optimum step size

- Step 6 : Update the matrix $\mathbf{A}^{(k)}$- approximation for the inverse of the Hessian matrix of the objective function – as

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \mathbf{B}^{(k)} + \mathbf{C}^{(k)} \quad ; \quad n \times n \text{ matrix}$$

where the correction matrices $\mathbf{B}^{(k)}$ and $\mathbf{C}^{(k)}$ are calculated as follows:

$$\mathbf{B}^{(k)} = \frac{\mathbf{s}^{(k)}(\mathbf{s}^{(k)})^T}{(\mathbf{s}^{(k)} \cdot \mathbf{y}^{(k)})} \quad ; \quad n \times n \text{ matrix} \qquad \mathbf{C}^{(k)} = \frac{-\mathbf{z}^{(k)}(\mathbf{z}^{(k)})^T}{(\mathbf{y}^{(k)} \cdot \mathbf{z}^{(k)})} \quad ; \quad n \times n \text{ matrix}$$

$$\mathbf{s}^{(k)} = \alpha_k \mathbf{d}^{(k)} \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{y}^{(k)} = \mathbf{c}^{(k+1)} - \mathbf{c}^{(k)} \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{c}^{(k+1)} = \nabla f(\mathbf{x}^{(k+1)}) \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{z}^{(k)} = \mathbf{A}^{(k)}\mathbf{y}^{(k)} \quad : \quad [n \times n][n \times 1] = [n \times 1] \text{ matrix}$$

- Step 7 : Set $k = k+1$ and go to Step 2.

3.1 Gradient Method
4. Davidon-Fletcher-Powell(DFP) Method(4): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

■ **1st Iteration: Find** $\mathbf{x}^{(1)}$

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \mathbf{A}^{(0)} = \mathbf{I}$$

$$\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 1 + 4\cdot 0 + 2\cdot 0 \\ -1 + 2\cdot 0 + 2\cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Check stopping criterion.

$$\left\|\mathbf{c}^{(0)}\right\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(0)} = -\mathbf{A}^{(0)}\mathbf{c}^{(0)} = -\mathbf{I}\,\mathbf{c}^{(0)} = -\mathbf{c}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

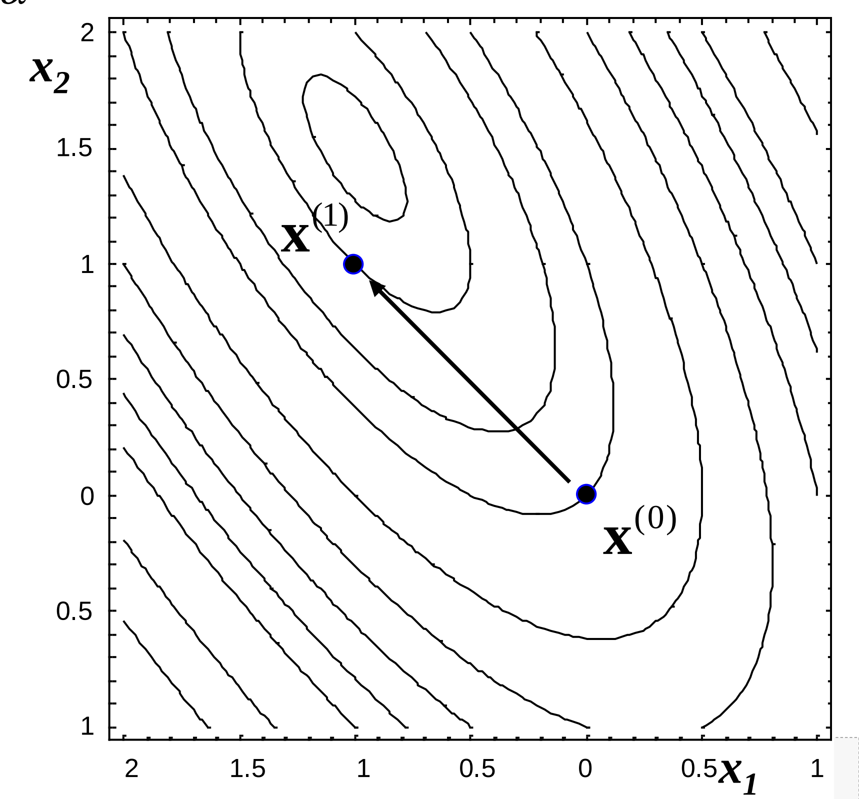
Replacing α_0 **to** α **for convenience**

Substitute $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$**,**

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \ \rightarrow \ \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

4. Davidon-Fletcher-Powell(DFP) Method(5): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2,$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ **2nd Iteration: Find** $\mathbf{x}^{(2)}$

Update the matrix $\mathbf{A}^{(1)}$ **- approximation for the inverse of the Hessian matrix of the objective function – as**

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \mathbf{B}^{(0)} + \mathbf{C}^{(0)}$$

$$\mathbf{B}^{(0)} = \frac{\mathbf{s}^{(0)}\mathbf{s}^{(0)^T}}{\mathbf{s}^{(0)} \cdot \mathbf{y}^{(0)}}$$

$$\mathbf{s}^{(0)} = \alpha \mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{s}^{(0)}\mathbf{s}^{(0)^T} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\mathbf{s}^{(0)} \cdot \mathbf{y}^{(0)} = 2$$

$$= \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}$$

$$\mathbf{C}^{(0)} = \frac{-\mathbf{z}^{(0)}\mathbf{z}^{(0)T}}{\mathbf{y}^{(0)} \cdot \mathbf{z}^{(0)}}$$

$$\mathbf{A}^{(0)} = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{z}^{(0)} = \mathbf{A}^{(0)}\mathbf{y}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)} \cdot \mathbf{z}^{(0)} = 4$$

$$\mathbf{z}^{(0)}\mathbf{z}^{(0)^T} = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \mathbf{B}^{(0)} + \mathbf{C}^{(0)}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}$$

3.1 Gradient Method
4. Davidon-Fletcher-Powell(DFP) Method(6): Example

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix}$$

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2,$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Check stopping criterion.

$$\left\|\mathbf{c}^{(1)}\right\| = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(1)} = -\mathbf{A}^{(1)}\mathbf{c}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1+\alpha \end{pmatrix}$$

Replacing α_1 **to** α **for convenience**

Substitute $\mathbf{x}^{(2)} = (-1, 1+\alpha)$ **into the objective function**

$$f(\mathbf{x}^{(2)}) = \alpha^2 - \alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$**,**

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 2\alpha - 1 = 0 \;\rightarrow\; \alpha = 0.5$$

$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

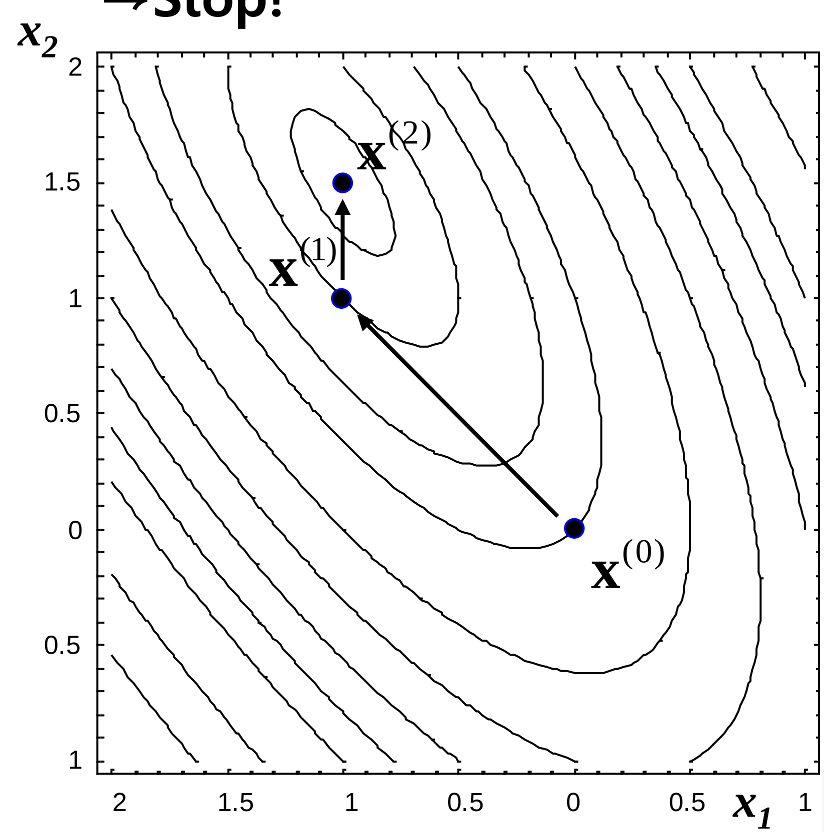
→Optimal design point

$$\mathbf{c}^{(2)} = \nabla f(\mathbf{x}^{(2)}) = \begin{pmatrix} 1+4\cdot(-1)+2\cdot 1.5 \\ -1+2\cdot(-1)+2\cdot 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Check stopping criterion.

$$\left\|\mathbf{c}^{(2)}\right\| = 0 \le \varepsilon$$

→Stop!



5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(1)

☑ This method updates the Hessian matrix rather than its inverse at every iteration.

- Step 1 : Estimate a starting design point $\mathbf{x}^{(0)}$.
  Choose a symmetric positive definite nxn matrix $\tilde{\mathbf{H}}^{(0)}$ as an approximation for the Hessian matrix of the objective function. In the absence of more information, let $\tilde{\mathbf{H}}^{(0)} = \mathbf{I}$. Specify a tolerance ε for the stopping criterion. Set $k = 0$, and compute the gradient vector as $\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)})$.

- Step 2 : Calculate the norm of the gradient vector as $\left\|\mathbf{c}^{(k)}\right\|$.
  If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$, then stop the iterative process. Otherwise, continue (note that Step 1 and 2 of this method and the steepest descent method are the same).

3.1 Gradient Method
5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(2)

- Step 3 : Solve the linear system of the following equation to obtain the search direction.

$$\boxed{\tilde{\mathbf{H}}^{(k)}\mathbf{d}^{(k)} = -\mathbf{c}^{(k)}}$$

This equation looks like $\mathbf{H}^{(k)}\mathbf{d}^{(k)} = -\mathbf{c}^{(k)}$ of Newton's Method, but $\tilde{\mathbf{H}}^{(k)}$ is an approximated Hessian matrix $\mathbf{H}^{(k)}$.

- Step 4 : Compute optimum step size $\alpha_k = \alpha$ to minimize $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$.

- Step 5 : Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$.

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(3)

- Step 6 : Update the matrix $\tilde{\mathbf{H}}^{(k)}$ - approximation for the Hessian matrix of the objective function - as

$$\tilde{\mathbf{H}}^{(k+1)} = \tilde{\mathbf{H}}^{(k)} + \mathbf{D}^{(k)} + \mathbf{E}^{(k)} \qquad : \quad n \times n \text{ matrix}$$

where the correction matrices $D^{(k)}$ and $E^{(k)}$ are give as follows:

$$\mathbf{D}^{(k)} = \frac{\mathbf{y}^{(k)}\mathbf{y}^{(k)^T}}{(\mathbf{y}^{(k)} \cdot \mathbf{s}^{(k)})}; \qquad \mathbf{E}^{(k)} = \frac{\mathbf{c}^{(k)}\mathbf{c}^{(k)^T}}{(\mathbf{c}^{(k)} \cdot \mathbf{d}^{(k)})};$$

$$\mathbf{s}^{(k)} = \alpha_k \mathbf{d}^{(k)} \qquad \text{: change in design}$$

$$\mathbf{y}^{(k)} = \mathbf{c}^{(k+1)} - \mathbf{c}^{(k)} \qquad \text{: change in gradient}$$

$$\mathbf{c}^{(k+1)} = \nabla f(\mathbf{x}^{(k+1)})$$

$\mathbf{d}^{(k)}$: search direction

$\alpha^{(k)}$: optimum step size

- Step 7 : Set $k = k+1$ and go to Step 2.

3.1 Gradient Method

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(4): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

■ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{H}}^{(0)} = \mathbf{I}$$

$$\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 1 + 4 \cdot 0 + 2 \cdot 0 \\ -1 + 2 \cdot 0 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Check stopping criterion.

$$\left\| \mathbf{c}^{(0)} \right\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(0)} = -(\tilde{\mathbf{H}}^{(0)})^{-1} \mathbf{c}^{(0)} = -\mathbf{I}\,\mathbf{c}^{(0)} = -\mathbf{c}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

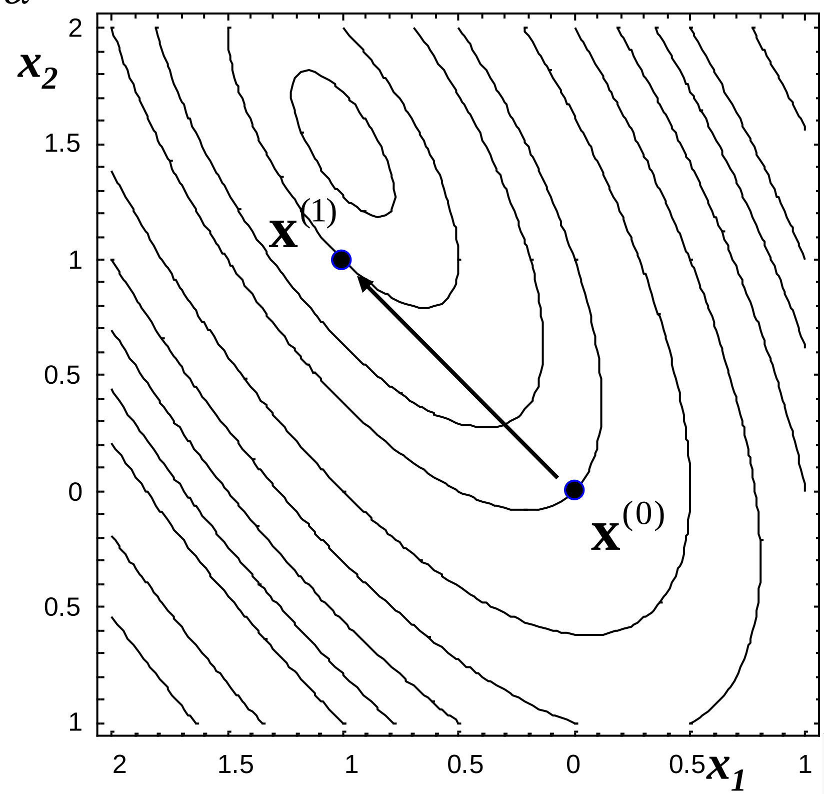
Replacing α_0 **to** α **for convenience**

Substitute $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$**,**

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \;\rightarrow\; \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(5): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2,$ **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Update the matrix $\tilde{\mathbf{H}}^{(0)}$- approximation for the Hessian matrix of the objective function - as

$$\tilde{\mathbf{H}}^{(1)} = \tilde{\mathbf{H}}^{(0)} + \mathbf{D}^{(0)} + \mathbf{E}^{(0)}$$

$$\mathbf{D}^{(0)} = \frac{\mathbf{y}^{(0)}\mathbf{y}^{(0)^T}}{\mathbf{y}^{(0)} \cdot \mathbf{s}^{(0)}}$$

$$\mathbf{s}^{(0)} = \alpha \mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)}\mathbf{y}^{(0)^T} = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)} \cdot \mathbf{s}^{(0)} = 2$$

$$= \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{E}^{(0)} = \frac{-\mathbf{c}^{(0)}\mathbf{c}^{(0)T}}{\mathbf{c}^{(0)} \cdot \mathbf{d}^{(0)}}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)}\mathbf{c}^{(0)^T} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} \cdot \mathbf{d}^{(0)} = -2$$

$$= \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

$$\tilde{\mathbf{H}}^{(1)} = \tilde{\mathbf{H}}^{(0)} + \mathbf{D}^{(0)} + \mathbf{E}^{(0)}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$
$$= \begin{pmatrix} 2.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

3.1 Gradient Method

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method(6): Example

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ **2nd Iteration: Find** $\mathbf{x}^{(2)}$

Check stopping criterion.

$$\left\|\mathbf{c}^{(1)}\right\| = \sqrt{2} > \varepsilon$$

$$\tilde{\mathbf{H}}^{(1)}\mathbf{d}^{(1)} = -\mathbf{c}^{(1)}$$

$$\tilde{\mathbf{H}}^{(1)} = \begin{pmatrix} 2.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, \quad \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 + 2\alpha \end{pmatrix}$$

Replacing α_1 **to** α **for convenience**

Substitute $\mathbf{x}^{(2)} = (-1, 1 + 2\alpha)$ **into the objective function**

$$f(\mathbf{x}^{(2)}) = 4\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$**,**

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 8\alpha - 2 = 0 \;\rightarrow\; \alpha = 0.25$$

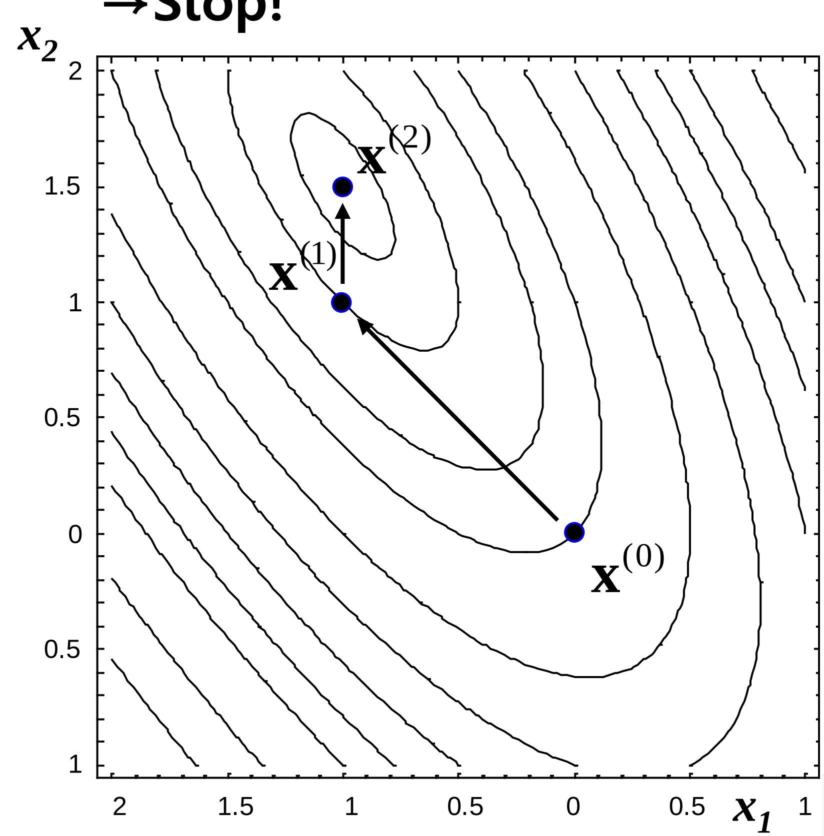
$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$ **→Optimal design point**

$$\mathbf{c}^{(2)} = \nabla f(\mathbf{x}^{(2)}) = \begin{pmatrix} 1 + 4\cdot(-1) + 2\cdot 1.5 \\ -1 + 2\cdot(-1) + 2\cdot 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Check stopping criterion.

$$\left\|\mathbf{c}^{(2)}\right\| = 0 \le \varepsilon$$

→Stop!



Ch.3 Unconstrained Optimization Method

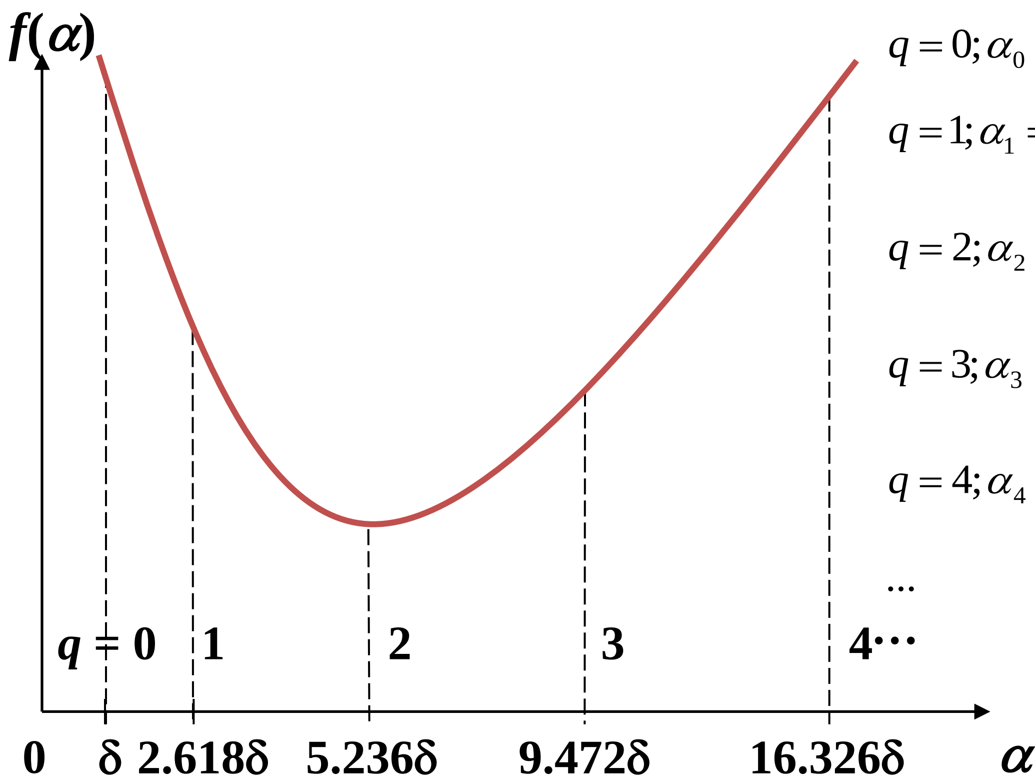
3.2 Golden Section Search Method (One Dimensional Search Method)

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

3.2 Golden Section Search method
- Phase 1: Global Search(1)

☑ **Search for the interval in which the minimum lies**

- **In the figure, starting at $q=0$, we evaluate $f(\alpha)$ at $\alpha=\delta$, where $\delta>0$ is a small number. If the value $f(\delta)$ is smaller than the value $f(0)$, we then take an increment of 1.618δ in the step size(i.e., the increment is 1.618 times the previous increment δ). (See Fibonacci sequence)**



$$q=0;\alpha_0=\delta$$

$$q=1;\alpha_1=\delta+1.618\delta=2.618\delta=\sum_{j=0}^{1}\delta(1.618)^j$$

$$q=2;\alpha_2=2.618\delta+1.618(1.618\delta)=5.236\delta=\sum_{j=0}^{2}\delta(1.618)^j$$

$$q=3;\alpha_3=5.236\delta+(1.618)^3\delta=9.472\delta=\sum_{j=0}^{3}\delta(1.618)^j$$

$$q=4;\alpha_4=9.472\delta+(1.618)^4\delta=16.326\delta=\sum_{j=0}^{4}\delta(1.618)^j$$

$$\cdots$$

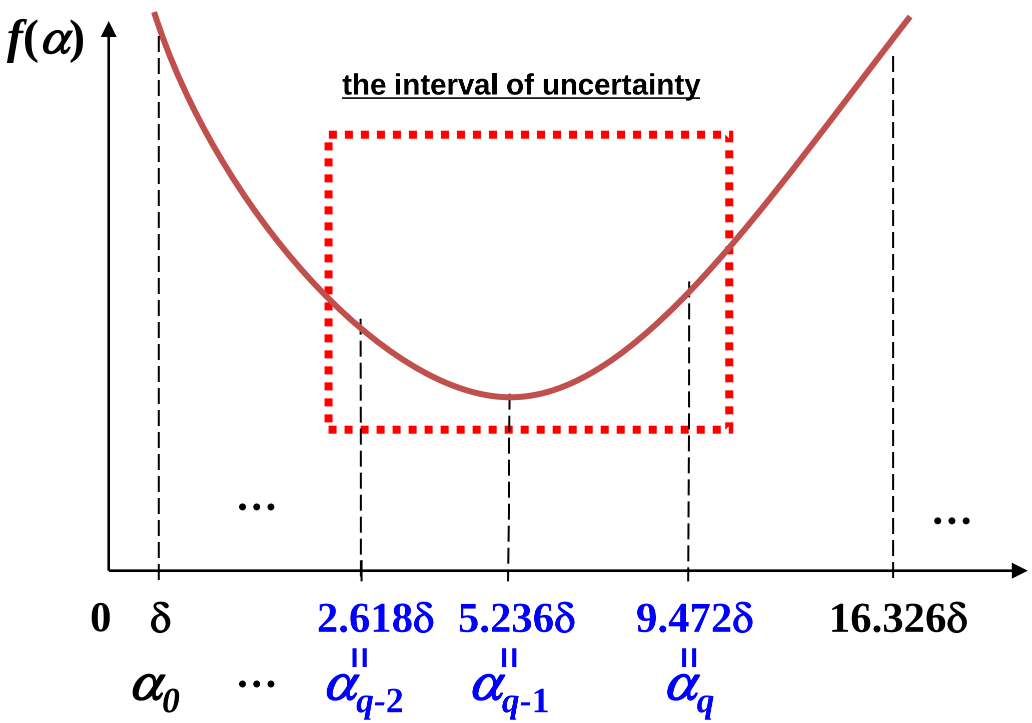
$$\boxed{\therefore \alpha_q=\sum_{j=0}^{q}\delta(1.618)^j,\ q=0,1,2,\ldots}$$

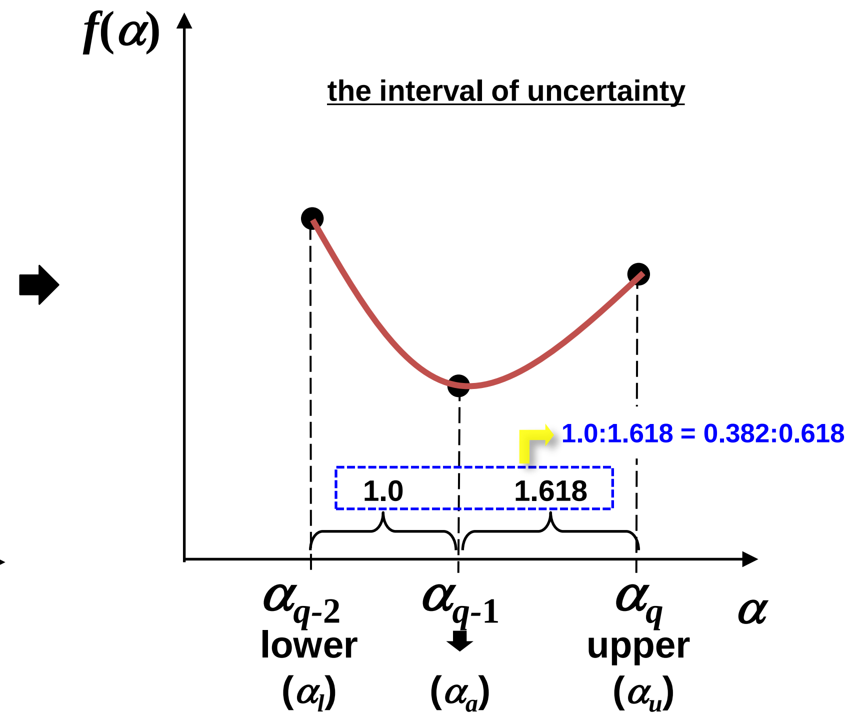
3.2 Golden Section Search method

- Phase 1: Global Search(2)

- If the function at α_{q-1} is smaller than that at the previous point α_{q-2} and the next point α_q, (i.e., $f(\alpha_{q-1}) < f(\alpha_{q-2})$, $f(\alpha_{q-1}) < f(\alpha_q)$) the minimum point lies between α_q and α_{q-2}.

(The interval in which the minimum lies is called the interval of uncertainty.)





- Therefore, upper and lower limits on the interval of uncertainty are

$$\underset{\text{upper}}{\alpha_u} \equiv \alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \quad \underset{\text{lower}}{\alpha_l} \equiv \alpha_{q-2} = \sum_{j=0}^{q-2} \delta(1.618)^j \quad , \alpha_a \equiv \alpha_{q-1} = \sum_{j=0}^{q-1} \delta(1.618)^j$$

Note: Fibonacci sequence

Fibonacci sequence defined as

$$F_0 = 0; \quad F_1 = 1; \qquad F_n = F_{n-1} + F_{n-2}, \quad n = 2,\ 3,\ \cdots$$

Any number of the Fibonacci sequence for n>1 is obtained by adding the previous two numbers, so the sequence is given as follows.

➔ **0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, …**

General term: $F_n = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}, \quad \varphi = \dfrac{1+\sqrt{5}}{2} \approx 1.6180339887\cdots$

Property: $\lim\limits_{n\to\infty} \dfrac{F_n}{F_{n-1}} = \varphi, \quad 1-\varphi = -\dfrac{1}{\varphi}$

$$\left(\because \frac{1-\varphi}{\varphi} < 1\right)$$

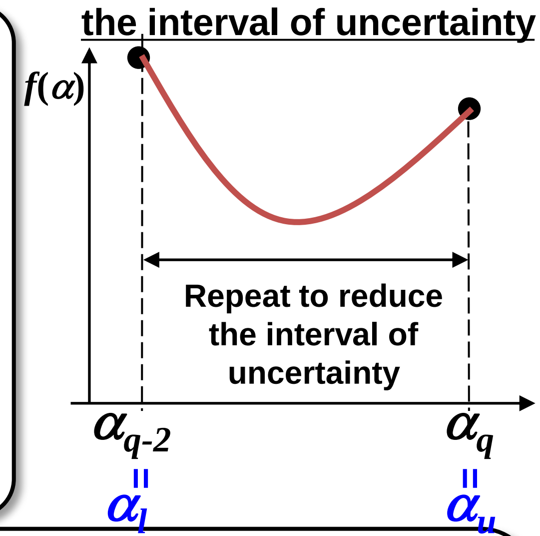
$$\lim_{n\to\infty}\frac{F_n}{F_{n-1}} = \lim_{n\to\infty}\frac{\varphi^n - (1-\varphi)^n}{\varphi^{n-1} - (1-\varphi)^{n-1}} = \lim_{n\to\infty}\frac{\dfrac{\varphi^n - (1-\varphi)^n}{\varphi^{n-1}}}{\dfrac{\varphi^{n-1} - (1-\varphi)^{n-1}}{\varphi^{n-1}}} = \lim_{n\to\infty}\frac{\varphi - (1-\varphi)\cancel{\left(\dfrac{1-\varphi}{\varphi}\right)^{n-1}}}{1 - \cancel{\left(\dfrac{1-\varphi}{\varphi}\right)^{n-1}}} = \varphi$$

3.2 Golden Section Search method

- Phase 2: Local Search(1)

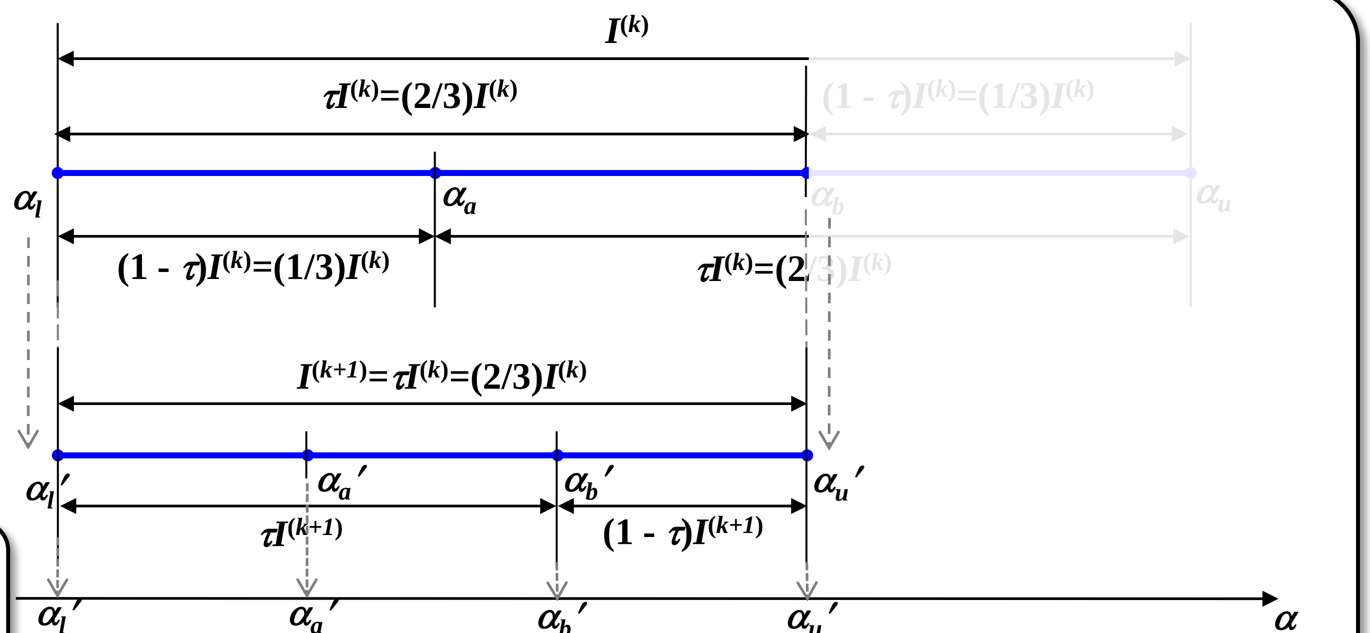
☑ **Reduction of interval of uncertainty by comparing function values at α_a and α_b**

- We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau I^{(k)}$ from either end of the interval.
- Comparing function values at α_a and α_b, either the left (α_l, α_a) or the right (α_b, α_u) portion of the interval gets discarded because the minimum cannot lie there.



< If $\tau = 2/3$ >

If $f(\alpha_a) < f(\alpha_b)$, then minimum point lies between α_l and α_b.



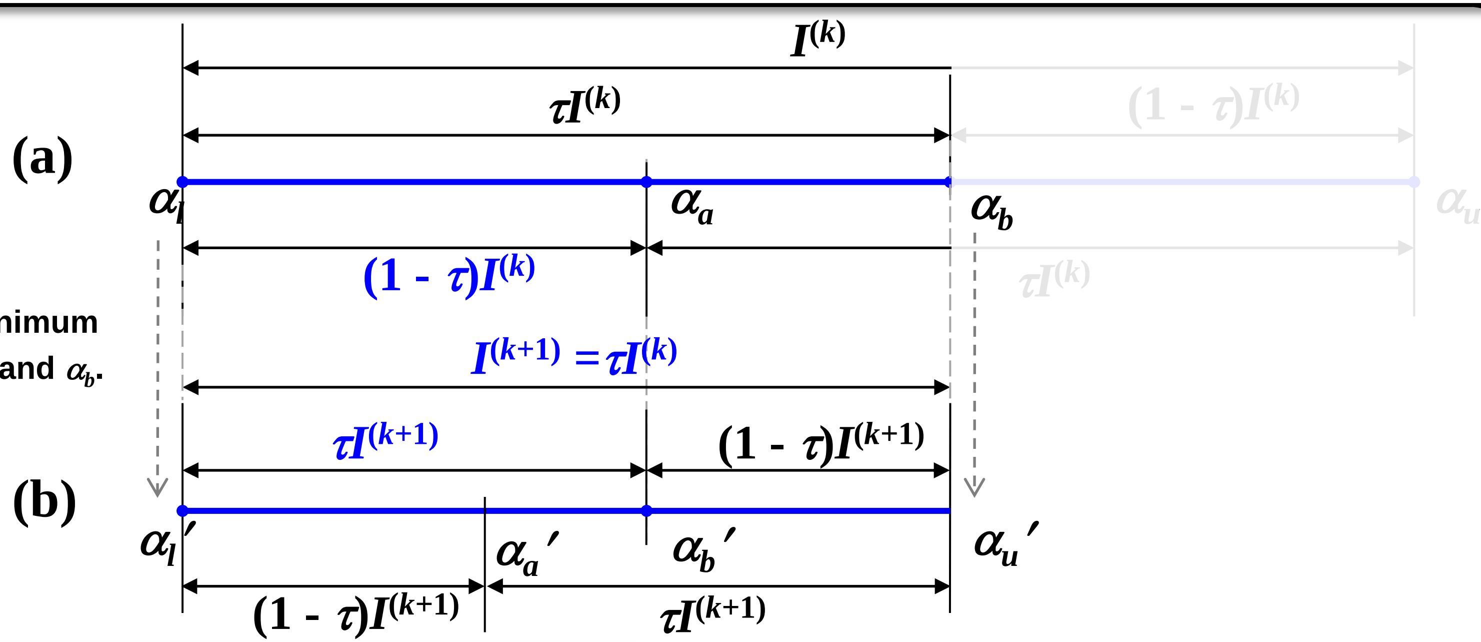
For new interval of uncertainty, we always have to compute $f(\alpha_a')$, $f(\alpha_b')$.

<Question>

Is there any method to use the previous function values?

3.2 Golden Section Search method
- Phase 2: Local Search(2)

- **Reduction of interval of uncertainty by comparing function values at α_a and α_b**
  - **We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau I^{(k)}$ from either end of the interval.**



If $f(\alpha_a) < f(\alpha_b)$, then minimum point lies between α_l and α_b.

1. $f(\alpha_b)$ will be used for the next interval of uncertainty $I^{(k+1)}$.
2. α_a is determined to equal to α_a' or α_b' of the next interval of uncertainty $I^{(k+1)}$.

3-1. Assume that α_a is equal to α_a'.

$$\alpha_a = \alpha_a'$$
$$(1-\tau)I^{(k)} = (1-\tau)I^{(k+1)}$$
$$(1-\tau)I^{(k)} = (1-\tau)\tau I^{(k)}$$
$$I^{(k)} = \tau I^{(k)}$$

Because τ=1, this assumption is wrong.

3-2. Assume that α_a is equal to α_b'.

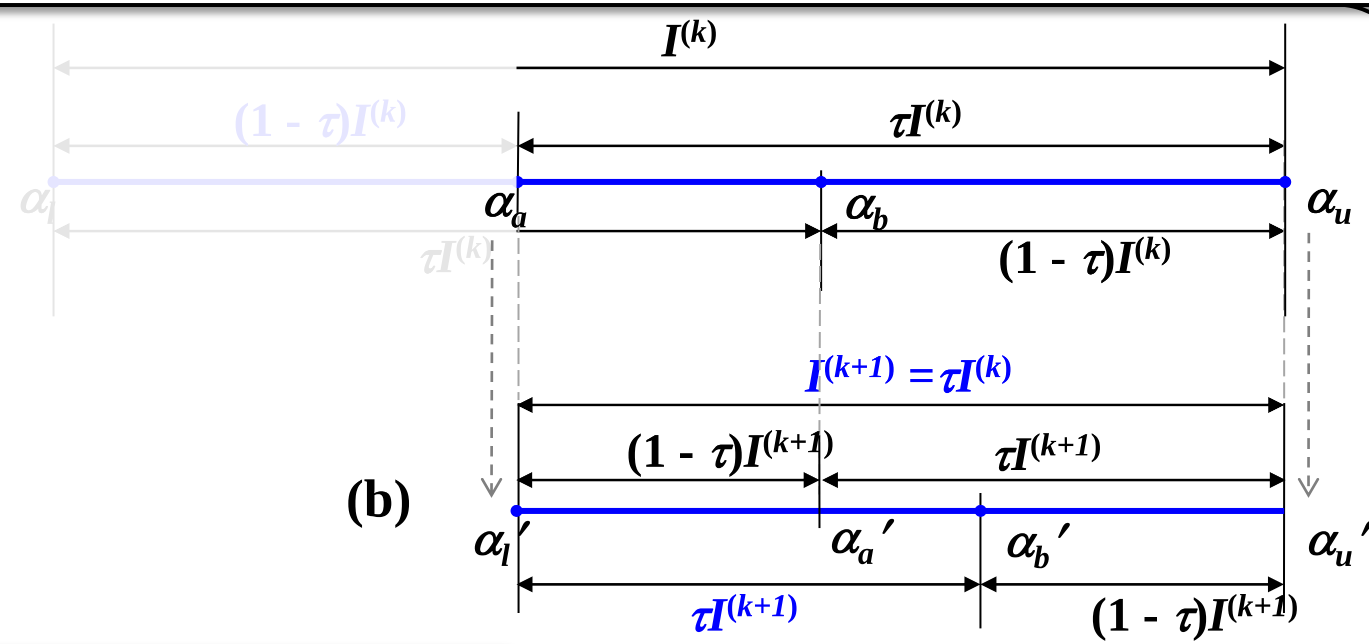
$$\alpha_a = \alpha_b'$$
$$(1-\tau)I^{(k)} = \tau I^{(k+1)}$$
$$(1-\tau)I^{(k)} = \tau \cdot \tau \cdot I^{(k)}$$
$$\tau \cdot \tau I^{(k)} - (1-\tau)I^{(k)} = 0$$
$$\tau^2 + \tau - 1 = 0$$
$$\Rightarrow \tau = 0.618, -1.618 \Rightarrow 0.618$$

3.2 Golden Section Search method

- Phase 2: Local Search(3)

- **Reduction of interval of uncertainty by comparing function values at α_a and α_b**
  - **We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau I^{(k)}$ from either end of the interval.**

(a)



If $f(\alpha_a) > f(\alpha_b)$, then minimum point lies between α_a and α_u.

1. $f(\alpha_a)$ will be used for the next interval of uncertainty $I^{(k+1)}$.
2. α_b is determined to equal to $\alpha_a{}'$ or $\alpha_b{}'$ of the next interval of uncertainty $I^{(k+1)}$.

3-1. Assume that α_b is equal to $\alpha_b{}'$.

$$\alpha_b = \alpha_b{}'$$
$$(1-\tau)I^{(k)} = (1-\tau)I^{(k+1)}$$
$$(1-\tau)I^{(k)} = (1-\tau)\tau I^{(k)}$$
$$I^{(k)} = \tau I^{(k)}$$

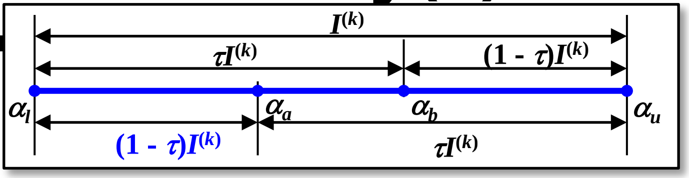
Because τ=1, this assumption is wrong.

3-2. Assume that α_b is equal to $\alpha_a{}'$.

$$\alpha_b = \alpha_a{}'$$
$$(1-\tau)I^{(k)} = \tau I^{(k+1)}$$
$$(1-\tau)I^{(k)} = \tau \cdot \tau \cdot I^{(k)}$$
$$\tau \cdot \tau I^{(k)} - (1-\tau)I^{(k)} = 0$$
$$\tau^2 + \tau - 1 = 0$$

$\Rightarrow \tau = 0.618, -1.618 \Rightarrow 0.618$

3.2 Golden Section Search method: Summary(1)



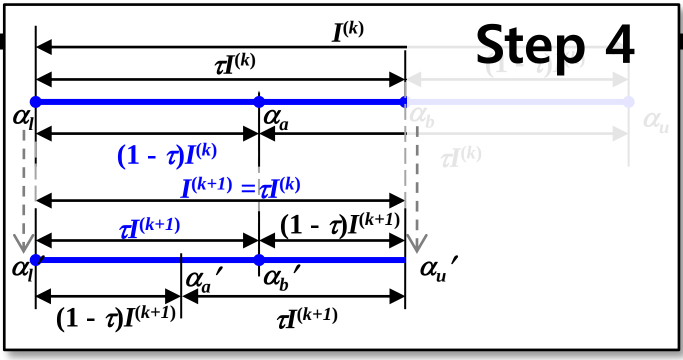
- **Step 1: For a chosen small number δ , let q be the smallest integer to satisfy $f(\alpha_{q-1}) < f(\alpha_{q-2}), f(\alpha_{q-1}) < f(\alpha_q)$ where α_q, α_{q-1} *and* α_{q-2} are calculated from $\alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \; (q = 0,1,2,\ldots)$. The upper and lower bounds on α^* (the optimum value for α) are given as follows.**

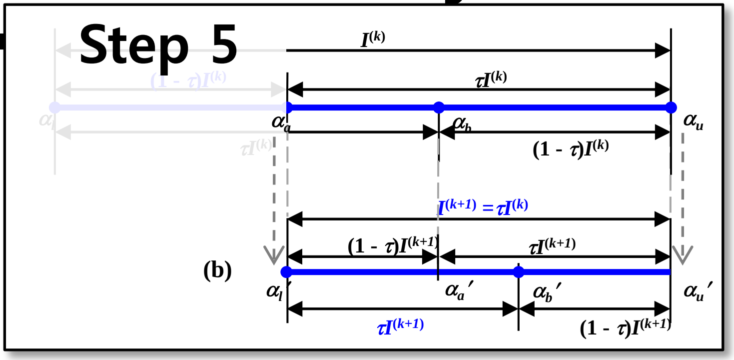
$$\alpha_u \equiv \alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \alpha_l \equiv \alpha_{q-2} = \sum_{j=0}^{q-2} \delta(1.618)^j$$

- **Step 2 : Compute $f(\alpha_a)$ and $f(\alpha_b)$ where $\alpha_a = \alpha_l + 0.382I$ and $\alpha_b = \alpha_l + 0.618I$ (interval of uncertainty $I = \alpha_u - \alpha_l$).**

- **Step 3 : Compute $f(\alpha_a)$ and $f(\alpha_b)$, and go to Step 4, Step 5 or Step 6.**

3.2 Golden Section Search method : Summary(2)



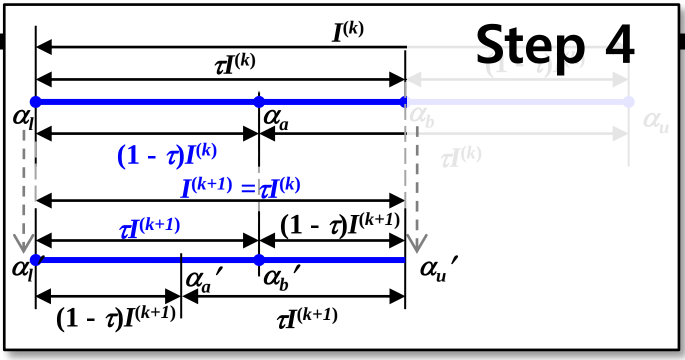


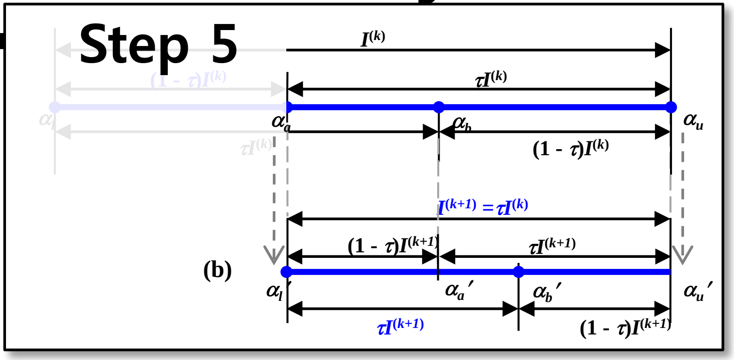
- **Step 4 : If $f(\alpha_a) < f(\alpha_b)$, then minimum point α^* lies between α_l and α_b, i.e., $\alpha_l \le \alpha^* \le \alpha_b$. The new limits for the reduced interval of uncertainty are $\alpha_l' = \alpha_l$ and $\alpha_u' = \alpha_b$. Also, $\alpha_b' = \alpha_a$. Compute $f(\alpha_a')$, where $\alpha_a' = \alpha_l' + 0.382(\alpha_u' - \alpha_l')$ and go to Step 7.**

- **Step 5 : If $f(\alpha_a) > f(\alpha_b)$, then minimum point α^* lies between α_a and α_u , i.e., $\alpha_a \le \alpha^* \le \alpha_u$. Similar to the procedure in Step 4, let $\alpha_l' = \alpha_a$ and $\alpha_u' = \alpha_u$, so that $\alpha_a' = \alpha_b$. Compute $f(\alpha_b')$, where $\alpha_b' = \alpha_l' + 0.618(\alpha_u' - \alpha_l')$ and go to Step 7.**

- **Step : If $f(\alpha_a) = f(\alpha_b)$, let $\alpha_l = \alpha_a$ and $\alpha_u = \alpha_b$ and return to Step 2.**

3.2 Golden Section Search method: Summary(3)





- **Step 7 : If the new interval of uncertainty $I' = \alpha_u' - \alpha_l'$ is small enough to satisfy a stopping criterion (i.e., $I' < \varepsilon$), let $\alpha^* = (\alpha_u' - \alpha_l')/2$ and stop. Otherwise, delete the primes(') on $\alpha_l', \alpha_a', \alpha_b'$ and α_u' and return to Step 3.**

Ch.3 Unconstrained Optimization Method

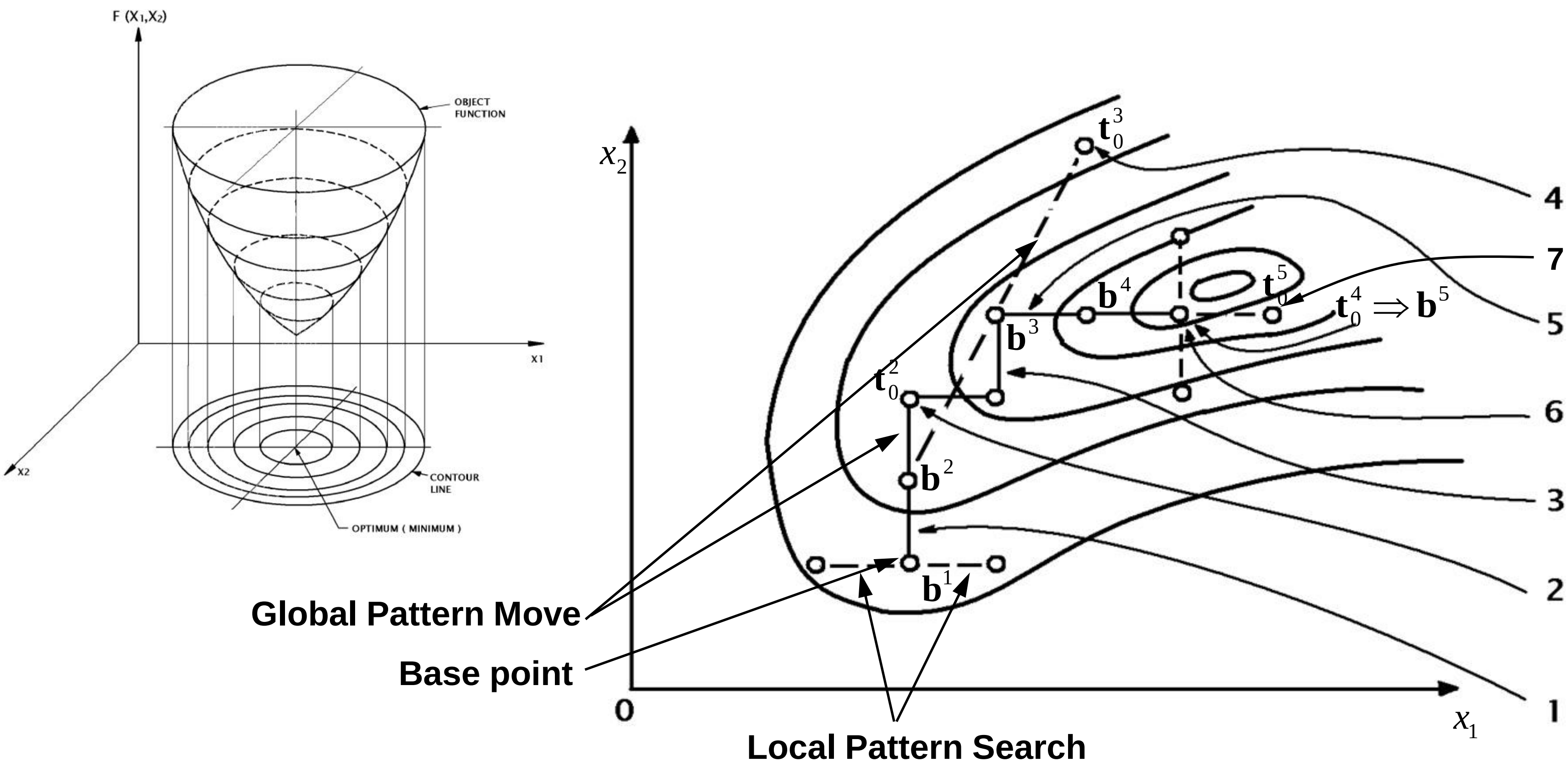
3.3 Direct Search Method

1. Hooke & Jeeves Direct Search Method
2. Nelder & Mead Simplex Method

3.3 Direct Search Method
1. Hooke & Jeeves Direct Search Method(1)

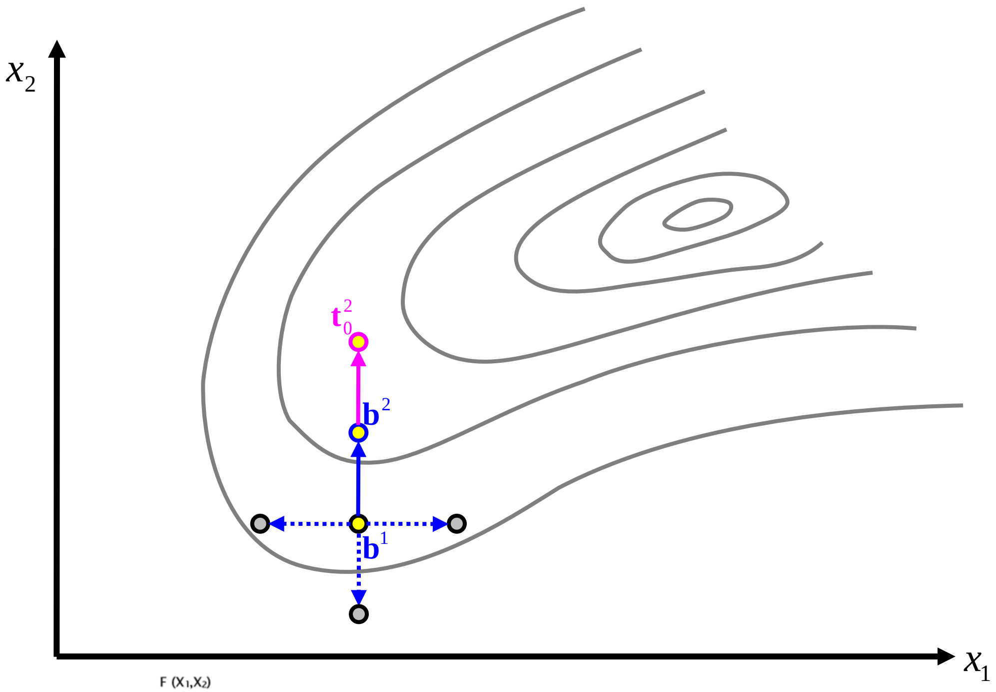
1. Base Point
2. Global Pattern Move
3. Local Pattern Search

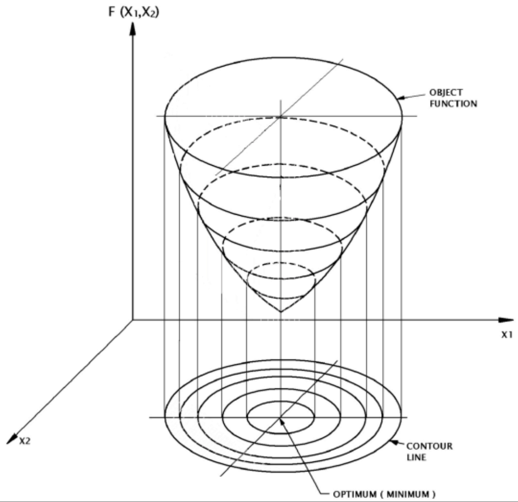
☑ This method is a sequential technique each step of which consists of two kinds of move, the **'Local Pattern Search'** at a base point and **'Global Pattern Move'** to the optimal design point.



3.3 Direct Search Method
1. Hooke & Jeeves Method(2): Example

1. Base Point
2. Global Pattern Move
3. Local Pattern Search





1. 'Local Pattern Search' at the base point b^1

- Search in x_1 direction.

  - No improvement of the value of the objective function in x_1 direction →No movement in x_1 direction

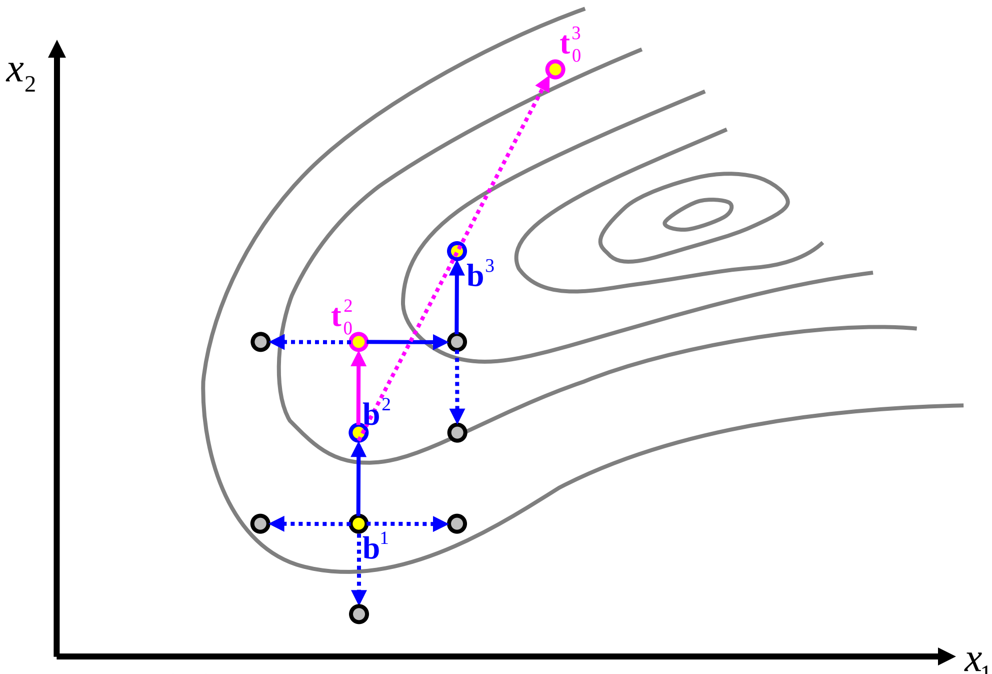
- Search in x_2 direction.

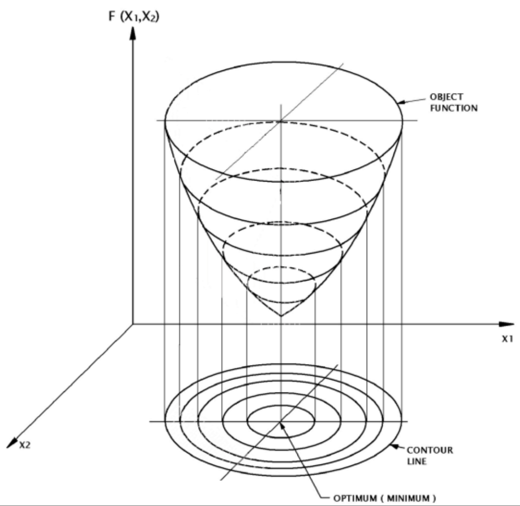
  - Improvement of the value of the objective function in x_2 direction →Movement in the positive x_2 direction

- Move and define the base point b^2.

2. 'Global Pattern Move' at the base point b^2

- Find a temporary base point t_0^2 by symmetrical displacement of b^1 to b^2.
- Because the value of the objective function at t_0^2 is better than that at b^2, do the 'Local Pattern Search' at t_0^2.





3. 'Local Pattern Search' at the temporary base point t_0^2

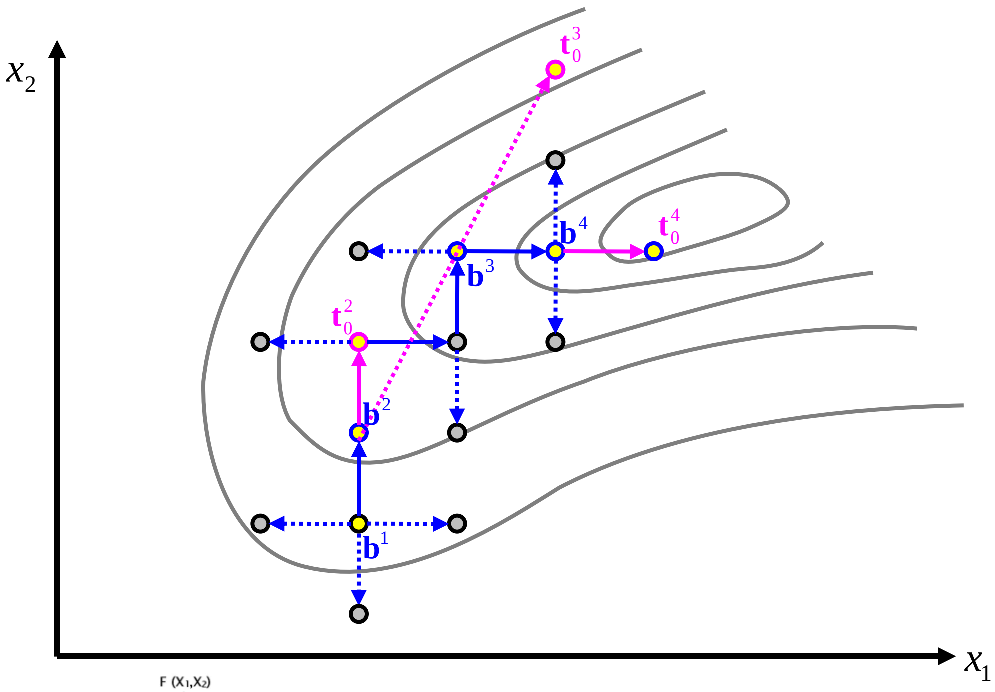
- Search in x_1 direction.
  - Improvement of the value of the objective function in x_1 direction → Movement in the positive x_1 direction
- Search in x2 direction.
  - Improvement of the value of the objective function in x2 direction → Movement in the positive x2 direction
- Move and define the base point b^3.

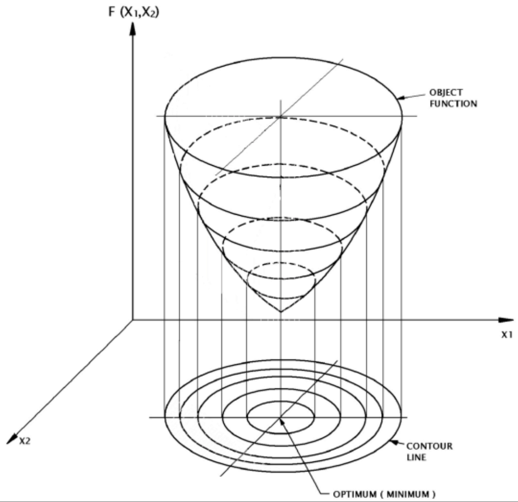
4. 'Global Pattern Move' at the base point b^3

- Find a temporary base point t_0^3 by symmetrical displacement of b^2 to b^3.
- Because the value of the objective function at t_0^3 is not better than that at b^3, perform the 'Local Pattern Search' at b^3.

3.3 Direct Search Method
1. Hooke & Jeeves Method(4)

1. Base Point
2. Global Pattern Move
3. Local Pattern Search



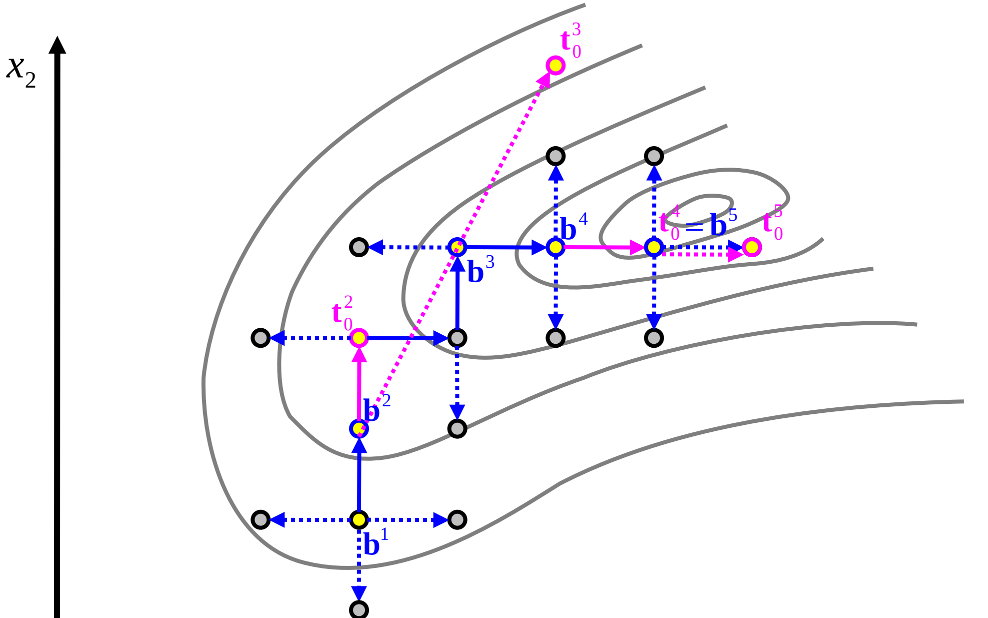


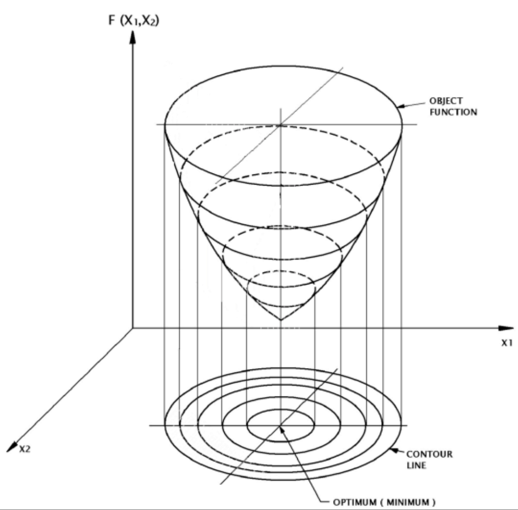
5. 'Local Pattern Search' at the base point b^3

- Search in x_1 direction.
  - Improvement of the value of the objective function in x_1 direction → Movement in the positive x_1 direction
- Search in x_2 direction.
  - No improvement of the value of the objective function in x_2 direction → No movement in x_2 direction
- Move and define the base point b^4.

6. 'Global Pattern Move' at the base point b^4

- Find a temporary base point t_0^4 by symmetrical displacement of b^3 to b^4.
- Because the value of the objective function at t_0^4 is better than that at b^4, perform the 'Local Pattern Search' at t_0^4 .





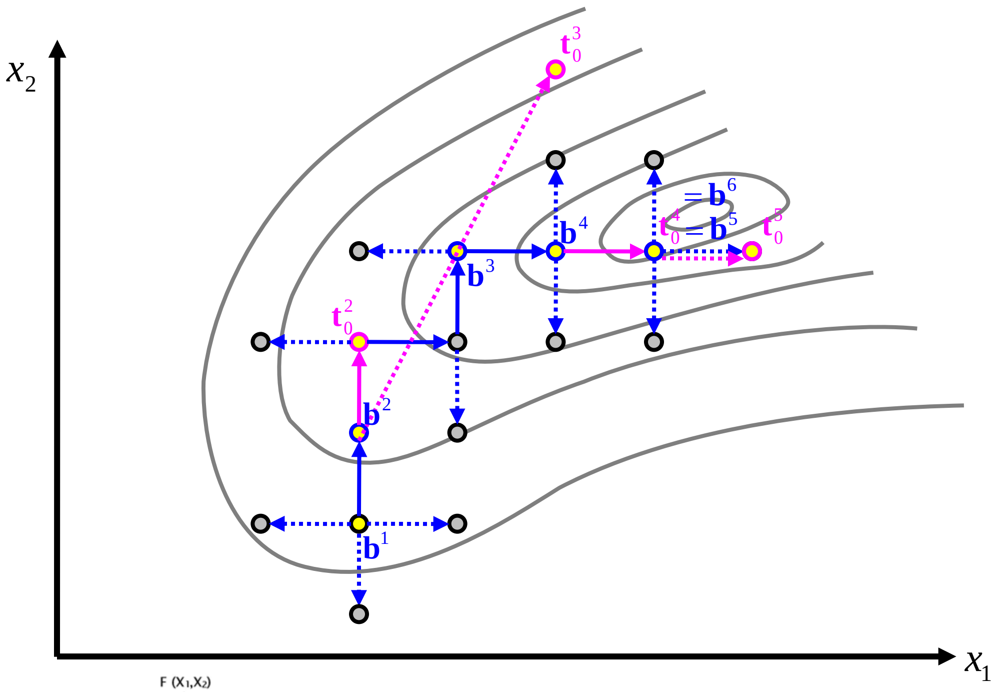
7. 'Local Pattern Search' at the temporary base point t_0^4

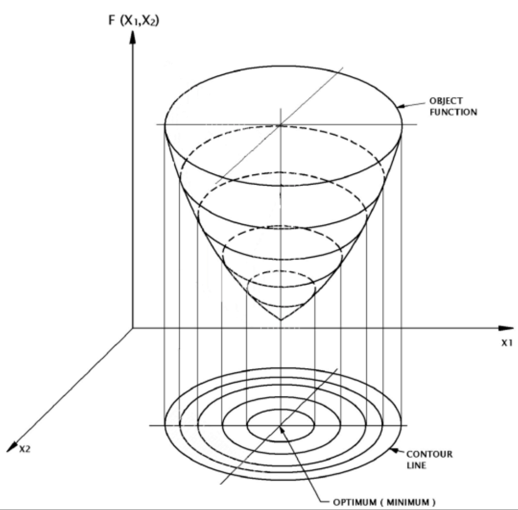
- Search in x_1 direction.
  - No improvement of the value of the objective function in x_1 direction → No movement in x_1 direction
- Search in x_2 direction.
  - No improvement of the value of the objective function in x_2 direction → No movement in x_2 direction
- Because there is no improvement of the value of the objective function in x_1 and x_2 direction, the current base point is defined as the base point b^5.

8. 'Global Pattern Move' at the base point b^5

- Find a temporary base point t_0^5 by symmetrical displacement of b^4 to b^5.
- Because the value of the objective function at t_0^5 is not better than at b^5, perform the 'Local Pattern Search' at b^5.

1. Base Point
2. Global Pattern Move
3. Local Pattern Search





9. 'Local Pattern Search' at the base point b^5

- Search in x_1 direction.
  - No improvement of the value of the objective function in x_1 direction → No movement in x_1 direction

- Search in x_2 direction.
  - No improvement of the value of the objective function in x_2 direction → No movement in x_2 in x_2 direction

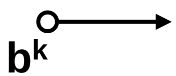
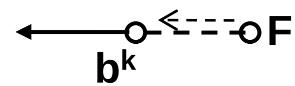
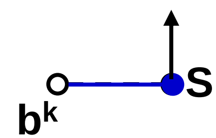
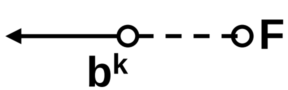
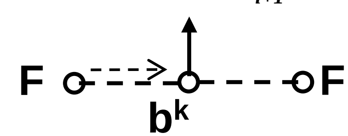
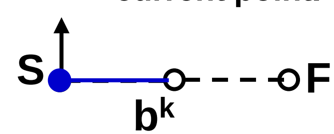
- Because there is no improvement of the value of the objective function in x1 and x2 direction, the current base point defined as base point b^6.

- Because $b^5 = b^6$, reduce the step size by half and perform the 'Local Pattern Search' at b^6.

3.3 Direct Search Method
1. Hooke & Jeeves Method(7): Rule of the 'Local Pattern Search'(1)

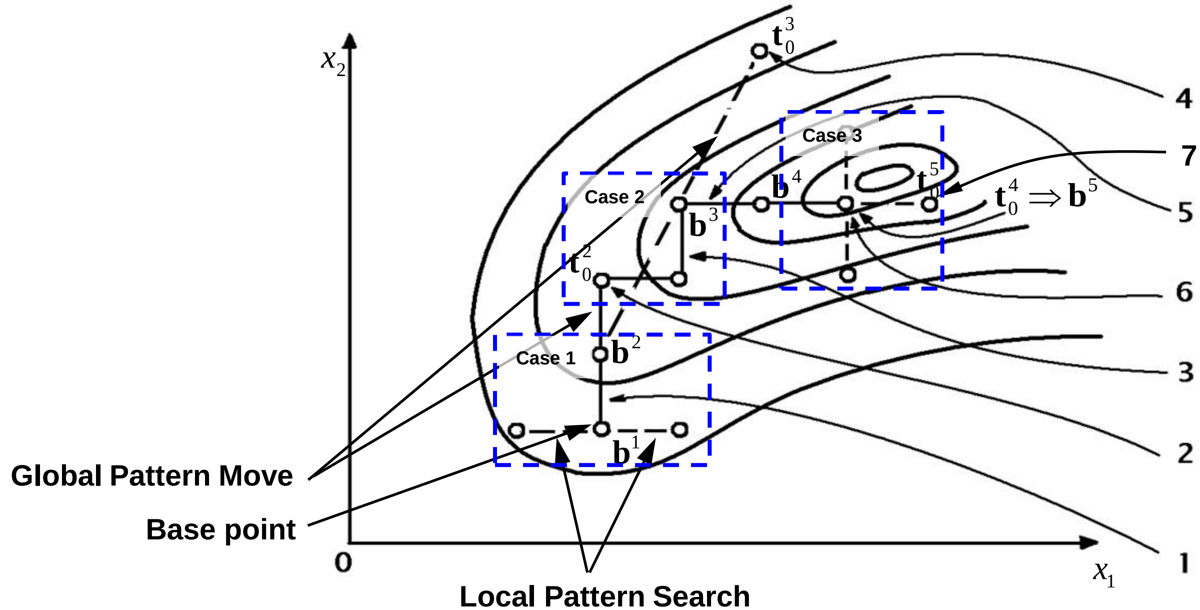
Rule of the 'Local Pattern Search'

(F: Fail, S: Success)

Case ① Search in the positive x_i direction.		
- Move the exploratory point in the positive x_i direction and evaluate the value of the objective function at that point.  bk	- If the value of the objective function is increased(Fail)	- Come back to the previous point and search in the negative x_i direction.  bk F
	- If the value of the objective function is decreased(Success)	- Search in the x_{i+1} direction at the current point.  bk S
Case ② Search in the negative x_i direction.		
- If the search in the positive x_i direction is failed, move the exploratory point in the negative x_i direction and evaluate the value of the objective function at that point.  bk F	- If the value of the objective function is increased(Fail)	- Come back to the previous point and search in x_{i+1} direction.  F bk F
	- If the value of the objective function is decreased(Success)	- Search in the x_{i+1} direction at the current point.  S bk F

- This process of the 'Local Pattern Search' is continued for $i = 1,\ldots, n$.
- After searching in x_n direction, the current point is defined as new base point b^{k+1}.

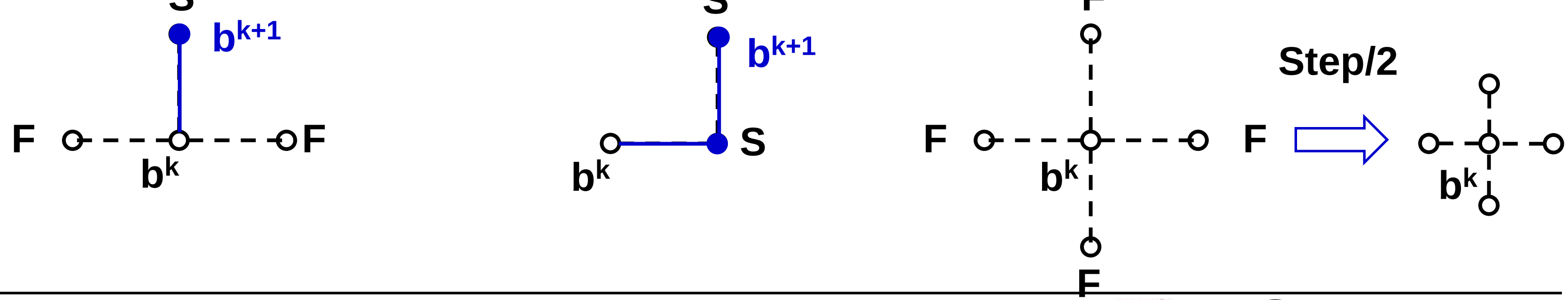
3.3 Direct Search Method
1. Hooke & Jeeves Method(8): Rule of the 'Local Pattern Search'(2)



* Super script 'k' means the number of step.

- **Rule of the Local Pattern Search**(F: Fail, S: Success)

<Case 1> <Case 2> <Case 3>



3.3 Direct Search Method
1. Hooke & Jeeves Method(9): Algorithm Summary(1)

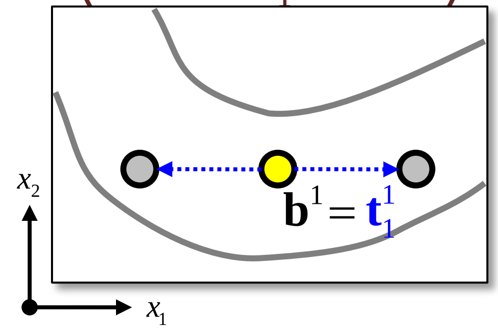
1) Local Pattern Search (Problem with n independent variables)

1. Compute the value of the objective function at the starting base point $\mathbf{b}^1$.

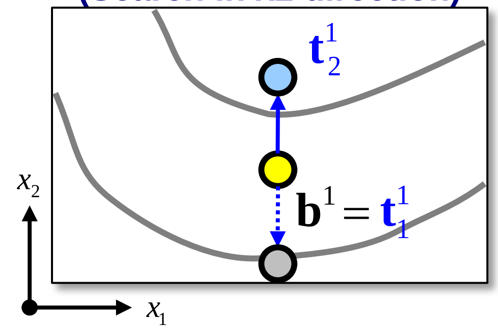
2. Compute the value of the objective function at $\mathbf{b}^1 \pm \boldsymbol{\delta}_1$, where $\boldsymbol{\delta}_1$ is input step size and a vector with n elements($\boldsymbol{\delta}_1 = [\delta_1, 0, 0, \ldots, 0]^T$). If the value of the objective function is decreased, $\mathbf{b}^1 \pm \boldsymbol{\delta}_1$ is adopted as $\mathbf{t}_1^1$(and the search is continued.

3. Compute the value of the objective function at $\mathbf{t}_1^1 \pm \boldsymbol{\delta}_2$, where $\boldsymbol{\delta}_2$ is also input step size and a vector with n elements($\boldsymbol{\delta}_2 = [0, \delta_2, 0, \ldots, 0]^T$). If the value of the function is decreased, $\mathbf{t}_1^1 \pm \boldsymbol{\delta}_2$ is adopted as $\mathbf{t}_2^1$.

Example of the 'Local Pattern Search' in the problem with two independent variables(x_1, x_2) (Search in x_1 direction)



Example of the 'Local Pattern Search' in the problem with two independent variables(x_1, x_2) (Search in x2 direction)



3.3 Direct Search Method
1. Hooke & Jeeves Method(10): Algorithm Summary(2)

1) Local Pattern Search (Problem with n independent variables)

4. After the 'Local Pattern Search' for all independent variables, new base point is defined. (new base point $b^2 = t_n^1$)

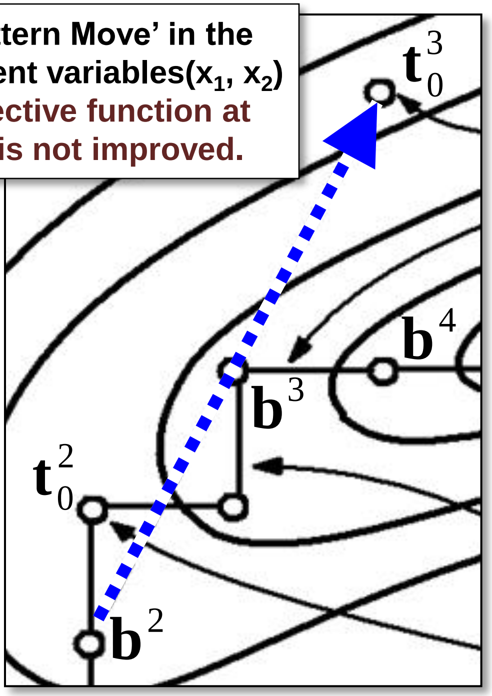
5. Perform the 'Global Pattern Move' from the previous base point along the line from the previous to current base point.

2) Global Pattern Move

1. Define the temporary base point located the same distance between the previous and current base point(obtained from 'Local Pattern Search') from the current base point ('Global Pattern Move'), and calculate the value of the objective function at this point. The temporary base point is calculated by 'Global Pattern Move' as follows.

$$\mathbf{t}_0^{k+1} = \mathbf{b}^k + 2(\mathbf{b}^{k+1} - \mathbf{b}^k) = 2\mathbf{b}^{k+1} - \mathbf{b}^k$$

Example of the 'Global Pattern Move' in the problem with two independent variables(x_1, x_2)
When the value of the objective function at the temporary base point is not improved.



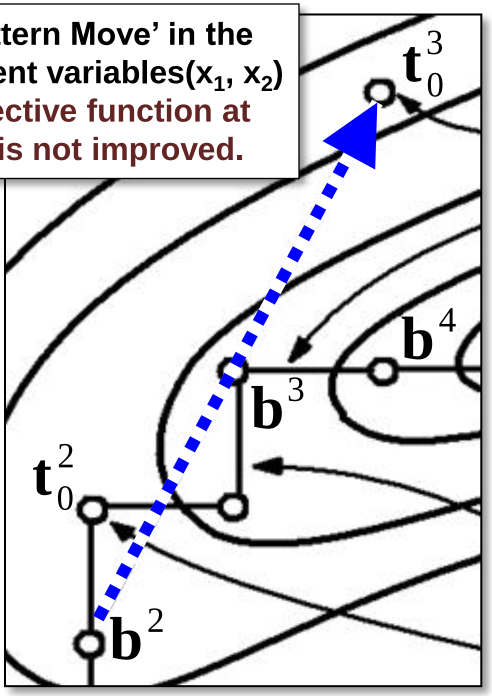
2. If the result of the temporary base point is a better point than the previous base point, perform the 'Local Pattern Search' at the temporary base point. Otherwise, come back to the previous base point and perform the 'Local Pattern Search'.

3.3 Direct Search Method
1. Hooke & Jeeves Method(12): Algorithm Summary(4)

3) Closing Conditions

1. When even this 'Local Pattern Search' fails($b^{k+1} = b^k$, there is no improvement), reduce the step sizes δ_i by halt, $\delta_i/2$, and resume the 'Local Pattern Search'.

Example of the 'Global Pattern Move' in the problem with two independent variables(x_1, x_2)
When the value of the objective function at the temporary base point is not improved.



2. If the step size δ_i is smaller than ε_i, stop the iteration and current base point is the optimal design point.

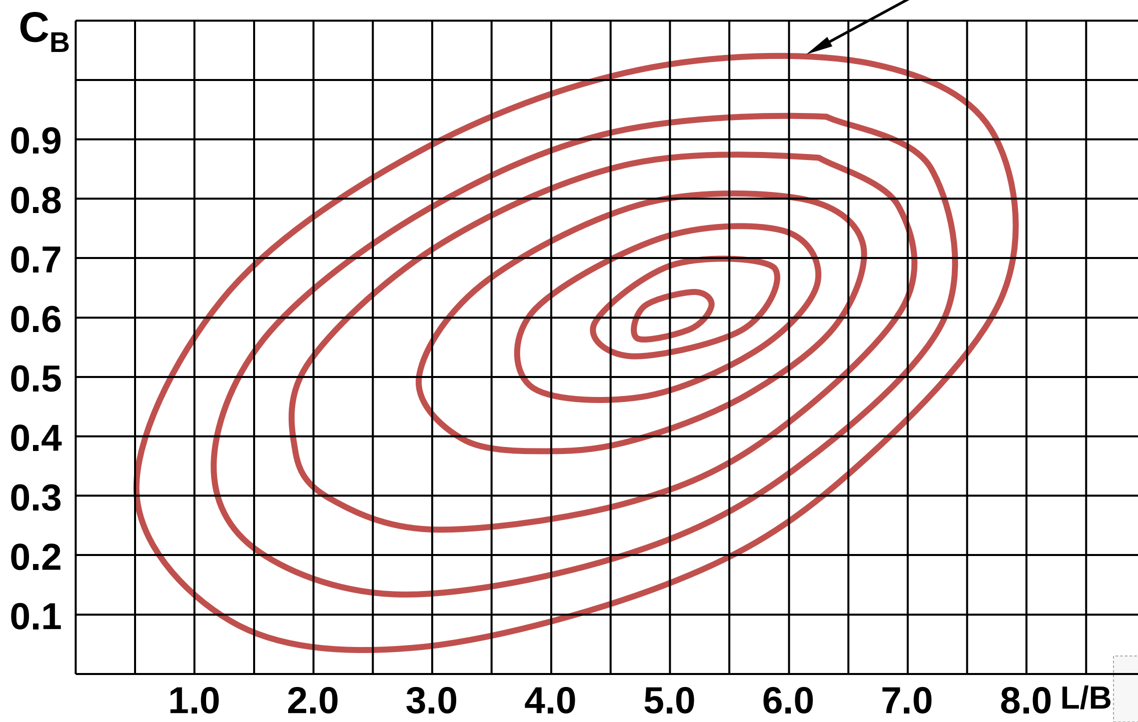
3.3 Direct Search Method
1. Hooke & Jeeves Method(13): Example

☑ **If the contour line of the objective function of shipbuilding cost with two independent variables, L/B and C_B, is given as shown in the Figure, find the optimal value of the L/B and C_B to minimize the shipbuilding cost by using the 'Hooke & Jeeves Direct Search Method' and plot the procedures in the graph.**

- **Hooke & Jeeves Direct Search Method**
  - **Starting design point: L/B = 7.0, C_B = 0.2**
  - **Step size at the starting design point: Δ(L/B) = 0.5, $\Delta(C_B)$ = 0.1**

Optimization problem with two unknown variables

Contour line of the objective function(f = const.)



3.3 Direct Search Method
1. Hooke & Jeeves Method(14): Example

$$x_1 = L/B,\ \ x_2 = C_B$$

- Iteration 1 : Local Pattern Search 1

$\mathbf{b}^0 = (7,\ \ 0.2),\ \ \Delta x_1 = 0.5,\ \ \Delta x_2 = 0.1,$

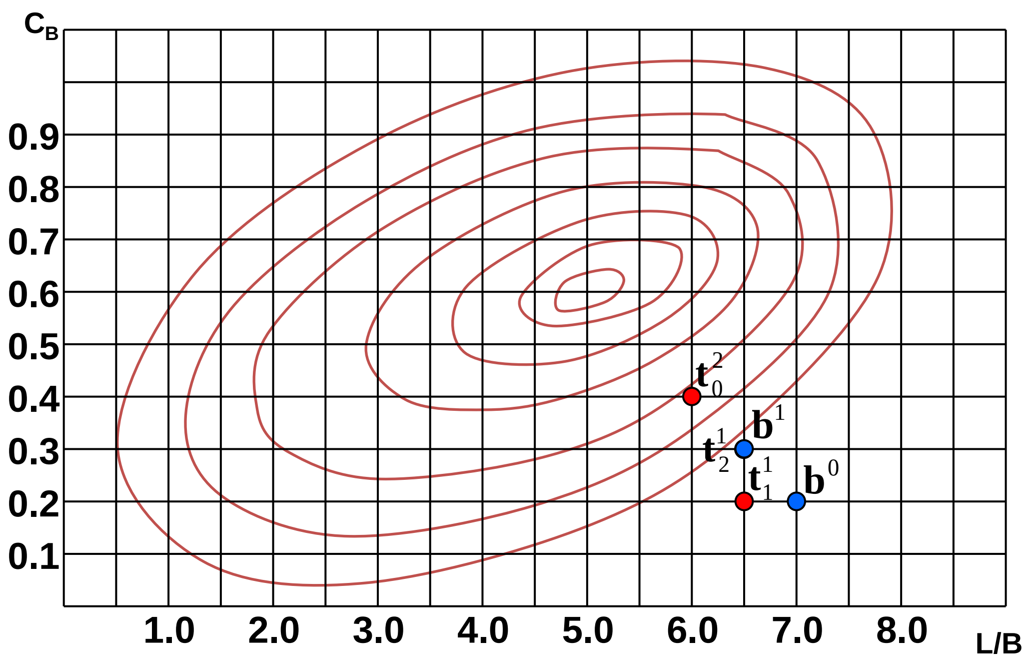
$\mathbf{t}_0^1 = \mathbf{b}^0$

Search from $\mathbf{t}_0^1$ in $-x_1$ direction $\rightarrow \mathbf{t}_1^1 = (6.5,\ \ 0.2)$

Search from $\mathbf{t}_1^1$ in $+x_2$ direction $\rightarrow \mathbf{t}_2^1 = (6.5,\ 0.3)$

Because the value of the objective function at $\mathbf{t}_2^1$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^1 = \mathbf{t}_2^1$$



- Iteration 2 : Global Pattern Move 1

Define the temporary base point by using $\mathbf{b}^0$ and $\mathbf{b}^1$

$\rightarrow \mathbf{t}_0^2 = (6,\ \ 0.4)$

Because the value of the objective function at $\mathbf{t}_0^2$ is improved, perform the 'Local Pattern Search' at this point.

3.3 Direct Search Method
1. Hooke & Jeeves Method(15): Example

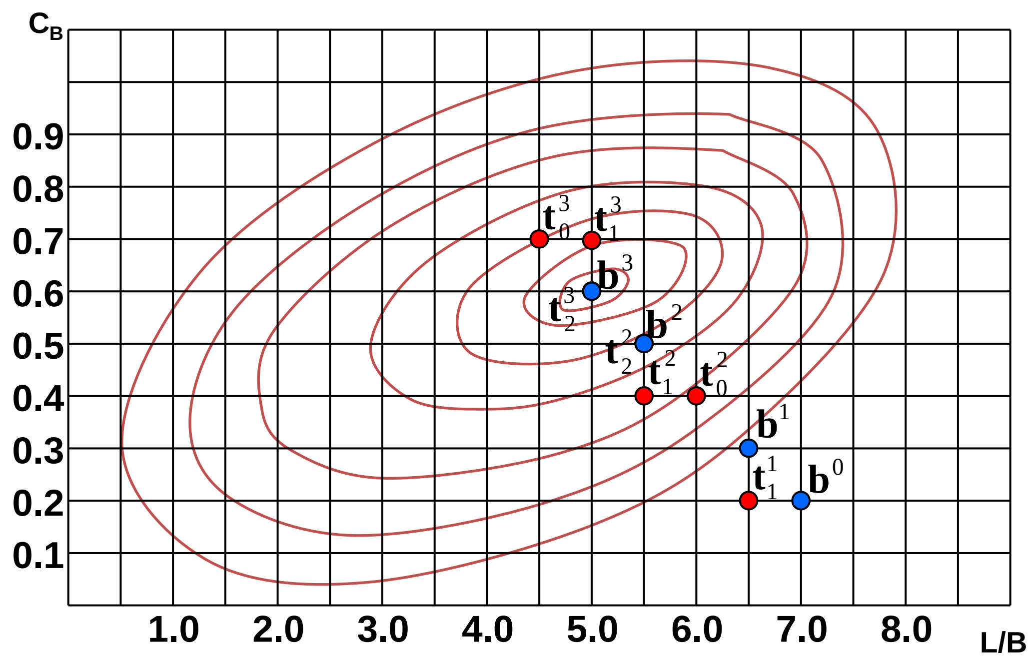
- Iteration 3 : Local Pattern Search 2

Search from $\mathbf{t}_0^2$ in $-x_1$ direction $\rightarrow \mathbf{t}_1^2 = (5.5,\ \ 0.4)$

Search from $\mathbf{t}_1^2$ in $+x_2$ direction $\rightarrow \mathbf{t}_2^2 = (5.5,\ \ 0.5)$

Because the value of the objective function at $\mathbf{t}_2^2$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^2 = \mathbf{t}_2^2$$



- Iteration 4 : Global Pattern Move 2

Define the temporary base point by using $\mathbf{b}^1$ and $\mathbf{b}^2$

$\rightarrow \mathbf{t}_0^3 = (4.5,\ \ 0.7)$

- Iteration 5 : Local Pattern Search 3

Search from $\mathbf{t}_0^3$ in $+x_1$ direction $\rightarrow \mathbf{t}_1^0 = (5,\ \ 0.7)$

Search from $\mathbf{t}_1^3$ in $-x_2$ direction $\rightarrow \mathbf{t}_2^3 = (5,\ \ 0.6)$

Because the value of the objective function at $\mathbf{t}_2^2$ $\mathbf{t}_2^3$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^3 = \mathbf{t}_2^3$$

3.3 Direct Search Method
1. Hooke & Jeeves Method(16): Example

- Iteration 6 : Global Pattern Move 3

  Define the temporary base point by using $\mathbf{b}^2$ and $\mathbf{b}^3$

  $\rightarrow \mathbf{t}_0^4 = (4.5,\ \ 0.7)$

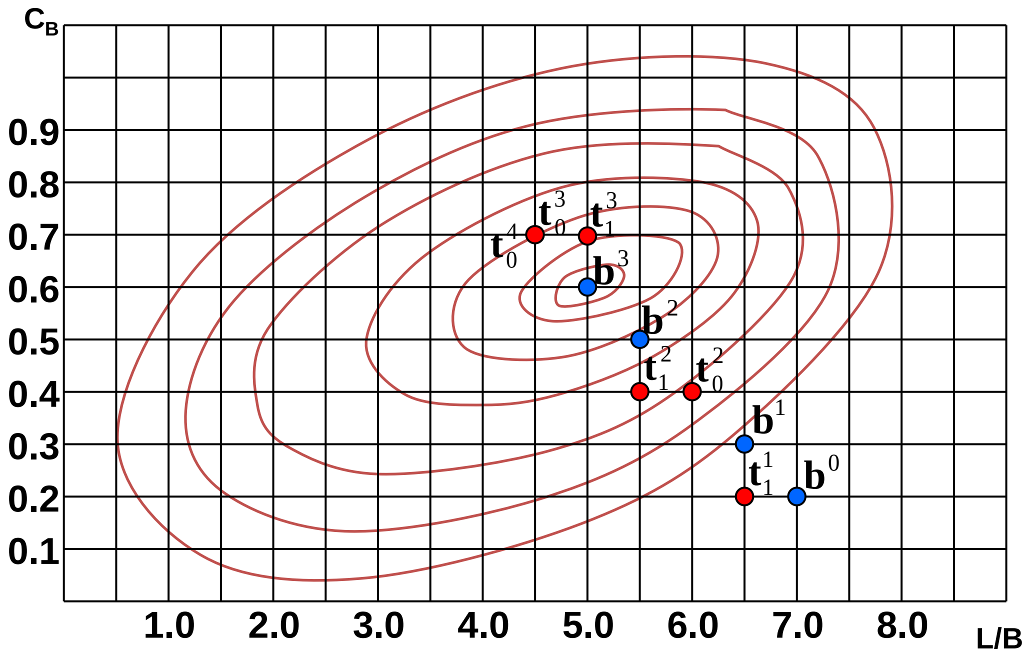
  Because the value of the objective function at $\mathbf{t}_0^4$ is not improved,

  $$\mathbf{t}_0^4 = \mathbf{b}^3$$

- Iteration 7 : Local Pattern Search 4

  Because there is no improvement of the value of the objective function from the temporary base design point $\mathbf{t}_0^4$ in x_1 direction and x_2 direction,

  $$\mathbf{t}_2^4 = \mathbf{t}_1^4 = \mathbf{t}_0^4$$



- Iteration 8 : Global Pattern Move 4

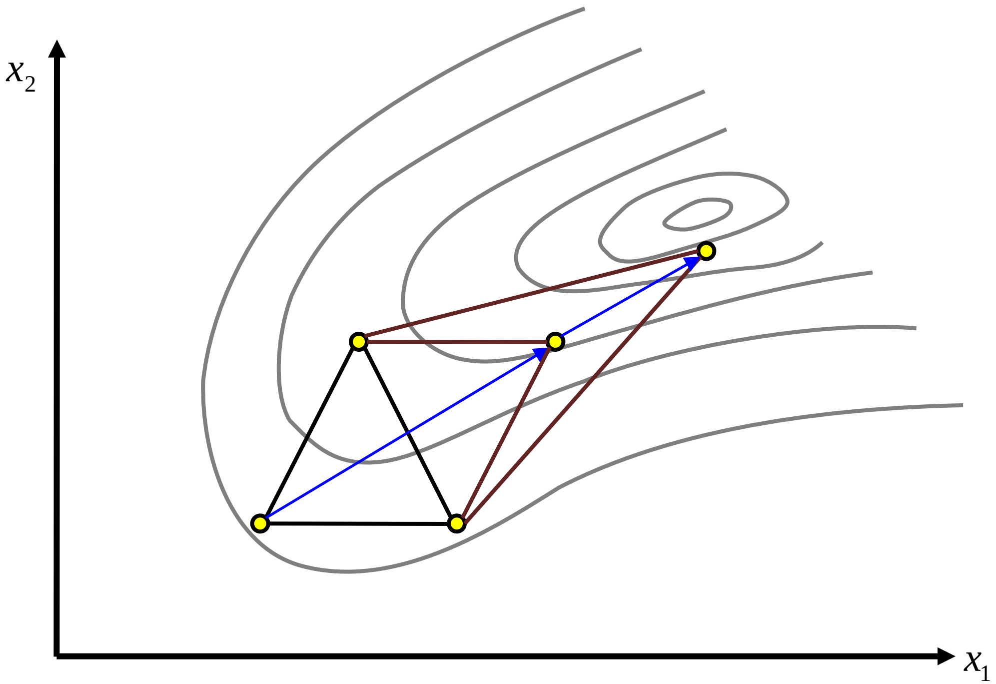
  $\mathbf{b}^4 = \mathbf{b}^3 \rightarrow\ \ \Delta x_1 = 0.25,\ \ \Delta x_2 = 0.05,$

  $\mathbf{t}_0^5 = \mathbf{b}^4$

- Iteration 9 : Stopping the iteration of search

  Because there is no improvement of the value of the objective function from base design point $(x_1,\ x_2) = (L/B,\ C_B) = (5.0,\ 0.6)$ in x_1 direction and x_2 direction by performing the 'Local Pattern Search' and 'Global Pattern Move', the optimal design point is $L/B = 5.0,\ \ C_B = 0.6$.

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(1)



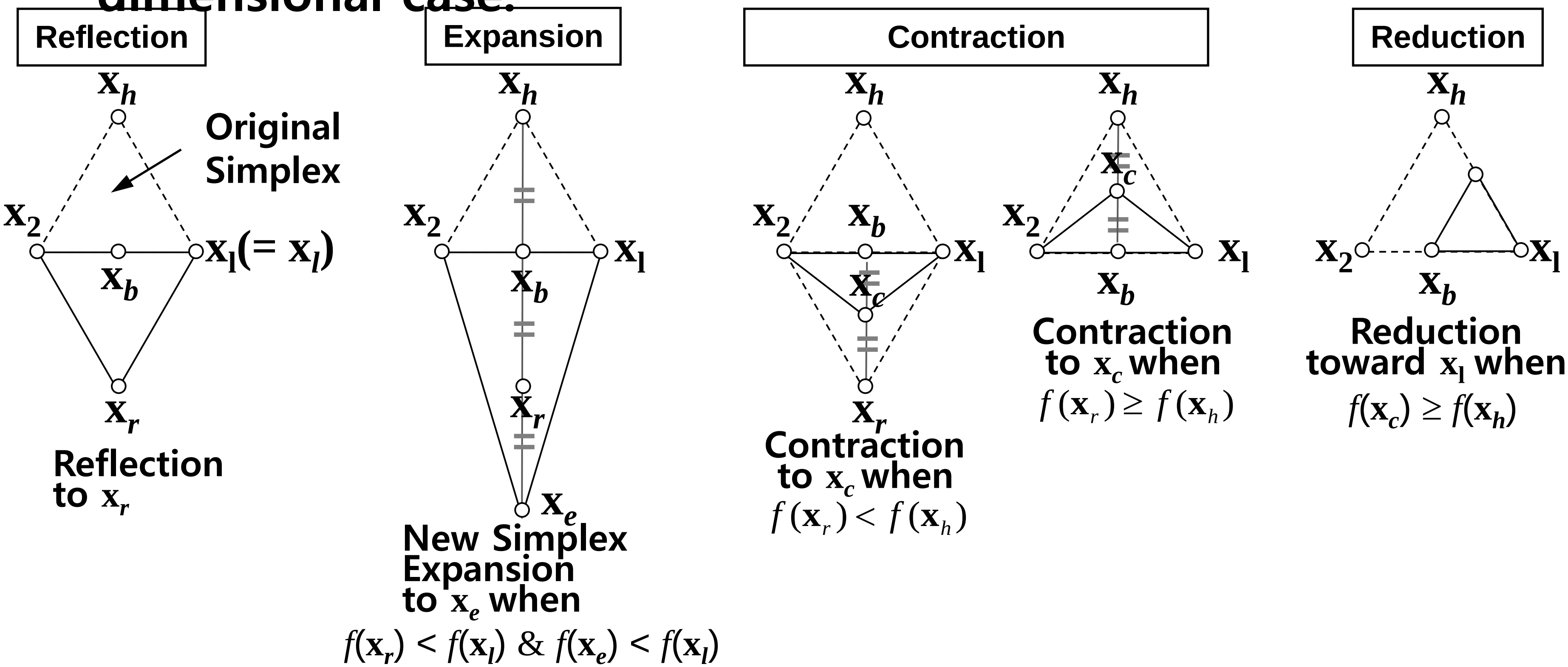
1. **This method uses n+1 points in the function of n design variables.**
   (ex) If the number of the design variables is two, this method use three points.)

2. **The simplex is reflected in the direction where the value of the objective function is improved.**

3. **If the value of the objective function is improved, the simplex is expanded. Otherwise, the simplex is reduced.**

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(2)

☑ **This method is used to find optimal design point by successively reflecting, expanding, contracting and reducing the simplex with (n+1) corners in the function of n design variables. Following figure shows an example of 2-dimensional case.**



$\mathbf{x}_h$: Simplex point having the largest objective function value
$\mathbf{x}_b$: Center point between $\mathbf{x}_1$ and $\mathbf{x}_2$

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(3)

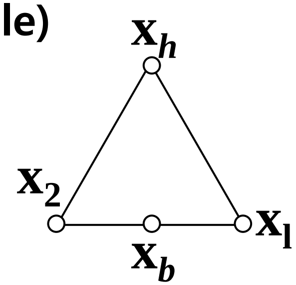
☑ **Step 1 : Calculate the value of the objective function f at the $n+1$ corners of the simplex.**

☑ **Step 2 : Establish the corners which yield the highest, $\mathbf{x}_h$, and lowest, $\mathbf{x}_l$, $f(\mathbf{x})$ in the current simplex.**

☑ **Step 3 : Calculate the value of the objective function f at the centroid($\mathbf{x}_b$) of all $\mathbf{x}_i$ except $\mathbf{x}_h$, i.e.,**

$$\mathbf{x}_b = \frac{1}{n}\sum_{i=1}^{n+1}\mathbf{x}_i \text{(with } \mathbf{x}_h \text{ excluded)}$$

Example)

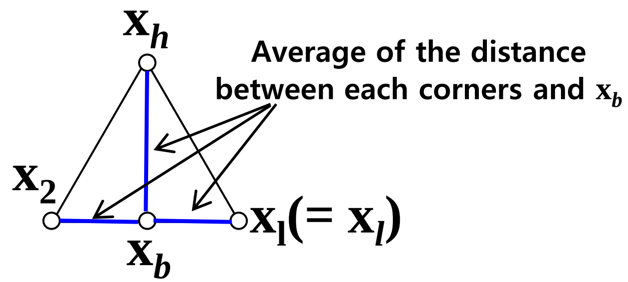


$$\mathbf{x}_b = \frac{\mathbf{x}_1 + \mathbf{x}_2}{2}$$

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(4)

☑ **Step 4 : Test stopping condition:**

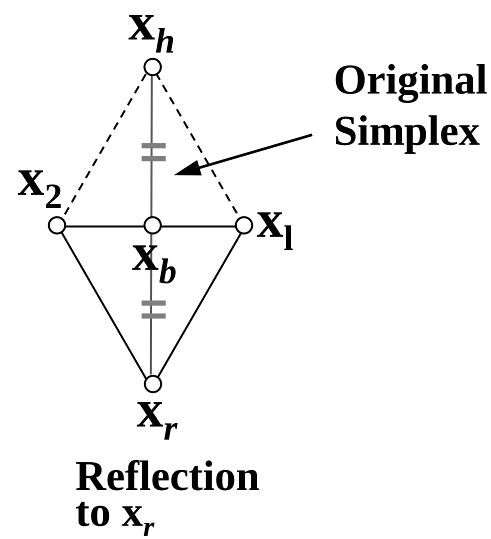
$$\left\{\frac{1}{n+1}\sum_{i=1}^{n+1}[f(\mathbf{x}_i)-f(\mathbf{x}_b)]^2\right\}^{1/2}\le\varepsilon$$



- **If the stopping condition is satisfied, stop and return $f(\mathbf{x}_l)$ as minimum. Otherwise, continue.**

☑ **Step 5 : Reflection**

- **Reflect $\mathbf{x}_h$ through $\mathbf{x}_b$ to give $\mathbf{x}_r = 2\mathbf{x}_b - \mathbf{x}_h$.**
  Calculate the value of the objective function f at $\mathbf{x}_r$ and change the simplex as following conditions.

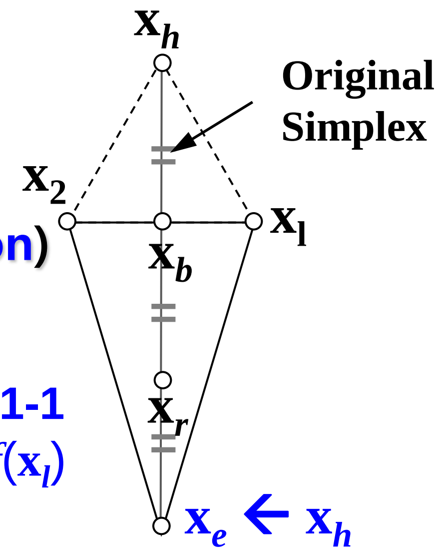


☑ **Step 6 : Expansion**

- **Step 6-1 : If $f(\mathbf{x}_r) < f(\mathbf{x}_l)$, reflect $\mathbf{x}_b$ through $\mathbf{x}_r$ to give $\mathbf{x}_e = 2\mathbf{x}_r - \mathbf{x}_b$. And then, calculate $f(\mathbf{x}_e)$ and compare $f(\mathbf{x}_e)$ and $f(\mathbf{x}_l)$.**

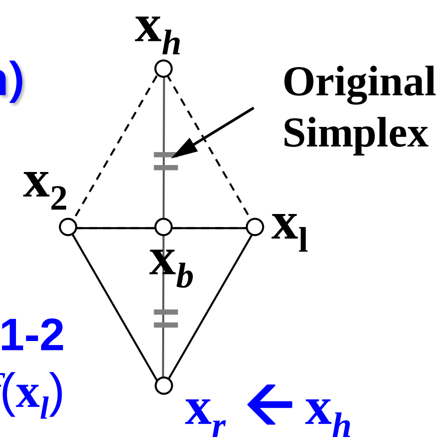
  - **Step 6-1-1 : If $f(\mathbf{x}_e) < f(\mathbf{x}_l)$, replace $\mathbf{x}_h$ by $\mathbf{x}_e$(expansion) and return to Step 2.**

➡ Step 6-1-1
$f(\mathbf{x}_e) < f(\mathbf{x}_l)$



  - **Step 6-1-2 : If $f(\mathbf{x}_e) \geq f(\mathbf{x}_l)$, replace $\mathbf{x}_h$ by $\mathbf{x}_r$(reflection) and return to Step 2.**

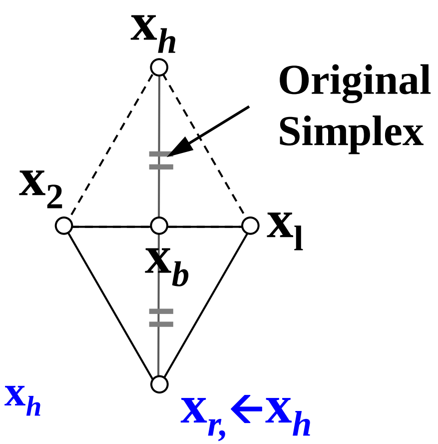
➡ Step 6-1-2
$f(\mathbf{x}_e) \geq f(\mathbf{x}_l)$



☑ **Step 6 : Expansion**

■ **Step 6-2 : If $f(\mathbf{x}_r) \geq f(\mathbf{x}_l)$,**

- **Step 6-2-1 : test $f(\mathbf{x}_r) < f(\mathbf{x}_i)$ for all $\mathbf{x}_i$ except $\mathbf{x}_h$.**
  If true, replace $\mathbf{x}_h$ by $\mathbf{x}_r$(reflection)
  and return to Step 2.



➡ **Step 6-2-1**
For all $\mathbf{x}_i$ except $\mathbf{x}_h$
$f(\mathbf{x}_r) < f(\mathbf{x}_i)$

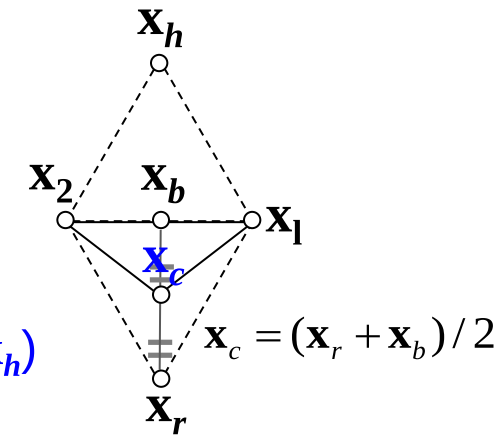
- **Step 6-2-2 : If false, continue.**

☑ **Step 7 : Contraction**

■ **Step 7-1 : If $f(\mathbf{x}_r) < f(\mathbf{x}_h)$,**

calculate the value of the objective function f at $\mathbf{x}_c = (\mathbf{x}_r + \mathbf{x}_b)/2$.

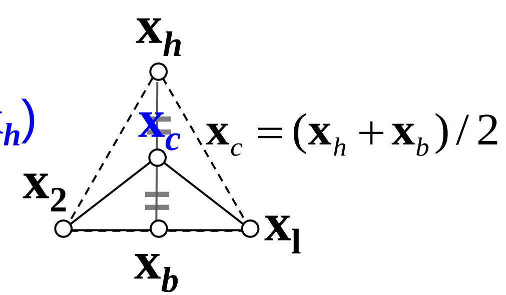
➡ **Step 7-1**
$f(\mathbf{x}_r) < f(\mathbf{x}_h)$



■ **Step 7-2 : If $f(\mathbf{x}_r) \geq f(\mathbf{x}_h)$,**

calculate the value of the objective function f at $\mathbf{x}_c = (\mathbf{x}_h + \mathbf{x}_b)/2$.

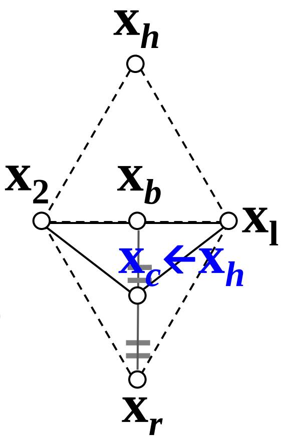
➡ **Step 7-2**
$f(\mathbf{x}_r) \geq f(\mathbf{x}_h)$



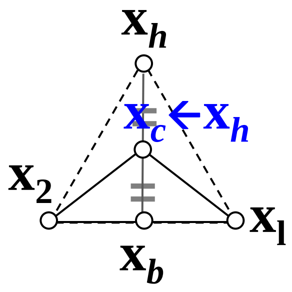
☑ **Step 8 : Reduction**

- **Step 8-1 : If $f(\mathbf{x}_c) < f(\mathbf{x}_h)$,**
  replace $\mathbf{x}_h$ by $\mathbf{x}_c$ (contraction)
  and return to Step 2.

➡ Step 8-1
$f(\mathbf{x}_c) < f(\mathbf{x}_h)$

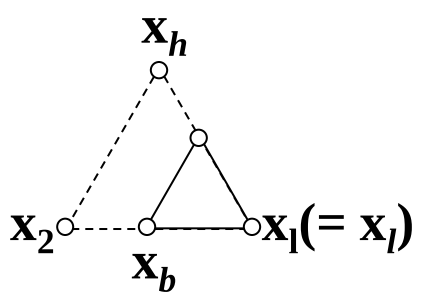


or



- **Step 8-2 : If $f(\mathbf{x}_c) \geq f(\mathbf{x}_h)$,**
  reduce the simplex toward $\mathbf{x}_l$ using $\mathbf{x}_i = (\mathbf{x}_i + \mathbf{x}_l)/2$
  (reduction) and return to Step 2.

➡ Step 8-2
$f(\mathbf{x}_c) \geq f(\mathbf{x}_h)$



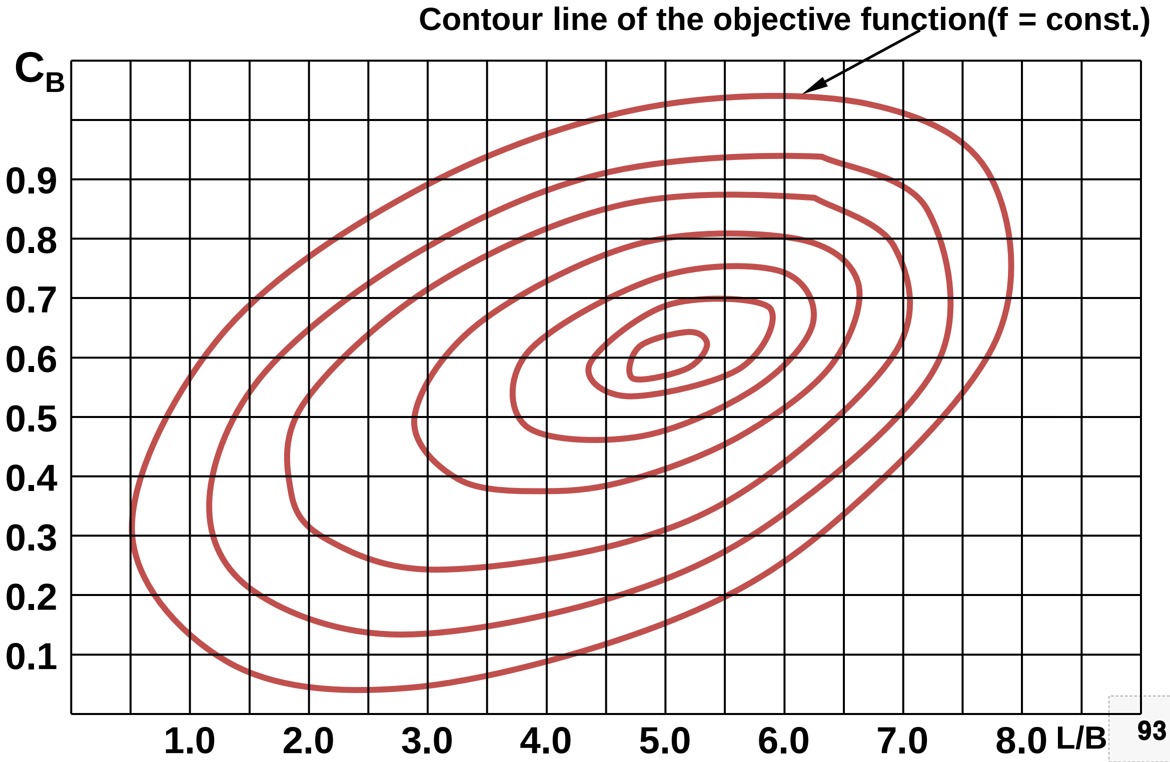
Reduction toward $\mathbf{x}_l$

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(9): Example

☑ **If the contour line of the objective function of shipbuilding cost with two independent variables, L/B and C_B, is given as shown in Fig, find the value of the L/B and C_B to minimize the shipbuilding cost by using the 'Nelder & Mead Simplex Method' and plot the procedures in the graph.**

- **Nelder & Mead Simplex Method**
  - **Starting corners of the simplex: (L/B, CB) = (1, 0.1), (1.5, 0.1), (1.5, 0.2)**
  - **Stopping criterion: 0.01**

Optimization problem with two unknown variables



3.3 Direct Search Method
2. Nelder & Mead Simplex Method(10): Example

$$x_1 = L/B,\ x_2 = C_B$$

Triangle 1 : $x_1,\ x_2,\ x_3$

Iteration 1) Because x_2 is x_h, reflect x_2 through the center between x_1 and x_3. $\rightarrow x_r$

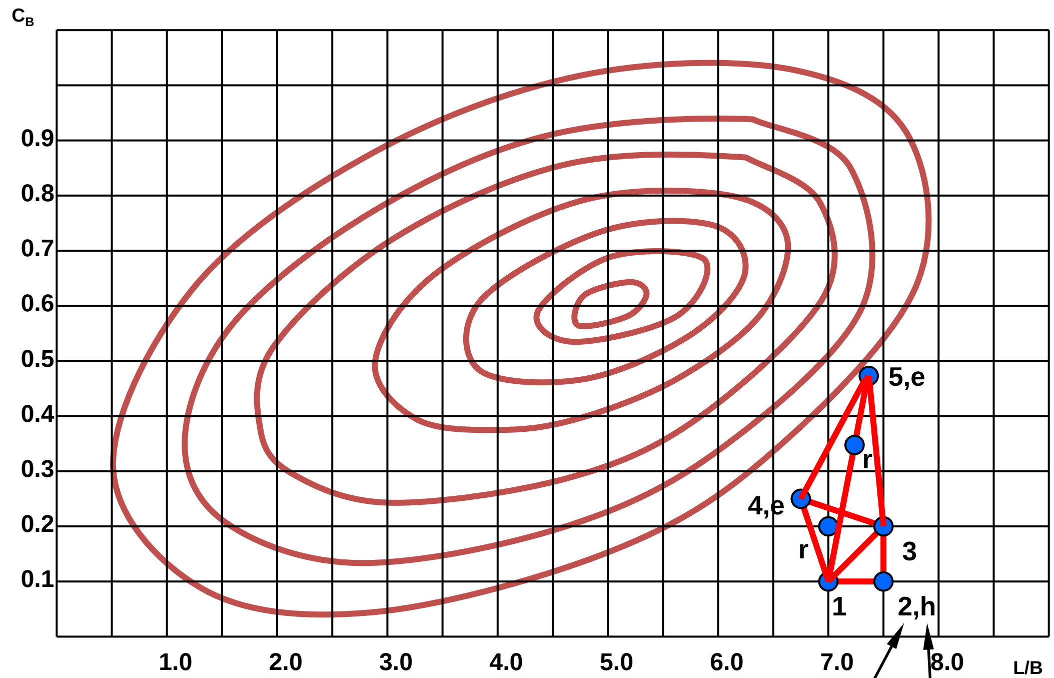
Because $f(x_r) < f(x_1)$ and $f(x_3)$, perform the expansion $\rightarrow x_{4,e}$

$\rightarrow$ Triangle 2 : $x_1,\ x_3,\ x_4$

Iteration 2) Because x_1 is x_h, reflect x_1 through the center between x_3 and x_4. $\rightarrow x_r$

Because $f(x_r) < f(x_3)$ and $f(x_4)$, perform the expansion $\rightarrow x_{5,e}$

$\rightarrow$ Triangle 3 : $x_3,\ x_4,\ x_5$



Number means the index 'i' of X_i.

Alphabet means the kind of X_i .

h: maximum point of the corner in the simplex(triangle)

r: reflection

e: expansion

c: contraction

3.3 Direct Search Method
2. Nelder & Mead Simplex Method(11): Example

$$x_1 = L/B,\ \ x_2 = C_B$$

Iteration 3) Because x_3 is x_h, reflect x_3 through the center between x_4 and x_5. $\rightarrow x_r$

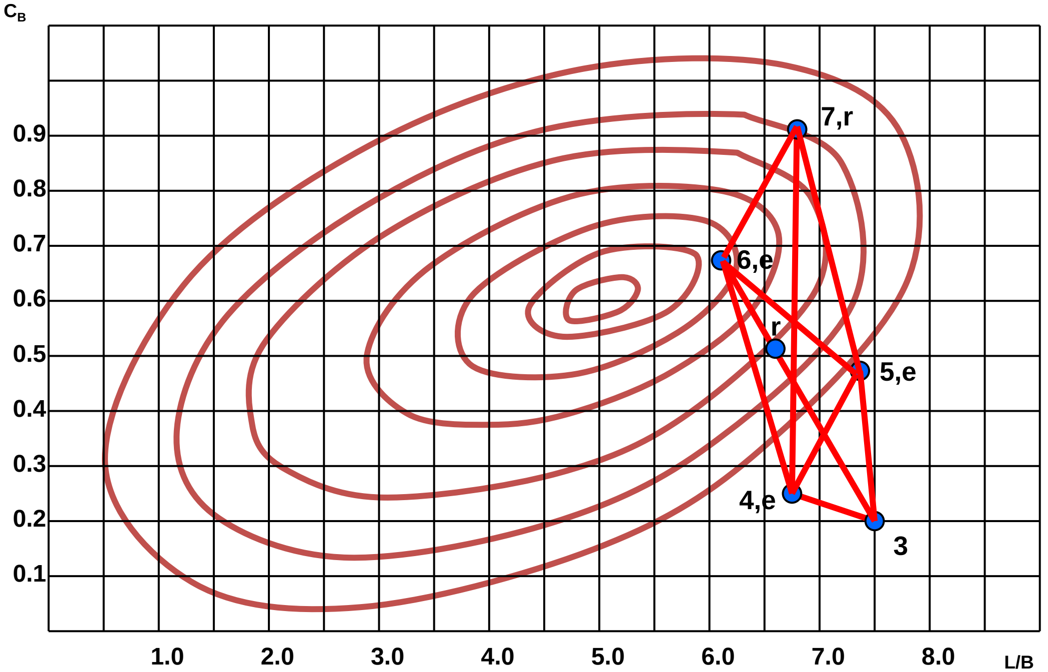
Because $f(x_r) < f(x_4)$ and $f(x_5)$, perform the expansion $\rightarrow x_{6,e}$

$\rightarrow$ Triangle 4 : $x_4,\ x_5,\ x_6$

Iteration 4) Because x_4 is x_h, reflect x_4 through the center between x_5 and x_6. $\rightarrow x_{7,r}$

Because $f(x_{7,r}) > f(x_6)$, go to the next iteration.

$\rightarrow$ Triangle 5 : $x_5,\ x_6,\ x_7$



2. Nelder & Mead Simplex Method(12): Example

Iteration 5) Because x_5 is x_h, reflect x_5 through the center between x_6 and x_7. $\rightarrow x_r$

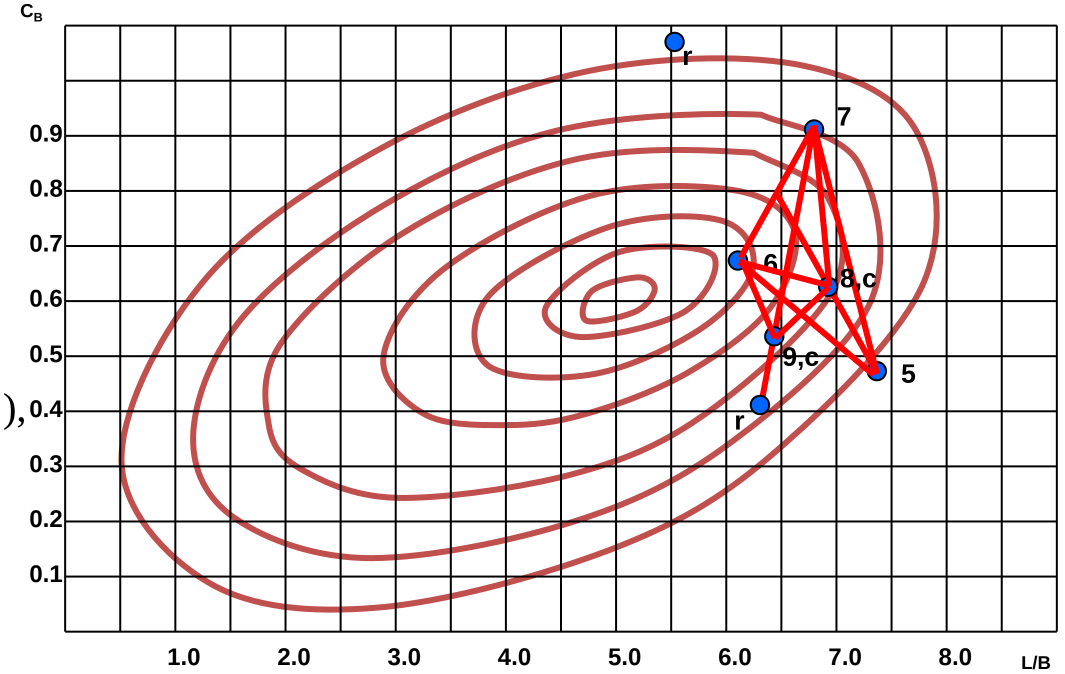
Because $f(x_r) >$ $f(x_5)$, $f(x_6)$ and $f(x_7)$, perform the constraction. $\rightarrow$ $x_{8,c}$

$\rightarrow$ Triangle 6 : x_6, x_7, x_8

Iteration 6) Because x_7 is x_h, reflect x_7 through the center between x_6 and x_8. $\rightarrow x_r$

Because $f(x_r) >$ $f(x_6)$, $f(x_8)$ and $f(x_r) < f(x_7)$, contract the simplex toward x_r $\rightarrow$ $x_{9,c}$

$\rightarrow$ Triangle 7 : x_6, x_8, x_9



2. Nelder & Mead Simplex Method(13): Example

Iteration 7) Because x_8 is x_h, reflect x_8 through the center between x_6 and x_9. $\rightarrow x_r$

Because $f(x_r) < f(x_6), f(x_9)$,

preforme the expansion $\rightarrow x_{10,c}$

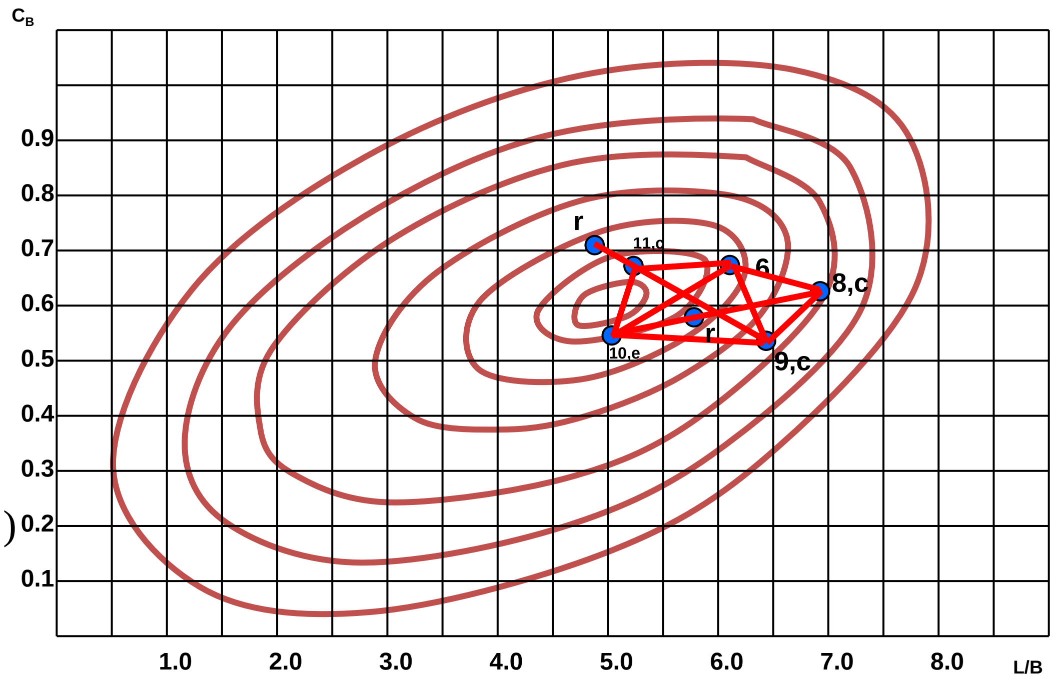
$\rightarrow$ Triangle 8 : x_6, x_9, x_{10}

Iteration 8) Because $x_{9,c}$ is x_h, reflect $x_{9,c}$ through the center between x_6 and x_{10}. $\rightarrow x_r$

Because $f(x_r) > f(x_6), f(x_{10})$ and $f(x_r) < f(x_9)$

contract the simplex toward $x_r \rightarrow x_{11,c}$

$\rightarrow$ Triangle 9 : x_6, x_{10}, x_{11}

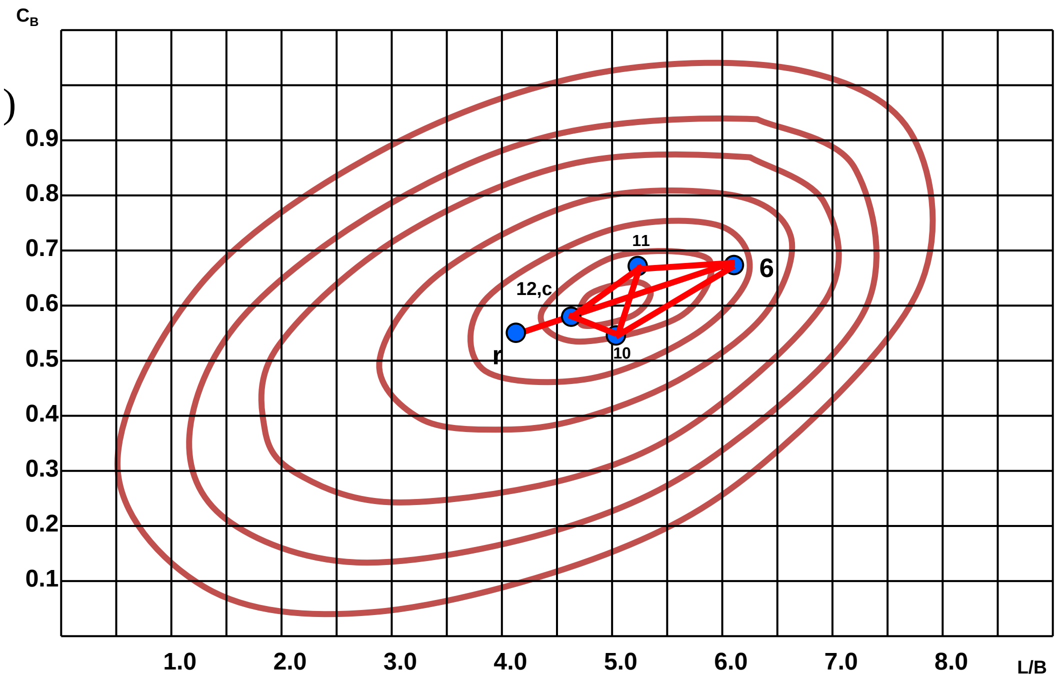


3.3 Direct Search Method
2. Nelder & Mead Simplex Method(14): Example

Iteration 9) Because x_6 is x_h, reflect x_6 through the center between x_{10} and x_{11}. $\rightarrow x_r$

Because $f(x_r) > f(x_{10}), f(x_{11})$ and $f(x_r) < f(x_6)$ contract the simplex toward $x_r \rightarrow x_{12,c}$

$\rightarrow$ Triangle 10 : $x_{10}, \ x_{11}, \ x_{12}$



$x_1(7, \ 0.1)$	$x_2(7.5, \ 0.1)$
$x_3(7.5, \ 0.2)$	$x_4(6.75, \ 0.25)$
$x_5(7.375, \ 0.475)$	$x_6(6.1875, \ 0.6875)$
$x_7(6.8125, \ 0.9125)$	$x_8(6.9375, \ 0.6375)$
$x_9(6.4375, \ 0.5375)$	$x_{10}(5.0625, \ 0.5625)$
$x_{11}(5.21875, \ 0.66875)$	$x_{12}(4.6171875, \ 0.5796875)$

Performing 10 times iterations, we can recognize that the simplex(triangle) has the tendency to approach the result obtained by the 'Hooke & Jeeves direct search method'.

03 Unconstrained Optimization Method

Computer Aided Ship Design

Part I. Optimization Method

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

4.1 Optimal Solution Using Optimality Condition

4.2 Lagrange Multiplier for Equality Constraints

4.3 Kuhn-Tucker Necessary Condition for Inequality Constraints

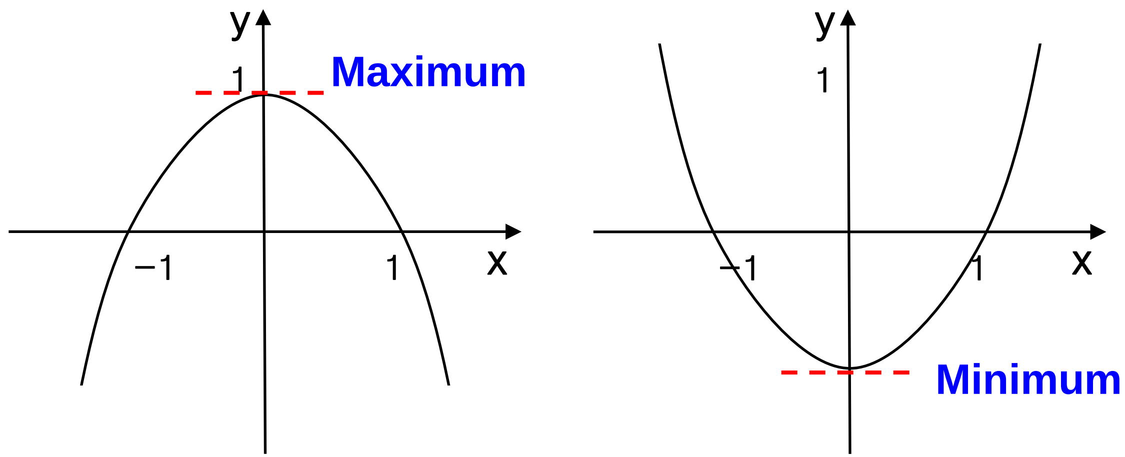
Seoul National Univ.
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

4.1 Optimal Solution Using Optimality Condition

- **Review of the Mathematics for the course of high school**

- **"수학의 정석"(Mathematics II) Review "6. Maximum, Minimum and Differentials" (p. 104)**



1) **Maximum value:** **The increase of the value of the continuous function $f(x)$ is changed to the decrease of that at $x = x^*$.**

2) **Minimum value:** **The decrease of the value of the continuous function $f(x)$ is changed to the increase of that at $x = x^*$.**

$$f'(x^*) = 0$$

(Necessary condition for $x = x^*$ to be a maximum or minimum)

▪ **First-order necessary condition for the function of a single variable:** $f'(x^*) = 0$

pf) The Taylor series expansion of $f(x)$ at the point x^* is as follows.

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 \boxed{+R}$$

Remainder
: If the difference between x and x^* is small, the value of the remainder is also very small.

Let $x - x^* = d$, the equation is as follows.

$$f(x) = f(x^*) + f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$$

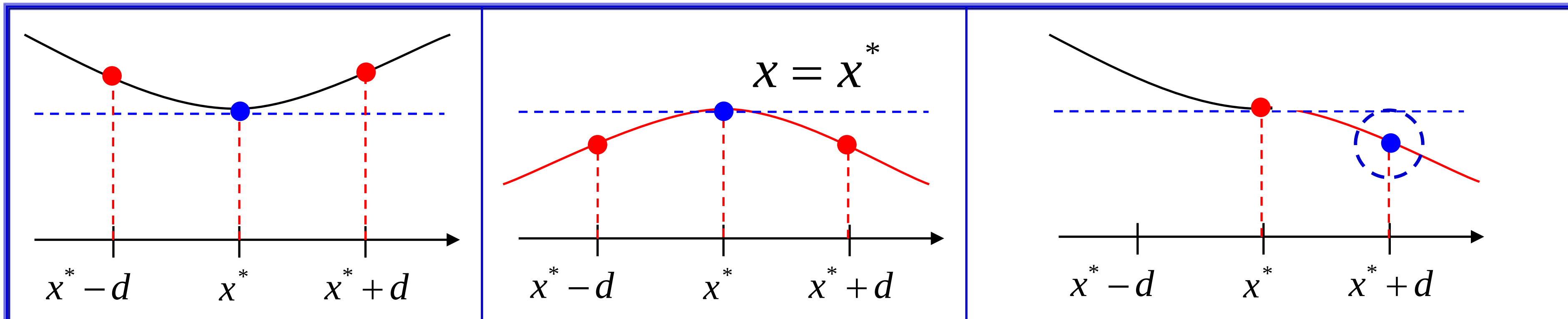
From this equation, the change in the function at x^* , i.e., $f(x) - f(x^*) = \Delta f(x)$ is given as

$$\Delta f(x) = f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$$

4.1 Optimal Solution Using Optimality Condition
- First-Order Necessary Conditions(2)

$$\Delta f(x) = f(x) - f(x^*) = f'(x^*)d + \frac{1}{2} f''(x^*)d^2 + R$$

Δf **must be positive,** **if** x^* **is a local minimum point.**



$f(x^*)$ **is minimum.**

$f(x^*)$ **is maximum.**

$f(x^*)$ **is neither minimum nor maximum.**

Since $d(= x - x^*)$ **is small, the first-order term** $f'(\mathbf{x}^*)d$ **dominates other terms.**

And the sign if the term $f'(\mathbf{x}^*)d$ **depends on the sing of** d **.**

Thus, the only way Δf **can be positive regardless of the sign of** d **in a neighborhood of** x^* **is** $f'(x^*) = 0$ **.**

In the same way, Δf **must be negative** **if** x^* **is a local maximum point. So, the only way** Δf **can be positive regardless of the sign of** d **in a neighborhood of** x^* **is** $f'(x^*) = 0$

$$\Delta f(x) = f(x) - f(x^*) = f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$$

- Now, we need a sufficient condition to determine which of the stationary points are actually minimum for the function.

Since stationary points satisfy the necessary condition $f'(x^*) = 0$, the change in function $\Delta f(x) = f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$ becomes as follows.

$$\Delta f(x) = \frac{1}{2}f''(x^*)d^2 + R$$

Since the second-order term dominates all other higher-order terms, the term can be positive for all $d \neq 0$, if

$$f''(x^*) > 0 \quad \text{(Sufficient condition)}$$

Summary

- First-order necessary condition

If x^* is a local minimum point, $f'(x^*) = 0$.

cf) If $f'(x^*) = 0$, x^* is a stationary point(minimum, maximum and inflection point).

- Sufficient condition

If x^* is a stationary point($f'(x^*) = 0$) and $f''(x^*) > 0$, x^* is a local minimum point.

4.1 Optimal Solution Using Optimality Condition
[Review] Taylor Series Expansion for the Function of Two Variables

Taylor series expansion for the function of two variables $f(x_1, x_2)$ **at** (x_1^*, x_2^*)

$$f(x_1,x_2) = f(x_1^*,x_2^*) + \frac{\partial f}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f}{\partial x_2}(x_2 - x_2^*)$$

$$+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(x_1 - x_1^*)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) + \frac{\partial^2 f}{\partial x_2^2}(x_2 - x_2^*)^2\right) + R$$

Each terms can be represented as follows:

$$\frac{\partial f}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f}{\partial x_2}(x_2 - x_2^*) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}^T \begin{bmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{bmatrix} = \underline{\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*)}$$

$$\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(x_1 - x_1^*)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) + \frac{\partial^2 f}{\partial x_2^2}(x_2 - x_2^*)^2\right) = \frac{1}{2}\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_1 - x_1^*) + \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_2 - x_2^*) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 - x_1^*) + \frac{\partial^2 f}{\partial x_2^2}(x_2 - x_2^*) \end{bmatrix}\begin{bmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} x_1 - x_1^* & x_2 - x_2^* \end{bmatrix}\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}\begin{bmatrix} x_1 - x_1^* \\ x_2 - x_2^* \end{bmatrix}$$

$$= \underline{\frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)}$$

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \underline{\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*)} + \underline{\frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)} + R$$

Element of the 2x2 Matrix

$$\left(\mathbf{x} = (x_1, x_2)^T,\ \mathbf{x}^* = (x_1^*, x_2^*)^T,\ \boxed{\mathbf{H} \in M_{2\times 2}}\right)$$

4.1 Optimal Solution Using Optimality Condition
- Optimality Conditions for Function of Several Variables (1)

- **Matrix form of the Taylor series expansion for the function of two variables**

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{x}-\mathbf{x}^*) + \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^*\right)^T \mathbf{H}(\mathbf{x}^*)\left(\mathbf{x}-\mathbf{x}^*\right) + R$$

Element of the 2x2 Matrix

$$\left(\mathbf{x} = (x_1, x_2)^T,\ \mathbf{x}^* = (x_1^*, x_2^*)^T,\ \mathbf{H} \in M_{2\times 2} \right)$$

- **Matrix form of the Taylor series expansion for the function of the several variables**

: It has the same form of the function of two variables.

$\mathbf{x},\ \mathbf{x}^*, \nabla f$ **:** n **dimension Vector**

$\mathbf{H} \in M_{n\times n}$

- **By defining** $\mathbf{x} - \mathbf{x}^* = \mathbf{d}$**, the Taylor series expansion for the function of the several variables is as follows.**

$$f(\mathbf{x}^* + \mathbf{d}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} + R$$

$$\nabla f(\mathbf{x}^*)^T = 0,\quad \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$$

➡ **Sufficient conditions for** $\mathbf{x} = \mathbf{x}^*$ **to be a local minimum**

4.1 Optimal Solution Using Optimality Condition
[Review] Hessian Matrix

- **Hessian matrix : Differentiating the gradient vector once again, we obtain a matrix of second partial derivatives for the function $f(x)$ called the Hessian matrix.**

That is, differentiating each component of the gradient vector with respect to $x_1, x_2, \cdots, x_n$, we obtain

$$\frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^T$$

: *n*-column Vector

- **Hessian matrix is denoted as H or $\nabla^2 f$.**

$$\mathbf{H} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right] \qquad (i = 1,2,\cdots,n \; ; \; j = 1,2,\cdots,n)$$

- **Property of the Hessian matrix**

$$\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}$$

Therefore, the Hessian matrix is always a symmetric matrix.

[Review] Quadratic Form

- **Quadratic form: This is a special nonlinear function having only second-order terms.**

ex) $F(x_1, x_2, x_3) = \frac{1}{2}\left(2x_1^2 + 2x_1x_2 + 4x_1x_3 - 6x_2^2 - 4x_2x_3 + 5x_3^2\right)$

The quadratic form can be written in the following matrix notation.

A :Symmetric matrix

$$F(x_1, x_2, x_3) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 2 \\ 1 & -6 & -2 \\ 2 & -2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2}\mathbf{x}^T \boxed{\mathbf{A}} \mathbf{x} \Longleftrightarrow \frac{1}{2}\mathbf{d}^T \mathbf{H}\mathbf{d}$$

- **The elements of symmetric matrix A is defined as follows.(a_{ij}: element of the matrix A at (i,j))**

  1) The diagonal terms of the matrix are equal to the coefficient of the squared terms.

$$a_{ii} = \left(\text{coefficient of } x_i^2\right)$$

  2) The all terms except for diagonal terms(a_{ij}) are equal to a half of the coefficient of the $x_i x_j$.

$$a_{ij} = \left(\text{coefficient of } x_i x_j\right) \times \frac{1}{2}$$

A symmetric matrix A is often referred to as **a positive definite** if the quadratic form associated with **A** is positive definite

▪ **Form of a quadratic form**

1) **Positive Definite**

: $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for any x except for $\mathbf{x} = 0$.

2) **Positive Semidefinite**
: $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all x and there exists at least one $\mathbf{x} \neq 0$ with $\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$.

3) **Negative Definite**

: $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$ for all x except for $\mathbf{x} = 0$

4) **Negative Semidefinite**

: $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$ for all x

5) **Indefinite**
: The quadratic form is positive for some vectors x and negative for others.

▪ **Use of the form of a quadratic form**

① **Minimum condition for the function of the single variable**

If x^* is a stationary point($f'(x^*) = 0$) and $f''(x^*) > 0$, x^* is a local minimum point.

② **Minimum condition for the function of the several variables**

If $\mathbf{x}^*$ is a stationary point($\nabla f(\mathbf{x}^*) = 0$) and $\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$, i.e., the quadratic form is positive definite, $\mathbf{x}^*$ is a local minimum point.

To be $\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$ at $\mathbf{x}^*$,
$\mathbf{H}(\mathbf{x}^*)$ must be positive definite

Ref) KREYSZIG E., Advanced Engineering Mathematics, WILEY, 2006, 8.4. Eigenbasis. Diagonalization. Quadratic forms.

Theorem: Methods for checking positive definiteness or semidefiniteness of a quadratic form or a matrix :

Let $\lambda_i, i=1,\ldots,n$ be n eigenvalues of a symmetric $n \times n$ matrix $\mathbf{A}$ associated with the quadratic form $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$.

1) $F(\mathbf{x})$ is positive definite if and only if all eigenvalues of $\mathbf{A}$ are strictly positive, i.e.,

$$\lambda_i > 0, i=1,\ldots,n$$

2) $F(\mathbf{x})$ is positive semidefinite if and only if all eigenvalues of $\mathbf{A}$ are nonnegative, i.e.,

$$\lambda_i \geq 0, i=1,\ldots,n$$

3) $F(\mathbf{x})$ is negative definite if and only if all eigenvalues of $\mathbf{A}$ are strictly negative, i.e.,

$$\lambda_i < 0, i=1,\ldots,n$$

4) $F(\mathbf{x})$ is negative semidefinite if and only if all eigenvalues of $\mathbf{A}$ are nonpositive, i.e.,

$$\lambda_i \leq 0, i=1,\ldots,n$$

5) $F(\mathbf{x})$ is indefinite if some $\lambda_i < 0$ and some other $\lambda_i > 0$.

- Eigenvalue of a Symmetric Matrix A associated with the quadratic Form

For a given matrix A, the eigenvalue problem is defined as $\mathbf{Av} = \lambda\mathbf{v}$

, where λ is an eigenvalue and $\mathbf{v}$ is the corresponding eigenvector.

How to determine the eigenvalues:

$$\mathbf{Av} = \lambda\mathbf{v} \quad \Rightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = 0 \quad \Rightarrow \quad \det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

Determine the eigenvalues and the form of the following matrix.

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

$$\det\begin{bmatrix} 4-\lambda & 2 & 2 \\ 2 & 4-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{bmatrix} = (2-\lambda)^2(8-\lambda) = 0$$

$$\therefore \lambda = 2(\text{equal root}),\ 8$$

Since all eigenvalues of $\mathbf{A}$ are positive, this matrix is positive definite.

[Summary] Optimality Conditions for Function of Several Variables

- The Taylor series expansion of $f(\mathbf{x})$, which is the function of n variables gives

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} + R$$

- From this equation, the change in the function at $\mathbf{x}^*$, i.e., $\Delta f(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*)$, is given as

$$\Delta f = \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} + R$$

- If we assume a local minimum at $\mathbf{x}^*$ then Δf must be positive.

1) The first-order necessary condition:

If $\nabla f(\mathbf{x}^*) = 0$, i.e., $\dfrac{\partial f(\mathbf{x}^*)}{\partial x_i} = 0\ ,\ (i = 1,2,\cdots n)$, x^* is a stationary point(minimum, maximum and inflection point).

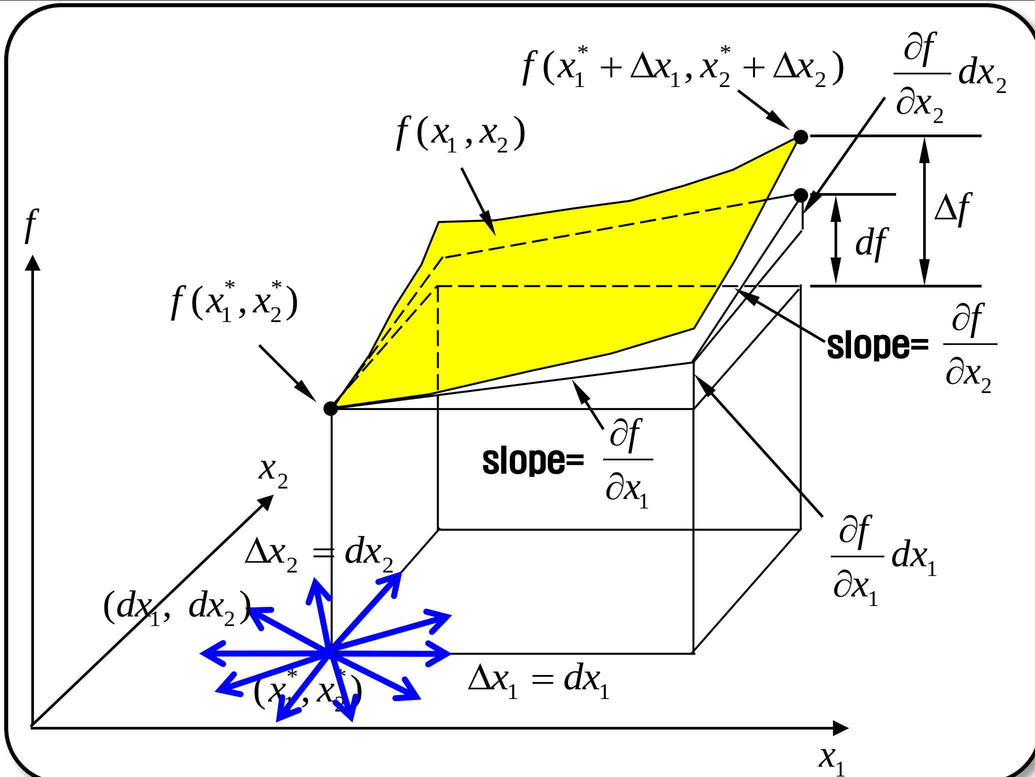
2) The sufficient condition:

If $\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$ a stationary point $(\nabla f(\mathbf{x}^*)^T = 0 \Rightarrow \nabla f(\mathbf{x}^*) = 0)$ is a local minimum.

To be $\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$,$\mathbf{H}(\mathbf{x}^*)$ must be positive definite.

4.1 Optimal Solution Using Optimality Condition

Necessary condition to be a stationary point : Total derivative d f = 0 -> grad f = 0.



The change in the function is defined as follows

$$\Delta f = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}\Delta x_1^{\ 2} + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}\Delta x_1 \Delta x_2 + \frac{\partial^2 f}{\partial x_2^2}\Delta x_2^{\ 2}\right) + R$$

If $\Delta x_1 \to 0, \Delta x_2 \to 0$, the first-order term $\frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2$ dominates other terms.

Therefore Δf can be approximated as $\Delta f \approx \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2$

The symbol *"d"* refers to the infinitesimal change. In accordance with the notation we write the change of the function f as follows

the change of the function in x_2 direction

$$df = \boxed{\frac{\partial f}{\partial x_1} dx_1} + \boxed{\frac{\partial f}{\partial x_2} dx_2}$$

the change of the function in x_1 direction

If $df = 0$, then x^* is a stationary point.

To be $df = 0$ regardless of the sign of dx_1 and dx_2, $\partial f / \partial x_1$ and $\partial f / \partial x_2$ must be zero.

It means that the gradient of function f is equal to zero.

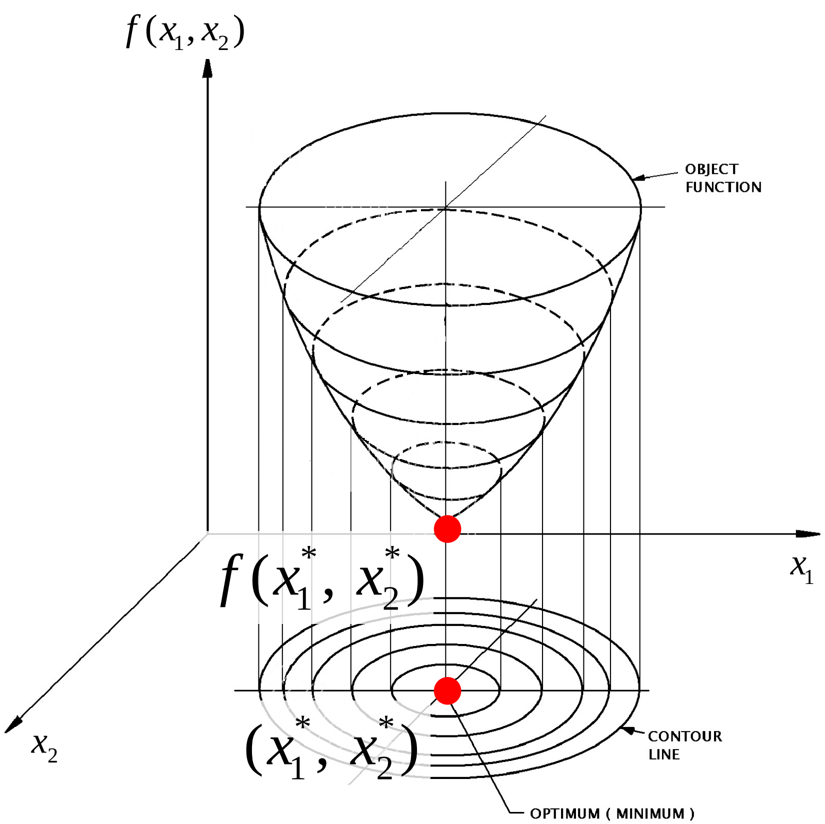
$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = 0 \Rightarrow \nabla f = 0$$

4.1 Optimal Solution Using Optimality Condition

Necessary condition to be a stationary point:Total derivative d f = 0 -> grad f = 0

Given: minimize $f(x_1, x_2)$

Find: Stationary point (x_1^*, x_2^*)



- The change in function(df) at the point(x_1^*, x_2^*) with the change in variables(dx_1, dx_2) is as follows.

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2$$

The point where the change in function(df) is zero is called stationary point . It includes the minimum, maximum and saddle point.

Note: In the general engineering optimization problem, the optimum point is more important than the optimum value.

[example] Main dimension of a ship (L, B, D, C_B) to minimize the shipbuilding cost is more important than the shipbuilding cost itself.

[Example] Solution of a Quadratic Programming problem

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Given: $$f(x_1, x_2, x_3) = x_1^{\ 2} + x_2^{\ 2} + x_3^{\ 2}$$

$$h(x_1, x_2, x_3) = x_1 + x_2 + x_3 + 1 = 0$$

Find: Stationary point (x_1^*, x_2^*, x_3^*)

Express h **(equality constraint) as an explicit function of** x_1**.**

$$x_1 = -x_2 - x_3 - 1$$

Substitute x_1 **into the function of** f

$$f = (-x_2 - x_3 - 1)^2 + x_2^{\ 2} + x_3^{\ 2}$$

$$= (x_2^{\ 2} + x_3^{\ 2} + 1 + 2x_2x_3 + 2x_2 + 2x_3) + x_2^{\ 2} + x_3^{\ 2}$$

$$= 2x_2^{\ 2} + 2x_3^{\ 2} + 1 + 2x_2x_3 + 2x_2 + 2x_3$$

Determine the stationary point in $(df = 0)$ **unconstrained optimization problem.**

$$df = \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0$$

$$\therefore \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0$$

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2x_3 + 2 = 0$$

$$\frac{\partial f}{\partial x_3} = 4x_3 + 2x_2 + 2 = 0$$

The equations are solved as $x_2 = -\frac{1}{3},\ x_3 = -\frac{1}{3}$

By substituting these value into the function of ***f, we obtain***

$$x_1 = -\frac{1}{3}$$

Therefore, the stationary point is $\left(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\right)$.

[Example] Solution of a Quadratic Programming problem

Given:

$$\textit{Minimize} \quad f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$

$$h(x_1, x_2) = x_1 + x_2 - 2 = 0$$

Find: Local minimum point(x_1*, x_2*)

1. Express h (equality constraint) as an explicit function of x_1.
2. Substitute x_1 into f and find the stationary point by using $df = 0$.

Solution

Express x_2 as an explicit function of x_1,

$$x_2 = \Phi(x_1) = -x_1 + 2$$

$$f(x_1, \Phi(x_1)) = (x_1 - 1.5)^2 + (-x_1 + 2 - 1.5)^2$$

$$\frac{df}{dx_1} = 2(x_1 - 1.5) - 2(-x_1 + 0.5) = 0$$

$$\Rightarrow x_1 = 1$$

$$\Rightarrow x_2 = -x_1 + 2 = 1$$

$$\frac{d^2 f}{dx_1^{\,2}} = 4 > 0$$

$\therefore (x_1^*, \ x_2^*) = (1, \ 1)$**: Local minimum point**

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

4.2 Lagrange Multiplier for equality constraints

4.2 Lagrange Multiplier for equality constraints
- Function and Stationary Point for Unconstrained Optimum Design Problem

Given: $minimize \;\; f(x_1, x_2, x_3)$

Find: Stationary point(x_1^*, x_2^*, x_3^*)

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3$$

At the stationary point, the change in the function(df) is zero.
The gradient of the function at stationary point must be zero for the change in the function(df) to be zero regardless of the sign of dx_1, dx_2, and dx_3.

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0$$

$$\Rightarrow \nabla f = 0$$

- Function and Stationary Point for Constrained Optimum Design Problem(1)

Given: $\text{minimize} \;\; f(x_1, x_2, x_3)$

$\text{Subject to} \;\; h(x_1, x_2, x_3) = 0$

Find: Stationary point(x_1*, x_2*, x_3*)

1. Express h (equality constraint) as an explicit function of x_1.
2. Substitute x_1 into f and find the stationary point by using $df = 0$.

In many problem, it may not be possible to express h (equality constraint) as an explicit function of x_1.

Example) It is difficult to express the following equality constraint as an explicit function.

$$ex) \;\; h(x_1, x_2, x_3) = \tan x_1 + \cos x_2 + e^{x_3} = 0$$

Is there any method to obtain the stationary point if the equality constraint can not be expressed as an explicit function?

$df = 0$ at the stationary point.

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \text{------} \;①$$

Since $h(x_1,x_2,x_3)=0$, dh is also zero.

$$dh = \frac{\partial h}{\partial x_1} dx_1 + \frac{\partial h}{\partial x_2} dx_2 + \frac{\partial h}{\partial x_3} dx_3 = 0 \quad \text{------} \;②$$

Since equation① and ② are equal to zero, the following equation is always satisfied.

$$df + \lambda \cdot dh = 0$$

λ: Undetermined Coefficient 'Lagrange multiplier'

4.2 Lagrange Multiplier for equality constraints
- Function and Stationary Point for Constrained Optimum Design Problem(2)

Given: $minimize\ f(x_1, x_2, x_3)$

Subject to $h(x_1, x_2, x_3) = 0$

Find: Stationary point(x_1*, x_2*, x_3*)

① $$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 = 0$$

② $$dh = \frac{\partial h}{\partial x_1}dx_1 + \frac{\partial h}{\partial x_2}dx_2 + \frac{\partial h}{\partial x_3}dx_3 = 0$$

- Because of the equality constraint h, dx_1, dx_2, and dx_3 are not linearly independent.

$$df + \lambda \cdot dh = 0$$

λ: **Undetermined Coefficient**
'Lagrange multiplier'

This equation can be rearranged as follows.

$$\frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 + \lambda\left(\frac{\partial h}{\partial x_1}dx_1 + \frac{\partial h}{\partial x_2}dx_2 + \frac{\partial h}{\partial x_3}dx_3\right) = 0$$

$$\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right)dx_1 + \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

4.2 Lagrange Multiplier for equality constraints - Function and Stationary Point for Constrained Optimum Design Problem(3)

Given: $minimize\ \ f(x_1, x_2, x_3)$

Subject to $h(x_1, x_2, x_3) = 0$

Find: Stationary point(x_1*, x_2*, x_3*)

① $df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 = 0$

② $dh = \frac{\partial h}{\partial x_1}dx_1 + \frac{\partial h}{\partial x_2}dx_2 + \frac{\partial h}{\partial x_3}dx_3 = 0$

- **Because of the equality constraint h, dx_1, dx_2, and dx_3 are not linearly independent.**

$$df + \lambda \cdot dh = 0$$

λ: **Undetermined Coefficient 'Lagrange multiplier'**

$$\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right)dx_1 + \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

If the dx_1, dx_2, and dx_3 were all independent of each other, all terms in the brackets will be zero. This however, is not the case because of the equality constraint h. Therefore, we should make the first term to be zero by determining a proper value of λ, so that the following equation is satisfied without considering the dx_1.

$$\left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

Since dx_2 and dx_3 are independent, the terms in the brackets must be equal to zero to satisfy the equation.

$$\therefore \left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right) = 0, \quad \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right) = 0, \quad \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right) = 0$$

Therefore, the point(λ, x_1, x_2, x_3) that satisfies the following equations is a stationary point.

$$\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1} = 0, \qquad \frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3} = 0, \qquad h(x_1, x_2, x_3) = 0$$

4 Unknown variables: (x_1, x_2, x_3, λ)

4 Equations

There exists a unique solution.

4.2 Lagrange Multiplier for equality constraints

the point(λ, x_1, x_2, x_3) that satisfies the following equations is a stationary point.

$$\frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0, \qquad \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial f}{\partial x_3} + \lambda \frac{\partial h}{\partial x_3} = 0, \qquad h(x_1, x_2, x_3) = 0$$

It is convenient to write these conditions in terms of a Lagrange function, L , defined as

$$L(x_1, x_2, x_3, \lambda) = f(x_1, x_2, x_3) + \lambda h(x_1, x_2, x_3)$$

$$\nabla L(x_1, x_2, x_3, \lambda) = 0$$

Constrained optimal design problem is transformed to the unconstrained optimal design problem.

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0 \qquad \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial L}{\partial x_3} = \frac{\partial f}{\partial x_3} + \lambda \frac{\partial h}{\partial x_3} = 0 \qquad \frac{\partial L}{\partial \lambda} = h(x_1, x_2, x_3) = 0$$

λ : Lagrange Multiplier
L : Lagrange Function

4.2 Lagrange Multiplier for equality constraints
- [Summary] Function and Stationary Point for Constrained Optimum Design Problem
- Solution of the Constrained Optimum Design by using the Lagrange Multiplier(1)

Optimization Problem

Minimize $f(x_1, x_2, x_3)$ .... ①

Subject to $h_1(x_1, x_2, x_3) = 0$ .... ②

$h_2(x_1, x_2, x_3) = 0$ .... ③

Number of variables: 3

Number of equation : 2

Necessary condition that minimize f is $df = 0$.
$df = 0$ is eq①' as following

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \cdots ①'$$

Subject to $h_1(x_1, x_2, x_3) = 0$ ... ②

$h_2(x_1, x_2, x_3) = 0$ ... ③

Since dx_1, dx_2, dx_3 are not independent because of the equality constraints h_1, h_2, we cannot set

$$\frac{\partial f}{\partial x_1} = 0, \frac{\partial f}{\partial x_2} = 0, \frac{\partial f}{\partial x_3} = 0$$

Number of variables: 3

Number of equations : 3

How could we generate more equations from the indeterminate equation?

Because 'Minimize f' is formulated as an equation($df = 0$), the number of equations is equal to the number of unknown variables.

We can solve it! (Num of Equation = Num of Variables)

Optimization Problem

Minimize $f(x_1, x_2, x_3)$ ···· ①

Subject to $h_1(x_1, x_2, x_3) = 0$ ···· ②

$h_2(x_1, x_2, x_3) = 0$ ···· ③

Number of variables: 3

Number of equation : 2

Necessary condition that minimize f is df = 0.
df = 0 is eq①' as following

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \cdots ①'$$

Subject to $h_1(x_1, x_2, x_3) = 0$ ··· ②

$h_2(x_1, x_2, x_3) = 0$ ··· ③

Since dx_1, dx_2, dx_3 are not independent because of the equality constraints h_1, h_2, we cannot set

$$\frac{\partial f}{\partial x_1} = 0, \frac{\partial f}{\partial x_2} = 0, \frac{\partial f}{\partial x_3} = 0$$

To find relationship between dx_1, dx_2, dx_3,
we modify the equation(s) ② and ③ to the form of total derivative dh_1, dh_2 .

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \cdots ①'$$

$$dh_1 = \frac{\partial h_1}{\partial x_1} dx_1 + \frac{\partial h_1}{\partial x_2} dx_2 + \frac{\partial h_1}{\partial x_3} dx_3 = 0 \quad \cdots ②'$$

$$dh_2 = \frac{\partial h_2}{\partial x_1} dx_1 + \frac{\partial h_2}{\partial x_2} dx_2 + \frac{\partial h_2}{\partial x_3} dx_3 = 0 \quad \cdots ③'$$

- [Summary] Function and Stationary Point for Constrained Optimum Design Problem
- Solution of the Constrained Optimum Design by using the Lagrange Multiplier(3)

Optimization Problem

Minimize $f(x_1, x_2, x_3)$ ···· ①

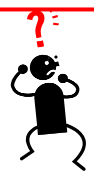
Subject to $h_1(x_1, x_2, x_3) = 0$ ···· ②

$h_2(x_1, x_2, x_3) = 0$ ···· ③

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \cdots ①'$$

$$dh_1 = \frac{\partial h_1}{\partial x_1} dx_1 + \frac{\partial h_1}{\partial x_2} dx_2 + \frac{\partial h_1}{\partial x_3} dx_3 = 0 \quad \cdots ②'$$

$$dh_2 = \frac{\partial h_2}{\partial x_1} dx_1 + \frac{\partial h_2}{\partial x_2} dx_2 + \frac{\partial h_2}{\partial x_3} dx_3 = 0 \quad \cdots ③'$$



Are the equation ①', ②' and ③' the differential equations with respect to f, h_1, h_2 ?

→ If the problem is given as following

- Given: $\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0,\ \frac{\partial h_1}{\partial x_1} dx_1 + \frac{\partial h_1}{\partial x_2} dx_2 + \frac{\partial h_1}{\partial x_3} dx_3 = 0,\ \frac{\partial h_2}{\partial x_1} dx_1 + \frac{\partial h_2}{\partial x_2} dx_2 + \frac{\partial h_2}{\partial x_3} dx_3 = 0,$

- Find: Function f, h_1, h_2

Then the equation ①', ②', and ③' are differential equations.

However, since the function f, h_1 and h_2(equation ①, ②, ③) are given and differential quantities to dx_1, dx_2 and dx_3 are finds, the equation ①', ②' and ③' are the algebraic equations of the variables x_1, x_2, x_3 .

4.2 Lagrange Multiplier for equality constraints

Optimization Problem

Minimize	$f(x_1, x_2, x_3)$	···· ①
Subject to	$h_1(x_1, x_2, x_3) = 0$	···· ②
	$h_2(x_1, x_2, x_3) = 0$	···· ③

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 = 0 \quad \cdots ①'$$

$$dh_1 = \frac{\partial h_1}{\partial x_1}dx_1 + \frac{\partial h_1}{\partial x_2}dx_2 + \frac{\partial h_1}{\partial x_3}dx_3 = 0 \quad \cdots ②'$$

$$dh_2 = \frac{\partial h_2}{\partial x_1}dx_1 + \frac{\partial h_2}{\partial x_2}dx_2 + \frac{\partial h_2}{\partial x_3}dx_3 = 0 \quad \cdots ③'$$

To eliminate dx_1, dx_2 in the equation ①', we multiply the equation②' and ③' by λ_1 and λ_2 respectively and add it to the equation ①'.

$$df + \lambda_1 dh_1 + \lambda_2 dh_2 = 0$$

$$\underbrace{\left(\frac{\partial f}{\partial x_1} + \lambda_1\frac{\partial h_1}{\partial x_1} + \lambda_2\frac{\partial h_2}{\partial x_1}\right)}_{=0 \ \cdots ④}dx_1 + \underbrace{\left(\frac{\partial f}{\partial x_2} + \lambda_1\frac{\partial h_1}{\partial x_2} + \lambda_2\frac{\partial h_2}{\partial x_2}\right)}_{=0 \ \cdots ⑤}dx_2 + \underbrace{\left(\frac{\partial f}{\partial x_3} + \lambda_1\frac{\partial h_1}{\partial x_3} + \lambda_2\frac{\partial h_2}{\partial x_3}\right)}_{=0 \ \cdots ⑥}dx_3 = 0$$

→ Determine λ_1, λ_2 so that the first term in the brackets becomes zero* (to eliminate dx_1)

→ Determine λ_1, λ_2 so that the second term in the brackets becomes zero* (to eliminate dx_2)

→ Since dx_3 is an independent variable

*** Since dx_1, dx_2, dx_3 are not independent because of the equality constraints h_1, h_2**

5 variables:

5 equations: 2,3,4,5,6

There exists a unique solution.

4.2 Lagrange Multiplier for equality constraints

the point($\lambda_1, \lambda_2, x_1, x_2, x_3$) that satisfies the following equations is a stationary point.

$$\frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial h_1}{\partial x_1} + \lambda_2 \frac{\partial h_2}{\partial x_1} = 0, \qquad \frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial h_1}{\partial x_2} + \lambda_2 \frac{\partial h_2}{\partial x_2} = 0$$

$$\frac{\partial f}{\partial x_3} + \lambda_1 \frac{\partial h_1}{\partial x_3} + \lambda_2 \frac{\partial h_2}{\partial x_3} = 0, h_1(x_1, x_2, x_3) = 0, \quad h_2(x_1, x_2, x_3) = 0$$

It is convenient to write these conditions in terms of a Lagrange function, L , defined as

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = f(x_1, x_2, x_3) + \lambda_1 h_1(x_1, x_2, x_3) + \lambda_2 h_2(x_1, x_2, x_3)$$

$$\nabla L(x_1, x_2, x_3, \lambda_1, \lambda_2) = 0$$

λ : Lagrange Multiplier

L : Lagrange Function

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial h_1}{\partial x_1} + \lambda_2 \frac{\partial h_2}{\partial x_1} = 0 \; \cdots\cdots\; ④$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda_1 \frac{\partial h_1}{\partial x_2} + \lambda_2 \frac{\partial h_2}{\partial x_2} = 0 \; \cdots \; ⑤$$

$$\frac{\partial L}{\partial x_3} = \frac{\partial f}{\partial x_3} + \lambda_1 \frac{\partial h_1}{\partial x_3} + \lambda_2 \frac{\partial h_2}{\partial x_3} = 0 \; \cdots \; ⑥$$

$$\frac{\partial L}{\partial \lambda_1} = h_1(x_1, x_2, x_3) = 0 \; \cdots \; ②$$

$$\frac{\partial L}{\partial \lambda_2} = h_2(x_1, x_2, x_3) = 0 \; \cdots \; ③$$

The Lagrange Function gives us a simple way of stating and remembering how to get the equations, which are satisfied at a stationary point.

4.2 Lagrange Multiplier for equality constraints
Example: Quadratic Programming Problem

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

$$L(x_1, x_2, x_3, \lambda) = f(x_1, x_2, x_3) + \lambda h(x_1, x_2, x_3)$$
$$\nabla L(x_1, x_2, x_3, \lambda) = 0$$

Original Problem

Minimize $f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

Subject to $h(x_1, x_2) = x_1 + x_2 - 2 = 0$

Lagrange Function

Minimize
$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2)$$
$$= (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$
$$+ \lambda(x_1 + x_2 - 2)$$

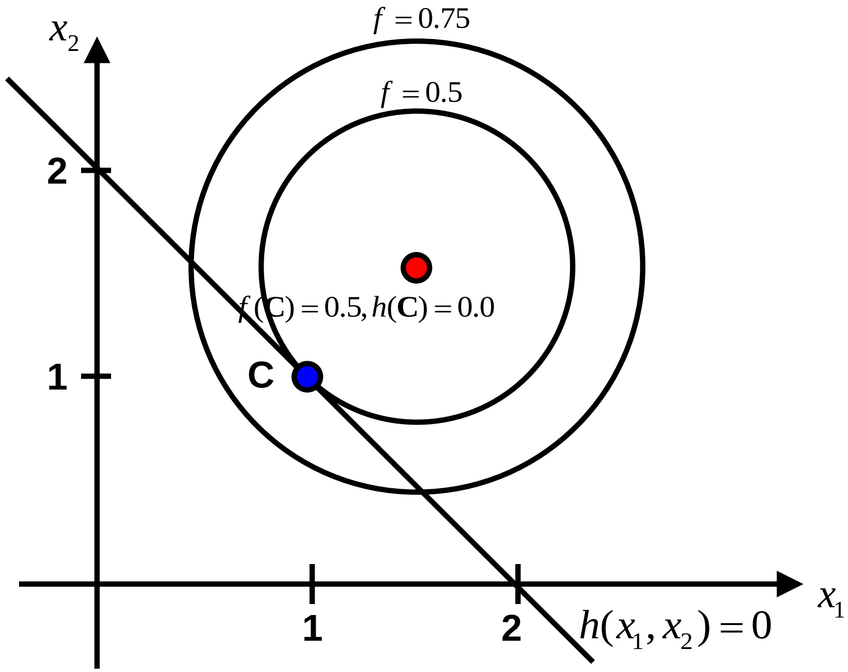
Necessary Condition: $\nabla L(x_1, x_1, \lambda) = 0$

$$\frac{\partial L}{\partial x_1} = 2(x_1 - 1.5) + \lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 2(x_2 - 1.5) + \lambda = 0$$

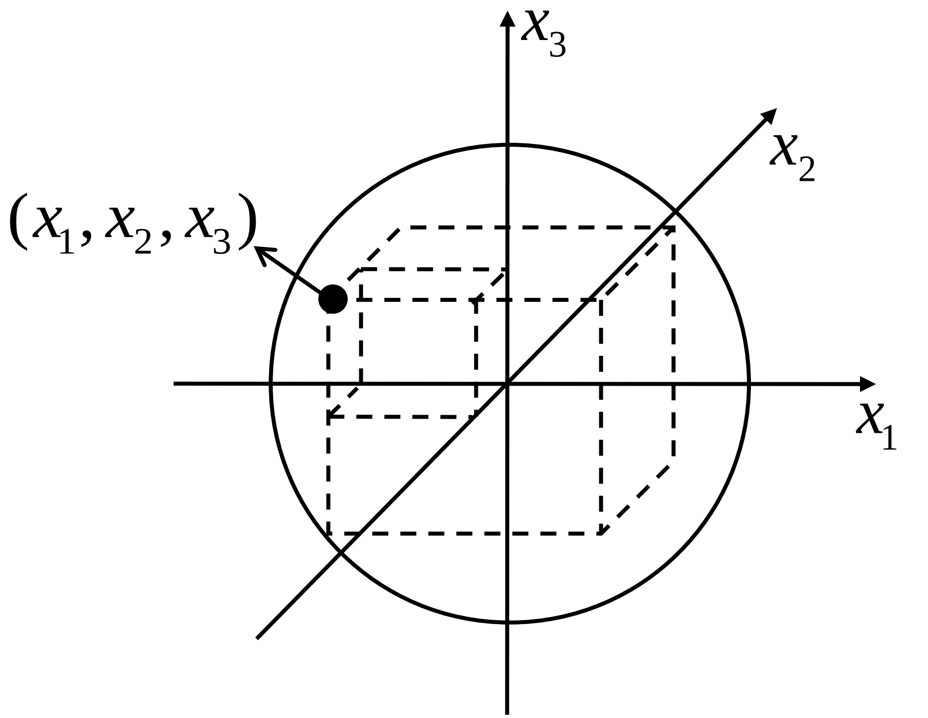
$$\frac{\partial L}{\partial \lambda} = x_1 + x_2 - 2 = 0$$

$\Rightarrow x_1^* = x_2^* = 1, \lambda^* = 1$ **(The point C is a stationary point.)**



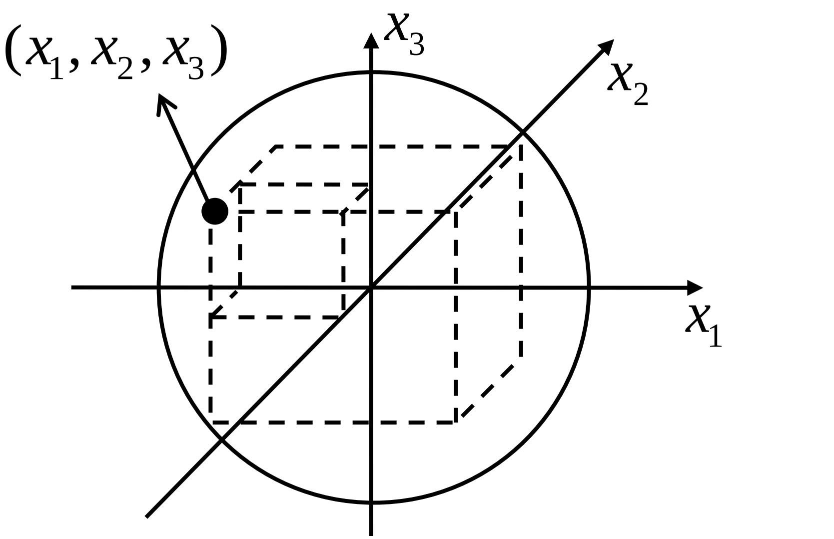
☑ There is a sphere whose center is (0,0,0) and radius is c.

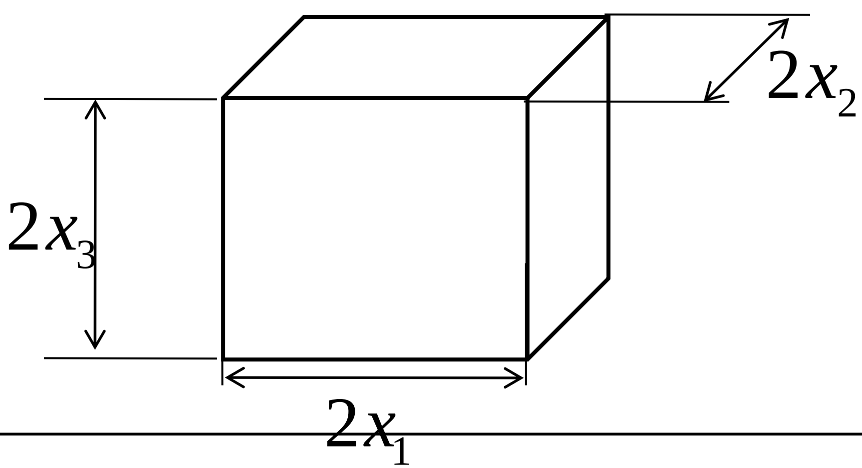
☑ Determine the maximum volume of the rectangular solid which is circumscribed* in the sphere.



*to draw a geometric figure around another figure so that the two are in contact but do not intersect

☑ Mathematical Modeling





The volume of the rectangular sold f is

$$f(x_1, x_2, x_3) = 2x_1 \cdot 2x_2 \cdot 2x_3$$

Because the vertices of the rectangular solid are on the surface of the sphere,

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - c^2 = 0$$

$$cf)\ equation\ \ for\ \ a\ \ sphere: x^2 + y^2 + z^2 = r^2$$

$$maximize: f(x_1, x_2, x_3) = 2x_1 \cdot 2x_2 \cdot 2x_3$$
$$= 8x_1x_2x_3$$

$constraint:$

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - c^2 = 0$$

$$minimize: f(x_1, x_2, x_3) = -8x_1x_2x_3$$

$constraint:$

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - c^2 = 0$$

☑ Solution(1/2)

$$minimize: f(x_1, x_2, x_3) = -8x_1x_2x_3$$
$$constraint:$$
$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - c^2 = 0$$

Lagrange function of this problem is as follow.

$$L(x_1, x_2, x_3, \lambda) = f(x_1, x_2, x_3) + \lambda h(x_1, x_2, x_3)$$
$$= -8x_1x_2x_3 + \lambda(x_1^2 + x_2^2 + x_3^2 - c^2)$$

$$\nabla L(x_1, x_2, x_3, \lambda) = 0 \Rightarrow$$

$$\frac{\partial L}{\partial x_1} = -8x_2x_3 + \lambda 2x_1 = 0 \qquad \frac{\partial L}{\partial x_2} = -8x_1x_3 + \lambda 2x_2 = 0$$

$$\frac{\partial L}{\partial x_3} = -8x_1x_2 + \lambda 2x_3 = 0 \qquad \frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 + x_3^2 - c^2 = 0$$

☑ Solution(2/2)

$$-8x_2x_3+\lambda 2x_1=0 \quad \text{------ ①}$$

$$-8x_1x_3+\lambda 2x_2=0 \quad \text{------ ②}$$

$$-8x_1x_2+\lambda 2x_3=0 \quad \text{------ ③}$$

$$x_1^2+x_2^2+x_3^2-c^2=0 \quad \text{------ ④}$$

Equation ① X x_1 $\quad -8x_1x_2x_3+\lambda 2x_1^2=0 \quad \Rightarrow \quad x_1^2=\dfrac{4x_1x_2x_3}{\lambda}$

Equation ② X x_2 $\quad -8x_1x_2x_3+\lambda 2x_2^2=0 \quad \Rightarrow \quad x_2^2=\dfrac{4x_1x_2x_3}{\lambda}$

Equation ③ X x_3 $\quad -8x_1x_2x_3+\lambda 2x_3^2=0 \quad \Rightarrow \quad x_3^2=\dfrac{4x_1x_2x_3}{\lambda}$

Substitute these into the equation ④

$$\frac{4x_1x_2x_3}{\lambda}+\frac{4x_1x_2x_3}{\lambda}+\frac{4x_1x_2x_3}{\lambda}-c^2=0$$

$$\frac{12x_1x_2x_3}{\lambda}=c^2$$

$$\frac{12x_1x_2x_3}{c^2}=\lambda \quad \text{------ ⑤}$$

Substitute the equation ⑤ into the equation ①

$$-8x_2x_3+\frac{12x_1x_2x_3}{c^2}2x_1=0$$

$$-8x_2x_3+\frac{24x_1^2x_2x_3}{c^2}=0$$

$$-8x_2x_3\left(1-\frac{3x_1^2}{c^2}\right)=0$$

If x_2 and x_3 are zero 0, the volume of the rectangular solid is zero and the result is not correct.

$$1-\frac{3x_1^2}{c^2}=0$$

$$\frac{3x_1^2}{c^2}=1$$

$$x_1^2=\frac{c^2}{3}$$

$$x_1=\pm\frac{c}{\sqrt{3}}$$

Because x_1 is the length, it is positive. x_2 and x_3 are obtained in the same way.

$$x_1=\frac{c}{\sqrt{3}},\quad x_2=\frac{c}{\sqrt{3}},\quad x_3=\frac{c}{\sqrt{3}}$$

So, the maximum volume is

$$8x_1x_2x_3=\frac{8c^3}{3\sqrt{3}}$$

4.2 Lagrange Multiplier for equality constraints

At the candidate minimum C, the meaning of is

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

$\nabla f(\mathbf{x})$: The direction where $f(\mathbf{x})$ is increased

$\nabla h(\mathbf{x})$: The direction where $h(\mathbf{x})$ is increased

Original Problem

Minimize $f(\mathbf{x}) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

Subject to $h(\mathbf{x}) = x_1 + x_2 - 2 = 0$

Lagrange Function

Minimize $L(\mathbf{x}, \nu) = f(\mathbf{x}) + \nu h(\mathbf{x})$

$$= (x_1 - 1.5)^2 + (x_2 - 1.5)^2 + \nu(x_1 + x_2 - 2)$$

Necessary Condition: $\nabla L(\mathbf{x}^*, \nu^*) = 0$

$$\nabla f(\mathbf{x}^*) + \nu^* \nabla h(\mathbf{x}^*) = 0$$

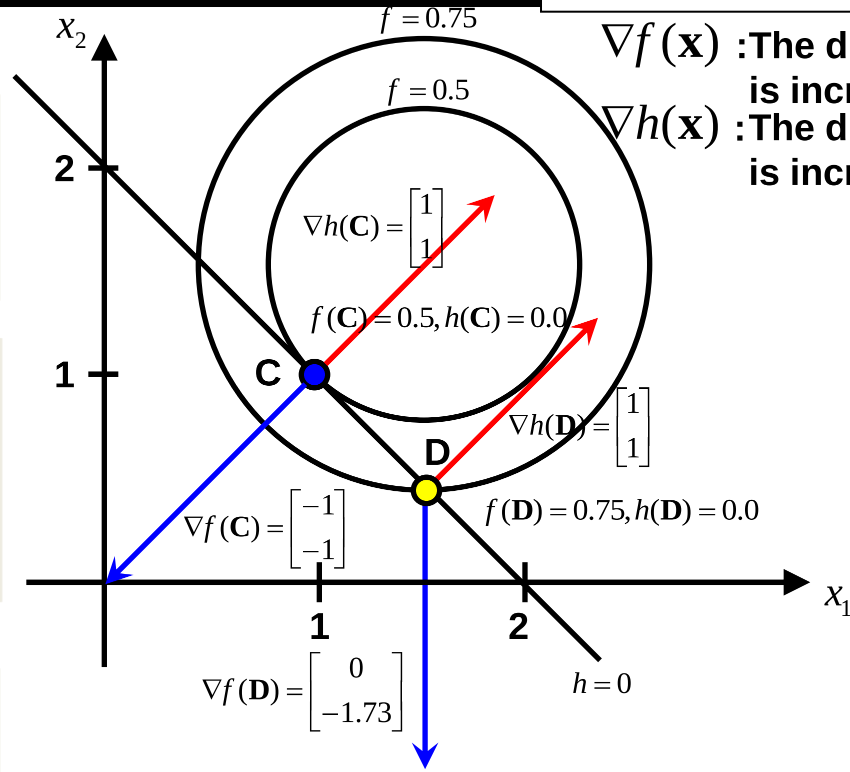
$$\therefore -\nabla f(\mathbf{x}^*) = \nu^* \nabla h(\mathbf{x}^*)$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 1.5) \\ 2(x_2 - 1.5) \end{bmatrix}, \quad \nabla h(\mathbf{x}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$-2(x_1^* - 1.5) = \nu^*, \; -2(x_2^* - 1.5) = \nu^*$$

$$x_1^* + x_2^* - 2 = 0$$

$\Rightarrow x_1^* = x_2^* = 1, \nu^* = 1$ **(point C)**



At the candidate minimum C, the meaning of $-\nabla f(\mathbf{x}^*) = \nu^* \nabla h(\mathbf{x}^*)$ **is**

The gradient vector of the objective function and constraint are on the same line and proportional to each other, and the Lagrange multiplier ν^* is the proportionality constant.

$$\nabla f(\mathbf{C}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla h(\mathbf{C}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \nu^* = 1$$

But point D is not a candidate minimum, because the gradient vector of the objective function and constraint are not on the same line.

☑Constrained Optimization Problem

$$\textit{Minimize} \quad f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$$

$$\textit{Subject to} \quad h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p$$

- Determination of the propeller main dimensions by using the Lagrange multiplier
- Determination of the main dimension of a ship by using the Lagrange multiplier

☑Definition of the Lagrange function(L)

$$L(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x})$$

$$= f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x})$$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \end{bmatrix}$$

v_i are the Lagrange multipliers for the equality constraints and are free in sign, i.e., they can be positive, negative, or zero.

<Reason>

The solution does not change, even if the equality constraint is multiplied by the minus sign,

Newton' Method for Unconstrained Optimization Problem

Given: $\mathit{Minimize} \quad f(\mathbf{x})$

Find: Local minimum design point

- By defining $\mathbf{x}-\mathbf{x}^* = \mathbf{d}$, the Taylor series expansion for the function of multi variables is as follows.

$$f(\mathbf{x}^* + \mathbf{d}) = f(\mathbf{x}^*) + \boxed{\nabla f(\mathbf{x}^*)^T}\,\mathbf{d} + \boxed{\frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d}} + R$$

Necessary condition for $\mathbf{x} = \mathbf{x}^*$ to be a candidate local minimum (stationary point)

$$\boxed{\nabla f(\mathbf{x}^*)^T = 0,}\ \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$$

Sufficient conditions for $\mathbf{x} = \mathbf{x}^*$ to be a local minimum

Method of Lagrange Multipliers for Constrained Optimization Problem

Given: $\mathit{Minimize} \quad f(\mathbf{x})$

$h(x_1, x_2, x_3) = 0$

Find: Local candidate minimum design point

$$\boxed{df + \lambda \cdot dh} = 0$$

$\lambda :$ Undetermined Coefficient 'Lagrange multiplier'

Define Lagrange function , $L = df + \lambda \cdot dh$
Necessary condition for $\mathbf{x} = \mathbf{x}^*$
to be a candidate local minimum-> grad L =0
(stationary point)

Minimize $f(x_1, x_2)$, ***Subject to*** $h(x_1, x_2) = 0$

By using $h(x_1,\ x_2)=0$, x_2 **can be expressed as the function of** x_1, **i.e.,** $f(x_1, x_2) = f(x_1, \phi(x_1))$

To determine the local candidate minimum of the function of the single variable,

$df(x_1, x_2)/dx_1 = 0$, **But, because** $df(x_1, x_2) = \frac{\partial f(x_1, x_2)}{\partial x_1}dx_1 + \frac{\partial f(x_1, x_2)}{\partial x_2}dx_2$ **,** $\frac{\partial f(x_1, x_2)}{\partial x_1} + \frac{\partial f(x_1, x_2)}{\partial x_2}\frac{dx_2}{dx_1} = 0$**.**

If we assume that $\mathbf{x}^* = (x_1^*,\ x_2^*)$ **is the local candidate minimum,**

$$\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2}\frac{d\phi(x_1)}{dx_1} = 0 \quad \cdots \text{ Equation (1)}$$

$x_2 = \phi(x_1)$ **is the explicit form, in general, it is impossible to represent the constraint as this from.**

Form the equality constraint $\boxed{h(x_1^*,\ x_2^*) = 0}$**,**

$$\Rightarrow \frac{dh(x_1^*, x_2^*)}{dx_1} = \frac{\partial h(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial h(x_1^*, x_2^*)}{\partial x_2}\frac{d\phi(x_1)}{dx_1} = 0$$

$$\therefore \frac{d\phi(x_1)}{dx_1} = -\frac{\partial h(x_1^*, x_2^*)/\partial x_1}{\partial h(x_1^*, x_2^*)/\partial x_2} \quad \cdots \text{ Equation (2)}$$

Substitute the equation (2) into the equation (1)

$$\therefore \frac{\partial f(x_1^*, x_2^*)}{\partial x_1} - \frac{\partial f(x_1^*, x_2^*)}{\partial x_2}\frac{\partial h(x_1^*, x_2^*)/\partial x_1}{\partial h(x_1^*, x_2^*)/\partial x_2} = 0 \quad \cdots \text{ Equation (3)}$$

If we assume that $\nu^* = -\frac{\partial f(x_1^*, x_2^*)/\partial x_2}{\partial h(x_1^*, x_2^*)/\partial x_2}$ $\cdots$ **Equation (4)**

The equation (3) becomes $\boxed{\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \nu^*\frac{\partial h(x_1^*, x_2^*)}{\partial x_1} = 0}$

Equation (4) can be rearranged as follows.

$$\boxed{\frac{\partial f(x_1^*, x_2^*)}{\partial x_2} + \nu^*\frac{\partial h(x_1^*, x_2^*)}{\partial x_2} = 0}$$

In summary, for $\mathbf{x}^* = (x_1^*,\ x_2^*)$ **to become the local candidate minimum, the following three conditions have to be satisfied.**

$$\boxed{h(x_1^*,\ x_2^*) = 0}$$

$$\boxed{\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \nu^*\frac{\partial h(x_1^*, x_2^*)}{\partial x_1} = 0}$$

$$\boxed{\frac{\partial f(x_1^*, x_2^*)}{\partial x_2} + \nu^*\frac{\partial h(x_1^*, x_2^*)}{\partial x_2} = 0}$$

ν^* is called the Lagrange multiplier.

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints

Ch.4 Optimality Condition Using Kuhn-Tucker Necessary Condition

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Quadratic Programming Problem with Inequality Constraint

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

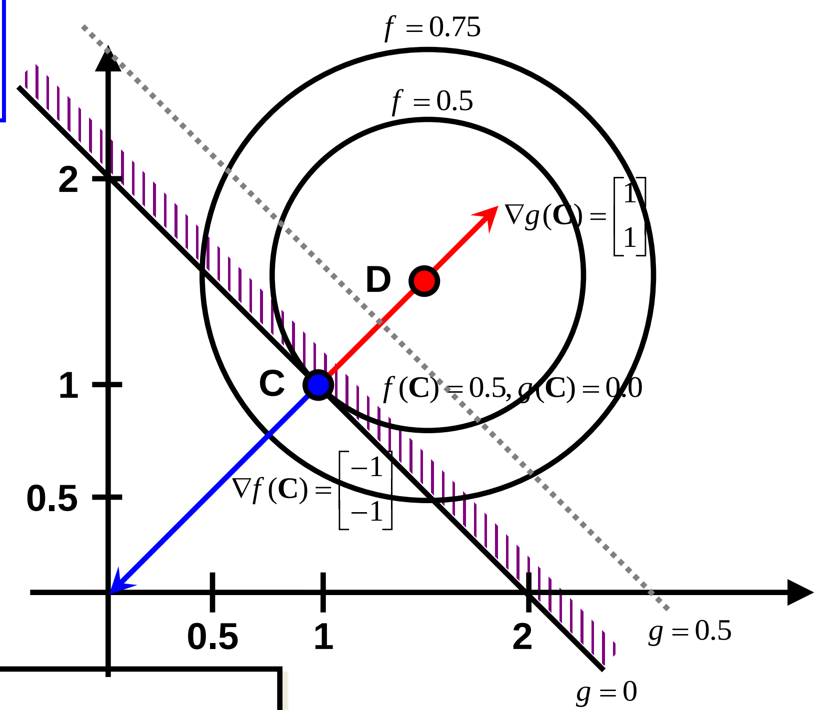
We can transform an inequality constraint to by adding a new variable to it, called the slack variable.

Original Problem

Minimize $f(\mathbf{x}) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$

$\Rightarrow g(\mathbf{x}) + s^2 = x_1 + x_2 - 2 + \boxed{s^2} = 0$



Lagrange Function

Minimize $L(\mathbf{x}, u, s) = f(\mathbf{x}) + u\left[g(\mathbf{x}) + s^2\right]$

$= (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

$+ u(x_1 + x_2 - 2 + s^2)$

Necessary Condition: $\nabla L(\mathbf{x}^*, u^*, s^*) = 0$

$\frac{\partial L}{\partial x_1} = 2(x_1 - 1.5) + u = 0, \quad \frac{\partial L}{\partial x_2} = 2(x_2 - 1.5) + u = 0$ — Linear indeterminate equation

$\frac{\partial L}{\partial u} = x_1 + x_2 - 2 + s^2 = 0,$

$\frac{\partial L}{\partial s} = 2us = 0$ — Nonlinear indeterminate equation

$u \ge 0$

(1) If s = 0,(Inequality constraint is transformed to the equality constraint.)

$x_1^* = x_2^* = 1, u^* = 1 \Rightarrow$ Candidate minimum point(point C)

(2) If u = 0,(the inequality constraint is not active)

$x_1^* = x_2^* = 1.5, u^* = 0, s^2 = -1$ (Point D: the constraint is violated)

- At first, we obtain the solution which satisfies the nonlinear indeterminate equation.

$(u = 0 \; or \; s = 0)$

- And then, we check whether each solution satisfies the linear indeterminate equation.

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- The Necessary Condition for a Candidate Local Optimal Solution in the Inequality Constrained Problem (1)

[Ref] Lagrange function for the equality constrained problem

$$L(\mathbf{x},\mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h(\mathbf{x}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x})$$

v_i are the Lagrange multipliers for the equality constraints and are free in sign.

Inequality constraint

$$g_i(\mathbf{x}) \le 0, \quad i = 1, \ldots, m$$

To transform the inequality constraint s to the equality constraints , the slack variable s s_i^2 are introduced :

$$g_i(\mathbf{x}) + s_i^2 = 0, \quad i = 1, \ldots, m$$

Lagrange function in the inequality constrained problem

Since the inequality constraint is transformed to the equality constraint by introducing the slack variable, the Lagrange function is defined as

$$L(\mathbf{x},\mathbf{u},\mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i (g_i(\mathbf{x}) + s_i^2) = f(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2), \quad \underline{u_i \ge 0}$$

u_i are the Lagrange multiplier for the inequality constraints and have to be nonnegative.
s_i are the slack variables to transform the inequality constraints to the equality.

Lagrange function in the inequality constrained problem

$$L(\mathbf{x},\mathbf{u},\mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{m} u_i (g_i(\mathbf{x}) + s_i^2) = f(\mathbf{x}) + \mathbf{u}^T(\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)$$

u_i are the Lagrange multiplier for the inequality constraints and have to be nonnegative.

s_i are the slack variables to transform the inequality constraints to the equality.

The Necessary condition for the candidate local optimal solution of the inequality constrained problem

$$\nabla L(\mathbf{x}^*,\mathbf{u}^*,\mathbf{s}^*) = \mathbf{0}$$

$$\Downarrow$$

$$\frac{\partial L}{\partial x_j} \equiv \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0, \qquad j = 1, \dots, n$$

$$\frac{\partial L}{\partial u_i} \equiv g_i(\mathbf{x}^*) + s_i^{*2} = 0, \qquad i = 1, \dots, \mathrm{m}$$

$$\frac{\partial L}{\partial s_i} \equiv u_i^* s_i^* = 0, \qquad i = 1, \dots, m$$

$$u_i^* \geq 0, \qquad i = 1, \dots, m$$

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints

Optimization Problem

Minimize $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$

Subject to $h_i(\mathbf{x}) = 0, \quad i = 1,\ldots,p$ **Equality constraints**

$g_i(\mathbf{x}) \le 0, \quad i = 1,\ldots,m$ **Inquality constraints**

Definition of the Lagrange function

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) + \sum_{i=1}^{m} u_i (g_i(\mathbf{x}) + s_i^2)$$

$$= f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)$$

v_i are the Lagrange multipliers for the equality constraints and are free in sign.

u_i are the Lagrange multiplier for the inequality constraints and have to be nonnegative.

s_i are the slack variables to transform the inequality constraints to the equality.

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_j} \equiv \frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0, \qquad j = 1, \ldots, n$$

$$\frac{\partial L}{\partial v_i} \equiv h_i(\mathbf{x}^*) = 0, \qquad i = 1, \ldots, p$$

$$\frac{\partial L}{\partial u_i} \equiv g_i(\mathbf{x}^*) + s_i^{*2} = 0, \qquad i = 1, \ldots, \mathrm{m}$$

$$\frac{\partial L}{\partial s_i} \equiv u_i^* s_i^* = 0, \qquad i = 1, \ldots, m$$

$$u_i^* \ge 0, \qquad i = 1, \ldots, m$$

If x^* is the candidate local minimum point, these equations have to be satisfied. That is, the Kuhn-Tucker necessary condition, which composed of these equations, is the necessary condition for x^* to be the candidate local minimum point.

Therefore, K.-T. condition can be used to find the candidate local minimum point in the equality and inequality constrained problem.

The value of the objective function and gradient vector are calculated at $\mathbf{x}^*$.

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints

[Example] Nonlinear Constrained Optimization Problem (1)

①

$$Minimize\ \ f(\mathbf{x}) = x_1^{\ 2} + x_2^{\ 2} - 3x_1x_2$$

$$g(\mathbf{x}) = x_1^{\ 2} + x_2^{\ 2} - 6 \le 0$$

②

$$L(\mathbf{x}, u, s) = x_1^{\ 2} + x_2^{\ 2} - 3x_1x_2 + u(x_1^{\ 2} + x_2^{\ 2} - 6 + s^2)$$

③

$$\frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2ux_1 = 0 \ \text{--------} \ ①$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2ux_2 = 0 \ \text{--------} \ ②$$

$$\frac{\partial L}{\partial u} = x_1^{\ 2} + x_2^{\ 2} - 6 + s^2 = 0,\ s^2 \ge 0,\ u \ge 0 \ \text{--------} \ ③$$

$$\frac{\partial L}{\partial s} = \boxed{2us = 0}$$ ➡ **There are two cases.**

CASE #1 : $u = 0$ **(The inequality constraint is considered as inactive at the solution point.)**

$$2x_1 - 3x_2 = 0$$

$$-3x_1 + 2x_2 = 0$$

➡ **Point A:** $x_1^* = 0,\ x_2^* = 0,\ f(x_1^*, x_2^*) = 0$

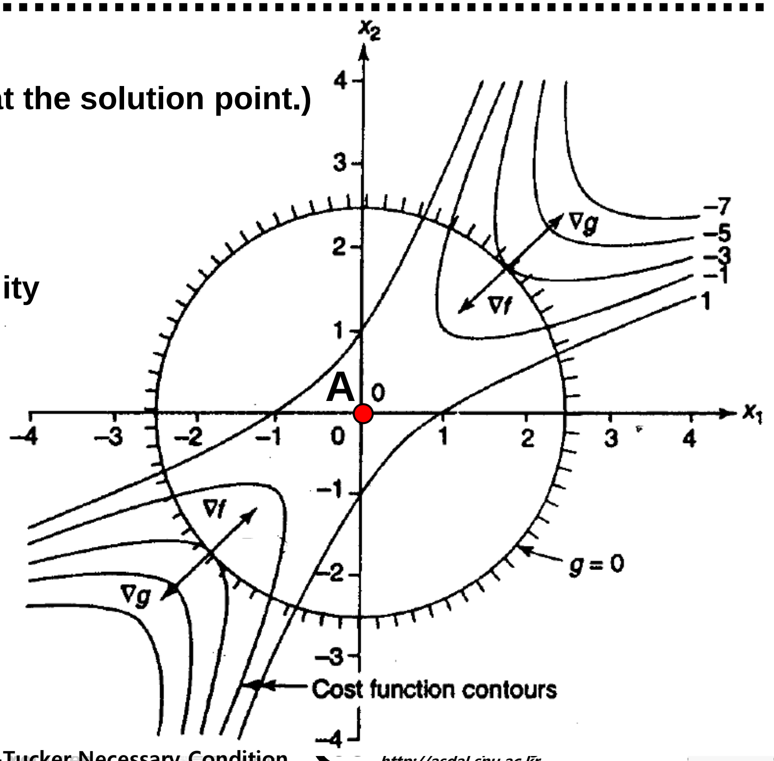
CASE #2 : $s = 0$ **(The solution point is on the boundary of the inequality constraint.)**

Rearrange the equation① $2x_1 - 3x_2 + 2ux_1 = 0,\ \ u = -1 + \frac{3}{2}\frac{x_2}{x_1}$

Substitute u into the equation ② $2x_2 - 3x_1 + 2(-1 + \frac{3}{2}\frac{x_2}{x_1})x_2 = 0$

$$2x_2 - 3x_1 - 2x_2 + 3\frac{x_2^2}{x_1} = 0,\quad 3\frac{x_2^2}{x_1} = 3x_1,\quad x_2^2 = x_1^2$$

Substitute x_2 into the equation ③ $2x_1^{\ 2} - 6,\quad x_1 = \pm\sqrt{3}$



4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Finding the Candidate Local Optimal Solution by using the Kuhn-Tucker Necessary Condition -
Nonlinear Constrained Optimization Problem (2)

①

$$\textit{Minimize} \;\; f(\mathbf{x}) = x_1^{\,2} + x_2^{\,2} - 3x_1x_2$$

$$g(\mathbf{x}) = x_1^{\,2} + x_2^{\,2} - 6 \le 0$$

②

$$L(\mathbf{x}, u, s) = x_1^{\,2} + x_2^{\,2} - 3x_1x_2 + u(x_1^{\,2} + x_2^{\,2} - 6 + s^2)$$

③

$$\frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2ux_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2ux_2 = 0$$

$$\frac{\partial L}{\partial u} = x_1^{\,2} + x_2^{\,2} - 6 + s^2 = 0,\; s^2 \ge 0,\; u \ge 0$$

$$\frac{\partial L}{\partial s} = \boxed{2us = 0}$$

➡ **There are two cases.**

CASE #1 : $u = 0$ **(The inequality constraint is considered as inactive at the solution point.)**

$$2x_1 - 3x_2 = 0$$
$$-3x_1 + 2x_2 = 0$$

➡ **Point A:** $x_1^* = 0,\; x_2^* = 0,\; f(x_1^*, x_2^*) = 0$

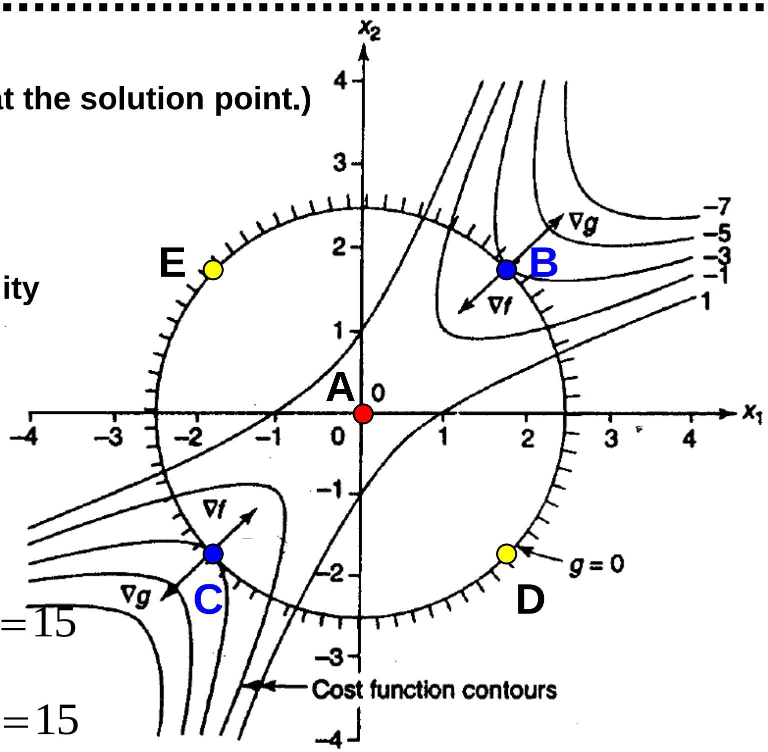
CASE #2 : $s = 0$ **(The solution point is on the boundary of the inequality constraint.)**

$x_1 = x_2 = \sqrt{3},\; u = \dfrac{1}{2}$ ➡ **Point B:** $x_1^* = x_2^* = \sqrt{3},\; f(x_1^*, x_2^*) = -3$

$x_1 = x_2 = -\sqrt{3},\; u = \dfrac{1}{2}$ ➡ **Point C:** $x_1^* = x_2^* = -\sqrt{3},\; f(x_1^*, x_2^*) = -3$

$x_1 = -x_2 = \sqrt{3},\; u = -\dfrac{5}{2}$ ➡ **Point D:** $x_1^* = \sqrt{3},\; x_2^* = -\sqrt{3},\; f(x_1^*, x_2^*) = 15$

$x_1 = -x_2 = -\sqrt{3},\; u = -\dfrac{5}{2}$ ➡ **Point E:** $x_1^* = -\sqrt{3},\; x_2^* = \sqrt{3},\; f(x_1^*, x_2^*) = 15$



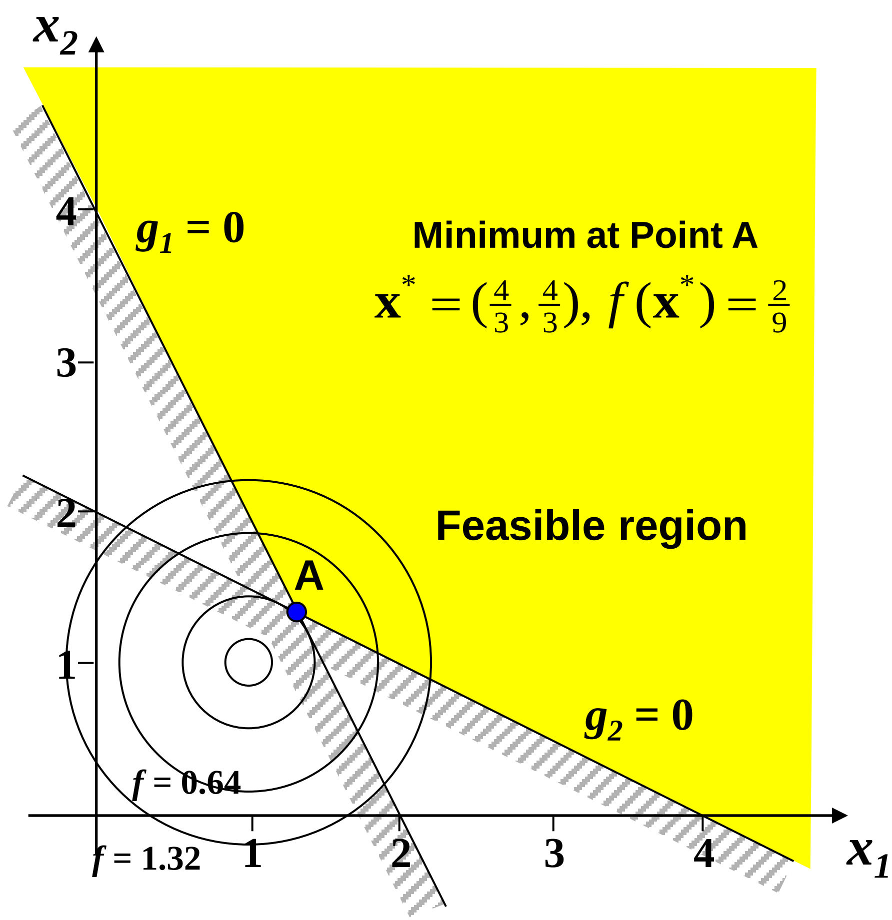
- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are free in sign (1)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$

$$\textit{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$

$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$



Lagrange function

$$L(\mathbf{x},\mathbf{u},\mathbf{s}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$

- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are free in sign (2)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Lagrange function

$$L(\mathbf{x},\mathbf{u},\mathbf{s}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 = 0 \qquad \frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 = 0$$

$$\frac{\partial L}{\partial u_1} = -2x_1 - x_2 + 4 + s_1^2 = 0 \qquad \frac{\partial L}{\partial u_2} = -x_1 - 2x_2 + 4 + s_2^2 = 0$$

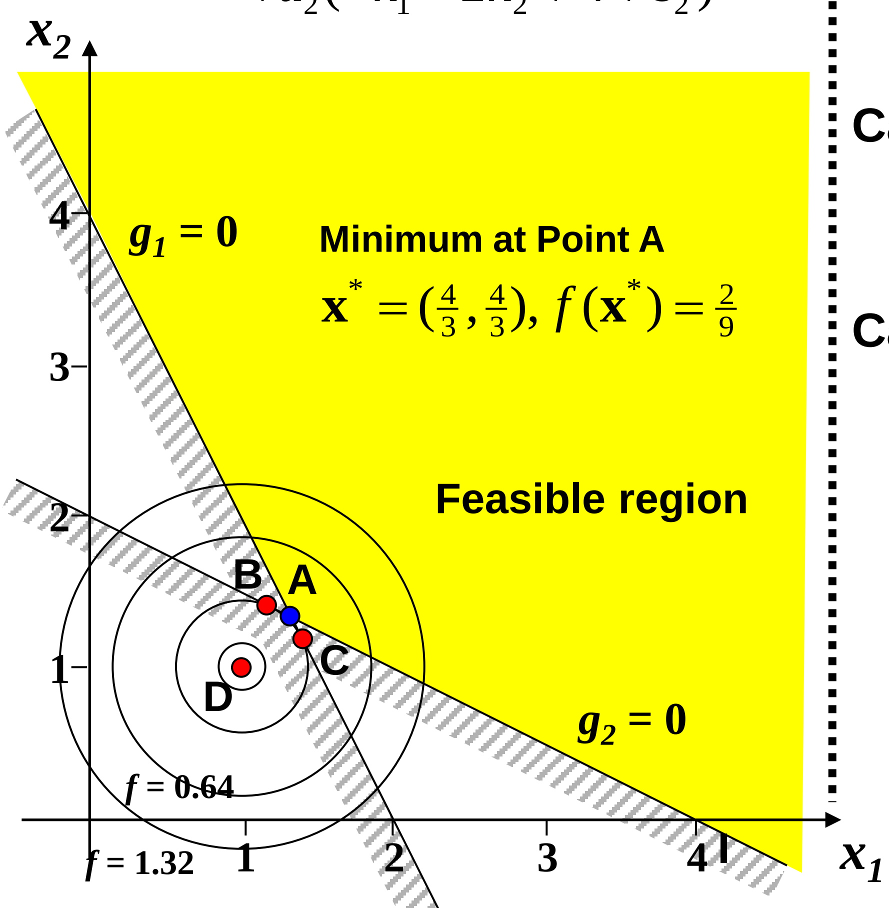
$$\frac{\partial L}{\partial s_1} = 2u_1 s_1 = 0 \qquad \frac{\partial L}{\partial s_2} = 2u_2 s_2 = 0 \qquad u_i \ge 0,\ i = 1, 2$$

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are free in sign (3)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Lagrange function

$$L(\mathbf{x},\mathbf{u},\mathbf{s}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$



Case #1: s_1=s_2=0, (Minimum at Point A)

$$x_1 = x_2 = \tfrac{4}{3},\ u_1 = u_2 = \tfrac{2}{9}$$

Case #2: u_1=s_2=0, (Point B)

$$x_1 = \tfrac{6}{5},\ x_2 = \tfrac{7}{5},\ u_2 = \tfrac{2}{5},\ s_1^2 = -\tfrac{1}{5}$$

It has to be nonnegative(g_1).

Case #3: u_2=s_1=0, (Point C)

$$x_1 = \tfrac{7}{5},\ x_2 = \tfrac{6}{5},\ u_1 = \tfrac{2}{5},\ s_2^2 = -\tfrac{1}{5}$$

It has to be nonnegative(g_2).

Case #4: u_1=u_2=0, (Point D)

$$x_1 = x_2 = 1,\ s_1^2 = s_2^2 = -1$$

It has to be nonnegative(g_2, g_2).

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are nonnegative (1)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

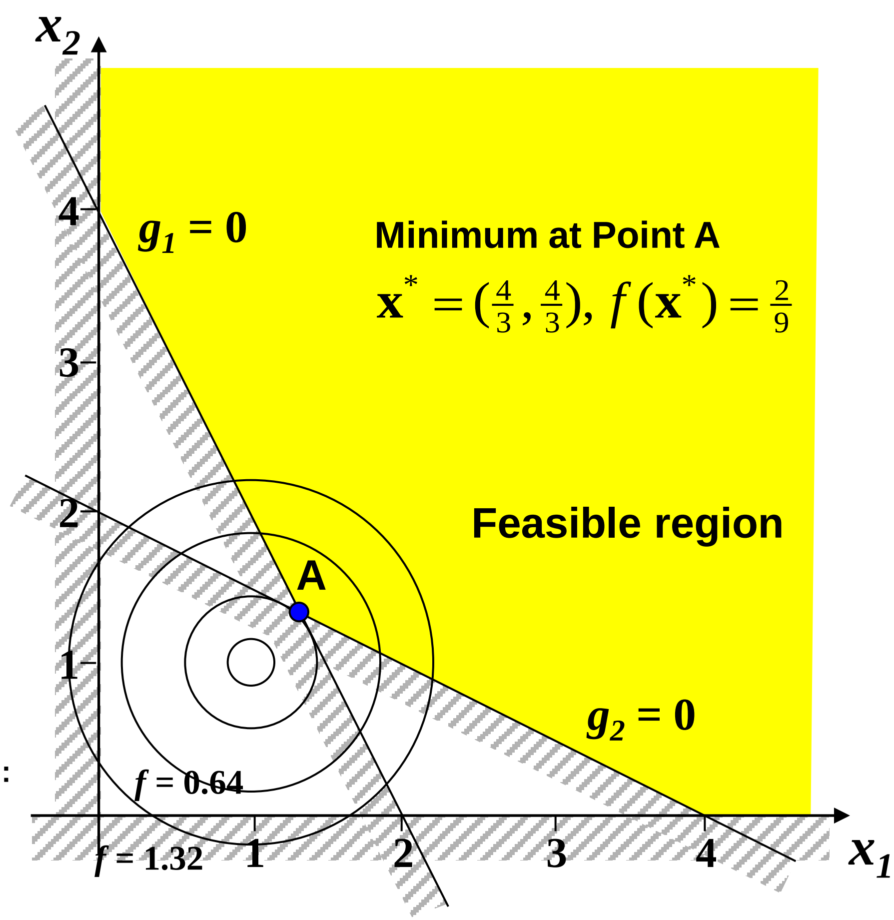
$$\textit{Minimize}\quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to}\quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$
$$x_1 \ge 0, x_2 \ge 0$$

Minimum point: $\mathbf{x}^* = (\frac{4}{3}, \frac{4}{3}),\ f(\mathbf{x}^*) = \frac{2}{9}$

$$\textit{Minimize}\quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to}\quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$
$$-x_1 \le 0, -x_2 \le 0$$

Inequality constraints whose form are"≤":
Introducing the slack variable

$$\textit{Minimize}\quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to}\quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 + s_1^2 = 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 + s_2^2 = 0$$
$$-x_1 + \delta_1^2 = 0, -x_2 + \delta_2^2 = 0$$



4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are nonnegative (2)

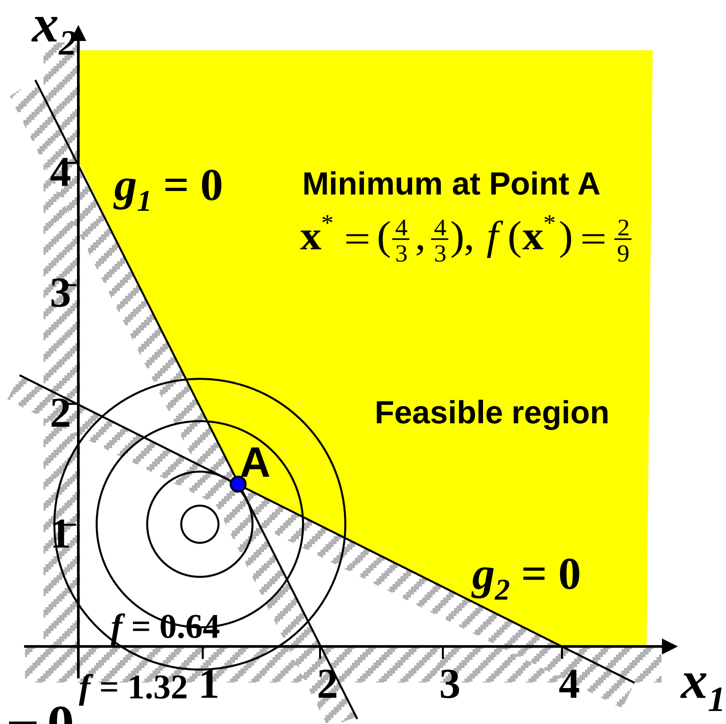
Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$

$$\textit{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 + s_1^2 = 0$$

$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 + s_2^2 = 0$$

$$-x_1 + \delta_1^2 = 0, -x_2 + \delta_2^2 = 0$$



Lagrange function

$$L(\mathbf{x}, \mathbf{u}, \mathbf{s}, \boldsymbol{\zeta}, \boldsymbol{\delta}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$
$$+\zeta_1(-x_1 + \delta_1^2) + \zeta_2(-x_2 + \delta_2^2)$$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x}, \mathbf{u}, \mathbf{s}, \boldsymbol{\zeta}, \boldsymbol{\delta}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 - \zeta_1 = 0 \qquad \frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 - \zeta_2 = 0$$

$$\frac{\partial L}{\partial u_1} = -2x_1 - x_2 + 4 + s_1^2 = 0 \qquad \frac{\partial L}{\partial u_2} = -x_1 - 2x_2 + 4 + s_1^2 = 0$$

$$\frac{\partial L}{\partial s_1} = 2u_1 s_1 = 0 \qquad \frac{\partial L}{\partial s_2} = 2u_2 s_2 = 0$$

$$\frac{\partial L}{\partial \zeta_1} = \delta_1^2 - x_1 = 0 \qquad \frac{\partial L}{\partial \zeta_2} = \delta_2^2 - x_2 = 0 \qquad \frac{\partial L}{\partial \delta_1} = 2\zeta_1\delta_1 = 0 \qquad \frac{\partial L}{\partial \delta_1} = 2\zeta_1\delta_1 = 0$$

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta})=\mathbf{0}$

$$\frac{\partial L}{\partial x_1}=2x_1-2-2u_1-u_2-\zeta_1=0 \qquad \frac{\partial L}{\partial x_2}=2x_2-2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1^2=0 \qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2^2=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1=0 \qquad \frac{\partial L}{\partial s_2}=2u_2s_2=0$$

$$\frac{\partial L}{\partial \zeta_1}=\delta_1^2-x_1=0\rightarrow\delta_1^2=x_1 \qquad \frac{\partial L}{\partial \zeta_2}=\delta_2^2-x_2=0\rightarrow\delta_2^2=x_2$$

Substitute   Substitute

$$\frac{\partial L}{\partial \delta_1}=2\zeta_1\delta_1=0\rightarrow 2\zeta_1\delta_1^2=0 \qquad \frac{\partial L}{\partial \delta_2}=2\zeta_1\delta_1=0\rightarrow 2\zeta_2\delta_2^2=0$$

Multiply δ_1 to the both sides.   Multiply δ_2 to the both sides.

$$u_i,\zeta_i\ge 0,\ i=1,2$$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta})=\mathbf{0}$

$$\frac{\partial L}{\partial x_1}=2x_1-2-2u_1-u_2-\zeta_1=0 \qquad \frac{\partial L}{\partial x_2}=2x_2-2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1^2=0 \qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2^2=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1=0 \qquad \frac{\partial L}{\partial s_2}=2u_2s_2=0$$

$$2\zeta_1x_1=0 \qquad 2\zeta_2x_2=0$$

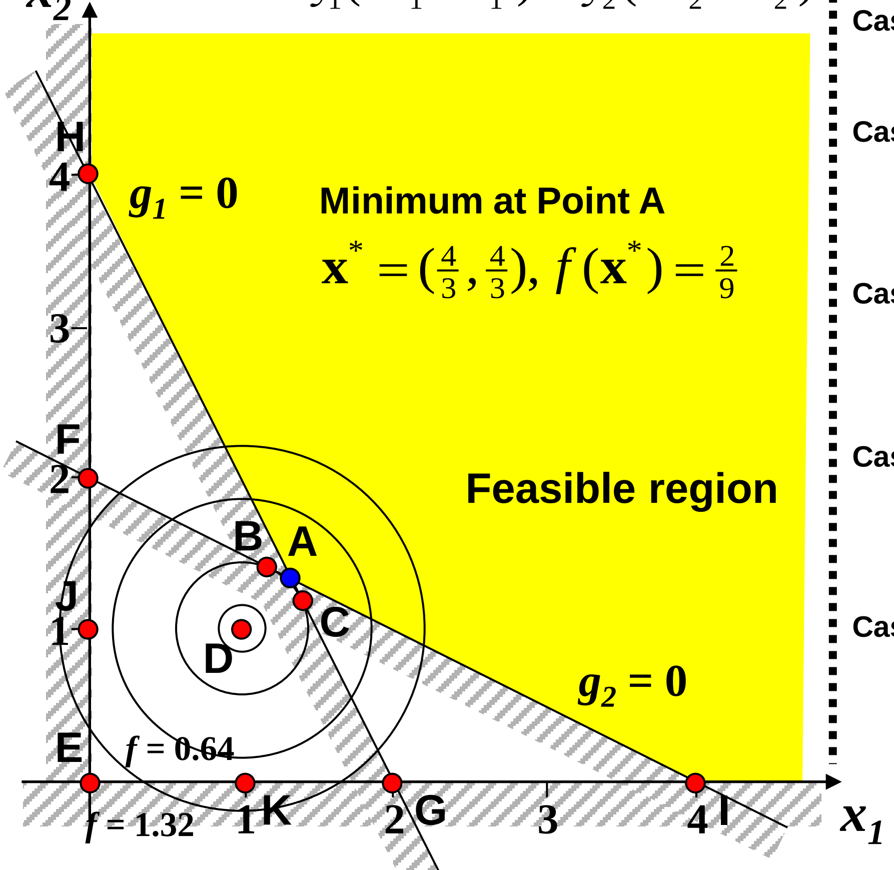
$$u_i,\zeta_i,\delta_i\ge 0,\ i=1,2$$

4.3 Kuhn-Tucker Necessary Condition for Inequality constraints
- Finding the Optimal Solution in the Quadratic Programming Problem by using the Kuhn-Tucker Necessary Condition – xi are nonnegative (4)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Lagrangian function

$$L(\mathbf{x},\mathbf{u},\mathbf{s},\zeta,\boldsymbol{\delta}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$
$$+\zeta_1(-x_1 + \delta_1^2) + \zeta_2(-x_2 + \delta_2^2)$$



Case #1: $s_1=s_2=\zeta_1=\zeta_2=0$, (Point A)

$$x_1 = x_2 = \tfrac{4}{3}, u_1 = u_2 = \tfrac{2}{9}$$

Case #2: $u_1=s_2=\zeta_1=\zeta_2=0$, (Point B)

$$x_1 = \tfrac{6}{5}, x_2 = \tfrac{7}{5}, u_2 = \tfrac{2}{5}, s_1^2 = -\tfrac{1}{5}$$

It has to be nonnegative.

Case #3: $u_2=s_1=\zeta_1=\zeta_2=0$, (Point C)

$$x_1 = \tfrac{7}{5}, x_2 = \tfrac{6}{5}, u_1 = \tfrac{2}{5}, s_2^2 = -\tfrac{1}{5}$$

It has to be nonnegative.

Case #4: $u_1=u_2=\zeta_1=\zeta_2=0$, (Point D)

$$x_1 = x_2 = 1, s_1^2 = s_2^2 = -1$$

It has to be nonnegative.

Case #5: $u_1=u_2=x_1=x_2=0$, (Point E)

$$x_1 = x_2 = 0, s_1^2 = s_2^2 = -4,$$

It has to be nonnegative.

$$\zeta_1 = \zeta_2 = -2$$

Case #6: $u_1=s_2=x_1=x_2=0$, (Point E)

$$x_1 = x_2 = 0, s_1^2 = -4,$$

It has to be nonnegative.

$$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$$

The constraint is violated.

Case #7: $u_2=s_1=x_1=x_2=0$, (Point E)

$$x_1 = x_2 = 0, s_2^2 = -4,$$

It has to be nonnegative.

$$-2x_1 - x_2 + 4 + s_1^2 \neq 0$$

The constraint is violated.

Case #8: $s_1=s_2=x_1=x_2=0$, (Point E)

$$x_1 = x_2 = 0, -2x_1 - x_2 + 4 + s_1^2 \neq 0,$$
$$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$$

The constraint is violated.

Case #9: $u_1=s_2=\zeta_2=x_1=0$, (Point F)

$$x_1 = 0, x_2 = 2, u_2 = 1,$$
$$s_1^2 = -2, \zeta_1 = -3$$

It has to be nonnegative.

Case #10: $u_2=s_1=\zeta_1=x_2=0$, (Point G)

$$x_1 = 2, x_2 = 0, u_1 = 1, s_2^2 = -2,$$
$$\zeta_2 = -3$$

It has to be nonnegative.

Case #11: $s_1=s_2=\zeta_1=x_2=0$, (Point G)

$$x_1 = 2, x_2 = 0,$$

The constraint is violated.

$$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$$

Case #12: $u_2=s_1=\zeta_2=x_1=0$, (Point H)

$$x_1 = 0, x_2 = 4, u_1 = 6,$$
$$s_2^2 = 4, \zeta_1 = -14$$

It has to be nonnegative.

Case #13: $s_1=s_2=\zeta_2=x_1=0$, (Point H)

$$x_1 = 0, x_2 = 4,$$

The constraint is violated.

$$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$$

Case #14: $u_1=s_2=\zeta_1=x_2=0$, (Point I)

$$x_1 = 4, x_2 = 0, u_2 = 6,$$
$$s_1^2 = 4, \zeta_2 = -14$$

It has to be nonnegative.

Case #15: $u_1=u_2=\zeta_2=x_1=0$, (Point J)

$$x_1 = 0, x_2 = 1, s_1^2 = -3,$$
$$s_2^2 = -2, \zeta_1 = -2$$

It has to be nonnegative.

Case #16: $u_1=u_2=\zeta_1=x_2=0$, (Point K)

$$x_1 = 1, x_2 = 0, s_1^2 = -2,$$

It has to be nonnegative.

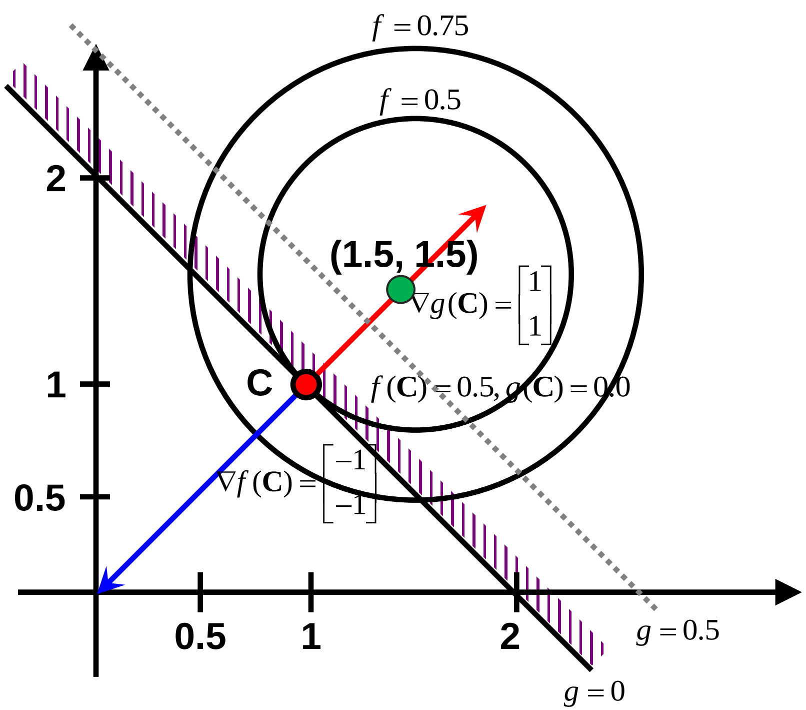
$$s_2^2 = -3, \zeta_2 = -2$$

-[Reference] The Reason Why Lagrange Multiplier for the Inequality Constraint has to be Positive

Original Problem

Minimize $f(\mathbf{x}) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$

Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$



→ : ∇f **Direction of the gradients of the objective function**

→ : ∇g **Direction of the gradients of the constraint**

If $u > 0$, the gradients of the objective and the constraint function point in opposite directions $-\nabla f = \nabla g$

To reduce the value of the objective function f, the design point steps in the negative gradient direction.

However, at the green point(1.5, 1.5), for examp

$$g(\mathbf{x}) = x_1 + x_2 - 2$$
$$= 1.5 + 1.5 - 2 = 1 \not\le 0$$

the constraint is violated.

Therefore, this way, f cannot be reduced any further by stepping in the negative gradient direction without violating the constraint

That is, the point C is the optimal solution satisfying the constraint and minimizing the objective function.

Example of a Constrained Nonlinear Optimization Method by using the Lagrange Multiplier

- Determination of the Optimum Main Dimensions of a Ship

- Determination of the Optimum Propeller Main Dimensions

- Determination of the Optimum Main Dimensions of a Ship (1)

- ***Given***: ***DWT, $V_{H.req}$, D, Ts, Td, $C_{B.d}$***
- ***Find*** : ***L, B, $C_{B.s}$***
  - **Hydrostatic equilibrium(Weight equation)**

$$L \cdot B \cdot T_s \cdot C_{B,d} \cdot \rho_{sw} \cdot C_\alpha = DWT_{given} + LWT(L,B,D,C_{B,d})$$

$$= DWT_{given} + \boxed{C_s \cdot L^{1.6} \cdot (B+D)} + C_o \cdot L \cdot B + \boxed{C_{power} \cdot (L \cdot B \cdot T_d \cdot C_{B,d})^{2/3} \cdot V^3} \quad ...(a)$$

Assumption①

$$\rightarrow C_s' \cdot L^{2.0} \cdot (B+D)$$

Assumption②,

$$\rightarrow C_{power}' \cdot (2 \cdot B \cdot T_d + 2 \cdot L \cdot T_d + L \cdot B) \cdot V^3$$

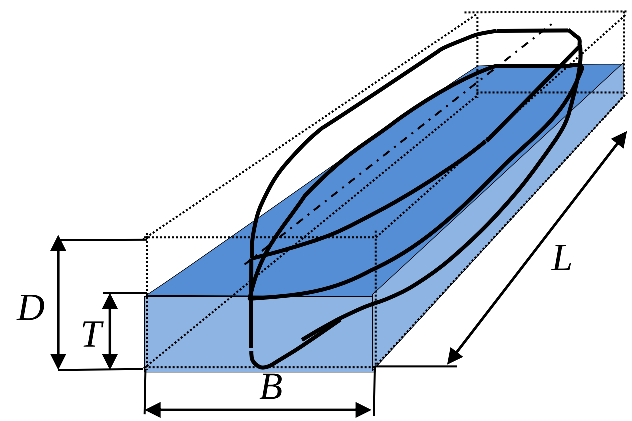
$(L \cdot B \cdot T \cdot C_B)^{2/3}$ **is Volume$^{2/3}$ and means the submerged area of the ship.**
So, we assume that the submerged area of the ship is equal to the submerged area of the rectangular box.

  - **Required cargo hold capacity(Volume equation)**

$$V_{H.req} = C_H \cdot L \cdot B \cdot D \quad ...(b)$$

  - **Recommended range of obesity coefficient with respect to the maneuverability**

$$\frac{C_{B,d}}{(L/B)} < 0.15 \quad ...(c)$$



➡ **Indeterminate Equation: 3 variables($L,B,C_{B,d}$), 2 equality constraints,((a), (b))**

⇨ **It can be solved as the optimization problem to minimize the objective function.**

- ***Given***: ***DWT, $V_{H.req}$, D, Ts, Td, $C_{B.d}$***
- ***Find*** : ***L, B, $C_{B.s}$***
- ***Minimize*** : **Building Cost**

$$f(L,B,C_{B.s}) = C_{PS} \cdot C_s{}' \cdot L^{2.0} \cdot (B+D) + C_{PO} \cdot C_o \cdot L \cdot B + C_{PM} \cdot C_{power}{}' \cdot (2 \cdot B \cdot T_d + 2 \cdot L \cdot T_d + L \cdot B) \cdot V^3 \quad \ldots(e)$$

- ***Subject to***
  - **Hydrostatic equilibrium(Weight equation)**

$$L \cdot B \cdot T_s \cdot C_B \cdot \rho_{sw} \cdot C_\alpha = DWT_{given} + LWT(L,B,D,C_B)$$
$$= DWT_{given} + C_s{}' \cdot L^{2.0} \cdot (B+D) + C_o \cdot L \cdot B + C_{power}{}' \cdot (2 \cdot B \cdot T_d + 2 \cdot L \cdot T_d + L \cdot B) \cdot V^3 \quad \ldots(b)$$

  - **Required cargo hold capacity(Volume equation)**

$$V_{H.req} = C_H \cdot L \cdot B \cdot D \quad \ldots(c)$$

  - **Recommended range of obesity coefficient with respect to the maneuverability**

$$\frac{C_{B.s}}{(L/B)} < 0.15 \quad \ldots(d)$$

- **By introducing the Lagrange multipliers λ_1, λ_2, u, formulate the Lagrange function H.**

$$H\left(L,B,C_{B.s},\lambda_1,\lambda_2,u,s\right)=f\left(L,B,C_{B.s}\right)+\lambda_1\cdot h_1\left(L,B,C_{B.s}\right)+\lambda_2\cdot h_2\left(L,B,D\right)+u\cdot g\left(L,B,C_{B.s},s\right)\quad ...(e)$$

$$f\left(L,B,C_{B.s}\right)=C_{PS}\cdot C_s'\cdot L^2\cdot(B+D)+C_{PO}\cdot C_o\cdot L\cdot B+C_{PM}\cdot C_{power}'\cdot\{2\cdot\left(B+L\right)\cdot T_d+L\cdot B\}\cdot V^3$$

$$h_1\left(L,B,C_{B.s}\right)=L\cdot B\cdot T_s\cdot C_B\cdot\rho_{sw}\cdot C_\alpha-DWT_{given}-C_s'\cdot L^{2.0}\cdot(B+D)-C_o\cdot L\cdot B-C_{power}'\cdot\{2\cdot\left(B+L\right)\cdot T_d+L\cdot B\}\cdot V^3$$

$$h_2\left(L,B,D\right)=C_H\cdot L\cdot B\cdot D-V_{H\_req}$$

$$g\left(L,B,C_{B.s},s\right)=\frac{C_{B.s}}{\left(L/B\right)}-0.15+s^2$$

$$L\rightarrow x_1,B\rightarrow x_2,C_B\rightarrow x_3$$

$$H\left(x_1,x_2,x_3,\lambda_1,\lambda_2,u,s\right)$$

$$=C_{PS}\cdot C_s'\cdot x_1{}^2(x_2+D)+C_{PO}\cdot C_o\cdot x_1\cdot x_2+C_{PM}\cdot C_{power}'\cdot\{2\cdot\left(x_2+x_1\right)\cdot T_d+x_1\cdot x_2\}\cdot V^3$$

$$+\lambda_1\cdot[x_1\cdot x_2\cdot T_s\cdot x_3\cdot\rho_{sw}\cdot C_\alpha-DWT_{given}-C_s\cdot x_1{}^2\cdot(x_2+D)-C_o\cdot x_1\cdot x_2-C_{power}'\cdot\{2\cdot\left(x_2+x_1\right)\cdot T_d+x_1\cdot x_2\}\cdot V^3]$$

$$+\lambda_2\cdot\left(C_H\cdot x_1\cdot x_2\cdot D-V_{H\_req}\right)$$

$$+u\cdot\left\{x_3/\left(x_1/x_2\right)-0.15+s^2\right\}\qquad ...(f)$$

Example of a Constrained Nonlinear Optimization Method by using the Lagrange Multiplier
- Determination of the Optimum Main Dimensions of a Ship (4)

[redacted] $_B \to x_3$

$$H\left(x_1,x_2,x_3,\lambda_1,\lambda_2,u,s\right)=C_{PS}\cdot C_s{}'\cdot x_1{}^2(x_2+D)+C_{PO}\cdot C_o\cdot x_1\cdot x_2+C_{PM}\cdot C_{power}{}'\cdot\{2\cdot\left(x_2+x_1\right)\cdot T_d+x_1\cdot x_2\}\cdot V^3$$
$$+\lambda_1\cdot[x_1\cdot x_2\cdot T_s\cdot x_3\cdot\rho_{sw}\cdot C_\alpha-DWT_{given}-C_s\cdot x_1{}^2\cdot(x_2+D)-C_o\cdot x_1\cdot x_2-C_{power}{}'\cdot\{2\cdot\left(x_2+x_1\right)\cdot T_d+x_1\cdot x_2\}\cdot V^3]$$
$$+\lambda_2\cdot\left(C_H\cdot x_1\cdot x_2\cdot D-V_{H\_req}\right)+u\cdot\left\{x_3/\left(x_1/x_2\right)-0.15+s^2\right\}\qquad...(f)$$

- **To determine the stationary point(x_1,x_2,x_3) of the Lagrangian function H(equation (f)), use the Kuhn-Tucker necessary condition $\nabla H\left(x_1,x_2,x_3,\lambda_1,\lambda_2,u,s\right)=0$.**

$$\frac{\partial H}{\partial x_1}=2C_{PS}\cdot C_s{}'\cdot x_1\cdot(x_2+D)+C_{PO}\cdot C_o\cdot x_2+C_{PM}\cdot C_{power}{}'\cdot\left(2\cdot T_d+x_2\right)\cdot V^3$$
$$+\lambda_1\cdot(x_2\cdot T_s\cdot x_3\cdot\rho_{sw}\cdot C_\alpha-[2\cdot C_s\cdot x_1\cdot(x_2+D)+C_o\cdot x_2+C_{power}{}'\cdot(2\cdot T_d+x_2)\cdot V^3])$$
$$+\lambda_2\cdot\left(C_H\cdot x_2\cdot D\right)+u\cdot\left(-x_3\cdot x_2/x_1{}^2\right)=0\qquad...(1)$$

$$\frac{\partial H}{\partial x_2}=C_{PS}\cdot C_s{}'\cdot x_1{}^2+C_{PO}\cdot C_o\cdot x_1+C_{PM}\cdot C_{power}{}'\cdot(2\cdot T_d+x_1)\cdot V^3$$
$$+\lambda_1\cdot[x_1\cdot T_s\cdot x_3\cdot\rho_{sw}\cdot C_\alpha-C_s{}'\cdot x_1{}^2-C_o\cdot x_1-C_{power}{}'(2\cdot T_d+x_1)\cdot V^3]$$
$$+\lambda_2\cdot\left(C_H\cdot x_1\cdot D\right)+u\cdot\left(x_3/x_1\right)=0\qquad...(2)$$

- Determination of the Optimum Main Dimensions of a Ship (5)

$L \rightarrow x_1, B \rightarrow x_2, C_B \rightarrow x_3$

$$H(x_1, x_2, x_3, \lambda_1, \lambda_2, u, s) = C_{PS} \cdot C_s' \cdot x_1^2 (x_2 + D) + C_{PO} \cdot C_o \cdot x_1 \cdot x_2 + C_{PM} \cdot C_{power}' \cdot \{2 \cdot (x_2 + x_1) \cdot T_d + x_1 \cdot x_2\} \cdot V^3$$

$$+ \lambda_1 \cdot [x_1 \cdot x_2 \cdot T_s \cdot x_3 \cdot \rho_{sw} \cdot C_\alpha - DWT_{given} - C_s \cdot x_1^2 \cdot (x_2 + D) - C_o \cdot x_1 \cdot x_2 - C_{power}' \cdot \{2 \cdot (x_2 + x_1) \cdot T_d + x_1 \cdot x_2\} \cdot V^3]$$

$$+ \lambda_2 \cdot (C_H \cdot x_1 \cdot x_2 \cdot D - V_{H\_req}) + u \cdot \{x_3 / (x_1 / x_2) - 0.15 + s^2\} \qquad ...(f)$$

- **Kuhn-Tucker necessary condition** $\nabla H(x_1, x_2, x_3, \lambda_1, \lambda_2, u, s) = 0$.

$$\frac{\partial H}{\partial x_3} = \lambda_1 . x_1 \cdot x_2 \cdot T_s \cdot \rho_{sw} \cdot C_\alpha + u \cdot (x_2 / x_1) = 0 \qquad ...(3)$$

$$\frac{\partial H}{\partial \lambda_1} = x_1 \cdot x_2 \cdot T_s \cdot x_3 \cdot \rho_{sw} \cdot C_\alpha - DWT_{given} - C_s \cdot x_1^2 \cdot (x_2 + D) - C_o \cdot x_1 \cdot x_2$$

$$- C_{power}' \cdot \{2 \cdot (x_2 + x_1) \cdot T_d + x_1 \cdot x_2\} \cdot V^3 \quad \cdots (4)$$

$$\frac{\partial H}{\partial \lambda_2} = C_H \cdot x_1 \cdot x_2 \cdot D - V_{H\_req} = 0 \qquad ...(5)$$

$$\frac{\partial H}{\partial u} = x_3 \cdot x_2 / x_1 - 0.15 + s^2 = 0 \qquad ...(6)$$

$$\frac{\partial H}{\partial s} = 2 \cdot u \cdot s = 0, \quad (u \geq 0) \qquad ...(7)$$

- $\nabla H(x_1, x_2, x_3, \lambda_1, \lambda_2, u, s)$ **: Nonlinear simultaneous equation having the 7 variables((1)~(7)) and 7 equations**

  → It can be solved by using the numerical method!

- Determination of the Optimum Propeller Main Dimensions (1)

Given $P, n, A_E / A_O, V$

Find $J, P_i / D_P$

Maximize $\eta_O = \dfrac{J}{2\pi} \cdot \dfrac{K_T}{K_Q}$ $\longrightarrow$ **Because K_T and K_Q are a function of J and P_i/D_p, the objective is also a function of J and P_i/D_p.**

Subject to $\dfrac{P}{2\pi n} = \rho \cdot n^2 \cdot D_P^{\ 5} \cdot K_Q$

: The propeller absorbs the torque delivered by Diesel Engine

Where, $J = \dfrac{V(1-w)}{n \cdot D_P}$

$$K_T = f(J, P_i / D_P)$$

$$K_Q = f(J, P_i / D_P)$$

P: Delivered Power to Propeller from the Main Engine, KW

n: Number of Revolutions, 1/sec

D_P: Propeller Diameter, m

P_i: Propeller Pitch, m

A_E/A_O: Expanded Area Ratio

V: Ship speed, m/s

η_O: Propeller efficiency(in open water)

➡ **Optimization problem having the two variables and one equality constraint**

- Determination of the Optimum Propeller Main Dimensions (2)

$$\frac{P}{2\pi n} = \rho \cdot n^2 \cdot D_P{}^5 \cdot K_Q \quad \cdots\cdots \quad \text{(a)}$$

: The propeller absorbs the torque delivered by Diesel Engine

If the propeller absorbs the torque delivered by Diesel Engine, the constraint is represented from the equation (a).

$$C = \frac{K_Q}{J^5} = \frac{P \cdot n^2}{2\pi\rho \cdot V_A{}^5}$$

$$G(J, P_i / D_P) = K_Q - C \cdot J^5 = 0 \quad \cdots\cdots \quad \text{(b)}$$

Propeller efficiency in open water η_0 is as follows.

$$F(J, P_i / D_P) = \eta_0 = \frac{J}{2\pi} \cdot \frac{K_T}{K_Q} \quad \cdots\cdots \quad \text{(c)}$$

The objective F is a function of J and P_i/D_p and we have to determine the optimal main dimensions(J and P_i/D_p) to maximize the propeller efficiency in open water satisfying the constraint (b) in this optimization problem.

- Determination of the Optimum Propeller Main Dimensions (3)

$$G(J, P_i/D_P) = K_Q - C \cdot J^5 = 0 \quad \cdots\cdots \quad \text{(b)}$$

$$F(J, P_i/D_P) = \eta_0 = \frac{J}{2\pi} \cdot \frac{K_T}{K_Q} \quad \cdots\cdots \quad \text{(c)}$$

Introduce the Lagrange multiplier λ to the equation (b) and (c).

$$H(J, P_i/D_P, \lambda) = F(J, P_i/D_P) + \lambda G(J, P_i/D_P) \quad \cdots\cdots \quad \text{(d)}$$

Determine the value of the P_i/D_P and λ to maximize the value of the function H.

$$\frac{\partial H}{\partial J} = \frac{1}{2\pi}\left(\frac{K_T}{K_Q}\right) + \frac{J}{2\pi} \frac{\left\{\left(\frac{\partial K_T}{\partial J}\right) \cdot K_Q - \left(\frac{\partial K_Q}{\partial J}\right) \cdot K_T\right\}}{{K_Q}^2} + \lambda\left\{\left(\frac{\partial K_Q}{\partial J}\right) - 5 \cdot C \cdot J^4\right\}$$

$$= 0 \quad \cdots\cdots \quad \text{(e)}$$

$$\frac{\partial H}{\partial (P_i/D_P)} = \frac{J}{2\pi} \frac{\left\{\left(\frac{\partial K_T}{\partial P_i/D_P}\right) \cdot K_Q - \left(\frac{\partial K_Q}{\partial P_i/D_P}\right) \cdot K_T\right\}}{{K_Q}^2} + \lambda\left(\frac{\partial K_Q}{\partial P_i/D_P}\right)$$

$$= 0 \quad \cdots\cdots \quad \text{(f)}$$

$$\frac{\partial H}{\partial \lambda} = K_Q - C \cdot J^5 = 0 \quad \cdots\cdots \quad \text{(g)}$$

- Determination of the Optimum Propeller Main Dimensions (4)

Eliminate λ in the equation (e), (f) and (g) rearrange as follows.

$$\left(\frac{\partial K_Q}{\partial (P_i / D_P)}\right)\{J \cdot \left(\frac{\partial K_T}{\partial J}\right) - 4K_T\}$$

$$+\left(\frac{\partial K_T}{\partial (P_i / D_P)}\right)\{5K_Q - J \cdot \left(\frac{\partial K_Q}{\partial J}\right)\} = 0 \quad \cdots\cdots \quad \text{(h)}$$

$$K_Q - C \cdot J^5 = 0 \quad \cdots\cdots \quad \text{(i)} \qquad\qquad P_i / D_P$$

By obtaining the solution of the equation (h) and (i), we can determine the value of the J and P_i/D_p to maximize the propeller efficiency absorbing the torque delivered by Diesel Engine.

Because $J = \dfrac{V(1-w)}{n \cdot D_P}$ **, if we obtain the value of J, we can find the value of D_p. And the value of P_i is obtained from the value of P_i/D_p.**

Therefore, we can obtain the value of the propeller diameter (D_P) and pitch(P_i).

- Determination of the Optimum Propeller Main Dimensions
- [Reference] Derivation of h from e, f, g (1)

$$\frac{1}{2\pi}\left(\frac{K_T}{K_Q}\right)+\frac{J}{2\pi}\frac{\left\{\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_T\right\}}{K_Q{}^2}+\lambda\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}=0 \quad \cdots\cdots \quad \text{(e)}$$

$$\frac{J}{2\pi}\frac{\left\{\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\right\}}{K_Q{}^2}+\lambda\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)=0 \quad \cdots\cdots \quad \text{(f)}$$

To eliminate λ, we calculate as follows.

$$\text{(e)}\times\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)-(f)\times\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}=0$$

$$\text{(e)}\times\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right):\ \frac{1}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)+\frac{J}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\frac{\left\{\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_T\right\}}{K_Q{}^2}+\cancel{\lambda\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}}=0$$

$$(f)\times\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}:\ \frac{J}{2\pi}\frac{\left\{\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\right\}}{K_Q{}^2}\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}+\cancel{\lambda\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}}=0$$

$$\text{(e)}\times\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)-(f)\times\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}$$

$$=\frac{1}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)+\frac{J}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\frac{\left\{\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_T\right\}}{K_Q{}^2}-\frac{J}{2\pi}\frac{\left\{\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\right\}}{K_Q{}^2}\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}=0$$

- Determination of the Optimum Propeller Main Dimensions

- [Reference] Derivation of h from e, f, g (2)

$$(\mathrm{e})\times\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)-(f)\times\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}$$

$$=\frac{1}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)+\frac{J}{2\pi}\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\frac{\left\{\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_T\right\}}{{K_Q}^2}-\frac{J}{2\pi}\frac{\left\{\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\right\}}{{K_Q}^2}\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}=0$$

Multiply 2π and the both side of the equation and rearrange the equation as follows.

$$\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)+\frac{J}{{K_Q}^2}\underline{\left[\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left\{\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_T\right\}-\left\{\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q-\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\right\}\left\{\left(\frac{\partial K_Q}{\partial J}\right)-5\cdot C\cdot J^4\right\}\right]}=0$$

The term underlined is rearranged as follows.

$$=\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\cancel{\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)}\cdot K_T-\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_Q+\cancel{\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)}\cdot K_T+5\cdot\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q\cdot C\cdot J^4-5\cdot\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\cdot C\cdot J^4$$

$$=\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_Q+5\cdot\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q\cdot C\cdot J^4-5\cdot\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\cdot C\cdot J^4$$

Substituting the rearranged term into the above equation.

$$\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)+\frac{J}{{K_Q}^2}\left[\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q-\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_Q+5\cdot\left(\frac{\partial K_T}{\partial(P_i/D_P)}\right)\cdot K_Q\cdot C\cdot J^4-5\cdot\left(\frac{\partial K_Q}{\partial(P_i/D_P)}\right)\cdot K_T\cdot C\cdot J^4\right]=0$$

- Determination of the Optimum Propeller Main Dimensions

- [Reference] Derivation of h from e, f, g (3)

$$\frac{\partial H}{\partial \lambda} = K_Q - C \cdot J^5 = 0 \quad \cdots\cdots \quad (g)$$

$$\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right) + \frac{J}{K_Q^{\,2}}\left[\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot K_Q - \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot K_Q + 5\cdot\left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\cdot K_Q \cdot C \cdot J^4 - 5\cdot\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\cdot K_T \cdot C \cdot J^4\right] = 0$$

Apply the distributive property.

$$\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right) + \left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot\frac{J}{K_Q} - \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot\frac{J}{K_Q} + 5\cdot\left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\cdot\cancel{\frac{C\cdot J^5}{K_Q}} - 5\cdot\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\cdot\frac{K_T}{K_Q}\cdot\cancel{\frac{C\cdot J^5}{K_Q}} = 0$$

By using the (g) $\dfrac{C \cdot J^5}{K_Q} = 1$

$$\underline{\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right)} + \left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot\frac{J}{K_Q} - \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot\frac{J}{K_Q} + 5\cdot\left(\frac{\partial K_T}{\partial (P_i/D_P)}\right) - 5\cdot\underline{\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\cdot\frac{K_T}{K_Q}} = 0$$

The underlined term is calculated as follows.

$$-4\cdot\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{K_T}{K_Q}\right) + \left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)\cdot\frac{J}{K_Q} - \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot\frac{J}{K_Q} + 5\cdot\left(\frac{\partial K_T}{\partial (P_i/D_P)}\right) = 0$$

Multiply K_Q and the both side of the equation.

$$-4\cdot\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)K_T + \left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_T}{\partial J}\right)J - \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left(\frac{\partial K_Q}{\partial J}\right)\cdot J + 5\cdot K_Q\left(\frac{\partial K_T}{\partial (P_i/D_P)}\right) = 0$$

Apply the distributive property. $\left(\dfrac{\partial K_Q}{\partial (P_i/D_P)}\right), \left(\dfrac{\partial K_T}{\partial (P_i/D_P)}\right)$

$$\left(\frac{\partial K_Q}{\partial (P_i/D_P)}\right)\left\{J\cdot\left(\frac{\partial K_T}{\partial J}\right) - 4K_T\right\} + \left(\frac{\partial K_T}{\partial (P_i/D_P)}\right)\left\{5K_Q - J\cdot\left(\frac{\partial K_Q}{\partial J}\right)\right\} = 0 \quad \cdots\cdots \quad (h)$$

Naval Architecture & Ocean Engineering

Computer Aided Ship Design

Part I. Optimization Method

Ch.5 Penalty Function Method

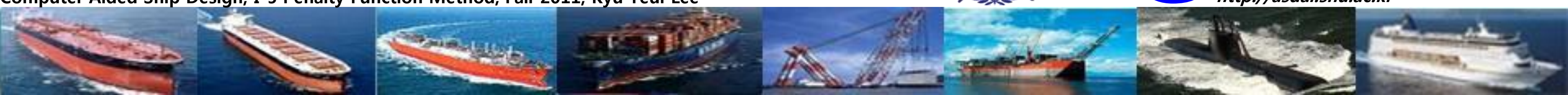
September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL Advanced Ship Design Automation Lab. http://asdal.snu.ac.kr



Ch.5 Penalty Function Method

Ch.5 Penalty Function Method

5.1 Interior Penalty Function Method

5.1 Interior Penalty function Method

- The Method of Transformation of Constrained Optimal Design Problem to Unconstrained Optimal Design Problem
- Lagrange Multiplier

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \le \mathbf{0}$ Inequality constraint

Transforming this problem to unconstrained optimal design problem by using the Lagrangian function

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)$$

By using the necessary condition for the candidate local optimal solution(∇L=0), are calculated.

1) If the constraints are satisfied at the current design point,

In case of the equality constraints: $\mathbf{h}(\mathbf{x}) = \mathbf{0}$

In case of the inequality constraints: $\mathbf{u} = \mathbf{0}$ (The constraints are inactive, i.e, the design point is in feasible re

$\mathbf{s} = \mathbf{0} \Rightarrow \mathbf{g}(\mathbf{x}) = \mathbf{0}$ (The constraints are active, i.e, the design point is on the constraints)

Therefore, $L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2) \Rightarrow f(\mathbf{x})$

➡ If all the constrains are satisfied, the Lagrange function is same with the original objective function.

2) If the constraints are violated at the current design point,

In case of the equality constraints: $\mathbf{v}^T \mathbf{h}(\mathbf{x}) \ne \mathbf{0}$

In case of the inequality constraints: $\mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2) > \mathbf{0}$

Therefore, $L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \underline{\mathbf{v}^T \mathbf{h}(\mathbf{x}) + \mathbf{u}^T (\mathbf{g}(\mathbf{x}) + \mathbf{s}^2)}$

➡ This term means augmenting a penalty to the original objective function when the constraints are violated.

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \le \mathbf{0}$ Inequality constraint

- Fiacco and McCormick suggested a method which transforms the constrained optimization problem into the unconstrained optimization problem by using the modified objective function in 1968.
The modified objective function is a function augmenting a penalty to the original objective function.

- SUMT: Sequential Unconstrained Minimization Technique

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) - r_k \sum_{j=1}^{m} \frac{1}{g_j(\mathbf{x})}$$

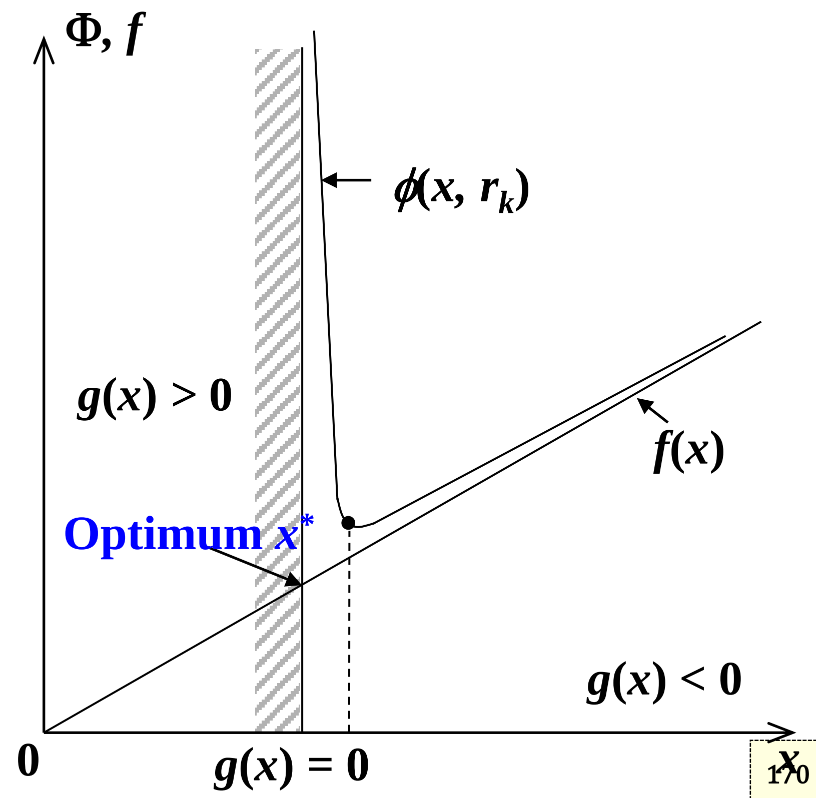
where r_k is given and positive value and getting smaller each iteration.

If the design point approaches to the boundary of the inequality constraints in the feasible region,

$g_j(\mathbf{x}) \le 0$, the absolute value of this is decreased.

$-r_k \dfrac{1}{g_j(\mathbf{x})} > 0$, the absolute value of this is increased.

Since the modified objective function is increased as the design point approaches to the boundary of the inequality constraint, this method prevents the design point violating the constraints.



5.1 Interior Penalty function Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem
- SUMT: Sequential Unconstrained Minimization Technique(Interior Penalty Function Method) (2)

- If the design point approaches to the boundary of the constraints in the feasible region, the objective function is augmented by a penalty.
- The starting design point has to be in the feasible region.

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) - r_k \sum_{j=1}^{m} \frac{1}{g_j(\mathbf{x})}$$

(r_k is decreased, when k is increased.)

[Example] Function of a single variable

$f(x) = \alpha x, \quad g(x) = \beta - x \le 0 \quad, (x \ge \beta \Rightarrow \beta - x \le 0)$

Transform the unconstrained optimization problem into the constrained optimization problem.

$$\Phi(x, r_k) = f(x) - r_k \frac{1}{g(x)} = \alpha x - r_k \frac{1}{\beta - x}$$

- k is the number of iteration.
- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke&Jeeves, Nelder&Mead.

$k = 1$, Starting design point: x^*_0

$$\Phi(x, r_1) = \alpha x - r_1 \frac{1}{\beta - x}$$ ⇒ Optimal design point: x^*_1

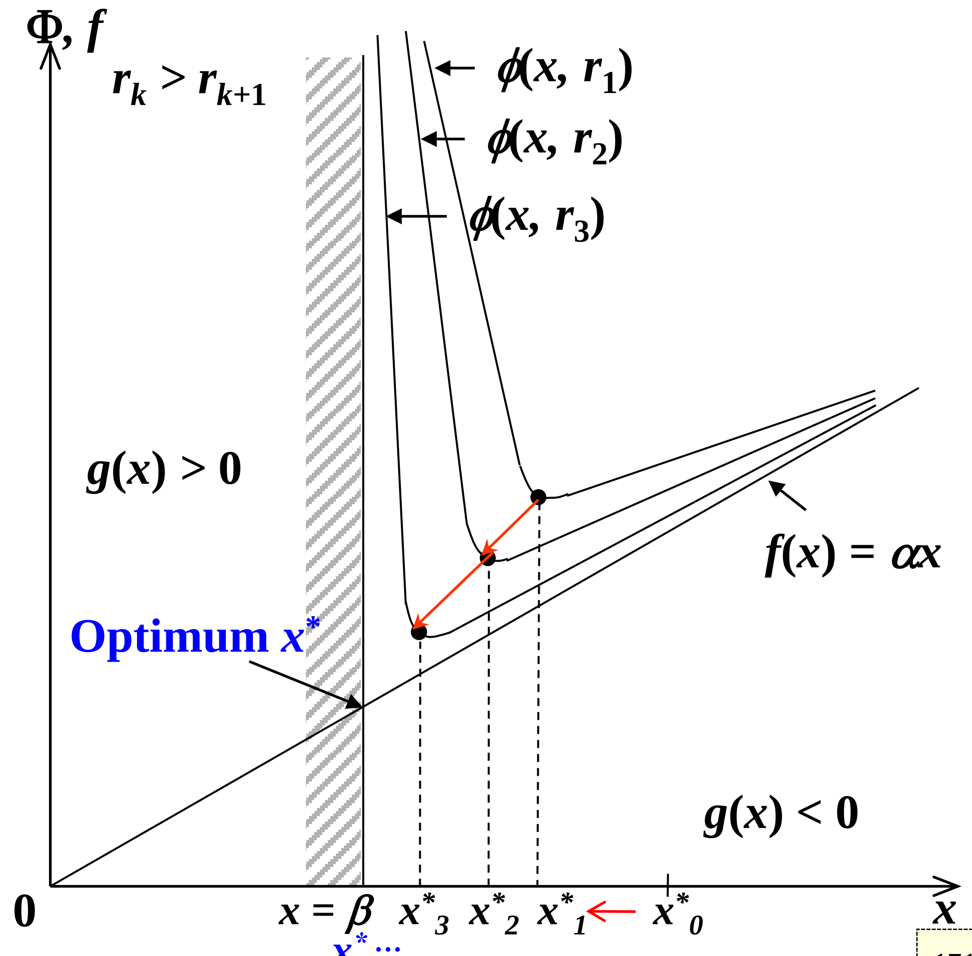
$k = 2$, Starting design point : x^*_1

$$\Phi(x, r_2) = \alpha x - r_2 \frac{1}{\beta - x}$$ ⇒ Optimal design point : x^*_2

$k = 3$, Starting design point : x^*_2

$$\Phi(x, r_3) = \alpha x - r_3 \frac{1}{\beta - x}$$ ⇒ Optimal design point : x^*_3

By iterating the above process, we find the optimal design point(x^*).



Naval Architecture & Ocean Engineering

Ch.5 Penalty Function Method

5.2 Exterior Penalty Function Method

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

5.2 Exterior Penalty Function Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem
- Exterior Penalty Function Method (1)

■ **There will be a penalty for only violating the constraints.**

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} \left[\max\{g_j(\mathbf{x}), 0\} \right]^2$$ **(r_k is increased, when k is increased.)**

[Example] Function of a single variable

$$f(x) = \alpha x, \quad g(x) = \beta - x \le 0, (x \ge \beta \Rightarrow \beta - x \le 0)$$

Transform the unconstrained optimization problem into the constrained optimization problem.

$$\Phi(x, r_k) = f(x) + r_k \max\{g(x), 0\}^2 = \alpha x + r_k \max\{g(x), 0\}^2$$

- k is the number of iteration.

- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke&Jeeves, Nelder&Mead.

$k = 1$, **Starting design point : x^*_0**

$$\Phi(x, r_1) = \alpha x + r_1 \left[\max\{g(x), 0\} \right]^2$$ ⇒ **Optimal design point : x^*_1**

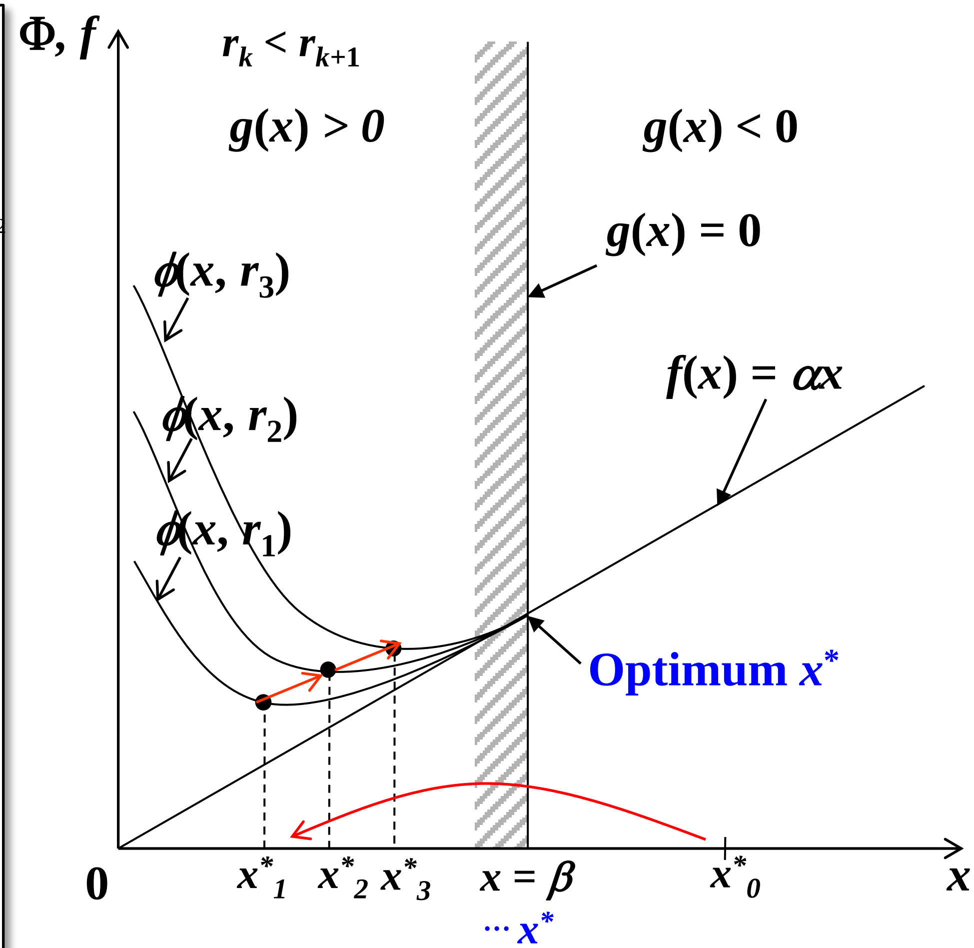
$k = 2$, **Starting design point : x^*_1**

$$\Phi(x, r_2) = \alpha x + r_2 \left[\max\{g(x), 0\} \right]^2$$ ⇒ **Optimal design point : x^*_2**

$k = 3$, **Starting design point : x^*_2**

$$\Phi(x, r_3) = \alpha x + r_3 \left[\max\{g(x), 0\} \right]^2$$ ⇒ **Optimal design point : x^*_3**

By iterating the above process, we find the optimal design point(x^*).



5.2 Exterior Penalty Function Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem

- Exterior Penalty Function Method (2)

■ **There will be a penalty for only violating the constraints.**

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} \max\{g_j(\mathbf{x}), 0\}$$ **(r_k is increased, when k is increased.)**

[Example] Function of a single variable

$$f(x) = \alpha x, \quad g(x) = \beta - x \le 0, (x \ge \beta \Rightarrow \beta - x \le 0)$$

Transform the unconstrained optimization problem into the constrained optimization problem.

$$\Phi(x, r_k) = f(x) + r_k \max\{g(x), 0\} = \alpha x + r_k \max\{g(x), 0\}$$

- k is the number of iteration.

- In each iteration, the optimal design point can be obtained by using the Gradient method, Hooke&Jeeves, Nelder&Mead.

$k = 1$, **Starting design point : x^*_0**

$$\Phi(x, r_1) = \alpha x + r_1 \max\{g(x), 0\}$$

⇨ **Optimal design point : we can not find it.**

$k = 2$, **Starting design point : x^*_1**

$$\Phi(x, r_2) = \alpha x + r_2 \max\{g(x), 0\}$$

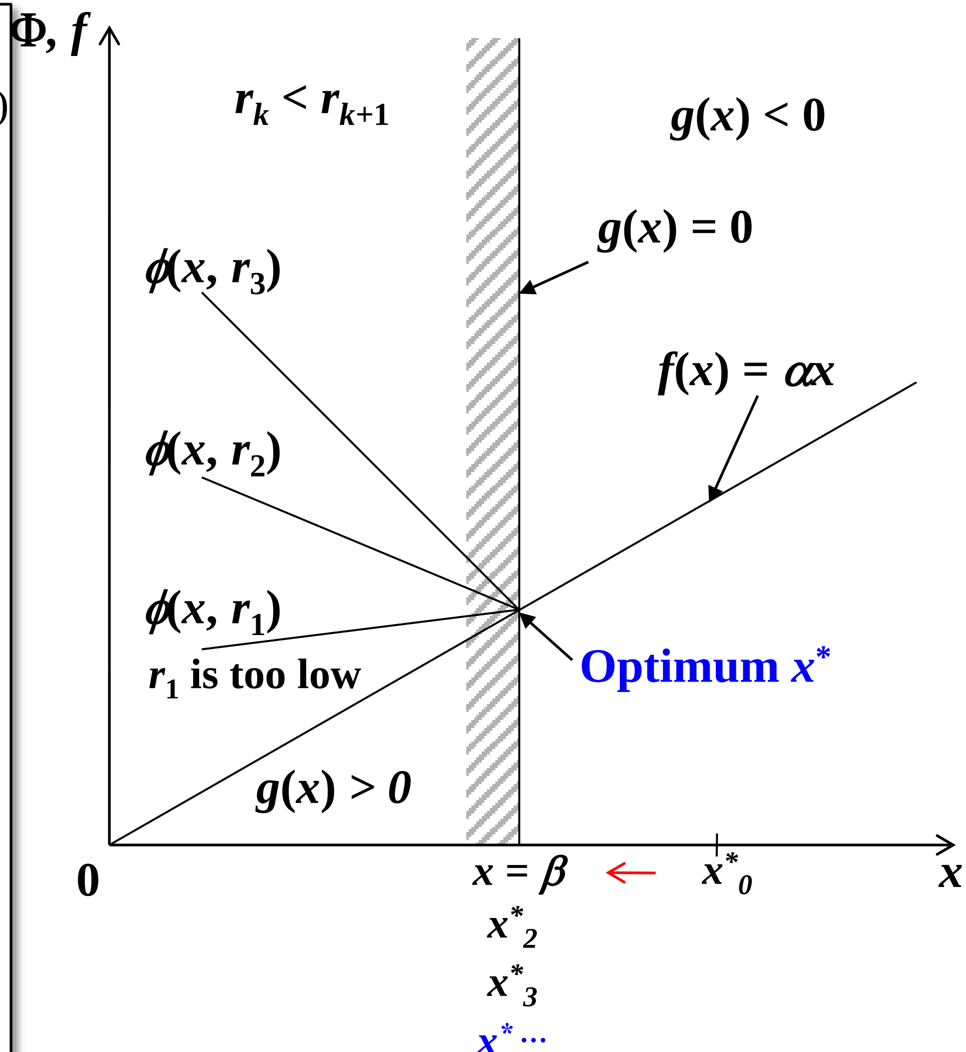
⇨ **Optimal design point : x^*_2**

$k = 3$, **Starting design point : x^*_2**

$$\Phi(x, r_3) = \alpha x + r_3 \max\{g(x), 0\}$$

⇨ **Optimal design point : x^*_3**

If r_k is determined properly, the optimal design point(x^*) is not changed.



5.2 Exterior Penalty Function Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem
- Relationship between External Penalty Function and Feasible Region (1)

■ Since there will be a penalty for only violating the constraints, **if the minimum design point is in the feasible region**, the result of the optimization method by using the exterior penalty function is the same with that only using the objective function.

[Example] Function of a single variable

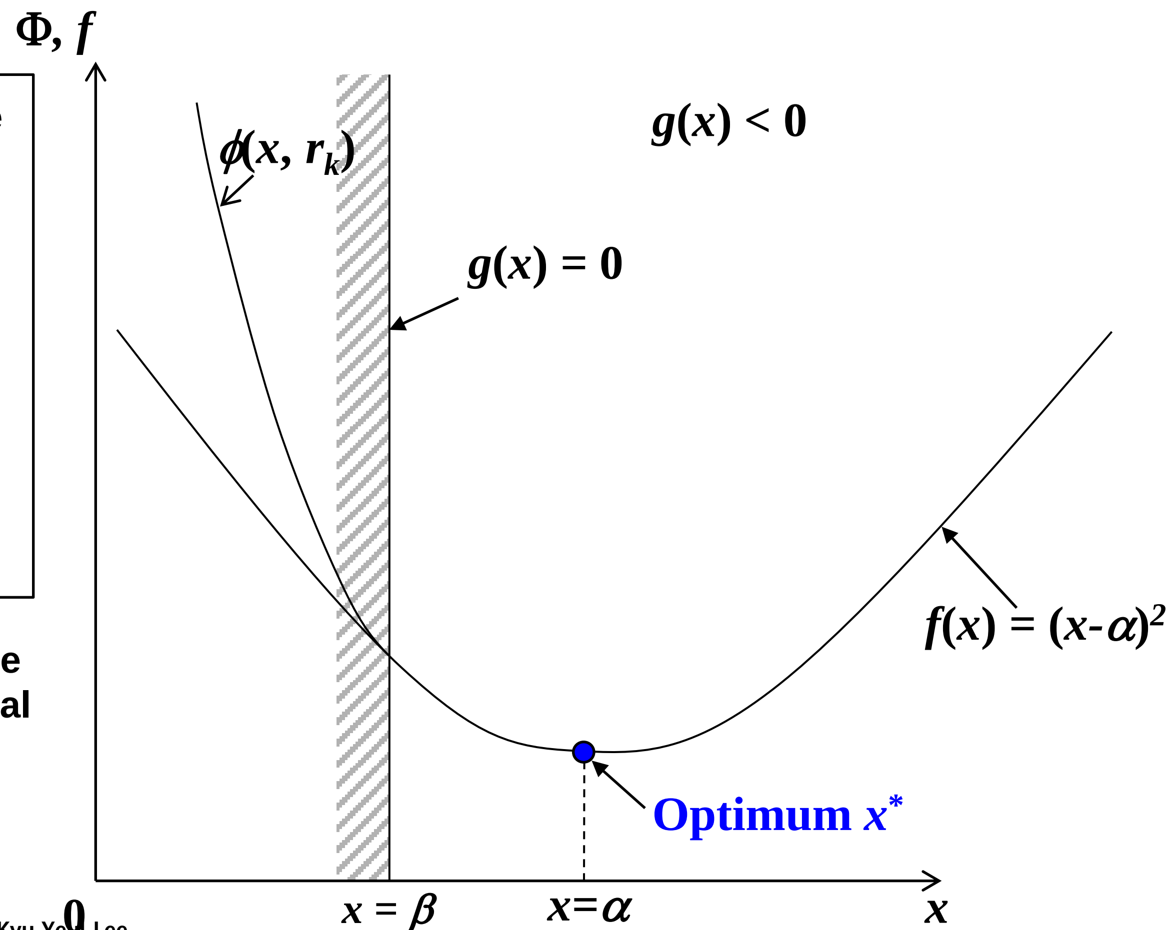
$$f(x) = (x-\alpha)^2, \; g(x) = \beta - x \le 0$$

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \underbrace{\sum_{j=1}^{m}\left[\max\left\{g_j(\mathbf{x}), 0\right\}\right]^2}_{\text{Penalty term}}$$

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x})$$

$$where,\; g(\mathbf{x}) \le 0,\; \max\left\{g_j(\mathbf{x}), 0\right\} = 0$$

If the minimum design point(x*) is in the feasible region, the penalty term is equal to zero. So, the objective function augmented by the penalty is the same with the original objective function.



- Since there will be a penalty for only violating the constraints, **if the minimum design point is not in the feasible region**, the result of the optimization method by using the exterior penalty function is different with that only using the objective function.

[Example] Function of a single variable

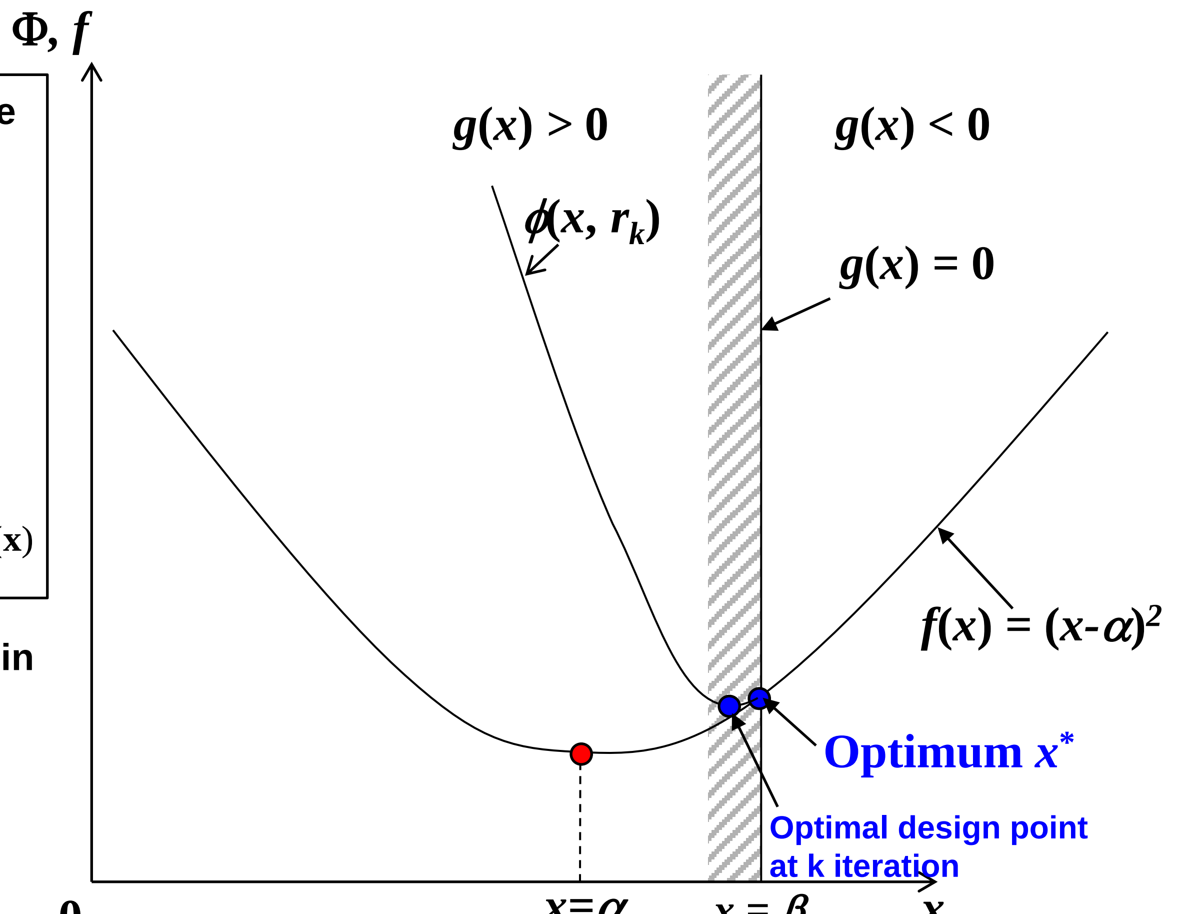
$$f(x) = (x-\alpha)^2,\ g(x) = \beta - x \le 0$$

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} \underbrace{\left[\max\left\{g_j(\mathbf{x}), 0\right\}\right]^2}_{\text{Penalty term}}$$

$$\Phi(\mathbf{x}, r_k) = f(\mathbf{x}) + r_k \sum_{j=1}^{m} g_j(\mathbf{x})^2$$

$$\text{where},\ g(\mathbf{x}) > 0,\ \max\left\{g_j(\mathbf{x}), 0\right\} = g_j(\mathbf{x})$$

If the minimum design point(x^*) is not in the feasible region, the penalty term is larger than zero. So, the objective function augmented by the penalty is different with the original objective function.



Ch.5 Penalty Function Method

5.3 Augmented Lagrange Multiplier Method

Seoul National Univ.
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

▪Augmented Lagrange multiplier method

- This method combines the Lagrange multiplier and the penalty function methods.
- There is no need for the penalty parameter r to go to infinity.
- Starting point does not have to be in feasible region.
- It has been proven that they possess a faster rate of convergence than interior and exterior penalty function method.

- Augmented Lagrange Multiplier Method in Equality Constrained Problem (1)

$$\begin{array}{ll} \textit{Minimize} & f(\mathbf{x}) \\ \textit{Subject to} & h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, \ldots, m \end{array}$$

Lagrangian function of this problem is as follows.

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x})$$

Augmented Lagrangian function of this problem is follows.

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k : **arbitrary constant**

5.3 Augmented Lagrange Multiplier Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem

- Augmented Lagrange Multiplier Method in Equality Constrained Problem (2)

$$\textbf{Minimize} \quad f(\mathbf{x})$$
$$\textbf{Subject to} \quad h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, \ldots, m$$

Lagrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x})$$

Augmented Lagrangian function

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

Necessary conditions for the minimum of Lagrangian function

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \boxed{\lambda_j^*} \frac{\partial h_j}{\partial x_i} = 0$$

Necessary conditions for the minimum of Augmented Lagrangian function

$$\frac{\partial \Phi}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \boxed{(\lambda_j + 2r_i h_j)} \frac{\partial h_j}{\partial x_i} = 0$$

Find iterative relation

$$\lambda_j^* = \lambda_j + 2r_k h_j \quad j = 1, 2, \ldots, m$$

$$\Downarrow$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, \ldots, m$$

- Augmented Lagrange Multiplier Method in Equality Constrained Problem (3)

$$\begin{array}{|ll|}\hline \textbf{\textit{Minimize}} & f(\mathbf{x}) \\ \textbf{\textit{Subject to}} & h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, \ldots, m \\ \hline \end{array}$$

Augmented Lagrangian function

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

Iterative relation

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2 r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, \ldots, m$$

---

1. In the first iteration(k=1), the values of $\lambda_j^{(1)}$ are chosen as zero, the value of r_k is set equal to an arbitrary constant.

2. Find the $\mathbf{x}^{(k)*}$ that minimize Φ by using any unconstrained optimization method and set $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)*}$.

Minimize	$f(\mathbf{x})$
Subject to	$h_j(\mathbf{x}) = \mathbf{0}, \quad j = 1, 2, ..., m$

Augmented Lagrangian function

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + \boxed{r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

Iterative relation

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2 r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, ..., m$$

---

3. The values of $\lambda_j^{(k)}$ and r_k are then updated by using the iterative relation to start the next iteration.

$$r_{k+1} = c r_k, \; c > 1$$

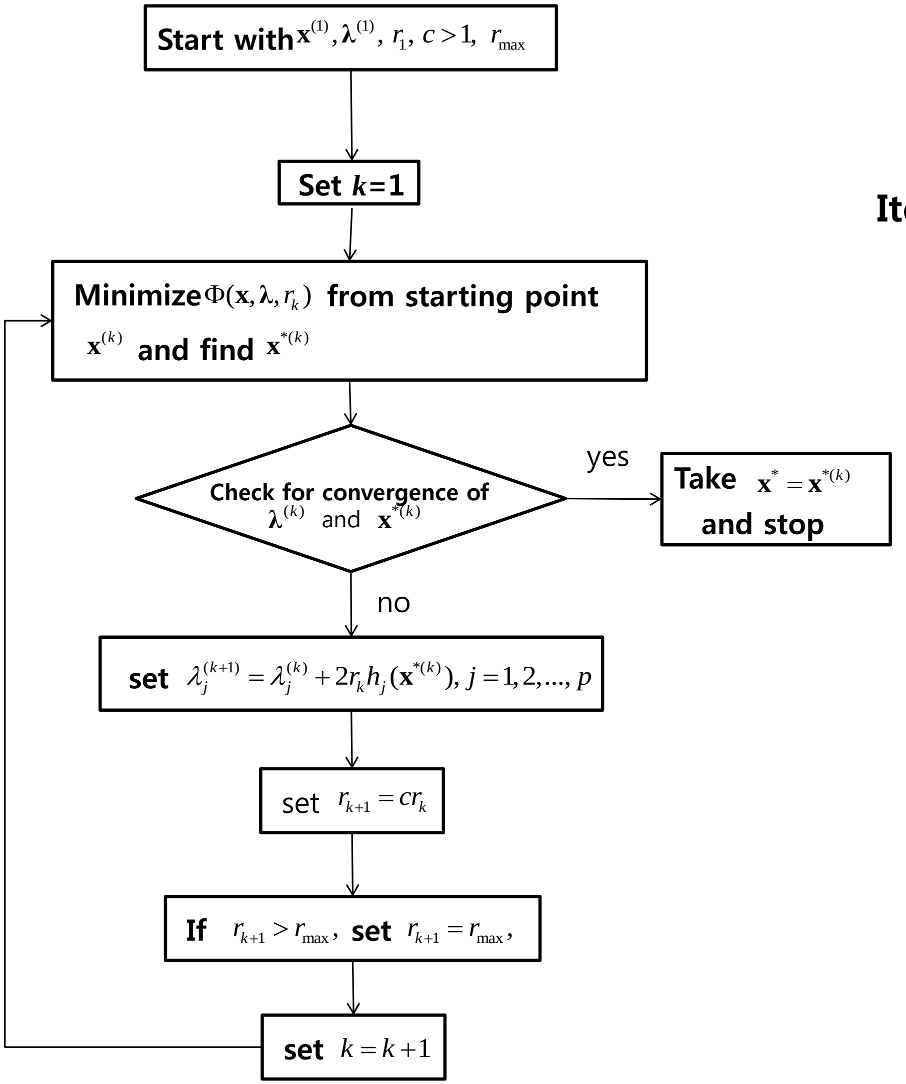
$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2 r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, ..., m$$

4. If $\left| \lambda_j^{(k+1)} - \lambda_j^{(k)} \right| < \varepsilon$, stop the iteration and take $\mathbf{x}^* = \mathbf{x}^{(k)*}$.

5.3 Augmented Lagrange Multiplier Method

- The Method of Transformation from Constrained Optimal Design Problem to Unconstrained Optimal Design Problem

- Algorithm of Augmented Lagrange Multiplier Method



Augmented Lagrangian function

$$\Phi(\mathbf{x}, \boldsymbol{\lambda}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x}) + r_k \sum_{j=1}^{m} h_j^2(\mathbf{x})$$

Iterative relation

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + 2r_k h_j(\mathbf{x}^{(k)}) \quad j = 1, 2, \dots, m$$

$$r_{k+1} = cr_k,\ c > 1$$

- Augmented Lagrange Multiplier Method in Inequality Constrained Problem

Minimize	$f(\mathbf{x})$
Subject to	$g_j(\mathbf{x}) \le \mathbf{0}, \quad j = 1, 2, \ldots, m$

Augmented Lagrangian function in the inequality constrained problem

$$\Phi(\mathbf{x}, \mathbf{u}, \mathbf{s}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} u_j [g_j(\mathbf{x}) + s_j^2] + r_k \boxed{\sum_{j=1}^{m} [g_j(\mathbf{x}) + s_j^2]^2}$$

Augmented term to Lagrangian function

r_k **: arbitrary constant**

s_j **: slack variable**

This function is equivalent to*

$$\Phi(\mathbf{x}, \mathbf{u}, r_k) = f(\mathbf{x}) + \sum_{j=1}^{m} u_j \alpha_j + r_k \sum_{j=1}^{m} \alpha_j^{\ 2}, \quad \alpha_j = \max\left\{ g_j(\mathbf{x}), -\frac{u_j}{2r_k} \right\}$$

Iterative relation

$$u_j^{(k+1)} = u_j^{(k)} + 2 r_k \alpha_j^{(k)}$$

*Rockafellar,R.T.,'The multiplier method of Hestenes and Powell applied to convex programming',*Journal of Optimization Theory and Applications*,1973

Ch.5 Penalty Function Method

5.4 Descent Function Method

Seoul National Univ.
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Constrained Optimal Design Problem

Minimize $f(\mathbf{x})$

Subject to $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ Equality constraint

$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ Inequality constraint

***Descent Function**

- Modified objective function by augmenting a penalty to the original objective function
- It has the same meaning with Penalty Function.

Pshenichny and Danilin suggested a method which transforms the constrained optimization problem into the unconstrained optimization by using the descent function* in 1978.

$$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$$: Maximum penalty by the constraints

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x})$$

$$R = \max\left\{\underline{R_0},\ r\left(= \sum_{i=1}^{p} |v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$: Penalty parameters which is the summation of the all Lagrange multipliers

The value defined by user

1) If the constraints are satisfied at the current design point,

$$V(\mathbf{x}) = 0 \Rightarrow R \cdot V(\mathbf{x}) = 0$$

$$\Rightarrow \Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x}) \Rightarrow f(\mathbf{x})$$

➡ If the constraints are satisfied at the current design point, the descent function is the same with the original objective function.

2) If the constraints are violated at the current design point,

$$R \cdot V(\mathbf{x}) > 0$$

$$\Rightarrow \Phi(\mathbf{x}) = f(\mathbf{x}) + \underline{R \cdot V(\mathbf{x})} > f(\mathbf{x})$$

➡ If the constraints are violated at the current design point, the value of the positive penalty is augmented to the original objective function.

- [Reference] The Meaning of the Constant 'R' in the Decent Function

Original Problem

Minimize $f(\mathbf{x}) = 100(x_1 - 1.5)^2 + 100(x_2 - 1.5)^2$

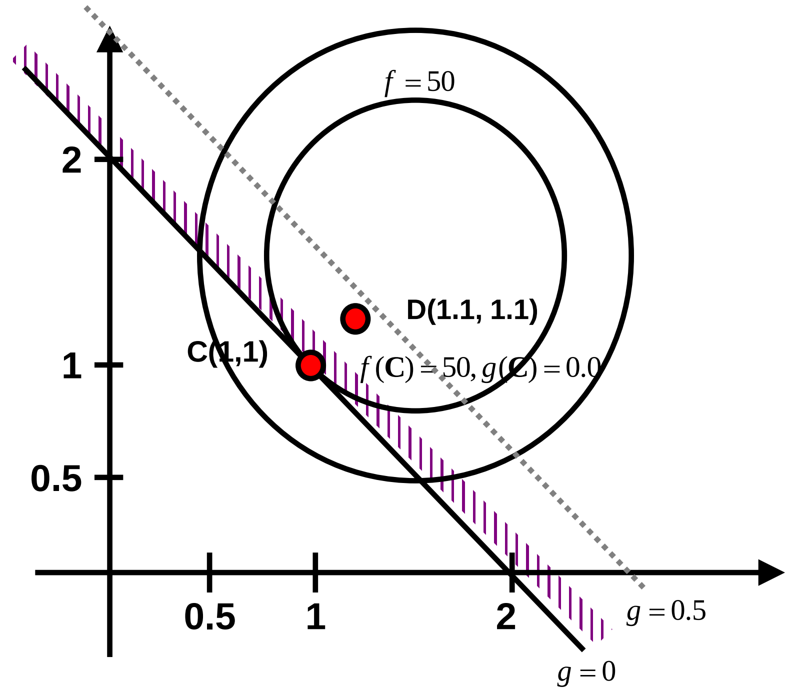
Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x})$$

$$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$$

$$R = \max\left\{R_0,\ r\left(= \sum_{i=1}^{p} |v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$

If 'R' is assumed as a constant '10',



Since the constraint is satisfied at the point C(1,1), the value of the decent function is as follows:

$\Phi(\mathbf{C}) = f(\mathbf{C}) + R \cdot V(\mathbf{C}) = 50 + R \cdot \max\{0, g(\mathbf{C})\}$

$= 50 + 10 \cdot \max\{0, 0\} = 50$

Since the constraint is violated at the point D(1.1, 1.1), the value of the decent function is as follows:

$\Phi(\mathbf{D}) = f(\mathbf{D}) + R \cdot V(\mathbf{D}) = 32 + R \cdot \max\{0, g(\mathbf{D})\}$

$= 32 + 10 \cdot \max\{0, 0.2\} = 32 + 2 = 34$

Although the constraint is violated, the value of the decent function is decreased.
Because the change in the original objective function f is larger than the change in the constraint g.
Therefore, if the decrease in the original objective function f is larger than the increase in the constraint g, the value of the penalty parameter 'R' has to be increased.

- [Reference] The Meaning of the Constant 'R' in the Decent Function

Original Problem

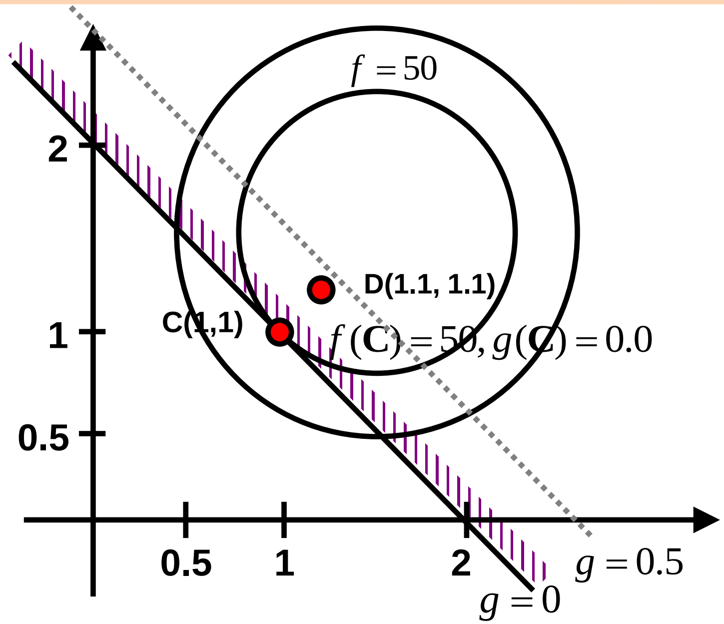
Minimize $f(\mathbf{x}) = 100(x_1 - 1.5)^2 + 100(x_2 - 1.5)^2$

Subject to $g(\mathbf{x}) = x_1 + x_2 - 2 \le 0$

$$\Phi(\mathbf{x}) = f(\mathbf{x}) + R \cdot V(\mathbf{x})$$

$$V(\mathbf{x}) = \max\{0; |\mathbf{h}|; \mathbf{g}\}$$

$$R = \max\left\{R_0,\ r\left(= \sum_{i=1}^{p} |v_i| + \sum_{i=1}^{m} u_i\right)\right\}$$



At point C, the value of $-\nabla f(\mathbf{x}^*) = u^* \nabla g(\mathbf{x}^*)$ is as follows.

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 200(x_1 - 1.5) \\ 200(x_2 - 1.5) \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} -100 \\ -100 \end{bmatrix}$$

$$\nabla g(\mathbf{x}) = \begin{bmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathbf{x}^*(1,1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$u^* = 100$$

If we use the value of the Lagrange Multiplier, 100, as the value of 'R', the value of the decent function at the point D increases by 52.

$$\Phi(\mathbf{D}) = f(\mathbf{D}) + R \cdot V(\mathbf{D}) = 32 + R \cdot \max\{0, g(\mathbf{D})\}$$
$$= 32 + 100 \cdot \max\{0, 0.2\} = 32 + 20 = 52$$

If the change in the objective function($\nabla f(\mathbf{x})$) is larger than the change in the constraint($\nabla g(\mathbf{x})$) respectively, the value of the Lagrange Multiplier is increased. Therefore, we use the value of the Lagrange Multiplier as the value of 'R'.

Naval Architecture & Ocean Engineering

Computer Aided Ship Design
Part I. Optimization Method

Ch.6 Linear Programming

September, 2011
Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.6 Linear Programming

6.1 Linear Programming Problem

6.1 Linear Programming Problem

☑ **Linear Programming(LP) Problem**

- This problem has linear objective function and linear constraint functions in the design variables.
- Since all functions are linear in an LP problem, the feasible set defined by linear equalities or inequalities is convex.
- Also, the objective function is linear, so it is convex.
- Therefore, the LP problem is convex, and if an optimum solution exists, it is global optimum solution.

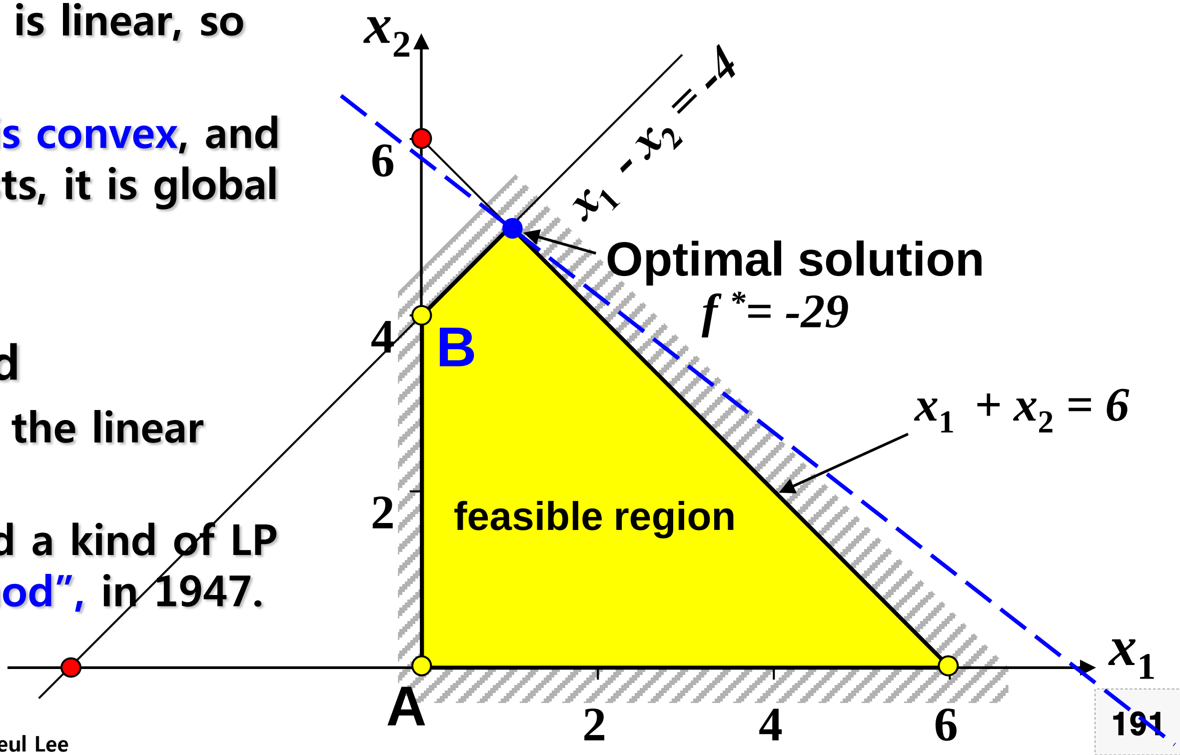
Objective function: *Minimize* $f = -4x_1 - 5x_2$

Constraints: *Subject to*

$$x_1 - x_2 \geq -4$$

$$x_1 + x_2 \leq 6$$

$$x_1, x_2 \geq 0$$



☑ **Linear Programming Method**

- This is the method to solve the linear programming problem.
- George B. Dantzig proposed a kind of LP method, "the Simplex method", in 1947.

- Property of the Linear Programming Problem

☑ **The objective function and constraints represent the linear relationship among the variables.**

- This problem has one objective function and constraints
- The objective function is minimum or maximum.

Objective function: $$\textit{Minimize} \quad f = -4x_1 - 5x_2$$

Constraints:
$$\textit{Subject to} \quad x_1 - x_2 \ge -4$$
$$x_1 + x_2 \le 6$$
$$x_1, x_2 \ge 0$$

☑ **The constraints are represented as the equality constraints(=) or inequality constraints(≥, ≤).**

☑ **To use the Simplex method, the variables have to be nonnegative in the LP problem.**

- If the variables are negative, the variable should be transformed to nonnegative.
  - Ex) x = -y (x is negative, y is positive)
- If a variable is unrestricted in sign, it can always be written as the difference of two nonnegative variables.
  - Ex) x = y – z(x is unrestricted in sign and y and z are nonnegative.)

✓ Example of problem which has nonnegative variables.
+ Distribution of the feed for animal : the amount of the feed can not be negative.
+ Distribution of the material for products : the amount of the material can not be negative.

✓ Example of variable which is unrestricted in sign.
+ Profit of the Shipyard = Price of a ship – Shipbuilding cost

6.1 Linear Programming Problem

- Example of the Linear Programming Problem: Problem with Two Variables and Inequality Constraint("<")

Objective function:

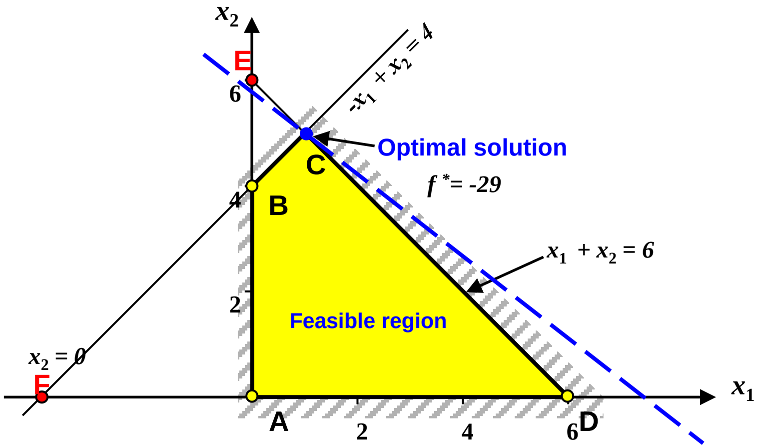
$$\textit{Minimize} \quad f = -4x_1 - 5x_2$$

Constraints:

$$\textit{Subject to} \quad x_1 - x_2 \ge -4$$

$$x_1 + x_2 \le 6$$

$$x_1, x_2 \ge 0$$



$$\textit{Minimize} \quad f = -4x_1 - 5x_2$$

$$\textit{Subject to} \quad -x_1 + x_2 \le 4$$

$$x_1 + x_2 \le 6$$

$$x_1, x_2 \ge 0$$

Maximization problem can be transformed to a minimization problem.

The right hand side of the constraints can always be made nonnegative by multiplying both side of the constraints by -1, if necessary.

Why should we transform the maximization problem to a minimization problem?
If the problem is not transformed to a minimization problem, we also have to find the method which can solve the maximization problem and minimization problem.

Ch.6 Linear Programming

6.2 Geometric Solution of Linear Programming Problem

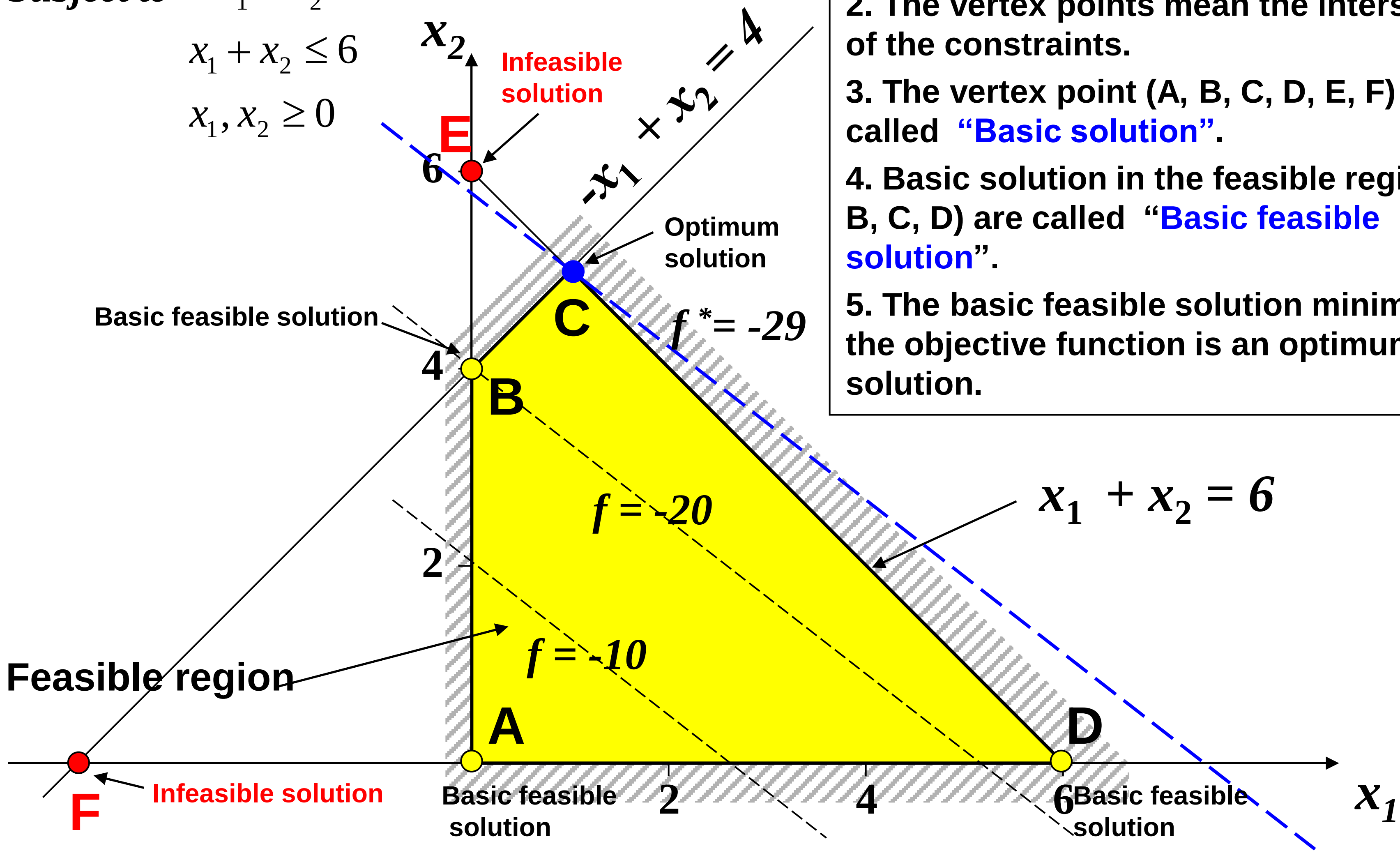
6.2 Geometric Solution of the Linear Programming Problem

$$\textit{Minimize} \quad f = -4x_1 - 5x_2$$

$$\textit{Subject to} \quad -x_1 + x_2 \le 4$$

$$x_1 + x_2 \le 6$$

$$x_1, x_2 \ge 0$$



1. The solution of LP problem lies on a vertex point of the polygon.
2. The vertex points mean the intersection of the constraints.
3. The vertex point (A, B, C, D, E, F) are called "Basic solution".
4. Basic solution in the feasible region (A, B, C, D) are called "Basic feasible solution".
5. The basic feasible solution minimizing the objective function is an optimum solution.

Ch.6 Linear Programming

6.3 Solution of Linear Programming Problem Using Simplex Method

Minimize $f = -4x_1 - 5x_2$

Subject to $\boxed{-x_1 + x_2 \le 4}$

$x_1 + x_2 \le 6$

$x_1, x_2 \ge 0$

For "≤" type inequality constraint we introduce a nonnegative slack variable.

$$-x_1 + x_2 \le 4 \quad \Rightarrow \quad -x_1 + x_2 + \underbrace{x_3}_{\text{Slack variable(nonnegative)}} = 4$$

Standard form of the Linear Programming Problem

1. Right hand side of the constraints should always be **nonnegative**.
2. Inequality constraint should be transformed to an **equality constraint**

6.3 Solution of Linear Programming Problem(1)

To transform "≤" type inequality constraints to the equality constraints, we introduce a nonnegative slack variable.

Minimize $f = -4x_1 - 5x_2$

Subject to $-x_1 + x_2 \le 4$

$x_1 + x_2 \le 6$

$x_1, x_2 \ge 0$

Transforming the inequality constraints to the equality constraints

Minimize $f = -4x_1 - 5x_2$

Subject to $-x_1 + x_2 + x_3 = 4$

$x_1 + x_2 + x_4 = 6$

$x_1, x_2, x_3, x_4 \ge 0$

$$-x_1 + x_2 + x_3 = 4$$
$$x_1 + x_2 + x_4 = 6$$

Because the number of variables(4) is larger than the number of equation(2), there are many sets of solution.

➡ If we assume the value of two(=4-2) unknown variables, we can obtain the solution.

➡ When we use the "Simplex method", the two unknown variables are assumed to be zero.

At this time, the variables set to zero are called "nonbasic variables", the remaining ones are called "basic variables".

When the number of unknown variables is n and the number of linear independent equations(constraints) is m,(n≥m)

- The degree of freedom is (n-m).
- If we assume the value of (n-m) unknown variables(degree of freedom), we can obtain the solution.
- In the "Simplex method", the (n-m) unknown variables are assumed to be zero.

6.3 Solution of Linear Programming Problem(2)

Minimize $f = -4x_1 - 5x_2$

Subject to $-x_1 + x_2 + x_3 = 4$

$x_1 + x_2 + x_4 = 6$

$x_1, x_2, x_3, x_4 \geq 0$

Nonbasic variables (assumed to be zero)	Basic variables	Solution (x_1, x_2, x_3, x_4)	Location of the solution ("Vertex point")	Objective function
(x_2, x_3)	(x_1, x_4)	(-4, 0, 0, 10)	F	16
(x_1, x_4)	(x_2, x_3)	(0, 6, -2, 0)	E	-30
(x_1, x_2)	(x_3, x_4)	(0, 0, 4, 6)	A	0
(x_2, x_4)	(x_1, x_3)	(6, 0, 10, 0)	D	-24
(x_1, x_3)	(x_2, x_4)	(0, 4, 0, 2)	B	-20
(x_3, x_4)	(x_1, x_2)	(1, 5, 0, 0)	C	-29

$-x_1 + x_2 + x_3 = 4$ ---- ①

$x_1 + x_2 + x_4 = 6$ ---- ②

$x_1, x_2, x_3, x_4 \geq 0$

Convert the inequality constraints to the equality constraint

Each vertex point is obtained by assuming the value of the two variables.

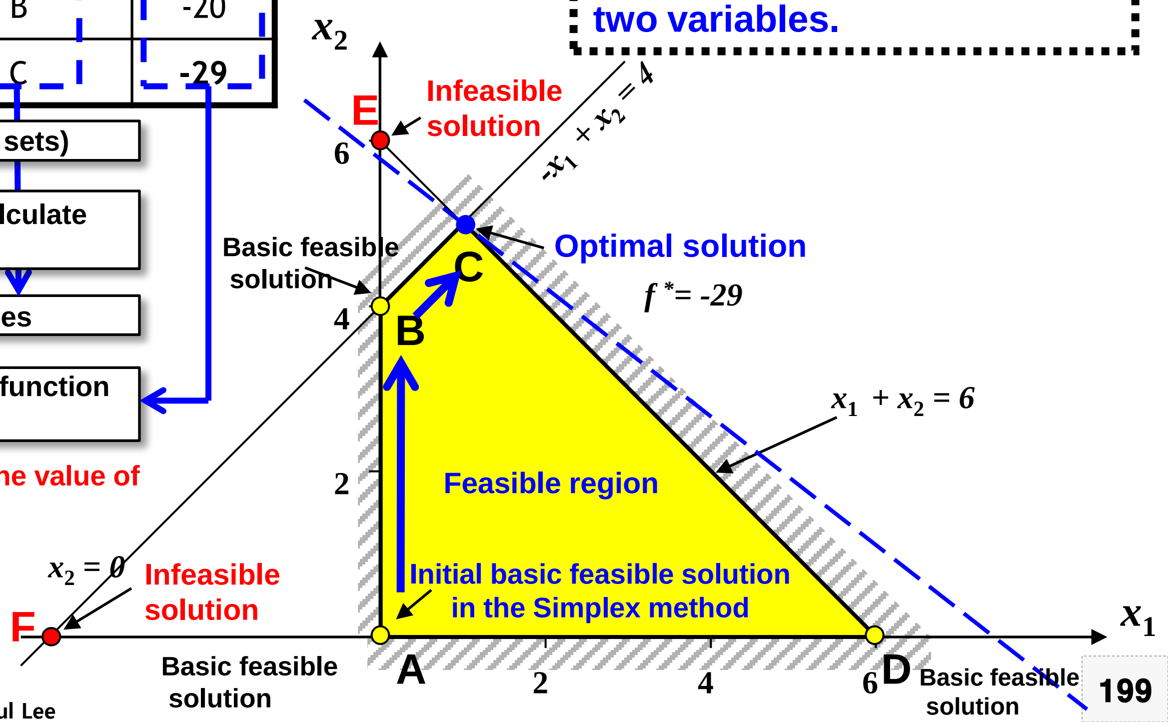
1) Select the two variables assumed to be zero(Total 6 sets)

2) Substitute the 6 sets into the equations ①, ② and calculate the value of the basic variables(vertex point)

3) Find the basic feasible solution in the 6 basic variables

4) The basic feasible solution minimizing the objective function is the optimum solution.

Q: Do we have to find all vertex points and calculate the value of the objective function?



General solution of LP problem:
"Simplex Method" starts at the initial basic feasible solution and finds the optimal solution by improving the objective function → We can minimize the number of calculating the vertex points.

Pivot: It is the same concept with Gauss-Jordan Elimination. This eliminates the selected variables from all the equations except one equation.

- In this example, we can solve this problem by **assuming the two variables as the nonbasic variables(=0)**.

.

(1) Transform the inequality constraints to the equality constraints

Minimize $f = -4x_1 - 5x_2$

Subject to $-x_1 + x_2 \le 4$

$x_1 + x_2 \le 6$

$x_1, x_2 \ge 0$

Mark the basic variable included in each row

Nonbasic variable | **Basic varialbe**

$\square$ **: Nonbasic variable(=0)**

$\bigcirc$ **: Basic vraible**

1row: x_3 | $\boxed{-x_1} \boxed{+x_2} + x_3 = 4$

2row: x_4 | $\boxed{x_1} \boxed{+x_2} + x_4 = 6$

3row: | $\boxed{-4x_1} \boxed{-5x_2} = f - 0$

$$x_1, x_2, x_3, x_4 \ge 0$$

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variable appears in only one row.

6.3 Solution of Linear Programming Problem by Using Simplex Method(2) - Interchange of Basic and Nonbasic Variables

1row: $x_3 \quad | \quad -x_1 + x_2 + x_3 = 4 \quad \longleftarrow$ **4/1 = 4** → Interchange the basic variable included in 1st row, i.e., x3 and the nonbasic variable, i.e., x2.

2row $x_4 \quad | \quad x_1 + x_2 + x_4 = 6 \quad \longleftarrow$ **6/1 = 6**

3row: $\quad | \quad -4x_1 - 5x_2 = f - 0$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)

○: Basic variable

Nonbasic variable: x1, x2, x3

Basic variable: x3, x4, x2

The greatest reduction in the objective function can be achieved by increasing x2, because its coefficient is most negative. → The nonbasic variable x2 should be replaced by a basic variable.

Because two variables should be the nonbasic variables(=0), x3 or x4 should be a nonbasic variable.

$$\frac{\text{Right hand side parameter in each row}}{\text{Positive coefficient of the element in the selected row}} =$$

Select the variable whose coefficient is positive and the row having the smallest positive ratio in the constraints → x3 is selected as the nonbasic variable.

<Ref.> What would be done if we do not select the row having the smallest positive ratio?

6.3 Solution of Linear Programming Problem by Using Simplex Method(3) - Pivot Operation

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variables appears in only one row.

$$
\begin{array}{ll|lllll}
\text{1row: } & x_3 & -x_1 & +x_2 & +x_3 & & =4 \quad \longleftarrow \mathbf{4/1 = 4} \\
\text{2row: } & x_4 & x_1 & +x_2 & & +x_4 & =6 \quad \longleftarrow \mathbf{6/1 = 6} \\
\text{3row: } & & -4x_1 & -5x_2 & & & =f-0
\end{array}
$$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)
○: Basic variable

Interchange the basic variable included in 1st row, i.e., x3 and the nonbasic variable, i.e., x2

Rearrange 1st row as: $x_2 = 4 + x_1 - x_3$

and substitute this into the 2 and 3 row.

$$x_1 + (4 + x_1 - x_3) + x_4 = 6$$
$$\Rightarrow 2x_1 - x_3 + x_4 = 2$$
$$-4x_1 - 5(4 + x_1 - x_3) = f$$
$$\Rightarrow -9x_1 + 5x_3 = f + 20$$

Nonbasic variable: x1, x2, x3

Basic variable: x3, x4, x2

Pivot on the selected variable(x_2 : 1st row, 2nd column)

$$
\begin{array}{ll|lllll}
\text{1st row: } & x_2 & -x_1 & +x_2 & +x_3 & & =4 \\
\text{2nd row: } & x_4 & 2x_1 & & -x_3 & +x_4 & =2 \\
\text{3rd row: } & & -9x_1 & & +5x_3 & & =f+20
\end{array}
$$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)
○: Basic variable

Pivot: It is the same concept with Gauss-Jordan Elimination. This eliminates the selected variables from all the equations except one equation.

6.3 Solution of Linear Programming Problem by Using Simplex Method(4)
- New Basic Variable("Vertex Point") after Pivot Operation

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variables appears in only one row.

$$
\begin{array}{l|l}
\textbf{1row: } x_2 & -x_1 + x_2 + x_3 \qquad = 4 \\
\textbf{2row: } x_4 & 2x_1 \qquad - x_3 + x_4 = 2 \\
\textbf{3row: } & -9x_1 \qquad + 5x_3 \qquad = f + 20
\end{array}
$$

$$x_1, x_2, x_3, x_4 \geq 0$$

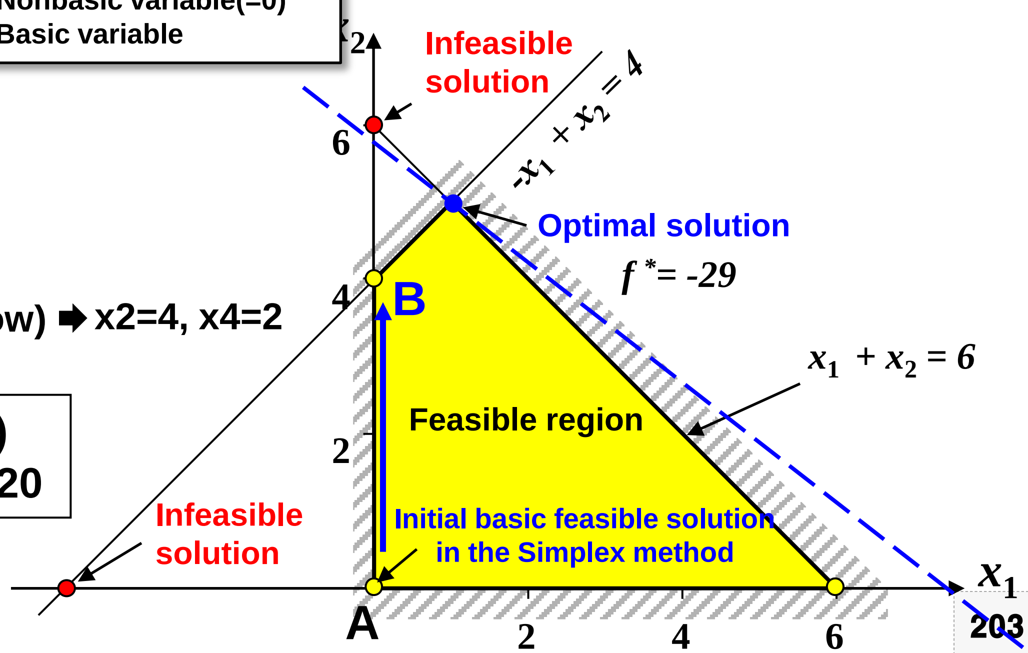
□ : Nonbasic variable(=0)
○ : Basic variable

Nonbasic variable: x1, x3

Basic variable: x2, x4

Substitute x1=x3=0 into the equations(1row, 2row) ➡ x2=4, x4=2

➡ **New solution B(x_1, x_2, x_3, x_4) = (0, 4, 0, 2)**
Value of the objective function at B = -20



6.3 Solution of Linear Programming Problem by Using Simplex Method(5) - Interchange of Basic and Nonbasic Variables

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variables appears in only one row.

$$
\begin{array}{lll}
\text{1row: } x_2 & -x_1 + x_2 + x_3 & = 4 \\
\text{2row: } x_4 & 2x_1 \qquad - x_3 + x_4 & = 2 \quad \longleftarrow \ 2/2 = 1 \\
\text{3row: } & -9x_1 \qquad + 5x_3 & = f + 20
\end{array}
$$

$$x_1, x_2, x_3, x_4 \geq 0$$

□ : Nonbasic variable(=0)
○ : Basic variable

→ Interchange the basic variable included in 2nd row, i.e., x4 and the nonbasic variable, i.e., x1.

Nonbasic variable: x1, x3 ,x4

Basic variable: x2, x4 ,x1

The greatest reduction in the objective function can be achieved by increasing x1, because its coefficient is most negative.
→ The nonbasic variable x1 should be replaced by a basic variable.

Because two variables should be the nonbasic variables(=0),
→ x2 or x4 should be the nonbasic variable.

$$\frac{\text{Right hand side parameter in each row}}{\text{Positive coefficient of the element in the selected row}} =$$

Select the variable whose coefficient is positive and row and the row having the smallest positive ratio in the constraints
→ x4 is selected as the nonbasic variable.

6.3 Solution of Linear Programming Problem by Using Simplex Method(6) - Pivot Operation

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variables appears in only one row.

1row: x_2 | $-x_1 + x_2 + x_3 = 4$

2row: x_4 | $2x_1 - x_3 + x_4 = 2$ ← 2/2 = 1

3row: | $-9x_1 + 5x_3 = f + 20$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)
○: Basic variable

→ **Interchange the basic variable included in the 2nd row, i.e., x4 and the nonbasic variable, i.e., x1.**

Nonbasic variable: x1, x3 ,x4

Basic variable: x2, x4 ,x1

Pivot on the selected variable(x_1 : 2nd row, 1st column)

(1 row + 0.5×2 row) → 1row: x_2 | $x_2 + 0.5x_3 + 0.5x_4 = 5$

(0.5×2 row) → 2row: x_1 | $x_1 - 0.5x_3 + 0.5x_4 = 1$

(3 row + 4.5×2 row) → 3row: | $+0.5x_3 + 4.5x_4 = f + 29$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)
○: Basic variable

6.3 Solution of Linear Programming Problem by Using Simplex Method(7) - New Basic Variable("Vertex Point") after Pivot Operation/ Stop to Simplex

Type of variables	Explanation	Method to classify
Nonbasic variables	A variable set to zero in variables	Objective function is only composed of the nonbasic variables.
Basic variables	A variable obtained by setting the nonbasic variable and solving the equations simultaneously	Each basic variables appears in only one row.

1row: x_2 | $x_2 + \boxed{0.5x_3} + \boxed{0.5x_4} = 5$

2row: x_1 | $x_1 \quad - \boxed{0.5x_3} + \boxed{0.5x_4} = 1$

3row: | $+ \boxed{0.5x_3} + \boxed{4.5x_4} = f + 29$

$$x_1, x_2, x_3, x_4 \geq 0$$

□ : Nonbasic variable(=0)
○ : Basic variable

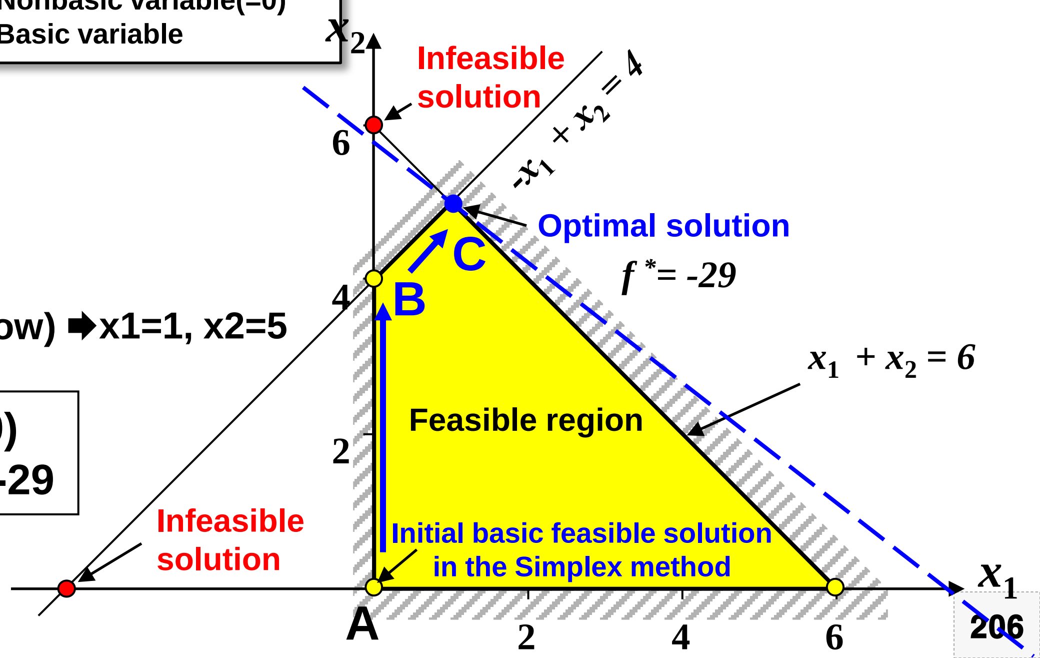
Because the coefficients of the objective function are nonnegative, the current solution is the optimal solution.
→ Stop the Simplex

Nonbasic variable: x3, x4

Basic variable: x1, x2

Substitute x3=x4=0 into the equations(1row, 2row) ➡x1=1, x2=5

➡ **New solution C(x_1, x_2, x_3, x_4) = (1, 5, 0, 0)**
Value of the objective function at B = -29



[Reference] The reason why the column which has the minimum coefficient of the objective function is selected for pivot.

1row: x_3	$-x_1$	$+x_2$	$+x_3$		$=4$
2row: x_4	x_1	$+x_2$		$+x_4$	$=6$
3row:	$-4x_1$	$-5x_2$			$=f-0$

$$x_1, x_2, x_3, x_4 \geq 0$$

□ : Nonbasic variable(=0)
○ : Basic variable

The nonbasic variables(x_1 and x_2) are equal to zero. $(x_3 = 4,\ x_4 = 6)$

If there are some variables whose coefficients are nonnegative in the objective function, the variables(x_1 and x_2) can be increased for decreasing the value of the objective function.

The greatest reduction in the value of the objective function can be achieved by increasing x_2 , because its coefficient is most negative.

[Reference] The reason why the row having the smallest positive ratio in the constraints is selected.

Select the variable whose coefficient is positive and the row having the smallest positive ratio in the constraints → x3 will be selected as the nonbasic variable.

1row:	x_3	$-x_1 + x_2 + x_3 = 4$	← 4/1 = 4	← The row having the smallest positive ratio(1 row)
2row:	x_4	$x_1 + x_2 + x_4 = 6$	← 6/1 = 6	
3row:		$-4x_1 - 5x_2 = f - 0$		

$$x_1, x_2, x_3, x_4 \geq 0$$

□ : Nonbasic variable(=0)
○ : Basic variable

The row 1 and 2 are rearranged as follows.

$$-x_1 + x_3 = 4 - x_2$$

$$x_1 + x_4 = 6 - x_2$$

1) If the 1st row is selected, then x_3 becomes nonbasic variable.

1st row: $x_1 = x_3 = 0, x_2 = 4 \quad (\because x_1, x_3 \text{ are nonbasic variables})$

2nd row: $x_1 = 0, x_2 = 4, x_4 = 2$

2) If the 2nd row is selected, then x_4 becomes nonbasic variable.

2nd row: $x_1 = x_4 = 0, x_2 = 6 \quad (\because x_1, x_4 \text{ are nonbasic variables})$

1st row: $x_1 = 0, x_2 = 6, x_3 = -2$ ➡ **The constraint, the variables have to be nonnegative, is violated.**

[Reference] The reason why the row having the negative coefficient in the selected column is not selected.(1)

1row: x_2 | $-x_1 + x_2 + x_3 \quad = 4$ ← The row having the negative coefficient in the selected column is not selected.

2row x_4 | $2x_1 \quad - x_3 + x_4 = 2$ ← 2/2 = 1

3row: | $-9x_1 \quad + 5x_3 \quad = f + 20$

$$x_1, x_2, x_3, x_4 \geq 0$$

□: Nonbasic variable(=0)
○: Basic variable

Nonbasic variable: x1, x3, x4

Basic variable: x2, x4, x1

1. The row 1 and 2 are rearranged as follows.

$$x_2 \quad + x_3 \quad = 4 + x_1 \qquad \text{------} \quad ①$$

$$-x_3 + x_4 = 2 - 2x_1 \qquad \text{------} \quad ②$$

2. x2 or x4 will become a nonbasic variable.

3. If x4 becomes a nonbasic variable,

3-1. Equation ② is changed as follows. (nonbasic variable x_3=0, x_4=0)

$$0 = 2 - 2x_1 \ \rightarrow \ 2 = 2x_1 \ \rightarrow \ 1 = x_1$$

3-2. Equation ① is changed as follows.(nonbasic variable x_3=0, x_4=0)

$$x_2 = 4 + x_1 \geq 0$$

In 3-1, any value of x_1 satisfies the equation ①

→ If the row having the positive coefficient in the selected column is selected, the row having the negative coefficient in the selected column is always satisfied.

[Reference] The reason why the row having the negative coefficient in the selected column is not selected.(2)

$$\begin{array}{llrrrrl}
\text{1row:} & x_2 & -x_1 & +x_2 & +x_3 & & =4 \\
\text{2row:} & x_4 & 2x_1 & & -x_3 & +x_4 & =2 \\
\text{3row:} & & -9x_1 & & +5x_3 & & =f+20
\end{array}$$

$$x_1, x_2, x_3, x_4 \geq 0$$

The row having the negative coefficient in the selected column is not selected.

2/2 = 1

□ : Nonbasic variable(=0)

○ : Basic variable

Nonbasic variable: x1, x3, x2

Basic variable: x2, x4, x1

1. The row 1 and 2 are rearranged as follows.

$$x_2 + x_3 = 4 + x_1 \quad \text{------} \quad ①$$

$$-x_3 + x_4 = 2 - 2x_1 \quad \text{------} \quad ②$$

2. x2 or x4 will become a nonbasic variable.

3. If x2 becomes a nonbasic variable,

3-1. Equation ① is changed as follows. (nonbasic variable x_2=0, x_3=0)

$$0 = 4 + x_1 \rightarrow x_1 = -4$$

➡ The constraint, the variables have to be nonnegative, is violated.

6.3 Solution of Linear Programming Problem by Using Simplex Tableau

Pivot: It is the same concept with Gauss-Jordan Elimination. This eliminates the selected variables from all the equations except one equation.

Basic variable | Nonbasic variable (=0) | Basic variable

1row: x_3 $\quad -x_1 + x_2 + x_3 = 4$ ← **4/1 = 4**

2row: x_4 $\quad x_1 + x_2 + x_4 = 6$ ← **6/1 = 6**

3row: $\quad -4x_1 - 5x_2 = f - 0$

Basic variable

		x1	x2	x3	x4	bi	bi/ai
1row:	x3	-1	1	1	0	4	4
2row:	x4	1	1	0	1	6	6
3row:	Obj.	-4	-5	0	0	f-0	-

Pivot on x2(1 row and 2 column)

New 2row = (2row - 1row)
New 3row = (3row + 5×1row)

Basic variable | Nonbasic variable (=0)

1row: x_2 $\quad -x_1 + x_2 + x_3 = 4$ ← **4/-1 = -4**

(If the coefficient of the variable is negative, the variable is not selected.)

2row: x_4 $\quad 2x_1 - x_3 + x_4 = 2$ ← **2/2 = 1**

3row: $\quad -9x_1 + 5x_3 = f + 20$

Basic variable

		x1	x2	x3	x4	bi	bi/ai
1row:	X2	-1	1	1	0	4	-4
2row:	x4	2	0	-1	1	2	1
3row:	Obj.	-9	0	5	0	f+20	-

Pivot on x1(2 row and 1 column)

New 1row = (1row + 0.5×2row)
New 2row = (0.5×2row)
New 3row = (3row + 4.5×2row)

Basic variable | Nonbasic variable (=0)

1row: x_2 $\quad x_2 + 0.5x_3 + 0.5x_4 = 5$

2row: x_1 $\quad x_1 - 0.5x_3 + 0.5x_4 = 1$

3row: $\quad +0.5x_3 + 4.5x_4 = f + 29$

Basic variable

		x1	x2	x3	x4	bi	bi/ai
1row:	x2	0	1	0.5	0.5	5	-
2row:	x1	1	0	-0.5	0.5	1	-
3row:	Obj.	0	0	0.5	4.5	f+29	-

Because all the coefficients of the objective function are nonnegative, the current solution is the optimal solution. (x_1=1, x_2=5, x_3=x_4=0, f=-29)

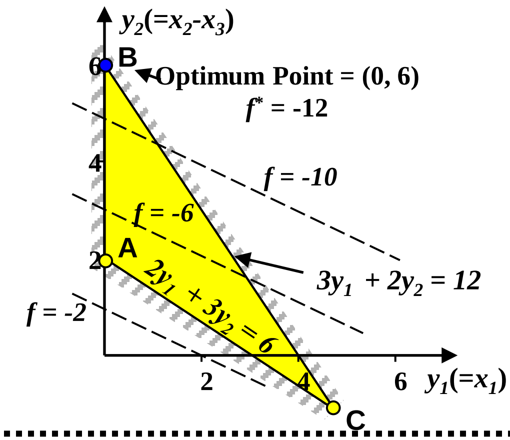
Maximize $z = y_1 + 2y_2$

Subject to $3y_1 + 2y_2 \le 12$

$2y_1 + 3y_2 \ge 6$

$y_1 \ge 0$

y_2 **is unrestricted in sign.**



Minimize $F = -y_1 - 2y_2$

Subject to $3y_1 + 2y_2 \le 12$

$2y_1 + 3y_2 \ge 6$

$y_1 \ge 0$

y_2 **is unrestricted in sign.**

Maximization problem can be transformed to a minimization problem.

The variable unrestricted in sign is expressed with two nonnegative variables.
($y_2 = y_2^+ - y_2^-$ **)**

Let be $x_1 = y_1, x_2 = y_2^+, x_3 = y_2^-$

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 \le 12$

$2x_1 + 3x_2 - 3x_3 \ge 6$

$x_1, x_2, x_3 \ge 0$

6.3 Solution of Linear Programming Problem Using Simplex Method - Transformation of "≥" Type Inequality Constraint

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 \le 12$

$\boxed{2x_1 + 3x_2 - 3x_3 \ge 6}$

$x_1, x_2, x_3 \ge 0$

[Review]For "≤" type inequality constraint: we introduce a nonnegative variable.

$$3x_1 + 2x_2 - 2x_3 + \underline{x_4} = 12$$

For "≥" type inequality constraint, we introduce a surplus variable and artificial variable.

$$2x_1 + 3x_2 - 3x_3 \ge 6 \quad ➡ \quad 2x_1 + 3x_2 - 3x_3 - \underbrace{x_5}_{\text{Surplus variable (nonnegative)}} + \underbrace{x_6}_{\text{Artificial variable(nonnegative)}} = 6$$

"The reason why we introduce the artificial variable"

At starting the Simplex method, we assume the original design variables (x_1, x_2, x_3) as "nonbasic variables" (x_1=x_2=x_3=0), -x_5 = 6.

➡ This violates the nonnegativity requirement. For satisfying the requirement, we introduce the variable x_6 artificially.

However, the artificial variable should be equal to zero in the feasible region, because x_6 is augmented artificially,

6.3 Solution of Linear Programming Problem Using Simplex Method (Simplex Tableau)
- Simplex Method for the Problem with "≥" Type Inequality Constraint (1)

①

Maximize $z = y_1 + 2y_2$

Subject to $3y_1 + 2y_2 \le 12$

$2y_1 + 3y_2 \ge 6$

$y_1 \ge 0$

y_2 is unrestricted in sign.

1. Transform to a minimization problem.
2. Since y_2 is unrestricted in sign, transform as $y_2 = y_2^+ - y_2^-$.
3. Let be $x_1 = y_1, x_2 = y_2^+, x_3 = y_2^-$
4. Transform the inequality constraints to the equality constraints (Introduce the slack and surplus variable.)

②

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 + x_4 = 12$

$2x_1 + 3x_2 - 3x_3 - x_5 = 6$

$x_i \ge 0; i = 1\ to\ 5$

Slack variable

Surplus variable

Assume the original variables (x_1, x_2, x_3) as nonbasic variables(=0) and calculate the basic variable (x_4, x_5).

$x_4 = 12, x_5 = -6$ ➡ This violates the nonnegativity requirement

Introduce an artificial variable x_6 in the "≥" type inequality constraints.

③

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 + x_4 = 12$

$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$

$x_i \ge 0; i = 1\ to\ 6$

Slack variable

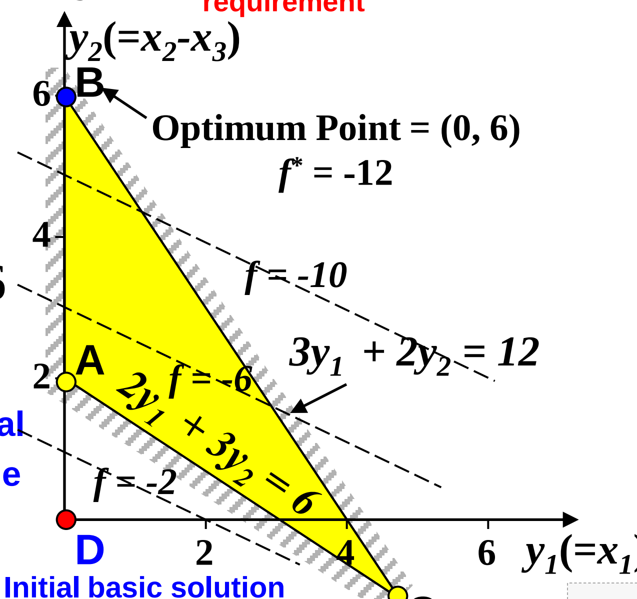
Surplus variable

Artificial variable

Assume the original design variables(x_1, x_2, x_3) and the surplus variable(x_5) as nonbasic variables(=0) and calculate the basic variable(x_4, x_6).

The result is x_4=12, x_6=6. ➡ Initial basic solution(Infeasible solution)

However, the artificial variable should be equal to zero in the feasible region, because x_6 is augmented artificially.



Initial basic solution
(Infeasible solution)

- Simplex Method for the Problem with "≥" Type Inequality Constraint (2)

3

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 + x_4 = 12$

$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$

$x_i \geq 0; i = 1 \, to \, 6$

Slack variable (x_4)

Surplus variable (x_5) **Artificial variable** (x_6)

Define an artificial objective function which is a sum of all the artificial variables(w=x_6)

4

$3x_1 + 2x_2 - 2x_3 + x_4 = 12$

$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$

$-x_1 - 2x_2 + 2x_3 = f$

$\boxed{-2x_1 - 3x_2 + 3x_3 + x_5 = w - 6}$

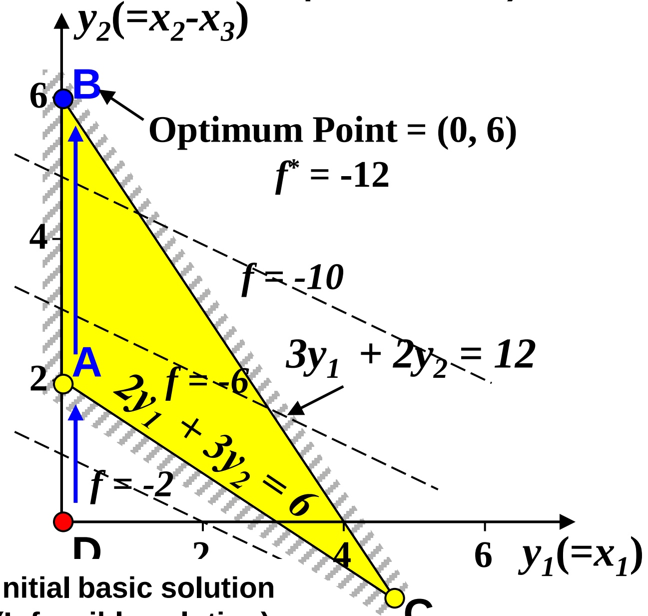
Artificial objective function

Designate x_6=w and rearrange $2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$

5 **Find the basic feasible solution(minimize the artificial objective function, w=x_6 ("w=0") (Phase 1 of the Simplex method)**

Since x_6 is augmented artificially, the artificial variable should be equal to zero in the feasible region.

6 **Find the optimal solution to minimize the original objective function(Phase 2 of the Simplex method)**



Initial basic solution (Infeasible solution)

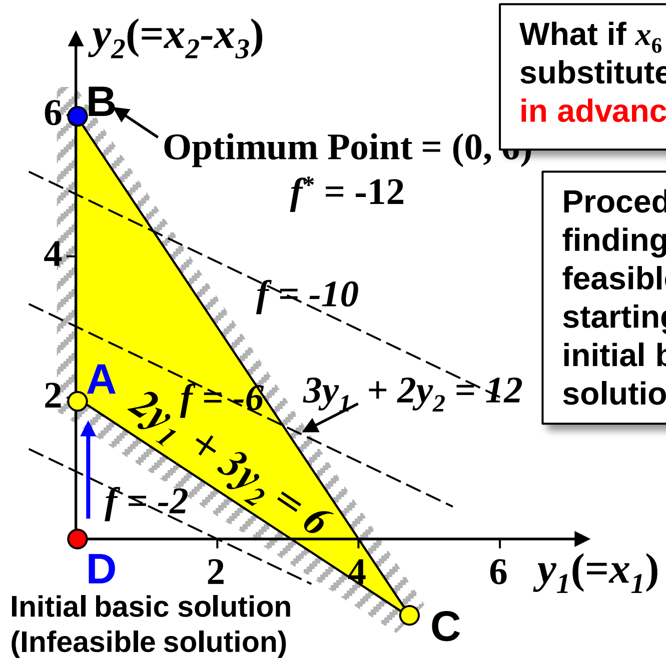
6.3 Solution of Linear Programming Problem Using Simplex Method(Simplex Tableau) - Simplex Method for the Problem with "≥" Type Inequality Constraint (3)

④

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$
$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$
$$-x_1 - 2x_2 + 2x_3 = f$$
$$-2x_1 - 3x_2 + 3x_3 + x_5 = w - 6$$

At first, **we assume the original design variables(x_1, …, x_3) and surplus variable(x_5) as nonbasic variables(=0), whereas the slack variable(x_4) and artificial variable(x_6) as basic variables**. Then solve the equation. ("Starting with the initial basic solution")

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	-
x6	2	3	-3	0	-1	1	6	-
Obj.	-1	-2	2	0	0	0	f-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-



What if x_6 is substituted for zero in advance?

Procedure of finding the basic feasible solution starting with the initial basic solution

⑤ Phase 1: Repeat Pivot operation until the artificial objective function w becomes zero.

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	6
x6	2	3	-3	0	-1	1	6	2
Obj.	-1	-2	2	0	0	0	f-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

➡

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	-
x2	2/3	1	-1	0	-1/3	1/3	2	-
Obj.	1/3	0	0	0	-2/3	2/3	f+4	-
A. Obj.	0	0	0	0	0	1	w-0	-

Since the artificial variable(x_6) is augmented artificially, the variable should be equal to zero in the feasible region.

New 1 row = 1 row - (2/3)×2 row
New 2 row = (1/3)×2 row
New 2 row = 3 row- (2/3)×2 row
New 4 row = 4 row + 2 row

Since the value of the artificial objective function becomes zero, the Phase 1 is completed.

Point A(x_1=x_3=x_5=x_6=0, x_2=2, x_4=8)

- Simplex Method for the Problem with "≥" Type Inequality Constraint (4)

5 Phase 1: Repeat Pivot operation until the artificial objective function *w* becomes zero.

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	6
x6	2	3	-3	0	-1	1	6	2
Obj.	-1	-2	2	0	0	0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

➡

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	-
x2	2/3	1	-1	0	-1/3	1/3	2	-
Obj.	1/3	0	0	0	-2/3	2/3	*f*+4	-
A. Obj.	0	0	0	0	0	1	w-0	-

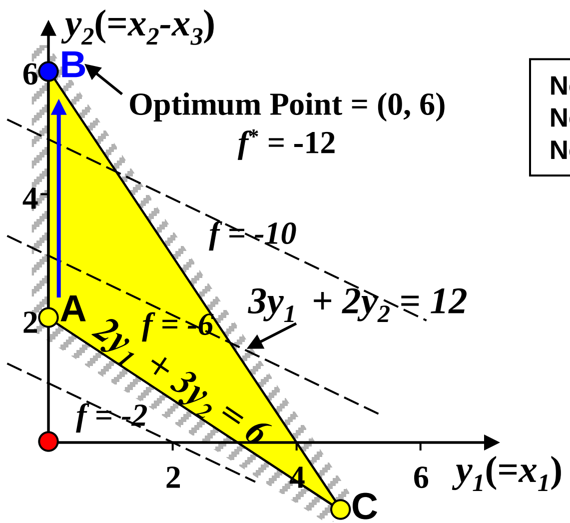
6 Phase 2: Repeat Pivot operation until all the coefficients of the original objective function *f* are nonnegative.

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	12
x2	2/3	1	-1	0	-1/3	1/3	2	-6
Obj.	1/3	0	0	0	-2/3	2/3	*f*+4	-

➡

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x5	5/2	0	0	3/2	1	-1	12	-
x2	3/2	1	-1	1/2	0	0	6	-
Obj.	2	0	0	1	0	0	*f*+12	-

New 1 row = 1 row × (2/3)
New 2 row = 2 row + (1/2)×1 row
New 3 row = 3 row + 1 row



Since all the coefficients of the objective function are nonnegative, the current solution is the optimal solution.
($x_1=x_3=x_4=0, x_2=6, x_5=12, f=-12$)

6.3 Solution of Linear Programming Problem Using Simplex Method
- Transformation of Equality("=") Constraint

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to

$3x_1 + 2x_2 - 2x_3 \le 12 \Rightarrow 3x_1 + 2x_2 - 2x_3 + \underline{x_4} = 12$

$2x_1 + 3x_2 - 3x_3 \ge 6 \Rightarrow 2x_1 + 3x_2 - 3x_3 - \underline{x_5} + \underline{x_6} = 6$

$\boxed{x_1 + x_2 + x_3 = 6}$

$x_1, x_2, x_3 \ge 0$

[Review] For "≤" type inequality constraint, we introduce a nonnegative slack variable.

[Review] For "≥" type inequality constraint, we introduce a surplus variable and artificial variable.

For "=" type equality constraint, we introduce an artificial variable.

$$x_1 + x_2 + x_3 = 6 \quad \Rightarrow \quad x_1 + x_2 + x_3 + \underline{x_7} = 6$$

Artificial variable(nonnegative)

"The reason why we introduce the artificial variable"

At starting the Simplex method, we assume the original design variables (x_1, x_2, x_3) as "nonbasic variables" (x_1=x_2=x_3=0). Then the equality constraint is violated(0 = 6).

➡ To satisfy the equality constraint , we introduce the variable x_7 artificially.
However, because x_7 is augmented artificially, the artificial variable should be equal to zero in the feasible region.

6.3 Solution of Linear Programming Problem Using Simplex Method - Method for Formulating the Artificial Objective Function

①

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to

$$3x_1 + 2x_2 - 2x_3 \le 12$$
$$2x_1 + 3x_2 - 3x_3 \ge 6$$
$$x_1 + x_2 + x_3 = 6$$
$$x_1, x_2, x_3 \ge 0$$

Transform the inequality constraints to equality constraints →

②

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$
$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$
$$x_1 + x_2 + x_3 + x_7 = 6$$
$$x_i \ge 0;\ i = 1\ to\ 7$$

$$x_6 - 6 = -2x_1 - 3x_2 + 3x_3 + x_5$$
$$x_7 - 6 = -x_1 - x_2 - x_3$$
$$w(= x_6 + x_7) - 12 = -3x_1 - 4x_2 + 2x_3 + x_5$$

Define an artificial objective function which is a sum of all the artificial variables (w=x_6+x_7)

③

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$
$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$
$$x_1 + x_2 + x_3 + x_7 = 6$$
$$-x_1 - 2x_2 + 2x_3 = f$$
$$-3x_1 - 4x_2 + 2x_3 + x_5 = w - 12$$

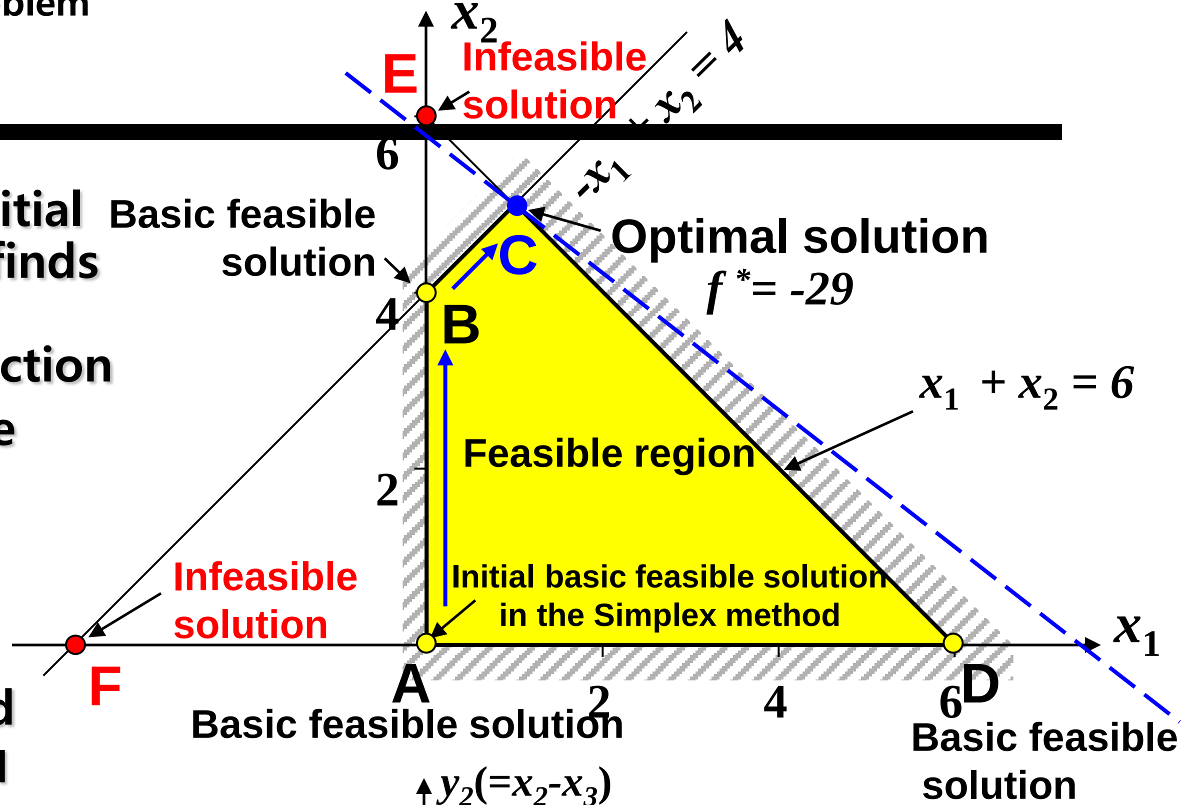
<Ref.> If we define the artificial objective functions for each artificial variable,

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$
$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$
$$x_1 + x_2 + x_3 + x_7 = 6$$
$$-x_1 - 2x_2 + 2x_3 = f$$
$$\left.\begin{aligned} -2x_1 - 3x_2 + 3x_3 + x_5 &= w_1 - 6 \\ -x_1 - x_2 - x_3 &= w_2 - 6 \end{aligned}\right\}$$

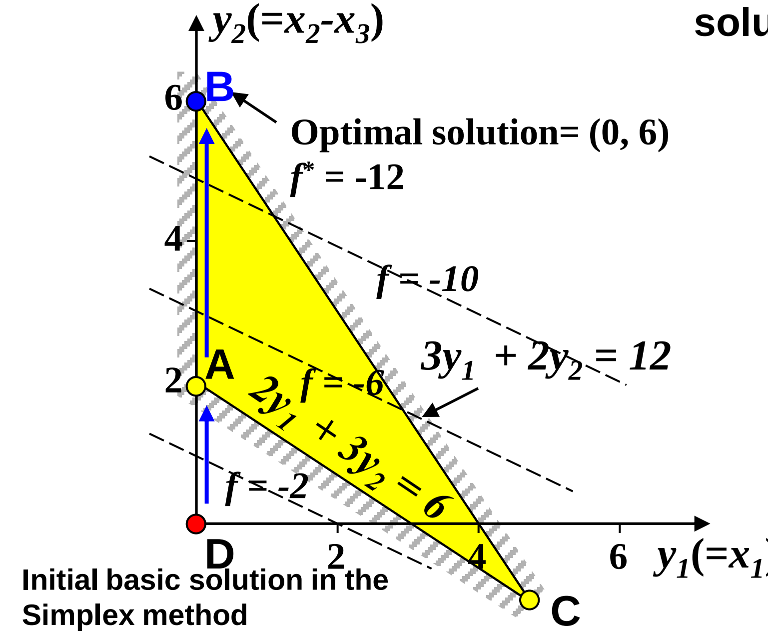
➡ **We have to minimize w_1(x_6=0) and w_2 (x_7=0).**

Since the artificial variables are nonnegative, solutions of minimizing the sum of all the artificial objective functions are the same as those of minimizing of each artificial objective function . Therefore, it is convenient to define the artificial objective function as a sum of all the artificial variables.

Find the basic feasible solution(minimize the artificial objective function, w=x_6+x_7("w=0"; x_6=x_7=0)



- ☑ This method starts at the initial basic feasible solution and finds the optimal solution by improving the objective function
- ☑ This method is based on the theory of the first-order simultaneous equations.
  - ■ Matrix calculation is used. (Gauss-Jordan Elimination)
- ☑ Type of the Simplex method
  - ■ One-phase Simplex method
    - ● The problem only having "≤" type inequality constraints
  - ■ Two-phase Simplex method
    - ● The problem having "≥" type inequality or equality ("=") constraint
    - ● Phase 1: Find the initial basic feasible solution to satisfy the artificial objective function(w) to be zero.
    - ● Phase 2: Find the optimal solution by starting with the initial basic feasible solution.



6.3 Solution of Linear Programming Problem Using Simplex Method - Summary of the Simplex Algorithm

- ☑ **Step 1: initial basic feasible solution**
  - ■ **“<=” type inequality constraints: Find the initial basic feasible variables by assuming the slack variables as basic and the original variables as nonbasic variables(=0).**
  - ■ **“>=” type inequality constraints: By using the Two-phase Simplex method, find the initial basic feasible variables to satisfy the artificial objective function to be zero in the Phase 1.**

- ☑ **Step 2: The objective function must be expressed with the nonbasic variables.**

- ☑ **Step 3: If all the reduced coefficient of the objective function for nonbasic variables are nonnegative, the current basic solution is the optimal solution. Otherwise, continue.**

- ☑ **Step 4: Determine the Pivot column and row. At this time, the nonbasic variable in the selected Pivot column should become the new basic variable and the basic variable in the selected Pivot row should become the new nonbasic variable.**

- ☑ **Step 5: Pivot operation by using the Gauss-Jordan Elimination**
- ☑ **Step 6: Calculate the value of the basic and nonbasic variable and go to Step 3.**

6.3 Solution of Linear Programming Problem Using Simplex Method
- [Reference] Time to transform the variables unrestricted in sign to the nonnegative variables

Mathematical Model

Minimize $z = -y_1 - 2y_2$

Subject to $3y_1 + 2y_2 \le 12$

$2y_1 + 3y_2 \ge 6$

$y_1 \ge 0$

y_2 **is unrestricted in sign.**

$\boxed{y_2 = y_2^+ - y_2^-}$

$y_2^+, y_2^+ \ge 0$

Transform the variable unrestricted in sign to nonnegative variable

Minimize $f = -y_1 \boxed{-2y_2^+ + 2y_2^-}$

Subject to $3y_1 + \boxed{2y_2^+ - 2y_2^-} \le 12$

$2y_1 + \boxed{3y_2^+ - 3y_2^-} \ge 6$

$y_1, y_2^+, y_2^- \ge 0$

Order (1)

"≤" type inequality constraint:
Introduce the slack variable.

"≥" type inequality constraint:
Introduce the surplus variable and the artificial variable.

"≤" type inequality constraint:
Introduce the slack variable.

"≥" type inequality constraint:
Introduce the surplus variable and the artificial variable.

Order (2)

Minimize $f = -y_1 - 2y_2$

Subject to $3y_1 + 2y_2 + \boxed{x_1} = 12$

$2y_1 + 3y_2 - \boxed{x_2} + \boxed{x_3} = 6$

$y_1, x_i \ge 0; i = 1 \ to \ 3$

y_2 **is unrestricted in sign.**

$\boxed{y_2 = y_2^+ - y_2^-}$

$y_2^+, y_2^+ \ge 0$

Transform the variable unrestricted in sign to nonnegative variable

Minimize $f = -y_1 \boxed{-2y_2^+ + 2y_2^-}$

Subject to $3y_1 + \boxed{2y_2^+ - 2y_2^-} + \boxed{x_1} = 12$

$2y_1 + \boxed{3y_2^+ - 3y_2^-} - \boxed{x_2} + \boxed{x_3} = 6$

$y_1, y_2^+, y_2^- \ge 0, \ x_i \ge 0; i = 1 \ to \ 3$

After formulating the mathematical model, there is no restriction in order between transforming the variables unrestricted in sign to the nonnegative variables and introducing the slack, surplus and artificial variables.

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$
$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$
$$-x_1 - 2x_2 + 2x_3 = f$$

When x_6 is substituted for zero,

the other variables(x_1, x_2, x_3, x_5) in the same equation should not be negative.

The procedure of the calculating the values of x_1, x_2, x_3, x_5 is identical with that of reducing the artificial objective function(x_6) to zero in the Simplex method.

- [Reference] Procedure of finding the basic feasible solution starting with the initial basic solution(1)

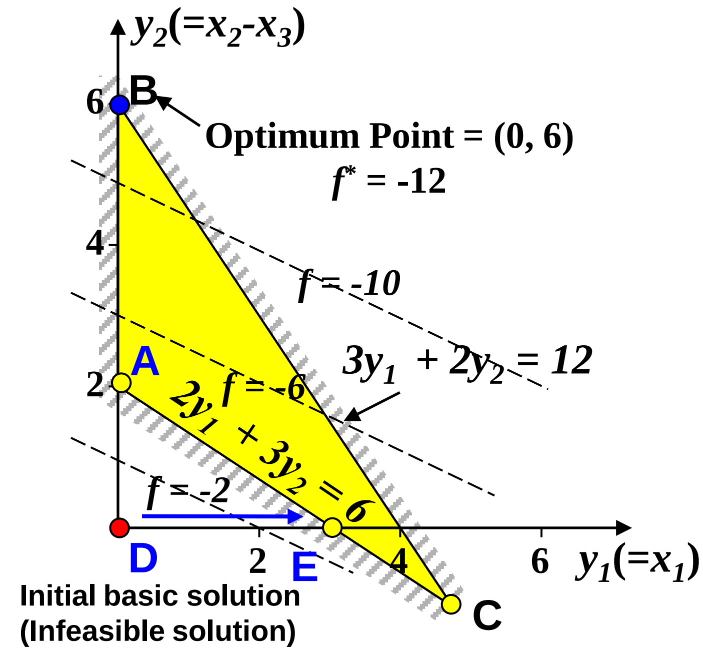
	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	4
x6	2	**3**	**-3**	**0**	**-1**	**1**	**6**	**3**
Obj.	-1	**-2**	2	0	0	0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

Select the first column and perform the Pivot.

(In the general Simplex method, the second column is selected.)

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	0	-5/2	5/2	1	3/2	-3/2	3	-
x1	1	3/2	-3/2	0	-1/2	1/2	3	-
Obj.	0	-1/2	1/2	0	-1/2	1/2	*f*+3	-
A. Obj.	0	0	0	0	0	1	w-0	-

Since the value of the artificial objective function becomes zero, the Phase 1 is completed. Point E(x_2=x_3=x_5=x_6=0, x_1=3, x_4=3)



The basic feasible solution can be found from the initial basic solution through the near corner.

→ It is similar with the procedure of finding the optimal solution from the initial basic feasible solution. (through the near corner)

- Since Phase1 is completed, Phase 2 is performed.

- Phase2: Pivot operation for the original objective function *f*

6.3 Solution of Linear Programming Problem Using Simplex Method
- [Reference] Procedure of finding the basic feasible solution starting with the initial basic solution(2)

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	0	-5/2	5/2	1	3/2	-3/2	3	-6/5
x1	1	3/2	-3/2	0	-1/2	1/2	3	2
Obj.	0	-1/2	1/2	0	-1/2	1/2	f+3	-
A. Obj.	0	0	0	0	0	1	w-0	-

Since the value of the artificial objective function becomes zero, the Phase 1 is completed.
Point E($x_2=x_3=x_5=x_6=0$, $x_1=3$, $x_4=3$)

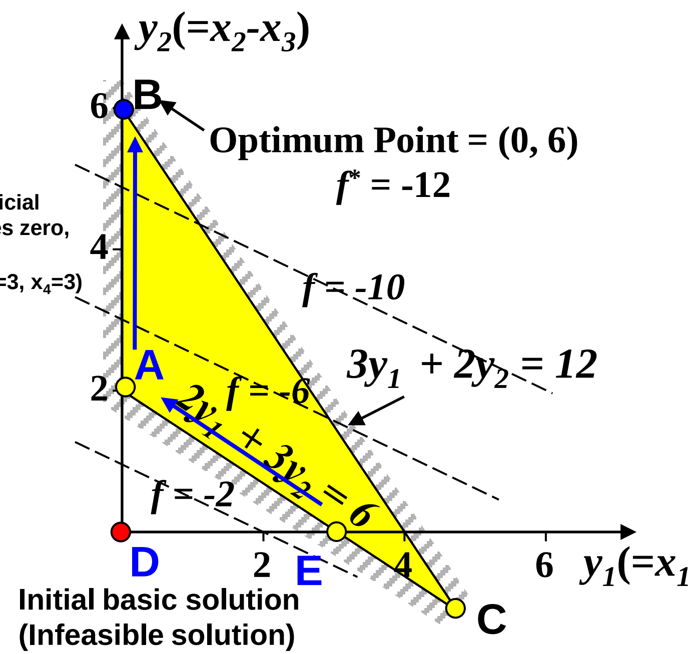
New 1 row = 1row + 2row × (5/3)
New 2row = 2row × (2/3)
New 3row = 3row + 2row × (1/3)

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	12
x2	2/3	1	-1	0	-1/3	1/3	2	-6
Obj.	1/3	0	0	0	-2/3	2/3	f+4	-

Point A($x_1=x_3=x_5=x_6=0, x_2=2, x_4=8$)

New 1row = 1row × (2/3)
New 2row = 2row + (1/2)×1row
New 3row = 3row + 1row

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x5	5/2	0	0	3/2	1	-1	12	-
x2	3/2	1	-1	1/2	0	0	6	-
Obj.	2	0	0	1	0	0	f+12	-



Since all the coefficients of the objective function are nonnegative, the current solution is the optimal solution.

Point B($x_1=x_3=x_4=x_6=0, x_2=6, x_5=12, f=-12$)

Consider a cargo ship departing from the port A to E via the ports B, C, D. The maximum cargo loading capacity of the ship is 50,000ton and the loadable cargo at each port is as follows. Formulate and find the optimum cargo transportation that maximizes the freight rate.

Type of cargo	Port of departure	Port of arrival	Loadable cargo at each port of departure (1,000ton)	Freight rate ($/ton)
1	A	B	100	5
2	A	C	40	10
3	A	D	25	20
4	B	C	50	8
5	B	D	100	12
6	C	D	50	6

☑ **Solve the linear programming problem only having the equality constraints(linear indeterminate equation).**

$$2x_1 + y - z - \zeta_1 = 3$$
$$2x_2 + y - z - \zeta_2 = 3$$
$$x_1 + x_2 = 2$$
$$where, x_1, x_2 y, z, \zeta_1, \zeta_2 \geq 0$$

Initial basic feasible solution: $x_1 = x_2 = 1, y = 1, z = 0, \zeta_1 = \zeta_2 = 0$

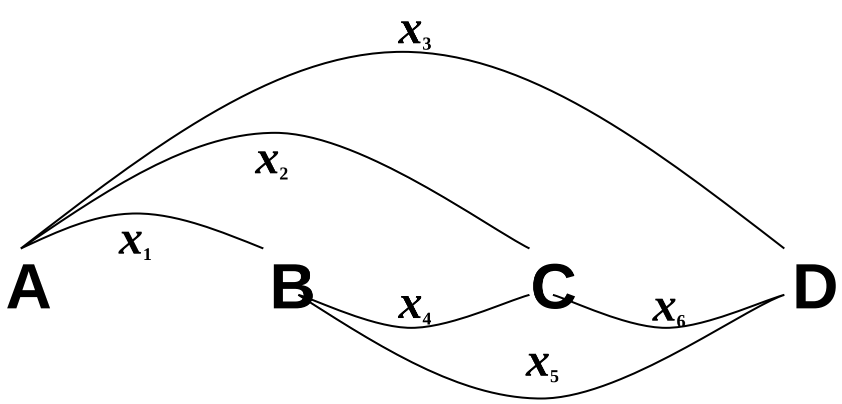
Solution

Consider a cargo ship departing from the port A to E via the ports B, C, D. The maximum cargo loading capacity of the ship is 50,000ton and the loadable cargo at each port is as follows. Formulate and find the optimum cargo transportation that maximizes the freight rate.

Type of cargo	Port of departure	Port of arrival	Loadable cargo at each port of departure (1,000ton)	Freight rate ($/ton)
1	A	B	100	5
2	A	C	40	10
3	A	D	25	20
4	B	C	50	8
5	B	D	100	12
6	C	D	50	6

Type of cargo	Port of departure	Port of arrival	Loadable cargo at the each ports of departure (1,000ton)	Shipping cost rate ($/ton)
1	A	B	100	5
2	A	C	40	10
3	A	D	25	20
4	B	C	50	8
5	B	D	100	12
6	C	D	50	6

The loadable cargo at each port (x_i, i type of cargo) by 1,000ton is as follows.



Design variables: $x_1, x_2 x_3, x_4, x_5, x_6$

Objective function: Maximization of the shipping cost

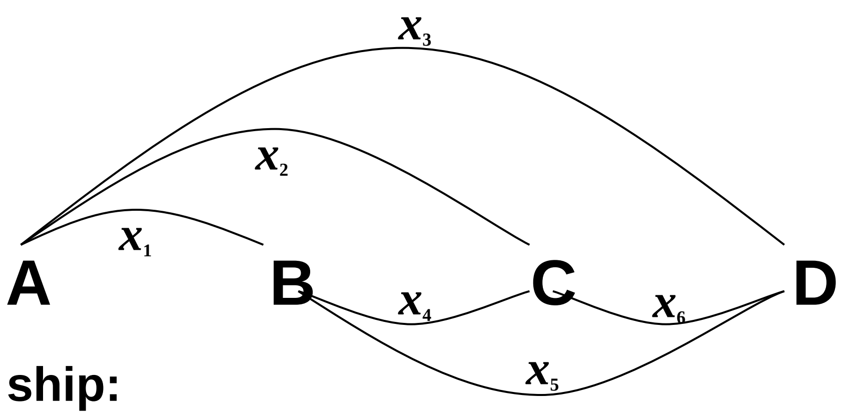
$$\textbf{Maximize } Z = 5x_1 + 10x_2 + 20x_3 + 8x_4 + 12x_5 + 6x_6$$

➡ **The maximization problem should be converted to a minimization problem by assuming *f* = -*Z***

$$\textbf{Minimize } f = -5x_1 - 10x_2 - 20x_3 - 8x_4 - 12x_5 - 6x_6$$

Type of cargo	Port of departure	Port of arrival	Loadable cargo at the each ports of departure (1,000ton)	Shipping cost rate ($/ton)
1	A	B	100	5
2	A	C	40	10
3	A	D	25	20
4	B	C	50	8
5	B	D	100	12
6	C	D	50	6

The loadable cargo at each port (x_i, i type of cargo) by 1,000ton is as follows.



Constraints:

The maximum cargo to be loaded in the ship:

$$A \Rightarrow B : x_1 + x_2 + x_3 \le 50 \qquad B \Rightarrow C : x_2 + x_3 + x_4 + x_5 \le 50$$

$$C \Rightarrow D : x_3 + x_5 + x_6 \le 50$$

The maximum cargo according to the type:

$$0 \le x_2 \le 40,\ 0 \le x_3 \le 25,\ 0 \le x_4 \le 50,\ 0 \le x_6 \le 50$$

The maximum loadable cargoes x_1, x_5 are larger than 50,000 ton, there are no upper limit related with x_1, x_5.

The maximum loadable cargoes x_4, x_6 are 50,000 ton, there are no upper limit related with x_4, x_6.

- [Example 1] Optimal Transportation of Cargo – Solution (3)

Find $x_1, x_2, x_3, x_4, x_5, x_6$

Minimize $f = -5x_1 - 10x_2 - 20x_3 - 8x_4 - 12x_5 - 6x_6$

Subject to

$$\left.\begin{aligned} & x_1 + x_2 + x_3 \le 50 \\ & x_2 + x_3 + x_4 + x_5 \le 50 \\ & x_3 + x_5 + x_6 \le 50 \end{aligned}\right\}$$

: Constraints related with the maximum cargo to be loaded in the ship

$$\left.\begin{aligned} & 0 \le x_2 \le 40,\ 0 \le x_3 \le 25, \\ & 0 \le x_4 \le 50,\ 0 \le x_6 \le 50 \end{aligned}\right\}$$

: Constraints related with the maximum cargo according to the type:

➡ **Optimization problem having the 6 unknown variables and 7 inequality constraints**

Constraints ① Convert to the standard form → ② Solve this problem by using the Simplex method →

$$x_1 + x_2 + x_3 \le 50$$
$$x_2 + x_3 + x_4 + x_5 \le 50$$
$$x_3 + x_5 + x_6 \le 50$$
$$0 \le x_2 \le 40,\ 0 \le x_3 \le 25,$$
$$0 \le x_4 \le 50,\ 0 \le x_6 \le 50$$

Objective function

$$f = -5x_1 - 10x_2 - 20x_3 - 8x_4 - 12x_5 - 6x_6$$

$$x_1 + x_2 + x_3 + x_7 = 50$$
$$x_2 + x_3 + x_4 + x_5 + x_8 = 50$$
$$x_3 + x_5 + x_6 + x_9 = 50$$
$$x_2 + x_{10} = 40,\ x_3 + x_{11} = 25,$$
$$x_4 + x_{12} = 50,\ x_6 + x_{13} = 50$$

Where, $x_7, x_8, x_9, x_{10}, x_{11}, x_{12} x_{13}$ **: slack variables**[1]

$$f = -5x_1 - 10x_2 - 20x_3 - 8x_4 - 12x_5 - 6x_6$$

→ ③

Perform the Simplex method.

starts at the initial basic feasible solution and finds the optimal solution by improving the objective function

1: Slack variable – The variables introduced for converting "≤" type inequality constraints.

6.3 Solution of Linear Programming Problem Using Simplex Method

- [Example 1] Optimal Transportation of Cargo – Solution (5)

positive ratio = $\dfrac{\text{Right hand side parameter in each column}}{\text{Positive coefficient of the element in the selected row}}$

1

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x7	1	1	1	0	0	0	1	0	0	0	0	0	0	50	50
x8	0	1	1	1	1	0	0	1	0	0	0	0	0	50	50
x9	0	0	1	0	1	1	0	0	1	0	0	0	0	50	50
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	-
x11	0	0	**1**	0	0	0	0	0	0	0	1	0	0	25	**25**
x12	0	0	0	1	0	0	0	0	0	0	0	1	0	50	-
x13	0	0	0	0	0	1	0	0	0	0	0	0	1	50	-
Obj.	-5	-10	**-20**	-8	-12	-6	0	0	0	0	0	0	0	f+0	-

Select the variable whose coefficient is positive and row has the smallest positive ratio in the constraints.

(1) Select the column which has the minimum coefficient of the objective function.

(3) Pivot on the selected variable(x_3 / 5 row, 3 column).

2

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x7	1	1	0	0	0	0	1	0	0	0	-1	0	0	25	-
x8	0	1	0	1	**1**	0	0	1	0	0	-1	0	0	25	**25**
x9	0	0	0	0	1	1	0	0	1	0	-1	0	0	25	25
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	-
x3	0	0	1	0	0	0	0	0	0	0	1	0	0	25	-
x12	0	0	0	1	0	0	0	0	0	0	0	1	0	50	-
x13	0	0	0	0	0	1	0	0	0	0	0	0	1	50	-
Obj.	-5	-10	0	-8	**-12**	-6	0	0	0	0	20	0	0	f+500	-

6.3 Solution of Linear Programming Problem Using Simplex Method
- [Example 1] Optimal Transportation of Cargo – Solution (6)

3

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x7	1	1	0	0	0	0	1	0	0	0	-1	0	0	25	-
x5	0	1	0	1	1	0	0	1	0	0	-1	0	0	25	-
x9	0	-1	0	-1	0	**1**	0	-1	1	0	0	0	0	0	**0**
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	-
x3	0	0	1	0	0	0	0	0	0	0	1	0	0	25	-
x12	0	0	0	1	0	0	0	0	0	0	0	1	0	50	-
x13	0	0	0	0	0	1	0	0	0	0	0	0	1	50	50
Obj.	-5	2	0	4	0	**-6**	0	12	0	0	8	0	0	f+800	-

4

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x7	**1**	1	0	0	0	0	1	0	0	0	-1	0	0	25	**25**
x5	0	1	0	1	1	0	0	1	0	0	-1	0	0	25	-
x6	0	-1	0	-1	0	1	0	-1	1	0	0	0	0	0	-
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	-
x3	0	0	1	0	0	0	0	0	0	0	1	0	0	25	-
x12	0	0	0	1	0	0	0	0	0	0	0	1	0	50	-
x13	0	1	0	1	0	0	0	1	-1	0	0	0	1	50	-
Obj.	**-5**	-4	0	-2	0	0	0	6	6	0	8	0	0	800	-

6.3 Solution of Linear Programming Problem Using Simplex Method

- [Example 1] Optimal Transportation of Cargo – Solution (7)

5

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x1	1	1	0	0	0	0	1	0	0	0	-1	0	0	25	
x5	0	1	0	1	1	0	0	1	0	0	-1	0	0	25	25
x6	0	-1	0	-1	0	1	0	-1	1	0	0	0	0	0	
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	
x3	0	0	1	0	0	0	0	0	0	0	1	0	0	25	
x12	0	0	0	1	0	0	0	0	0	0	0	1	0	50	50
x13	0	1	0	1	0	0	0	1	-1	0	0	0	1	50	50
Obj.	0	1	0	-2	0	0	5	6	6	0	3	0	0	f+925	

The row having the negative coefficient (-1) in the selected column is not selected.

6

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	bi	bi/ai
x1	1	1	0	0	0	0	1	0	0	0	-1	0	0	25	
x4	0	1	0	1	1	0	0	1	0	0	-1	0	0	25	
x6	0	0	0	0	1	1	0	0	1	0	-1	0	0	25	
x10	0	1	0	0	0	0	0	0	0	1	0	0	0	40	
x3	0	0	1	0	0	0	0	0	0	0	1	0	0	25	
x12	0	-1	0	0	-1	0	0	-1	0	0	1	1	0	25	
x13	0	0	0	0	-1	0	0	0	-1	0	1	0	1	25	
Obj.	0	3	0	0	2	0	5	8	6	0	1	0	0	f+975	

Because all the coefficients of the objective function are nonnegative, the current solution is the optimal solution(x_2=x_5=0,x_1=x_3=x_4=x_6=25,f=-975)

Therefore, the maximum shipping cost (975,000$) can be achieved by loading 25,000 tons per the cargo type(1, 3, 4, 6).

☑ **Solve the linear programming problem only having the equality constraints(linear indeterminate equation).**

$$2x_1 + y - z - \zeta_1 = 3$$
$$2x_2 + y - z - \zeta_2 = 3$$
$$x_1 + x_2 = 2$$
$$where, x_1, x_2 y, z, \zeta_1, \zeta_2 \geq 0$$

Initial basic feasible solution: $x_1 = x_2 = 1, y = 1, z = 0, \zeta_1 = \zeta_2 = 0$

- [Example 2] Linear Programming Program – Solution (1)

1. The problem is the linear programming problem only having the equality constraints(linear indeterminate equation).

2. To solve this problem, we introduce the artificial variables and artificial objective function to find the initial basic feasible solution in the Simplex method.

$$\mathbf{B}_{(3\times6)}\mathbf{X}_{(6\times1)} + \underbrace{\mathbf{Y}_{(3\times1)}}_{\text{Artificial variable}} = \mathbf{D}_{(3\times1)}$$

3. The artificial objective function is defined as follows.

$$w = \sum_{i=1}^{3} Y_i = \sum_{i=1}^{3} D_i - \sum_{j=1}^{6}\sum_{i=1}^{3} B_{ij} X_j = w_0 + \sum_{j=1}^{6} C_j X_j$$

where $C_j = -\sum_{i=1}^{3} B_{ij}$ **: Sum the all the elements at the j column in Matrix B and change the sign. (Relative objective coefficient)**

$$w_0 = \sum_{i=1}^{3} D_i = 3+3+2=8$$ **: Sum of all the elements in the Matrix D. (Initial basic solution for the artificial objective function)**

$$2x_1 + y - z - \zeta_1 = 3$$
$$2x_2 + y - z - \zeta_2 = 3$$
$$x_1 + x_2 = 2$$
$$where, x_1, x_2 y, z, \zeta_1, \zeta_2 \geq 0$$

$$\mathbf{B}_{(3\times6)}\mathbf{X}_{(6\times1)} + \underline{\mathbf{Y}_{(3\times1)}} = \mathbf{D}_{(3\times1)}$$

Artificial variable

$$\begin{bmatrix} 2 & 0 & 1 & -1 & -1 & 0 \\ 0 & 2 & 1 & -1 & 0 & -1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(=X_1) \\ x_2(=X_2) \\ y(=X_3) \\ z(=X_4) \\ \zeta_1(=X_5) \\ \zeta_2(=X_6) \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$$

1

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
Y1	**2**	0	1	-1	-1	0	1	0	0	3	**3/2**
Y2	0	2	1	-1	0	-1	0	1	0	3	-
Y3	1	1	0	0	0	0	0	0	1	2	2
A. Obj.	**-3**	-3	-2	2	1	1	0	0	0	w-8	-

Artificial objective function

Sum the all the elements at the each column in Matrix B and change the sign. (ex. 1 column: -(2+0+1)=-3)

- [Example 2] Linear Programming Program – Solution (3)

2

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	1/2	-1/2	-1/2	0	1/2	0	0	3/2	-
Y2	0	2	1	-1	0	-1	0	1	0	3	3/2
Y3	0	**1**	-1/2	1/2	1/2	0	-1/2	0	1	1/2	**1/2**
A. Obj.	0	**-3**	-1/2	1/2	-1/2	1	3/2	0	0	w-7/2	-

3

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	1/2	-1/2	-1/2	0	1/2	0	0	3/2	3
Y2	0	0	**2**	-2	-1	-1	1	1	-2	2	**1**
X2	0	1	-1/2	1/2	1/2	0	-1/2	0	1	1/2	-
A. Obj.	0	0	**-2**	2	1	1	0	0	3	w-2	-

4

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	0	0	-1/4	1/4	1/4	-1/4	1/2	1	
X3	0	0	1	-1	-1/2	-1/2	1/2	1/2	-1	1	
X2	0	1	0	0	1/4	-1/4	-1/4	1/4	1/2	1	-
A. Obj.	0	0	0	0	0	0	1	1	1	w-0	-

$$\mathbf{X}^{T}{}_{(1\times 5)} = \begin{bmatrix} x_1 & x_2 & y & z & \zeta_1 & \zeta_2 \end{bmatrix}$$

$\rightarrow$ $\boldsymbol{X_1=1, X_2=1, X_3=1, X_4=X_5=X_6=0}$

Since the value of the artificial objective function becomes zero, the initial basic feasible solution is obtained.

Therefore, one of the initial basic feasible solutions is $x_1 = x_2 = 1, v = y - z = 1, \zeta_1 = \zeta_2 = 0.$

Naval Architecture & Ocean Engineering

Computer Aided Ship design
-Part I. Optimal Ship Design-

Programming Assignment

2011, Fall

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Naval Architecture & Ocean Engineering

Programming Assignment #1

Golden Section Method
Programming Guide

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Programming Assignment #1

Write a program, which is applying the "Golden Section Method" and minimize following functions.

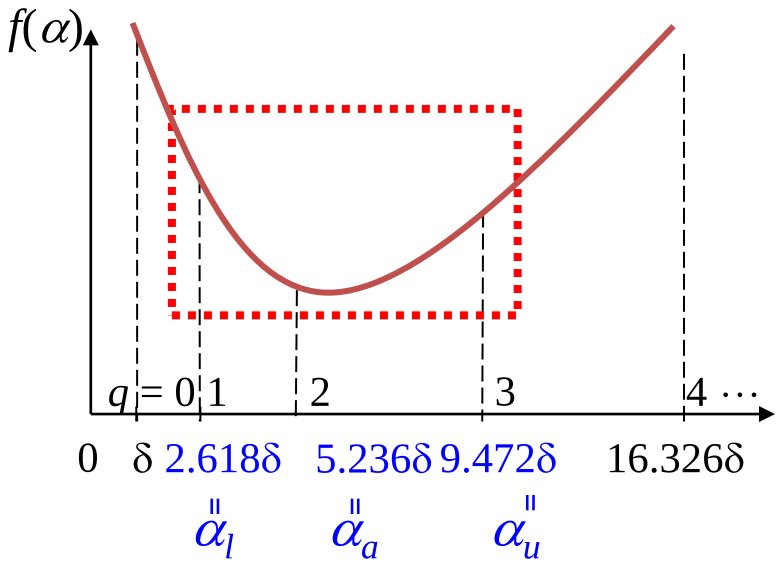
1. $f(x) = x^2$

2. $f(x) = \sin x$

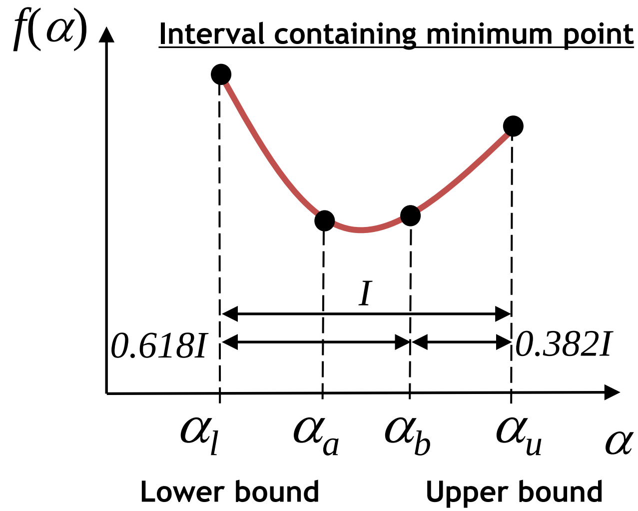
3. $f(x) = x^3 - x^2 + x - 1$

Golden Section Search method Program Guide

Step 1 : Find the interval where the minimum is located



Step 2: Calculate $f(\alpha_a)$ and $f(\alpha_b)$



Golden Section Search method Programming Guide

Step 3: Check in which interval we have the minimum value

① If $f(\alpha_a) < f(\alpha_b)$

Then the optimum point α^* is between α_l and α_b .

The new lower bound is $\alpha_l' = \alpha_l$ and the new $\alpha_b' = \alpha_a$

The upper bound $\alpha_u' = \alpha_b$ and $\alpha_a' = \alpha_l' + 0.382(\alpha_u' - \alpha_l')$.

② If else $f(\alpha_a) > f(\alpha_b)$

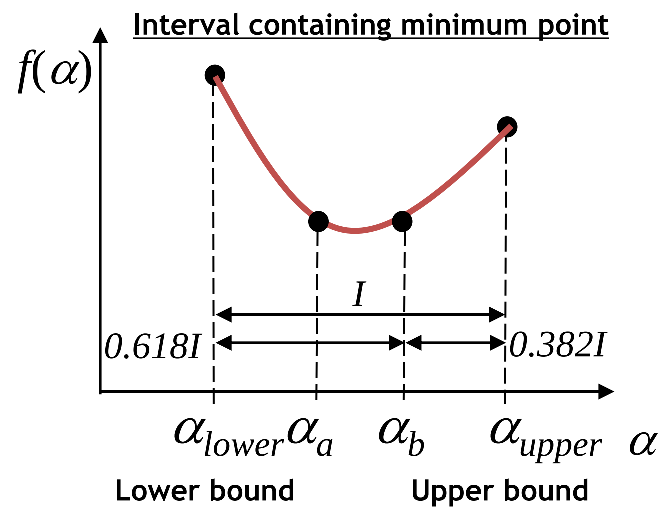
The optimum point α^* is between α_a and α_u .

The new lower bound is $\alpha_l' = \alpha_a$ and the new $\alpha_a' = \alpha_b$

The upper bound $\alpha_u' = \alpha_u$ and $\alpha_b' = \alpha_l' + 0.618(\alpha_u' - \alpha_l')$.

③ If $f(\alpha_a) = f(\alpha_b)$

Then $\alpha_l = \alpha_a, \alpha_u = \alpha_b$

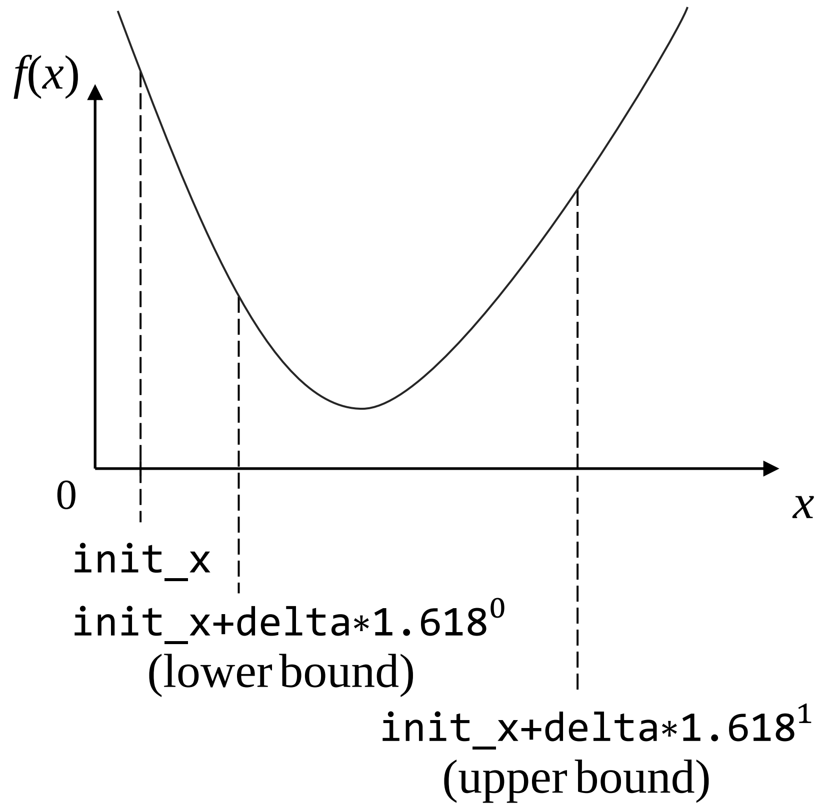


Step 4: Determine if the tolerance is acceptable, if not then enter the loop at step 2

The distance between the points left and right from the minimum value should be smaller than our tolerance $\| \alpha_b - \alpha_a \| < 10^{-6}$

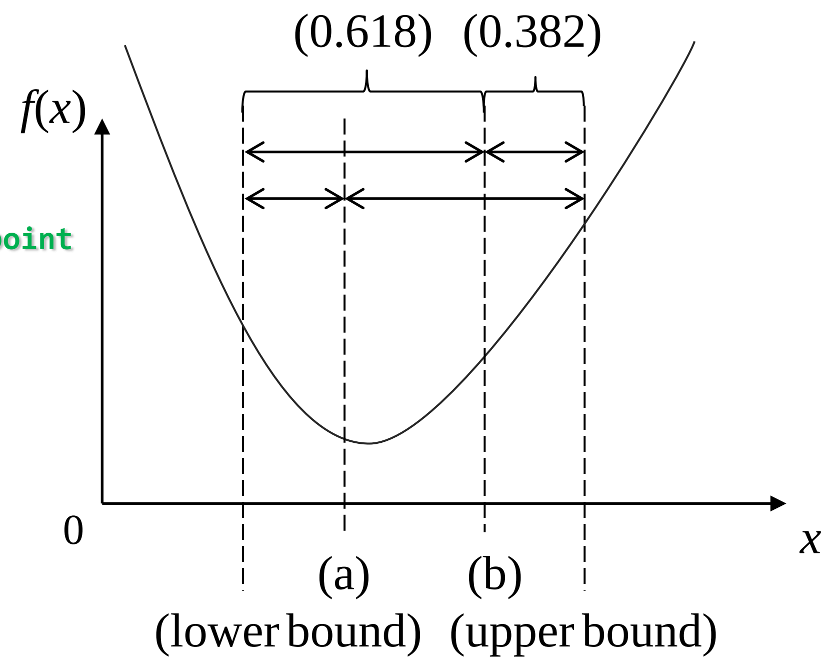
Golden Section Search method Programming Guide

```
#include "stdafx.h“
#include "math.h"

//Function which we want to optimize.

double f1(double x)
    {
          return x*x-4*x+4;
    }

//Main function and variables

int _tmain(int argc, _TCHAR* argv[])
{

    double init_x=0.0;
    double delta = 0.1;
    double final;

    double lower;
    double a;
    double b;
    double upper;
```



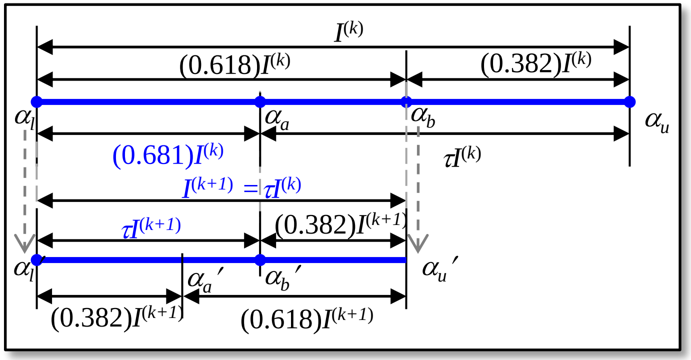
Golden Section Search method Programming Guide

```
//Divide the interval
while(true)
{

        if (f1(init_x)>f1(init_x+delta) && f1(init_x+delta)<f1(init_x+delta*2.618))
        {
                break;
        }
        init_x=init_x+delta;
        delta = 1.618*delta;
}

final = init_x+2.618 * delta;

//Check the interval in which we presume the minimum point

printf("%lf \n", init_x);
printf("%lf \n", final);

//Lower bound, upper bound and point a and b
lower=init_x;
upper=final;
a= ( upper-lower )*0.382+lower;
b= ( upper-lower )*0.618+lower;
```



Golden Section Search method Programming Guide

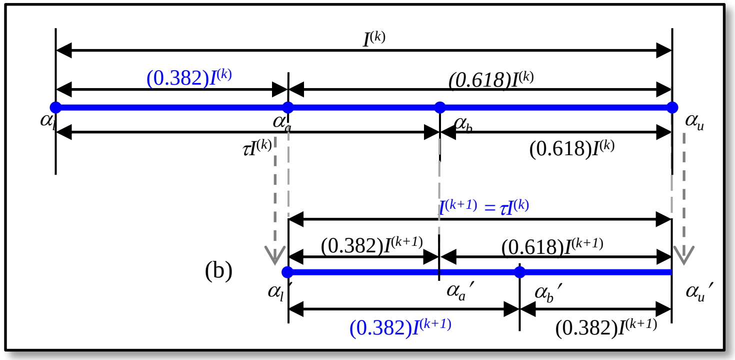
```
while(true)
{
        //If the tolerance is not reached keep executing
        if(fabs(b-a)<0.00000001)
        {
                break;
        }
        //If a is smaller than b, then the minimum point is in the left interval
        else if(f1(a)<f1(b))
        {
                lower=lower;
                upper=b;
                b=a;
                a=lower+(upper-lower)*0.382;
        }
```




```
        //If b is smaller than a, then the minimum point is in the right interval
        else if(f1(a)>f1(b))
        {
                ...
        }

        //If a and b are same, then the minimum point is in the interval between a and b
        else
        {
                ...
        }
}
```

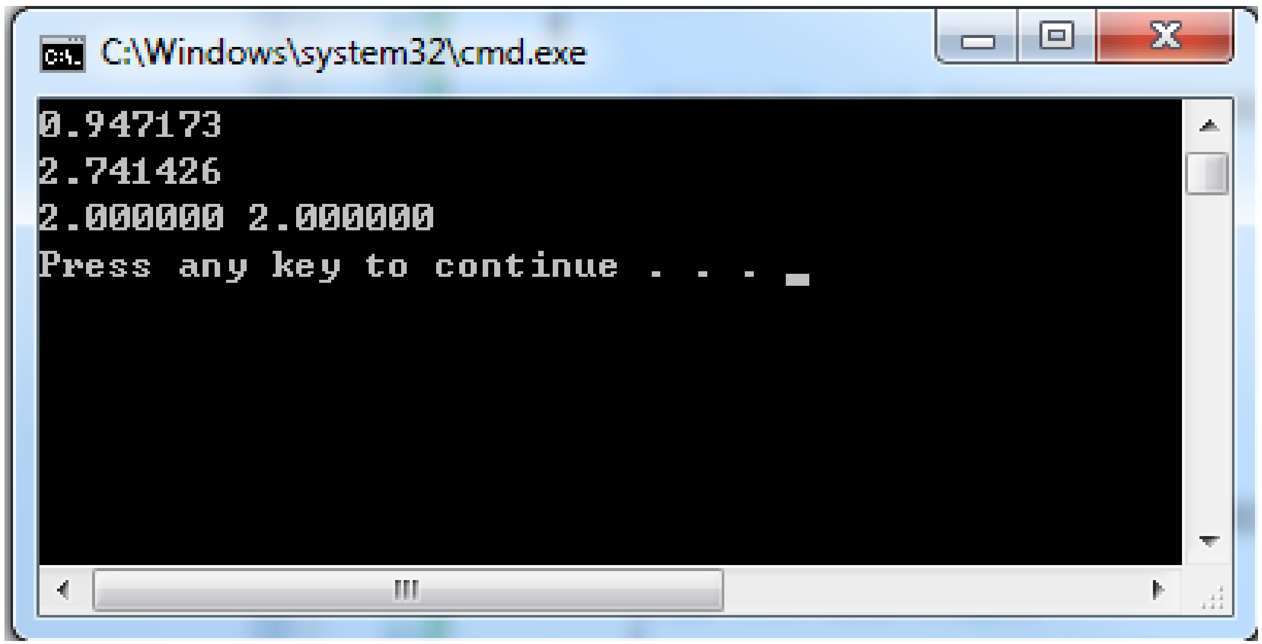
Golden Section Search method Programming Guide

```
while(true)
{
        //If the tolerance is not reached keep executing
        if(fabs(b-a)<0.00000001)
        {
                break;
        }
        //If a is smaller than b, then the minimum point is in the left interval
        else if(f1(a)<f1(b))
        {
                ...
        }
        //If b is smaller than a, then the minimum point is in the right interval
        else if(f1(a)>f1(b))
        {



                What should we define?



        }
        //If a and b are same, then the minimum point is in the interval between a and b
        else
        {
                ...
        }
}
```



Golden Section Search method Programming Guide

```
……………………
……………………
//Print the result
          printf("%lf %lf \n", a,b );
          return 0;
}
```



Naval Architecture & Ocean Engineering

Programming Assignment #2

Simplex Programming Guide

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Simplex Programming Assignment

☑ **Programming for the following optimization problems using Simplex methods**

- **Linear programming problem #1: refer page 13**
- **Linear programming problem #2: refer page 14**
- **Linear programming problem #3: refer page 15**

☑ **Caution**

- **Separate the procedures for minimizing the objective function, and the artificial objective function into two phases**
- **Output the simplex tables during the iteration into the console window or a file.**
- **Find out at list 2 solutions of indeterminate equations by using Roll-Back procedure.**

Linear programming problem #1

$$\text{Maximize} \quad z = 4x_1 + 5x_2$$
$$\text{Subject to} \quad -x_1 + x_2 \le 4$$
$$x_1 + x_2 \le 6$$
$$x_1, x_2 \ge 0$$

Optimal solution: $x_1=1$, $x_2=5$, $x_3=x_4=0$, $f=-29$

		x1	x2	x3	x4	bi	bi/ai
1st row:	x3	**-1**	**1**	**1**	**0**	**4**	**4**
2nd row:	x4	1	1	0	1	6	6
3rd row:	Obj.	-4	**-5**	0	0	f-0	-

		x1	x2	x3	x4	bi	bi/ai
1st row:	X2	-1	1	1	0	4	-4
2nd row:	x4	**2**	0	-1	1	2	**1**
3rd row:	Obj.	**-9**	0	5	0	f+20	-

		x1	x2	x3	x4	bi	bi/ai
1st row:	x2	0	1	0.5	0.5	5	-
2nd row:	x1	1	0	-0.5	0.5	1	-
3rd row:	Obj.	**0**	**0**	**0.5**	**4.5**	**f+29**	-

Linear programming problem #2

Minimize $f = -x_1 - 2x_2 + 2x_3$

Subject to $3x_1 + 2x_2 - 2x_3 \le 12$

$2x_1 + 3x_2 - 3x_3 \ge 6$

$x_1, x_2, x_3 \ge 0$

2) Phase 2: Repeat the pivot operation until all the coefficients of the original objective function f are nonnegative

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	**2/3**	-2/3	8	**12**
x2	2/3	1	-1	0	-1/3	1/3	2	-6
Obj.	1/3	0	0	0	**-2/3**	2/3	f+4	-

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x5	5/2	0	0	3/2	1	-1	**12**	-
x2	3/2	1	-1	1/2	0	0	**6**	-
Obj.	2	0	0	1	0	0	***f*+12**	-

Linear programming problem #3

$$\text{Minimize} \quad f = -x_1 - 2x_2 + 2x_3$$

$$\text{Subject to} \quad 3x_1 + 2x_2 - 2x_3 \le 12$$

$$2x_1 + 3x_2 - 3x_3 \ge 6$$

$$x_1, x_2, x_3 \ge 0$$

1) Phase 1: Repeat the pivot operation until the artificial objective function w becomes zero

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	-
x6	2	3	-3	0	-1	1	6	-
Obj.	-1	-2	2	0	0	0	f-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	6
x6	2	**3**	-3	0	-1	1	6	**2**
Obj.	-1	**-2**	2	0	0	0	f-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	-
x2	2/3	1	-1	0	-1/3	1/3	2	-
Obj.	1/3	0	0	0	-2/3	2/3	f+4	-
A. Obj.	0	0	0	0	0	1	**w-0**	-

2) Phase 2: Repeat the pivot operation until all the coefficients of the original objective function f are nonnegative

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	**2/3**	-2/3	8	**12**
x2	2/3	1	-1	0	-1/3	1/3	2	-6
Obj.	1/3	0	0	0	**-2/3**	2/3	f+4	-

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x5	5/2	0	0	3/2	1	-1	**12**	-
x2	3/2	1	-1	1/2	0	0	**6**	-
Obj.	2	0	0	1	0	0	***f*+12**	-

Linear programming problem #4

Solve the following indeterminate equations

$$2x_1 + y - z - \zeta_1 = 3$$

$$2x_2 + y - z - \zeta_2 = 3$$

$$x_1 + x_2 = 2$$

$$where, x_1, x_2 y, z, \zeta_1, \zeta_2 \geq 0$$

$$x_1 = x_2 = 1, v = y - z = 1, \zeta_1 = \zeta_2 = 0$$

1

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
Y1	**2**	0	1	-1	-1	0	1	0	0	3	**3/2**
Y2	0	2	1	-1	0	-1	0	1	0	3	-
Y3	1	1	0	0	0	0	0	0	1	2	2
A. Obj.	**-3**	-3	-2	2	1	1	0	0	0	w-8	-

2

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	1/2	-1/2	-1/2	0	1/2	0	0	3/2	-
Y2	0	2	1	-1	0	-1	0	1	0	3	3/2
Y3	0	**1**	-1/2	1/2	1/2	0	-1/2	0	1	1/2	**1/2**
A. Obj.	0	**-3**	-1/2	1/2	-1/2	1	3/2	0	0	w-7/2	-

3

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	1/2	-1/2	-1/2	0	1/2	0	0	3/2	3
Y2	0	0	**2**	-2	-1	-1	1	1	-2	2	**1**
X2	0	1	-1/2	1/2	1/2	0	-1/2	0	1	1/2	-
A. Obj.	0	0	**-2**	2	1	1	0	0	3	w-2	-

4

	X1	X2	X3	X4	X5	X6	Y1	Y2	Y3	bi	bi/ai
X1	1	0	0	0	-1/4	1/4	1/4	-1/4	1/2	1	
X3	0	0	1	-1	-1/2	-1/2	1/2	1/2	-1	1	
X2	0	1	0	0	1/4	-1/4	-1/4	1/4	1/2	1	-
A. Obj.	0	0	0	0	0	0	1	1	1	w-0	-

An example of solution for the Linear programming problem #1

$$\text{Maximize} \quad z = 4x_1 + 5x_2$$

$$\text{Subject to} \quad -x_1 + x_2 \le 4$$

$$x_1 + x_2 \le 6$$

$$x_1, x_2 \ge 0$$

Optimal Solution: $x_1=1$, $x_2=5$, $x_3=x_4=0$, $f=-29$

		x1	x2	x3	x4	bi	bi/ai
1st row:	x3	-1	**1**	1	0	4	**4**
2nd row:	x4	1	1	0	1	6	6
3rd row:	Obj.	-4	**-5**	0	0	f-0	-

		x1	x2	x3	x4	bi	bi/ai
1st row:	X2	-1	1	1	0	4	-4
2nd row:	x4	**2**	0	-1	1	2	**1**
3rd row:	Obj.	**-9**	0	5	0	f+20	-

		x1	x2	x3	x4	bi	bi/ai
1st row:	x2	0	1	0.5	0.5	**5**	-
2nd row:	x1	1	0	-0.5	0.5	**1**	-
3rd row:	Obj.	**0**	**0**	**0.5**	**4.5**	**f+29**	-

```
 -1.0000  1.0000  1.0000  0.0000  |   4.0000
  1.0000  1.0000  0.0000  1.0000  |   6.0000
----------------------------------|---------
 -4.0000 -5.0000  0.0000  0.0000  |   0.0000

Row = 1
Col = 2



 -1.0000  1.0000  1.0000  0.0000  |   4.0000
  2.0000  0.0000 -1.0000  1.0000  |   2.0000
----------------------------------|---------
 -9.0000  0.0000  5.0000  0.0000  |  20.0000

Row = 2
Col = 1



  0.0000  1.0000  0.5000  0.5000  |   5.0000
  1.0000  0.0000 -0.5000  0.5000  |   1.0000
----------------------------------|---------
  0.0000  0.0000  0.5000  4.5000  |  29.0000
```

Explanation of Simplex Class

```
class Simplex
{
public:
	Simplex();
	Simplex(const Simplex& rhs);
	virtual ~Simplex();

	//member variables
	Matrix m_matSimplexTable;
	Matrix m_matVar;
	Matrix m_matBiAi;
	int m_nPivotRow,m_nPivotCol;
	int m_nPhase;
	static std::vector<Simplex*> m_vSimplexStack;

	//member function
	void SetSimplexTable(Matrix& m_matSimplexTable);
	void SetPhase(int phase);
	void FindPivotColumn();
	void FindPivotRow();
	void Pivot();
	bool CheckEndCondition();
	void Solve();
};
```

Programming Guide for Simplex Method

1) Construct the Simplex Table using given objective function and constraints

	Basic variable	Nonbasic variable(=0)	Basic variable	
1st row:	x_4	$3x_1 + 2x_2 - 2x_3$	$+x_4$	$=12$
2nd row:	x_6	$2x_1 + 3x_2 - 3x_3$	$-x_5 + x_6$	$=6$
3rd row:		$-x_1 - 2x_2 + 2x_3$		$=f$
4th row:		$-2x_1 - 3x_2 + 3x_3$	$+x_5$	$=w-6$

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	-
x6	2	3	-3	0	-1	1	6	-
Obj.	-1	-2	2	0	0	0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

```
void SetSimplexTable
(Matrix& m_matSimplexTable);
```

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	-
x6						1	6	-
Obj.						0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

```
Matrix
m_matSimplexTable;
```

```
Matrix
m_matVar;
```

```
Matrix
m_biaiTable;
```

1) Construct the Simplex Table using given objective function and constraints

☑ **Caution for constructing Simplex Table**

1. **Elements in the column "bi" must be nonnegative. If there is negative element, then multiply "-1" to the row on which the negative element is.**

Programming Guide for Simplex Method

2) Phase 1. Minimize the artificial objective function

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	-
x6	2	3	-3	0	-1	1	6	-
Obj.	-1	-2	2	0	0	0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

1) Select the column whose element is the most negative value in the last row

void FindPivotColumn();

2) Select the row whose bi/ai is the smallest nonnegative value.

void FindPivotRow();

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	3	2	-2	1	0	0	12	6
x6	2	3	-3	0	-1	1	6	2
Obj.	-1	-2	2	0	0	0	*f*-0	-
A. Obj.	-2	-3	3	0	1	0	w-6	-

※ Caution for pivot operation

- **The row whose bi/ai is zero should be candidate for selecting row.**
- **Round off Error**

  Wrong example: if (x==0)
  Right example: if (fabs(x) < 10e-6)

- **Roll Back Function**

  When the column, whose element is most negative value in the last row, is selected, if several columns have same most negative element:
  → Save the matrix and pivot point.

  And it is same when the row is selected.

- **When all of the elements in the last row are nonnegative and w is not zero, go back to the matrix which is saved by Roll Back function.**

Programming Guide for Simplex Method

3) An example of function "FindPivotCol()"

```
void Simplex::FindPivotCol()
{
    int i = 0;                  // Initialize index for "iteration"
    double val = 0.0;           // Initialize variable to compare the coefficients of the
                                // objective function

    // Input the row number, which store the coefficients of the objective function
    int nRow = m_matSimplexTable.GetNumOfRows() - 1;

    // Select the column whose element is most negative value in the last row
    for (i=0; i<m_matSimplexTable.GetNumOfCols()-1; i++)
    {
        if (m_matSimplexTable.GetElement(nRow, i) < val)
        {
            val = m_matSimplexTable[nRow][i]; // save the most negative value
            m_nPivotCol = i;                  // save the index for most negative value
        }
    }
}
```

※ Implementation of Roll Back

1. Find the most negative value in the last row using "void FindPivotColumn();"
2. If several columns have same most negative element,
3. Then, save the simplex table into the variable "m_SimplexChild" with the pivot column.

```
std::vector<Simplex*> m_vSimplexStack;      // Initialize the stack to save simplex tables
……


for(int i=0;i<NumOfColumn;i++)
{
    double element = m_matSimplexTable.GetElement(nRow, i)

    if(fabs(val - element) < 10e-6)
    {
        // Copy this simplex table to the temporary variable "temp"
        Simplex* temp = new Simplex(*this);

        // Save the povot column
        temp->m_nPivotCol = i;

        // Save this simplex table
        m_SimplexChild.push_back(temp);
    }
}
```

Programming Guide for Simplex Method

5) End condition of "Phase 1"

① End condition of "Phase 1"

1. If all of the element in the last row are nonnegative and w is not zero

2. Then, start "Phase 2"

3. Else, go back to the matrix which is saved by Roll Back function and carry out the pivot operation for "Phase 1"

```
if (x==0) (X)
if (fabs(x) < 10e-6) (0)
```

② Phase 2

✓ **Since the artificial objective function is not used anymore, eliminate the last row.**

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	2/3	-2/3	8	-
x2	2/3	1	-1	0	-1/3	1/3	2	-
Obj.	1/3	0	0	0	-2/3	2/3	f+4	-
A. Obj.	0	0	0	0	0	1	w-0	-

✓ **carry out pivot operation for "Phase 2"**

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x4	5/3	0	0	1	**2/3**	-2/3	8	**12**
x2	2/3	1	-1	0	**-1/3**	1/3	2	-6
Obj.	1/3	0	0	0	**-2/3**	2/3	f+4	-

Programming Guide for Simplex Method
6) End condition of "Phase 2"

If the all element of the last row., i.e., the coefficients of the objective function, are nonnegative, then the current solution is the optimal solution.
→ Stop the simplex, and print out the result.

	x1	x2	x3	x4	x5	x6	bi	bi/ai
x5	5/2	0	0	3/2	1	-1	**12**	-
x2	3/2	1	-1	1/2	0	0	**6**	-
Obj.	2	0	0	1	0	0	***f+12***	-

Because all the coefficients of the objective function are nonnegative, the current solution is the optimal solution.
$(x_1=x_3=x_4=0, x_2=6, x_5=12, f=-12)$

Programming Guide for Simplex Method

7) An example of iteration procedure for "Phase 1"

※ Implementation of the function "Solve"

```
bool Simplex::Solve()
{
    while(1)
    {
        if (m_pivotColumn == -1)
            FindPivotColumn();
        if (m_pivotColumn == -1)
        {
            Simplex temp = m_SimplexChild[m_SimplexChild.size()-1]
            m_SimplexChild.pop_back();
            return temp.Solve();
        }
        if (m_pivotRow == -1)
            FindPivotRow();
        if (m_pivotRow == -1)
        {
            //Same as above
        }
        Pivot();
        if(CheckEndCondition())
            return true;
        if (m_NumOfInteration >= 100 && m_SimplexChild.size() > 0)
        {
            //Same as above
        }
    }
    return true;
}
```

All values of the last row, coefficients of the objective function, are positive

Go back to the matrix which is saved by Roll Back function

All values of bi/ai are negative

to prevent the infinite iteration

Programming Guide for Simplex Method

- How to use 'vector' library for implementation of Roll Back function

※ vector의 사용

1. definition:

```
#include <vector>
using namespace std;
...
std::vector<int> a;
std::vector<Simplex*> m_SimplexChild;
```

2. Member functions:

push_back(...) : save a variable

pop_back() : delete the variable which is saved at last

size() : the number of variables which are saved

3. examples :

```
std::vector<int> a;
a.push_back(1);
a.push_back(2);
int b = a.size();
a.pop_back();
b = a.size();
```

An example for use of Vector Library #1

```
#include <vector>
#include <iostream>
#include <string>
using namespace std;

void main()
{
  vector<string> sV;              // Declare a new vector
  sV.push_back("This");           // Adds an element to the end
  sV.push_back("is");
  sV.push_back("a");
  sV.push_back("test");

  for(vector<string>::iterator p=sV.begin(); p < sV.end(); ++p)
    cout << *p << endl;

}
```

An example for use of Vector Library #2

```
#include <algorithm>
#include <vector>
#include <iostream>
using namespace std;

int main()
{
        vector<char> vec;
        vec.push_back( 'e');
        vec.push_back( 'b');
        vec.push_back( 'a');
        vec.push_back( 'd');
        vec.push_back( 'c');

        sort( vec.begin(), vec.end() ); // sort the variables using "sort()"

        // print out the results.
        cout << "After sorting vector□n";
        for(vector<char>::iterator it= vec.begin(); it != vec.end(); ++it)
                cout << *it;

        return 0;
}
```

Naval Architecture & Ocean Engineering

Computer Aided Ship Design

Part I. Optimization Method

Ch.7 Constrained Nonlinear Optimization Method

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.7 Constrained Nonlinear Optimization Method

7.1 Quadratic Programming(QP)

[Review]4.3 : Finding the optimal solution for the quadratic objective function with linear inequality constraints problem by using the Kuhn-Tucker Necessary Condition, where xi are nonnegative (1)

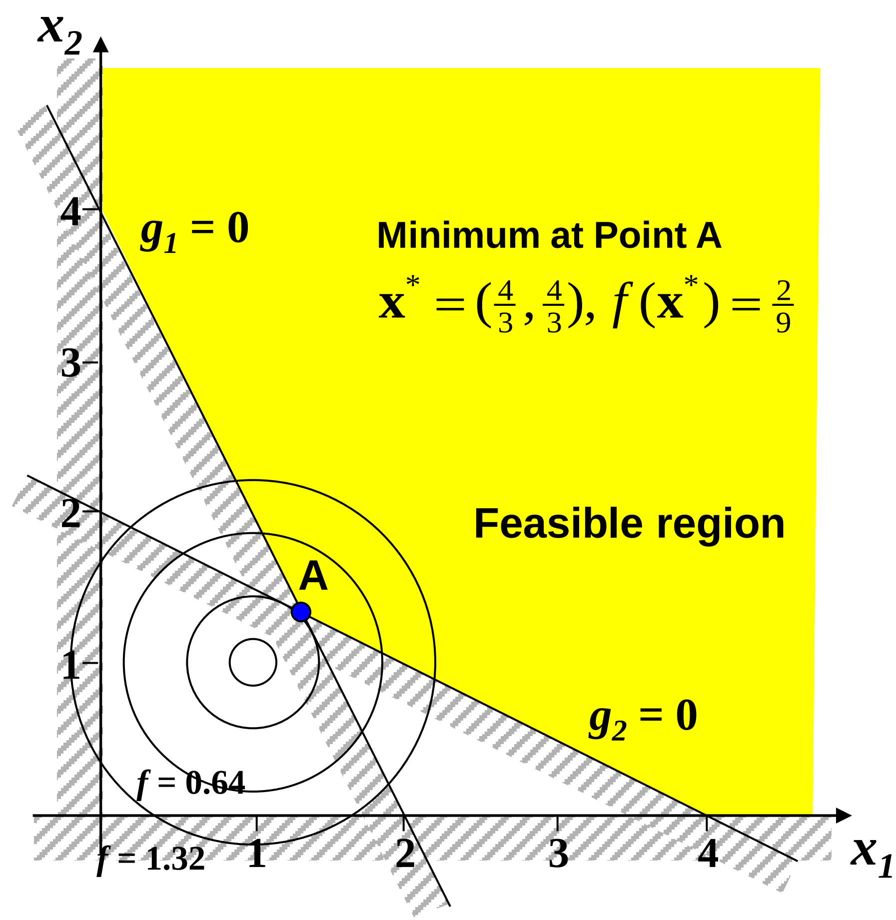
$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$
$$x_1 \ge 0, x_2 \ge 0$$

Minimum point: $\mathbf{x}^* = (\frac{4}{3}, \frac{4}{3}),\ f(\mathbf{x}^*) = \frac{2}{9}$

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 \le 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 \le 0$$
$$-x_1 \le 0, -x_2 \le 0$$

Inequality constraints are transformed to equality constraints by introducing the slack variable

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$\textit{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 + s_1^2 = 0$$
$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 + s_2^2 = 0$$
$$-x_1 + \delta_1^2 = 0, -x_2 + \delta_2^2 = 0$$



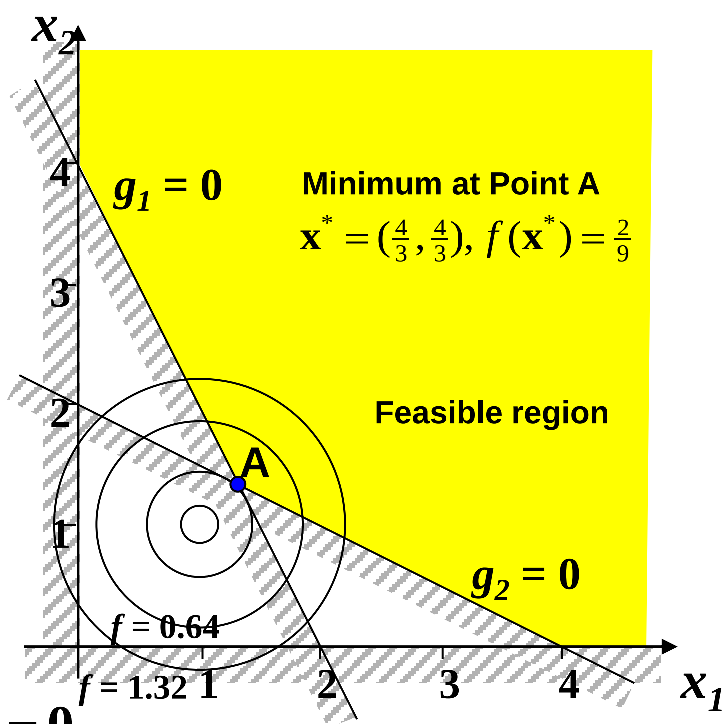
[Review]4.3 : Finding the optimal solution for the quadratic objective function with linear inequality constraints problem by using the Kuhn-Tucker Necessary Condition, where xi are nonnegative (2)

$$\textbf{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$

$$\textbf{Subject to} \quad g_1(\mathbf{x}) = -2x_1 - x_2 + 4 + s_1^2 = 0$$

$$g_2(\mathbf{x}) = -x_1 - 2x_2 + 4 + s_2^2 = 0$$

$$-x_1 + \delta_1^2 = 0, -x_2 + \delta_2^2 = 0$$



Lagrange function

$$L(\mathbf{x}, \mathbf{u}, \mathbf{s}, \boldsymbol{\zeta}, \boldsymbol{\delta}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2$$
$$+u_1(-2x_1 - x_2 + 4 + s_1^2)$$
$$+u_2(-x_1 - 2x_2 + 4 + s_2^2)$$
$$+\zeta_1(-x_1 + \delta_1^2) + \zeta_2(-x_2 + \delta_2^2)$$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x}, \mathbf{u}, \mathbf{s}, \boldsymbol{\zeta}, \boldsymbol{\delta}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 - \zeta_1 = 0 \qquad \frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 - \zeta_2 = 0$$

$$\frac{\partial L}{\partial u_1} = -2x_1 - x_2 + 4 + s_1^2 = 0 \qquad \frac{\partial L}{\partial u_2} = -x_1 - 2x_2 + 4 + s_1^2 = 0$$

$$\frac{\partial L}{\partial s_1} = 2u_1 s_1 = 0 \qquad \frac{\partial L}{\partial s_2} = 2u_2 s_2 = 0$$

$$\frac{\partial L}{\partial \zeta_1} = \delta_1^2 - x_1 = 0 \qquad \frac{\partial L}{\partial \zeta_2} = \delta_2^2 - x_2 = 0 \qquad \frac{\partial L}{\partial \delta_1} = 2\zeta_1\delta_1 = 0 \qquad \frac{\partial L}{\partial \delta_1} = 2\zeta_1\delta_1 = 0$$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta})=\mathbf{0}$

$$\frac{\partial L}{\partial x_1}=2x_1-2-2u_1-u_2-\zeta_1=0 \qquad \frac{\partial L}{\partial x_2}=2x_2-2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1^2=0 \qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2^2=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1=0 \qquad \frac{\partial L}{\partial s_2}=2u_2s_2=0$$

$$\frac{\partial L}{\partial \zeta_1}=\delta_1^2-x_1=0 \rightarrow \delta_1^2=x_1 \qquad \frac{\partial L}{\partial \zeta_2}=\delta_2^2-x_2=0 \rightarrow \delta_2^2=x_2$$

Substitute **Substitute**

$$\frac{\partial L}{\partial \delta_1}=2\zeta_1\delta_1=0 \rightarrow 2\zeta_1\delta_1^2=0 \qquad \frac{\partial L}{\partial \delta_2}=2\zeta_1\delta_1=0 \rightarrow 2\zeta_2\delta_2^2=0$$

Multiply both sides by δ_1 Multiply both sides by δ_2

$$u_i,\zeta_i \ge 0,\ i=1,2$$

We **eliminate two variables** δ_1 , δ_2 and two equations.

Reformulated Kuhn-Tucker necessary condition:

$$\frac{\partial L}{\partial x_1}=2x_1-2-2u_1-u_2-\zeta_1=0 \qquad \frac{\partial L}{\partial x_2}=2x_2-2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1^2=0 \qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2^2=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1=0 \qquad \frac{\partial L}{\partial s_2}=2u_2s_2=0$$

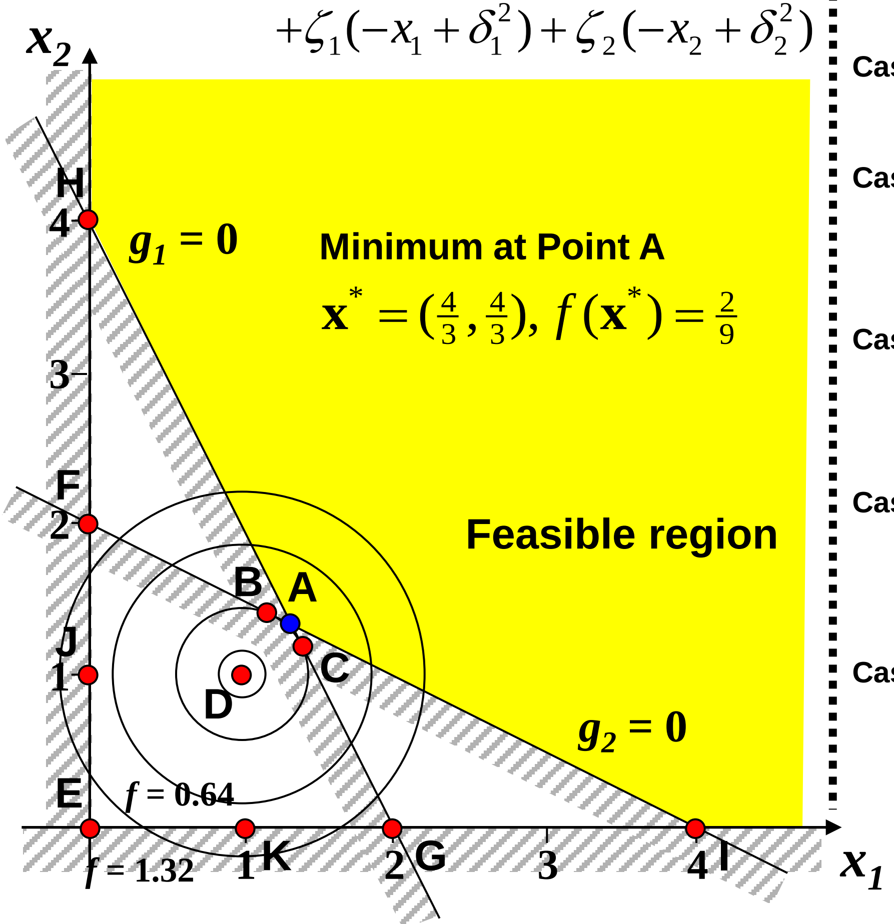
$$2\zeta_1x_1=0 \qquad 2\zeta_2x_2=0$$

$$u_i,\zeta_i,\delta_i \ge 0,\ i=1,2$$

[Review]4.3 : Finding the optimal solution for the quadratic objective function with linear inequality constraints problem by using the Kuhn-Tucker Necessary Condition, where xi are nonnegative (4)

Lagrange function

$$L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta}) = x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 + u_1(-2x_1 - x_2 + 4 + s_1^2) + u_2(-x_1 - 2x_2 + 4 + s_2^2) + \zeta_1(-x_1 + \delta_1^2) + \zeta_2(-x_2 + \delta_2^2)$$



Case #1: $s_1=s_2=\zeta_1=\zeta_2=0$, (Point A)

$x_1 = x_2 = \frac{4}{3}, u_1 = u_2 = \frac{2}{9}$

Case #2: $u_1=s_2=\zeta_1=\zeta_2=0$, (Point B)

$x_1 = \frac{6}{5}, x_2 = \frac{7}{5}, u_2 = \frac{2}{5}, s_1^2 = -\frac{1}{5}$ It has to be nonnegative.

Case #3: $u_2=s_1=\zeta_1=\zeta_2=0$, (Point C)

$x_1 = \frac{7}{5}, x_2 = \frac{6}{5}, u_1 = \frac{2}{5}, s_2^2 = -\frac{1}{5}$ It has to be nonnegative.

Case #4: $u_1=u_2=\zeta_1=\zeta_2=0$, (Point D)

$x_1 = x_2 = 1, s_1^2 = s_2^2 = -1$ It has to be nonnegative.

Case #5: $u_1=u_2=x_1=x_2=0$, (Point E)

$x_1 = x_2 = 0, s_1^2 = s_2^2 = -4,$ It has to be nonnegative.

$\zeta_1 = \zeta_2 = -2$

Case #6: $u_1=s_2=x_1=x_2=0$, (Point E)

$x_1 = x_2 = 0, s_1^2 = -4,$ It has to be nonnegative.

$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$ The constraint is violated.

Case #7: $u_2=s_1=x_1=x_2=0$, (Point E)

$x_1 = x_2 = 0, s_2^2 = -4,$ It has to be nonnegative.

$-2x_1 - x_2 + 4 + s_1^2 \neq 0$ The constraint is violated.

Case #8: $s_1=s_2=x_1=x_2=0$, (Point E)

$x_1 = x_2 = 0, -2x_1 - x_2 + 4 + s_1^2 \neq 0,$

$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$ The constraint is violated.

Case #9: $u_1=s_2=\zeta_2=x_1=0$, (Point F)

$x_1 = 0, x_2 = 2, u_2 = 1,$

$s_1^2 = -2, \zeta_1 = -3$ It has to be nonnegative.

Case #10: $u_2=s_1=\zeta_1=x_2=0$, (Point G)

$x_1 = 2, x_2 = 0, u_1 = 1, s_2^2 = -2,$

$\zeta_2 = -3$ It has to be nonnegative.

Case #11: $s_1=s_2=\zeta_1=x_2=0$, (Point G)

$x_1 = 2, x_2 = 0,$ The constraint is violated.

$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$

Case #12: $u_2=s_1=\zeta_2=x_1=0$, (Point H)

$x_1 = 0, x_2 = 4, u_1 = 6,$

$s_2^2 = 4, \zeta_1 = -14$ It has to be nonnegative.

Case #13: $s_1=s_2=\zeta_2=x_1=0$, (Point H)

$x_1 = 0, x_2 = 4,$ The constraint is violated.

$-x_1 - 2x_2 + 4 + s_2^2 \neq 0$

Case #14: $u_1=s_2=\zeta_1=x_2=0$, (Point I)

$x_1 = 4, x_2 = 0, u_2 = 6,$

$s_1^2 = 4, \zeta_2 = -14$ It has to be nonnegative.

Case #15: $u_1=u_2=\zeta_2=x_1=0$, (Point J)

$x_1 = 0, x_2 = 1, s_1^2 = -3,$

$s_2^2 = -2, \zeta_1 = -2$ It has to be nonnegative.

Case #16: $u_1=u_2=\zeta_1=x_2=0$, (Point K)

$x_1 = 1, x_2 = 0, s_1^2 = -2,$

It has to be nonnegative. $s_2^2 = -3, \zeta_2 = -2$

Summary

$$\textit{Minimize} \quad f(\mathbf{x}) = -2x_1 - 2x_2 + x_1^2 + x_2^2 + 2$$

$$\textit{Subject to} \quad \begin{aligned} &-2x_1 - x_2 \le -4 \\ &-x_1 - 2x_2 \le -4 \\ &x_1 \ge 0, x_2 \ge 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} &-2x_1 - x_2 + 4 + s_1^2 = 0 \\ &-x_1 - 2x_2 + 4 + s_2^2 = 0 \\ &-x_1 + \delta_1^2 = 0, \ -x_2 + \delta_2^2 = 0 \end{aligned}$$

Lagrange function

$$\begin{aligned} L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta}) = {} & x_1^2 + x_2^2 - 2x_1 - 2x_2 + 2 \\ & + u_1(-2x_1 - x_2 + 4 + s_1^2) + u_2(-x_1 - 2x_2 + 4 + s_2^2) \\ & + \zeta_1(-x_1 + \delta_1^2) + \zeta_2(-x_2 + \delta_2^2) \end{aligned}$$

where, $u_i, \zeta_i \ge 0$

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta}) = 0$

$$\frac{\partial L}{\partial x_1} = -2 + 2x_1 - 2u_1 - u_2 - \zeta_1 = 0, \qquad \frac{\partial L}{\partial x_2} = -2 + 2x_2 - u_1 - 2u_2 - \zeta_2 = 0$$

$$\frac{\partial L}{\partial u_1} = -2x_1 - x_2 + 4 + s_1^2 = 0, \qquad \frac{\partial L}{\partial u_2} = -x_1 - 2x_2 + 4 + s_2^2 = 0 \qquad \frac{\partial L}{\partial \zeta_1} = -x_1 + \delta_1^2 = 0$$

$$\underbrace{\frac{\partial L}{\partial s_1} = 2u_1 s_1 = 0,}_{\text{Equation ①}} \quad \underbrace{\frac{\partial L}{\partial s_2} = 2u_2 s_2 = 0}_{\text{Equation ②}} \qquad \underbrace{\frac{\partial L}{\partial \delta_1} = 2\zeta_1 \delta_1 = 0,}_{\text{Equation ③}} \quad \underbrace{\frac{\partial L}{\partial \delta_2} = 2\zeta_2 \delta_2 = 0}_{\text{Equation ④}} \qquad \frac{\partial L}{\partial \zeta_2} = -x_2 + \delta_2^2 = 0$$

where $u_i, \zeta_i \ge 0$

Multiply both side of each equation ①, ② by s_1, s_2 **, respectively**

multiply both side of each equation ③, ④ by δ_1, δ_2 **, respectively**

Another Method for solving the equations derived from K.-T. conditions:

Apply the Simplex Algorithm

Apply the Simplex Algorithm

- Eliminate variables using relevant equations and introduce new 'virtual' linear variables x'

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta})=0$

$$\frac{\partial L}{\partial x_1}=-2+2x_1-2u_1-u_2-\zeta_1=0,\qquad \frac{\partial L}{\partial x_2}=-2+2x_2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1^2=0,\qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2^2=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1^2=0,\ \frac{\partial L}{\partial s_2}=2u_2s_2^2=0\qquad \frac{\partial L}{\partial \delta_1}=2\zeta_1\delta_1^2=0,\ \frac{\partial L}{\partial \delta_2}=2\zeta_2\delta_2^2=0$$

Substitute

$$\boxed{\frac{\partial L}{\partial \zeta_1}=-x_1+\delta_1^2=0 \qquad \frac{\partial L}{\partial \zeta_2}=-x_2+\delta_2^2=0}$$

-> redefine $\mathbf{s}_i^2$ *as* $\mathbf{s}_i'$

where $u_i,\zeta_i,s_i'\geq 0$

Reformulated Kuhn-Tucker necessary condition:

$$\frac{\partial L}{\partial x_1}=-2+2x_1-2u_1-u_2-\zeta_1=0,\qquad \frac{\partial L}{\partial x_2}=-2+2x_2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1'=0,\qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2'=0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1'=0,\ \frac{\partial L}{\partial s_2}=2u_2s_2'=0\qquad \frac{\partial L}{\partial \delta_1}=2\zeta_1x_1=0,\ \frac{\partial L}{\partial \delta_2}=2\zeta_2x_2=0$$

where $u_i,\zeta_i,s_i',x_i\geq 0$

We eliminate two variables using relevant two equations and also introduce new variable s' instead of s.

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

<u>**Kuhn-Tucker necessary condition:**</u> $\nabla L(\mathbf{x},\mathbf{u},\mathbf{s},\boldsymbol{\zeta},\boldsymbol{\delta})=0$

Linear indeterminate equations

$$\frac{\partial L}{\partial x_1}=-2+2x_1-2u_1-u_2-\zeta_1=0,\qquad \frac{\partial L}{\partial x_2}=-2+2x_2-u_1-2u_2-\zeta_2=0$$

$$\frac{\partial L}{\partial u_1}=-2x_1-x_2+4+s_1'\ =0,\qquad \frac{\partial L}{\partial u_2}=-x_1-2x_2+4+s_2'\ =0$$

$$\frac{\partial L}{\partial s_1}=2u_1s_1'\ =0,\ \frac{\partial L}{\partial s_2}=2u_2s_2'\ =0\qquad \frac{\partial L}{\partial \delta_1}=2\zeta_1x_1=0,\ \frac{\partial L}{\partial \delta_2}=2\zeta_2x_2=0$$

where $u_i,\zeta_i,s_i',x_i\geq 0$

Solve the linear indeterminate equations by using the Simplex Algorithm of phase 1 .

Define the standard LP problem

$$\begin{bmatrix}2&0&-2&-1&-1&0&0&0\\0&2&-1&-2&0&-1&0&0\\-2&-1&0&0&0&0&1&0\\-1&-2&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\u_1\\u_2\\\zeta_1\\\zeta_2\\s_1'\\s_2'\end{bmatrix}=\begin{bmatrix}2\\2\\-4\\-4\end{bmatrix}\ \Rightarrow\ \begin{bmatrix}2&0&-2&-1&-1&0&0&0\\0&2&-1&-2&0&-1&0&0\\2&1&0&0&0&0&-1&0\\1&2&0&0&0&0&0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\u_1\\u_2\\\zeta_1\\\zeta_2\\s_1'\\s_2'\end{bmatrix}=\begin{bmatrix}2\\2\\4\\4\end{bmatrix}$$

Multiply both side of the constraints by -1

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

Introduce the artificial variables to treat the linear equality constraints

$$\begin{bmatrix} 2 & 0 & -2 & -1 & -1 & 0 & 0 & 0 \\ 0 & 2 & -1 & -2 & 0 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(=X_1) \\ x_2(=X_2) \\ u_1(=X_3) \\ u_2(=X_4) \\ \zeta_1(=X_5) \\ \zeta_2(=X_6) \\ s_1'(=X_7) \\ s_2'(=X_8) \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 4 \end{bmatrix}$$

Define the artificial objective function as sum of all the artificial variables (Y_1+Y_2+Y_3+Y_4)

$$5x_1 + 5x_2 - 3u_1 - 3u_2 - \zeta_1 - \zeta_2 - s_1' - s_2' + \underbrace{Y_1 + Y_2 + Y_3 + Y_4}_{w} = 12$$

$$-5x_1 - 5x_2 + 3u_1 + 3u_2 + \zeta_1 + \zeta_2 + s_1' + s_2' = w - 12$$ **: Artificial objective function**

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

$$\begin{bmatrix} 2 & 0 & -2 & -1 & -1 & 0 & 0 & 0 \\ 0 & 2 & -1 & -2 & 0 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(=X_1) \\ x_2(=X_2) \\ u_1(=X_3) \\ u_2(=X_4) \\ \zeta_1(=X_5) \\ \zeta_2(=X_6) \\ s_1'(=X_7) \\ s_2'(=X_8) \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 4 \end{bmatrix}$$

$$-5x_1 - 5x_2 + 3u_1 + 3u_2 + \zeta_1 + \zeta_2 + s_1' + s_2' = w - 12$$ **: Artificial objective function**

1

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
Y1	**2**	0	-2	-1	-1	0	0	0	1	0	0	0	2	**1**
Y2	0	2	-1	-2	0	-1	0	0	0	1	0	0	2	-
Y3	2	1	0	0	0	0	-1	0	0	0	1	0	4	2
Y4	1	2	0	0	0	0	0	-1	0	0	0	1	4	4
A. Obj.	**-5**	-5	3	3	1	1	1	1	0	0	0	0	w-12	-

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

2

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-1	-1/2	-1/2	0	0	0	1/2	0	0	0	1	-
Y2	0	**2**	-1	-2	0	-1	0	0	0	1	0	0	2	**1**
Y3	0	1	2	1	1	0	-1	0	-1	0	1	0	2	2
Y4	0	2	1	1/2	1/2	0	0	-1	-1/2	0	0	1	3	3/2
A. Obj.	0	**-5**	-2	1/2	-3/2	1	1	1	5/2	0	0	0	w-7	-

3

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-1	-1/2	-1/2	0	0	0	1/2	0	0	0	1	-
X2	0	1	-1/2	-1	0	-1/2	0	0	0	1/2	0	0	1	-
Y3	0	0	**5/2**	2	1	1/2	-1	0	-1	-1/2	1	0	1	**2/5**
Y4	0	0	2	5/2	1/2	1	0	-1	-1/2	-1	0	1	1	1/2
A. Obj.	0	0	**-9/2**	-9/2	-3/2	-3/2	1	1	5/2	5/2	0	0	w-2	-

4

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	0	3/10	-1/10	1/5	-2/5	0	1/10	-1/5	2/5	0	7/5	14/3
X2	0	1	0	-3/5	1/5	-2/5	-1/5	0	-1/5	2/5	1/5	0	6/5	-
X3	0	0	1	4/5	2/5	1/5	-2/5	0	-2/5	-1/5	2/5	0	2/5	1/2
Y4	0	0	0	**9/10**	-3/10	3/5	4/5	-1	3/10	-3/5	-4/5	1	1/5	**2/9**
A. Obj.	0	0	0	**-9/10**	3/10	-3/5	-4/5	1	7/10	8/5	9/5	0	w-1/5	-

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

5

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	0	0	0	0	-2/3	1/3	0	0	2/3	-1/3	4/3	-
X2	0	1	0	0	0	0	7/15	-2/3	2/5	0	-7/15	2/15	4/3	-
X3	0	0	1	0	2/3	-1/3	-10/9	8/9	-2/3	7/15	10/9	-8/45	2/9	-
X4	0	0	0	1	-1/3	2/3	8/9	-10/9	1/3	-2/3	-8/9	2/9	2/9	-
A. Obj.	0	0	0	0	0	0	0	0	1	1	1	1	w-0	-

Since the value of the objective function becomes zero, the initial basic feasible solution is obtained.

$$\mathbf{X}^T{}_{(1\times 8)} = \begin{bmatrix} x_1 & x_2 & u_1 & u_2 & \zeta_1 & \zeta_2 & s_1{}' & s_2{}' \end{bmatrix}$$

The one of the initial basic feasible solutions is $X_1 = X_2 = 4/3$, $X_3 = X_4 = 2/9$, $X_5 = X_6 = X_7 = X_8 = 0$.

$$x_1 = x_2 = \tfrac{4}{3}, u_1 = u_2 = \tfrac{2}{9}, \zeta_1 = \zeta_2 = s_1{}' = s_2{}' = 0$$

And this solution satisfy the all nonlinear indeterminate equation(constraints)

$$u_2 s_2' = 0, \quad u_1 s_1' = 0, \quad \zeta_1 x_1 = 0, \quad \zeta_2 x_2 = 0$$

Therefore, the optimal solution of this problem is $x_1 = x_2 = \tfrac{4}{3}, u_1 = u_2 = \tfrac{2}{9}, \zeta_1 = \zeta_2 = s_1' = s_2' = 0.$

This result is same with the method which solves the nonlinear indeterminate equation at first.

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

If we choose the second column whose coefficient of the objective function is same with the first column of that as the pivot column in the first table, what will happen?

1

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
Y1	2	0	-2	-1	-1	0	0	0	1	0	0	0	2	-
Y2	0	2	-1	-2	0	-1	0	0	0	1	0	0	2	1
Y3	2	1	0	0	0	0	-1	0	0	0	1	0	4	4
Y4	1	2	0	0	0	0	0	-1	0	0	0	1	4	2
A. Obj.	-5	-5	3	3	1	1	1	1	0	0	0	0	w-12	-

2

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
Y1	2	0	-2	-1	-1	0	0	0	1	0	0	0	2	1
X2	0	1	-1/2	-1	0	-1/2	0	0	0	1/2	0	0	1	-
Y3	2	0	1/2	1	0	1/2	-1	0	0	-1/2	1	0	3	3/2
Y4	1	0	1	2	0	1	0	-1	0	-1	0	1	2	2
A. Obj.	-5	0	1/2	-2	1	-3/2	1	1	0	5/2	0	0	w-7	-

3

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-1	-1/2	-1/2	0	0	0	1/2	0	0	0	1	-
X2	0	1	-1/2	-1	0	-1/2	0	0	0	1/2	0	0	1	-
Y3	0	0	5/2	2	1	1/2	-1	0	-1	-1/2	1	0	1	2/5
Y4	0	0	2	5/2	1/2	1	0	-1	-1/2	-1	0	1	1	1/2
A. Obj.	0	0	-9/2	-9/2	-3/2	-3/2	1	1	5/2	5/2	0	0	w-2	-

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

*** -9/10 is selected originally, but select -9/5.**

4

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	0	3/10	-1/10	1/5	-2/5	0	1/10	-1/5	2/5	0	7/5	-
X2	0	1	0	-6/10	1/5	-2/5	-1/5	0	-1/5	2/5	1/5	0	6/5	-
X3	0	0	1	4/5	2/5	1/5	-2/5	0	-2/5	-1/5	2/5	0	2/5	-
Y4	0	0	0	9/10	-3/10	3/5	4/5	-1	3/10	-3/5	-4/5	1	1/5	1/4
A. Obj.	0	0	0	-9/10	3/10	-3/5	-4/5	1	7/10	8/5	9/5	0	w-1/5	-

5

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	0	3/4	-1/4	1/2	0	-1/2	-1/4	-1/2	0	1/2	3/2	-
X2	0	1	0	-3/8	1/8	-1/4	0	-1/4	-1/8	1/4	0	1/4	5/4	-
X3	0	0	1	5/4	1/4	1/2	0	-1/2	-1/4	-1/2	0	1/2	1/2	-
X7	0	0	0	9/8	-3/8	3/4	1	-5/4	3/8	-3/4	-1	5/4	1/4	-
A. Obj.	0	0	0	0	0	0	0	0	1	1	1	1	w-0	-

$$\mathbf{X}^T_{(1\times 8)} = \begin{bmatrix} x_1 & x_2 & u_1 & u_2 & \zeta_1 & \zeta_2 & s_1' & s_2' \end{bmatrix}$$

Since the value of the objective function becomes zero, the initial basic feasible solution is obtained.

The another initial basic feasible solution is X_1=3/2, X_2=5/4, X_3=1/2, X_4=X_5=X_6=0, X_7=1/4, X_8=0.

$$x_1 = \tfrac{2}{3}, x_2 = \tfrac{5}{4}, u_1 = \tfrac{1}{2}, u_2 = \zeta_1 = \zeta_2 = 0, s_1' = \tfrac{1}{4}, s_2' = 0$$

But this solution does not satisfy the constraint($u_1 s_1' = 0$).

Therefore, this solution cannot be the optimal solution.

- **When the smallest(i.e., the most negative) coefficient of the artificial objective function or the smallest positive ratio "b_i/a_i" appears more than one entry, the initial basic feasible solution can be changed depending on the selection of the pivot element in the pivot operation.**
- **We have to check whether the solution obtained by the Simplex algorithm satisfies the nonlinear equation. (constraint, u_i*s_i'=0).**

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

In the tableau 3, if we choose the column 4 as a pivot column which has the same coefficient of the artificial objective function(column 3), what will happen?

3

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-1	-1/2	-1/2	0	0	0	1/2	0	0	0	1	-
X2	0	1	-1/2	-1	0	-1/2	0	0	0	1/2	0	0	1	-
Y3	0	0	5/2	2	1	1/2	-1	0	-1	-1/2	1	0	1	1/2
Y4	0	0	2	5/2	1/2	1	0	-1	-1/2	-1	0	1	1	5/2
A. Obj.	0	0	-9/2	-9/2	-3/2	-3/2	1	1	5/2	5/2	0	0	w-2	-

4

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-6/10	0	-2/5	1/5	0	-1/5	2/5	-1/5	0	1/5	6/5	-
X2	0	1	3/10	0	1/5	-1/10	0	-2/5	-1/5	1/10	0	2/5	7/5	-
Y3	0	0	9/10	0	3/5	-3/10	-1	4/5	-3/5	3/10	1	-4/5	1/5	1/4
X4	0	0	4/5	1	1/5	2/5	0	-2/5	-1/5	-2/5	0	2/5	2/5	-
A. Obj.	0	0	-9/10	0	-3/5	3/10	1	-4/5	8/5	7/10	0	9/5	w-1/5	-

5

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-3/8	0	-1/4	1/8	-1/4	0	1/4	-1/8	1/4	0	5/4	-
X2	0	1	3/4	0	1/2	-1/4	-1/2	0	-1/2	-1/4	1/2	0	3/2	-
X8	0	0	9/8	0	3/4	-3/8	-5/4	1	3/4	3/8	5/4	-1	1/4	-
X4	0	0	5/4	1	1/2	1/4	-1/2	0	-1/2	-1/4	1/2	0	1/2	-
A. Obj.	0	0	0	0	0	0	0	0	1	1	1	1	w-0	-

Apply the Simplex Algorithm of phase 1 for solving linear indeterminate equations

5

	X1	X2	X3	X4	X5	X6	X7	X8	Y1	Y2	Y3	Y4	bi	bi/ai
X1	1	0	-3/8	0	-1/4	1/8	-1/4	0	1/4	-1/8	1/4	0	5/4	-
X2	0	1	3/4	0	1/2	-1/4	-1/2	0	-1/2	-1/4	1/2	0	3/2	-
X8	0	0	9/8	0	3/4	-3/8	-5/4	1	3/4	3/8	5/4	-1	1/4	-
X4	0	0	5/4	1	1/2	1/4	-1/2	0	-1/2	-1/4	1/2	0	1/2	-
A. Obj.	0	0	0	0	0	0	0	0	1	1	1	1	w-0	-

Since the value of the objective function becomes zero, the initial basic feasible solution is obtained.

$$\mathbf{X}^T{}_{(1\times 8)} = \begin{bmatrix} x_1 & x_2 & u_1 & u_2 & \zeta_1 & \zeta_2 & s_1 & s_2 \end{bmatrix}$$

The another initial basic feasible solution is X_1=5/4, X_2=3/2, X_3=0, X_4=1/2, X_5=X_6=0=X_7=0, X_8=1/4.

$$x_1 = \tfrac{4}{3}, x_2 = \tfrac{5}{4}, u_1 = 0, u_2 = \tfrac{1}{2}, \zeta_1 = \zeta_2 = s_1' = 0, s_2' = \tfrac{1}{4}$$

But this solution does not satisfy the constraint($u_2 s_2' = 0$).

Therefore, this solution cannot be the optimal solution.

[Ref] Decomposition of the Unrestricted Variable into the Difference of Two Nonnegative Variables (1)

For using the Simplex method, the variables have to be nonnegative in the LP problem.

We can use the Simplex method only for the case that all the variables are nonnegative at the optimal point.

The variables unrestricted in sign at the optimal point should be decomposed into the difference of two nonnegative variables in the LP problem.

Minimize $z = -y_1 - 2y_2$

Subject to $3y_1 + 2y_2 \le 12$

$2y_1 + 3y_2 \ge 6$

$y_1 \ge 0$

y_2 **is unrestricted in sign.**

$\boxed{y_2 = y_2^+ - y_2^-}$

$y_2^+, y_2^- \ge 0$

$\longrightarrow$

Minimize $f = -y_1 \boxed{-2y_2^+ + 2y_2^-}$

Subject to $3y_1 + \boxed{2y_2^+ - 2y_2^-} \le 12$

$2y_1 + \boxed{3y_2^+ - 3y_2^-} \ge 6$

$y_1, y_2^+, y_2^- \ge 0$

Since y_2 is free in sign, it should be decomposed into the difference of two nonnegative variables.

For "≤" type inequality constraint, we Introduce the slack variable.

For "≥" type inequality constraint, we Introduce the surplus variable and the artificial variable.

$\longrightarrow$

Minimize $f = -y_1 \boxed{-2y_2^+ + 2y_2^-}$

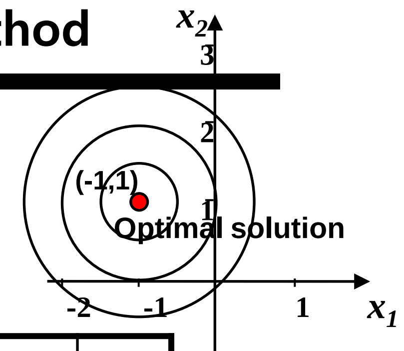
Subject to $3y_1 + \boxed{2y_2^+ - 2y_2^-} + \boxed{x_1} = 12$

$2y_1 + \boxed{3y_2^+ - 3y_2^-} - \boxed{x_2} + \boxed{x_3} = 6$

$y_1, y_2^+, y_2^- \ge 0,\ x_i \ge 0; i = 1\ to\ 3$

$\longrightarrow$ **Solve the problem by using the Simplex method.**

[Reference] Reason to Decompose the Unrestricted Variable into the Difference of Two Nonnegative Design Variables for Using the Simplex Method



Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$ **: Quadratic objective function**

Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$

Definition of the Lagrange function

$$L(x_1, x_2) = x_1^2 + x_2^2 + 2x_1 - 2x_2$$

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2 = 0 \quad \text{---} ①'$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 = 0 \quad \text{---} ②'$$

We try to solve this problem by using the Simplex method

$$-2x_1 = 2$$
$$2x_2 = 2$$

Since these constraints are the equality constraints, we must introduce artificial variables for the equality constraints and define an auxiliary minimization LP problem, and solve it .

$$-2x_1 + y_1 = 2 \quad \text{---}③$$
$$2x_2 + y_2 = 2 \quad \text{---}④$$

Since the artificial objective function is the sum of all the artificial variables, its minimum value must be zero.

Eq. ③+④ $\longrightarrow -2x_1 + 2x_2 + y_1 + y_2 = 4$

$$2x_1 - 2x_2 = \underbrace{y_1 + y_2}_{w} - 4$$

Redefine the variables as $x_1 = X_1, x_2 = X_2, y_1 = Y_1, y_2 = Y_2$ **and express in Matrix from.**

Basic variable	X1	X2	Y1	Y2	bi	bi/ai
Y1	**-2**	0	1	0	2	**-**
Y2	0	2	0	1	2	1
A. Obj.	**2**	-2	0	0	w-4	-

Basic variable	X1	X2	Y1	Y2	bi	bi/ai
Y1	**-2**	0	1	0	2	**2**
X2	0	1	0	1/2	1	-
A. Obj.	**2**	0	0	1	w-2	-

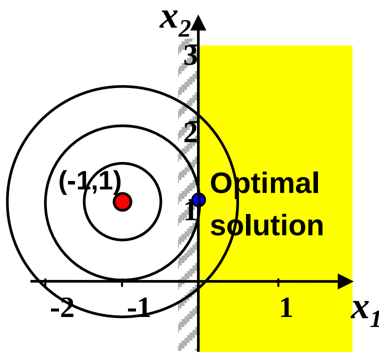
- All coefficients of the artificial objective function become nonnegative.
- However, the sum of all the artificial variables(w) does not become zero.

Stop the simplex

$$x_1 = 0, x_2 = 1, y_1 = 2, y_2 = 0$$

- **The simplex method does not give the optimum solution of x1=-1, x2=1, rather x1=0, x2=1.**
- **The reason is that the simplex method assume all the variables are nonnegative, whereas the variables x1,x2 of this example are free in sign. From this, we can see that to use the simplex method, the unrestricted variables must be decomposed into the two nonnegative variables.**

[Review]



Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$ **: Quadratic objective function**

Subject to $x_1 \ge 0$ **: Linearized inequality constraint**

Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$

Subject to $x_1 \ge 0 \rightarrow -x_1 \le 0 \rightarrow -x_1 + \delta^2 = 0$

Definition of the Lagrange function

$$L(x_1, x_2, \zeta, \delta) = x_1^2 + x_2^2 + 2x_1 - 2x_2 + \zeta(-x_1 + \delta^2)$$

Kuhn-Tucker necessary condition: $\nabla L(x_1, x_2, \zeta, \delta) = 0$

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2 - \zeta = 0 \quad \text{---} \; ①$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 = 0 \quad \text{---} \; ②$$

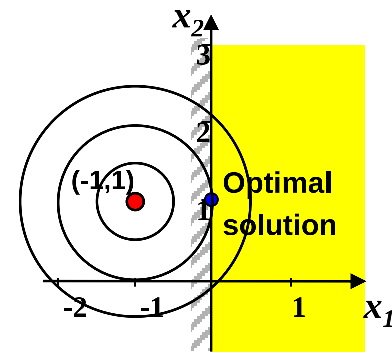
$$\frac{\partial L}{\partial \zeta} = -x_1 + \delta^2 = 0 \quad \text{---} \; ③$$

$$\frac{\partial L}{\partial \delta} = 2\zeta\delta = 0 \quad \text{---} \; ④$$

If we assume $\zeta = 0,\ x_1 = -1$ **→ The equation ③ is not satisfied.**

If we assume $\delta = 0, x_1 = 0, x_2 = 1, \zeta = 2$

Solving the problem by Using the Simplex Method(1/2)



Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$: Quadratic objective function

Subject to $x_1 \geq 0$: Linearized inequality constraint

Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$

Subject to $x_1 \geq 0 \rightarrow -x_1 \leq 0 \rightarrow -x_1 + \delta^2 = 0$

Definition of the Lagrange function

$$L(x_1, x_2, \zeta, \delta) = x_1^2 + x_2^2 + 2x_1 - 2x_2 + \zeta(-x_1 + \delta^2)$$

Kuhn-Tucker necessary condition: $\nabla L(x_1, x_2, \zeta, \delta) = 0$

We try to solve this problem by using the Simplex method.

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2 - \zeta = 0$$
$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 = 0$$
$$\frac{\partial L}{\partial \zeta} = -x_1 + \delta^2 = 0$$
$$\frac{\partial L}{\partial \delta} = 2\zeta\delta = 0$$

$$\boxed{x_2 = x_2^+ - x_2^-}$$
$$x_2^+, x_2^- \geq 0$$

Since x_2 is free in sign, we may decompose it as $x_2 = x_2^+ - x_2^-$

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2 - \zeta = 0 \text{ --- ①}$$
$$\frac{\partial L}{\partial x_2} = \boxed{2x_2^+ - 2x_2^-} - 2 = 0 \text{ – ②}$$
$$\frac{\partial L}{\partial \zeta} = -x_1 + \delta^2 = 0 \text{ --- ③}$$
$$\frac{\partial L}{\partial \delta} = 2\zeta\delta = 0 \text{ --- ④}$$

① → $2x_1 - \zeta = -2$

② → $2x_2^+ - 2x_2^- = 2$

③ → $x_1 = \delta^2$

④ → $2\zeta\delta^2 = 0$

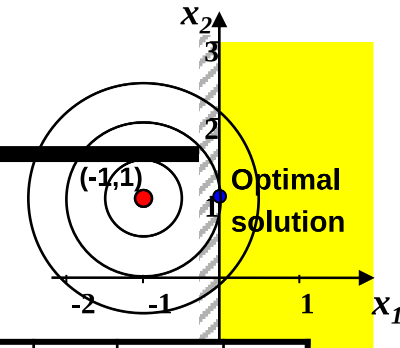
$\zeta x_1 = 0$ --- ⑤

Multiply the both side of equation ④ by δ and substitute the equation ③ into that. We eliminate one variable δ and one equation.

1st stage: Find the solution satisfying the equation ①, ② and ⑤.

2nd stage: Check whether the solution obtained in the 1st stage satisfies the nonlinear equation ⑤ or not.

Solving the problem by Using the Simplex Method(2/2)



Minimize $f(x) = x_1^2 + x_2^2 + 2x_1 - 2x_2$ **: Quadratic objective function**

Subject to $x_1 \ge 0$ **: Linearized inequality constraint**

$$2x_1 - \zeta = -2$$
$$2x_2^+ - 2x_2^- = 2$$

The right hand sides of the equations have to be nonnegative.

$$-2x_1 + \zeta = 2$$
$$2x_2^+ - 2x_2^- = 2$$

$$-2x_1 + \zeta = 2$$
$$2x_2^+ - 2x_2^- = 2$$

Since the constraints are the equality constraints, introduce the artificial variables.

$$-2x_1 + \zeta + y_1 = 2$$
$$2x_2^+ - 2x_2^- + y_2 = 2$$

The artificial variables have to be zero.
Since the artificial objective function is the sum of all the artificial variables, its minimum value is clearly zero.

$$-2x_1 + 2x_2^+ - 2x_2^- + \zeta + y_1 + y_2 = 4$$
$$2x_1 - 2x_2^+ + 2x_2^- - \zeta = \underbrace{\boxed{y_1 + y_2}}_{w} - 4$$

Change the variables as $x_1 = X_1, x_2^+ = X_2, x_2^- = X_3, \zeta = X_4, y_1 = Y_1, y_2 = Y_2$
and express these as the Matrix from.

Basic variable	X1	X2	X3	X4	Y1	Y2	bi	bi/ai
Y1	**-2**	0	0	1	1	0	2	-
Y2	0	2	-2	0	0	1	2	1
A. Obj.	**2**	-2	2	-1	0	0	w-4	-

Basic variable	X1	X2	X3	X4	Y1	Y2	bi	bi/ai
Y1	**-2**	0	0	1	1	0	2	-
Y2	0	2	-2	0	0	1	2	1
A. Obj.	**2**	-2	2	-1	0	0	w-4	-

Basic variable	X1	X2	X3	X4	Y1	Y2	bi	bi/ai
Y1	**-2**	0	0	1	1	0	2	**2**
X2	0	1	-1	0	0	1/2	1	-
A. Obj.	**2**	0	0	-1	0	1	w-2	-

Basic variable	X1	X2	X3	X4	Y1	Y2	bi	bi/ai
X4	**-2**	0	0	1	1	0	2	-
X2	0	1	-1	0	0	1/2	1	-
A. Obj.	**0**	0	0	0	1	1	w-0	-

$$X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 2, Y_1 = 0, Y_2 = 0$$
$$x_1 = 0, x_2^+ = 1, x_2^- = 0, \zeta = 2, y_1 = 0, y_2 = 0$$
$$x_2 = x_2^+ - x_2^- = 1 - 0 = 1$$
$$\zeta x_1 = 2 \cdot 0 = 0 \text{ -- ⑤}$$

Since the equation ⑤ is satisfied, this is the solution of this problem.

[Reference] Solution of the Problem Having the Design Variables whose sign is Unrestricted (1/2)

Matrix Form

$$\underbrace{\begin{bmatrix} \mathbf{H}_{(n\times n)} & -\mathbf{H}_{(n\times n)} & \mathbf{A}_{(n\times m)} & \mathbf{0}_{(n\times m)} & \mathbf{N}_{(n\times p)} & -\mathbf{N}_{(n\times p)} \\ \mathbf{A}^T_{(m\times n)} & -\mathbf{A}^T_{(m\times n)} & \mathbf{0}_{(m\times m)} & \mathbf{I}_{(m\times m)} & \mathbf{0}_{(m\times p)} & \mathbf{0}_{(m\times p)} \\ \mathbf{N}^T_{(p\times n)} & -\mathbf{N}^T_{(p\times n)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times p)} & \mathbf{0}_{(p\times p)} \end{bmatrix}}_{=\mathbf{B}_{((n+m+p)\times(2n+2m+2p))}} \underbrace{\begin{bmatrix} \mathbf{d}^+_{(n\times 1)} \\ \mathbf{d}^-_{(n\times 1)} \\ \mathbf{u}_{(m\times 1)} \\ \mathbf{s}'_{(m\times 1)} \\ \mathbf{y}_{(p\times 1)} \\ \mathbf{z}_{(p\times 1)} \end{bmatrix}}_{=\mathbf{X}_{((2n+2m+2p)\times 1)}} = \underbrace{\begin{bmatrix} -\mathbf{c}_{(n\times 1)} \\ \mathbf{b}_{(m\times 1)} \\ \mathbf{e}_{(p\times 1)} \end{bmatrix}}_{=\mathbf{D}_{((n+m+p)\times 1)}}$$

Number of design variable $2n+2m+2p$

Number of equation $n+2m+p$

$$u_i s'_i = 0; i = 1 \text{ to } m$$

$$\mathbf{B}_{((n+m+p)\times(2n+2m+2p))}\mathbf{X}_{((2n+2m+2p)\times 1)} = \mathbf{D}_{((n+m+p)\times 1)}$$

The number of the design variables is the same with that of the equations as $n+2m+p$ in the original problem.

Since the equations $\mathbf{v}_{(p\times 1)} = \mathbf{y}_{(p\times 1)} - \mathbf{z}_{(p\times 1)}$ and $\mathbf{d}_{(n\times 1)} = \mathbf{d}^+_{(n\times 1)} - \mathbf{d}^-_{(n\times 1)}$ are introduced, the number of the design variables is also increased by $n+p$.

The interesting variables v_i and d_i are determined by the equation $\mathbf{v}_{(p\times 1)} = \mathbf{y}_{(p\times 1)} - \mathbf{z}_{(p\times 1)}$.

Example

Equation

$$x+y+z=2$$
$$2x+2y+z=6$$
$$2x+y=5$$

Replace z as $z_1 - z_2$ $(z_1, z_2 \geq 0)$

$$x+y+z_1-z_2=2$$
$$2x+2y+z_1-z_2=6$$
$$2x+y=5$$

After replacing the variable, this problem becomes the indeterminate equation. The value of z_1-z_2 is always -2.

Solution $x=1, y=3, z=-2$ $\qquad$ $x=1, y=3, z_1=0, z_2=2$

[Reference] Solution of the Problem Having the Design Variables whose sign is Unrestricted (2/2)

Example	
Equation	$x+y+z=5$ $2x+3y+z=11$ $xz=0$

Replace z as $z_1 - z_2$ $(z_1, z_2 \geq 0)$ →

$$x+y+z_1-z_2=5$$
$$2x+3y+z_1-z_2=11$$
$$xz=0$$

Case #1

Introduce the artificial variables for using the Simplex method

$$x+y+z+Y_1=5$$
$$2x+3y+z+Y_2=11$$

←

- Number of the design variables: 5
- Number of the linear independent equation: 2

Solve this problem by assuming the three design variables as zero.
Stop the Simplex method, if the sum of all the artificial variables(Y1+Y2) zero.

	(x,	y,	z,	Y1,	Y2)
①	(4,	1,	0,	0,	0)
②	(6,	0,	-1,	0,	0)
③	(0,	3,	2,	0,	0)

Between the solution ① and ③ obtained by using the Simplex method, the final solution has to satisfy the equation $xz=0$.

If the solution whose value of z (z1-z2) is negative is excluded, the solution of the Case #1 is the same with that of the Case #2.

Case #2

Introduce the artificial variables for using the Simplex method

$$x+y+z_1-z_2+Y_1=5$$
$$2x+3y+z_1-z_2+Y_2=11$$

←

- Number of the design variables: 6
- Number of the linear independent equation: 2

Solve this problem by assuming the four design variables as zero.
Stop the Simplex method, if the sum of all the artificial variables(Y1+Y2) zero.

			z=z1-z2			
	(x,	y,	z1,	z2,	Y1,	Y2)
①	(4,	1,	0,	0,	0,	0)
②	(6,	0,	0,	1,	0,	0)
③	(6,	0,	-1,	0,	0,	0)
④	(0,	0,	-,	-,	0,	0)
⑤	(0,	3,	0,	-2,	0,	0)
⑥	(0,	3,	2,	0,	0,	0)

Among the solution ①,② and ⑥ obtained by using the Simplex method, the final solution has to satisfy the equation $xz=0$.

[Ref] Taylor Series Expansion for the Function of Two Variables (Review, 1)

The second-order Taylor series expansion of $f(x_1, x_2)$ **at** (x_1^*, x_2^*)

$$f(x_1,x_2) = f(x_1^*,x_2^*) + \frac{\partial f(x_1^*,x_2^*)}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f(x_1^*,x_2^*)}{\partial x_2}(x_2 - x_2^*)$$

$$+\frac{1}{2}\left(\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1^2}(x_1 - x_1^*)^2 + 2\frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) + \frac{\partial^2 f(x_1^*,x_2^*)}{\partial x_2^2}(x_2 - x_2^*)^2\right) \quad \cdots\cdots ①$$

define: $\mathbf{c} = \nabla f(\mathbf{x}^*),\ \mathbf{d} = (\mathbf{x} - \mathbf{x}^*)$

$$\Rightarrow f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = f(\mathbf{x}^*) + \mathbf{c}^T\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\,\mathbf{d} \quad \cdots\cdots ②$$

7.1 Quadratic Programming(QP)

- Approximate the original problem as a Quadratic Programming Problem

$$\textit{Minimize} \quad f(\mathbf{x}+\Delta\mathbf{x}) \cong f(\mathbf{x}) + \nabla f^T(\mathbf{x})\Delta\mathbf{x} + 0.5\Delta\mathbf{x}^T\mathbf{H}\Delta\mathbf{x}$$

The second-order Taylor series expansion of the objective function

$$\textit{Subject to} \quad h_j(\mathbf{x}+\Delta\mathbf{x}) \cong h_j(\mathbf{x}) + \nabla h_j^T(\mathbf{x})\Delta\mathbf{x} = 0;\ j = 1\ to\ p$$

The first-order(linear) Taylor series expansion of the equality constraints

$$g_j(\mathbf{x}+\Delta\mathbf{x}) \cong g_j(\mathbf{x}) + \nabla g_j^T(\mathbf{x})\Delta\mathbf{x} \le 0;\ j = 1\ to\ m$$

The first-order(linear) Taylor series expansion of the inequality constraints

Define:

$$\bar{f} = f(\mathbf{x}+\Delta\mathbf{x}) - f(\mathbf{x}),\ e_j = -h_j(\mathbf{x}),\ b_j = -g_j(\mathbf{x}),$$

$$c_i = \partial f(\mathbf{x})/\partial x_i,\ n_{ij} = \partial h_j(\mathbf{x})/\partial x_i,\ a_{ij} = \partial g_j(\mathbf{x})/\partial x_i,$$

$$d_i = \Delta x_i$$

<u>Matrix form</u>

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$

: Quadratic objective function

$$\textit{Subject to} \quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$

: Linear equality constraints

$$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$

: Linear inequality constraints

7.1 Quadratic Programming(QP)
- Construction of Lagrange Function

Minimize $\quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$

Subject to $\quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$

$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)} \Rightarrow \mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} - \mathbf{b}_{(m\times 1)} + \mathbf{s}_{(m\times 1)}{}^2 = \mathbf{0}$

Lagrange Function

$$L = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$
$$+\mathbf{u}^T{}_{(1\times m)}(\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} + \mathbf{s}_{(m\times 1)}{}^2 - \mathbf{b}_{(m\times 1)})$$
$$+\mathbf{v}^T{}_{(1\times p)}(\mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} - \mathbf{e}_{(p\times 1)})$$

7.1 Quadratic Programming(QP)
- Apply the K-T Necessary Condition to the Lagrange function

Lagrange Function

$$L(\mathbf{d},\mathbf{v},\mathbf{s},\mathbf{u}) = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$

$$+\mathbf{u}^T{}_{(1\times m)}(\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} + \mathbf{s}_{(m\times 1)}{}^2 - \mathbf{b}_{(m\times 1)})$$

$$+\mathbf{v}^T{}_{(1\times p)}(\mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} - \mathbf{e}_{(p\times 1)})$$

Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d},\mathbf{v},\mathbf{u},\mathbf{s},\mathbf{u}) = \mathbf{0}$

$$\frac{\partial L(\mathbf{d},\mathbf{v},\mathbf{s},\mathbf{u})}{\partial \mathbf{d}_{(n\times 1)}} = \mathbf{c}_{(n\times 1)} + \mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)} + \mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times 1)} + \mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times 1)} = \mathbf{0}$$

$$\frac{\partial L(\mathbf{d},\mathbf{v},\mathbf{s},\mathbf{u})}{\partial \mathbf{v}_{(p\times 1)}} = \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} - \mathbf{e}_{(p\times 1)} = \mathbf{0}$$

$$\frac{\partial L(\mathbf{d},\mathbf{v},\mathbf{s},\mathbf{u})}{\partial \mathbf{u}_{(n\times 1)}} = \mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} + \mathbf{s}_{(m\times 1)}{}^2 - \mathbf{b}_{(m\times 1)} = 0$$

$$\frac{\partial L(\mathbf{d},\mathbf{v},\mathbf{s},\mathbf{u})}{\partial s_i} = u_i s_i = 0, \quad i = 0 \ to\ m$$

7.1 Quadratic Programming(QP)
- Method 1: Direct Solving the Eqs. from the K._T. Conditions

Optimization problem

Minimize $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$

Subject to $h_i(\mathbf{x}) = 0, \quad i = 1,\ldots,p$ **Equality constraint**

$g_i(\mathbf{x}) \le 0, \quad i = 1,\ldots,m$ **Inequality constraint**

Definition of Lagrange function

$$L(\mathbf{x},\mathbf{v},\mathbf{u},\mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) + \sum_{i=1}^{m} u_i (g_i(\mathbf{x}) + s_i^2)$$

v_i : Lagrange multiplier for the equality constraint(It is free in sign.)

u_i : Lagrange multiplier for the inequality constraint(Nonnegative)

s_i : Slack variable transforming an inequality constraint to an equality constraint

---

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x},\mathbf{v},\mathbf{u},\mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0, \qquad j = 1, \ \ldots \ , n$$

$$\frac{\partial L}{\partial v_i} = h_i(\mathbf{x}^*) = 0, \qquad i = 1, \ \ldots \ , p$$

$$\frac{\partial L}{\partial u_i} = g_i(\mathbf{x}^*) + s_i^{*2} = 0, \qquad i = 1, \ \ldots \ , \mathrm{m}$$

Linear indeterminate equations

$$\frac{\partial L}{\partial s_i} = u_i^* s_i^* = 0, \qquad i = 1, \ \ldots \ , m$$

Nonlinear indeterminate equations

$$u_i^* \ge 0, \qquad i = 1, \ \ldots \ , m$$

Method 1.
- Find the solutions which satisfy the nonlinear indeterminate equations.
- Check whether the solutions satisfy the linear indeterminate equations and determine the solution of this problem.
- Human can find the solution of this problem easily by using this method.

7.1 Quadratic Programming(QP)
- Method 2: Formulate the Problem of the K.-T. Condition as a LP problem

Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d},\mathbf{v},\mathbf{u},\mathbf{s})=\mathbf{0}$

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}}=\mathbf{c}_{(n\times1)}+\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)}+\mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)}+\mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)}=\mathbf{0},\quad \frac{\partial L}{\partial \mathbf{v}_{(p\times1)}}=\mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)}-\mathbf{e}_{(p\times1)}=\mathbf{0}$$

$$\frac{\partial L}{\partial s_i}=u_i \boxed{s_i}=0,\ i=0\ to\ m \text{ ----①}\qquad \frac{\partial L}{\partial \mathbf{u}_{(n\times1)}}=\mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)}+\mathbf{s}_{(m\times1)}{}^2-\mathbf{b}_{(m\times1)}=0$$

Multiply s_i **both side of the equation①**

$$u_i s_i=0 \ \Rightarrow\ u_i s_i^2=0$$

Although the equation ① is multiplied by s_i **,**
the solution($u_i=0\ or\ s_i=0$**) is not changed.**

Transform Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d},\mathbf{v},\mathbf{u},\mathbf{s})=\mathbf{0}$

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}}=\mathbf{c}_{(n\times1)}+\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)}+\mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)}+\mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)}=\mathbf{0},\quad \frac{\partial L}{\partial \mathbf{v}_{(p\times1)}}=\mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)}-\mathbf{e}_{(p\times1)}=\mathbf{0}$$

$$\frac{\partial L}{\partial s_i}=u_i \boxed{s_i^2}=0,\ i=0\ to\ m \text{ ----①'}\qquad \frac{\partial L}{\partial \mathbf{u}_{(n\times1)}}=\mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)}+\mathbf{s}_{(m\times1)}{}^2-\mathbf{b}_{(m\times1)}=0$$

7.1 Quadratic Programming(QP)

Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}} = \mathbf{c}_{(n\times1)} + \mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)} + \mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)} + \mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)} = \mathbf{0},\quad \frac{\partial L}{\partial \mathbf{v}_{(p\times1)}} = \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)} - \mathbf{e}_{(p\times1)} = \mathbf{0}$$

$$\frac{\partial L}{\partial s_i} = u_i \boxed{s_i^2} = 0,\ i = 0\ to\ m \text{----}1' \qquad \frac{\partial L}{\partial \mathbf{u}_{(n\times1)}} = \mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)} + \boxed{\mathbf{s}_{(m\times1)}^{\ 2}} - \mathbf{b}_{(m\times1)} = 0$$

Replace s_i^2 with s_i' (where $s_i' \geq 0$)

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}} = \mathbf{c}_{(n\times1)} + \mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)} + \mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)} + \mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)} = \mathbf{0},\quad \frac{\partial L}{\partial \mathbf{v}_{(p\times1)}} = \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)} - \mathbf{e}_{(p\times1)} = \mathbf{0}$$

$$\frac{\partial L}{\partial s_i} = u_i s_i' = 0,\ i = 0\ to\ m$$

Nonlinear indeterminate equations

$$\frac{\partial L}{\partial \mathbf{u}_{(n\times1)}} = \mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)} + \mathbf{s}'_{(m\times1)} - \mathbf{b}_{(m\times1)} = 0$$

Linear indeterminate equations

Check whether the solution obtained from the linear indeterminate equations satisfies the nonlinear indeterminate equations and determine the solution.

Since these equations are linear in the variables $\mathbf{d}, \mathbf{s}', \mathbf{u}, \mathbf{v}$, this problem is a linear programming problem only having the equality constraints.

where $u_i, s_i' \geq 0;\ i = 1\ to\ m$

7.1 Quadratic Programming(QP)

Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d},\mathbf{v},\mathbf{u},\mathbf{s})=\mathbf{0}$

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}}=\mathbf{c}_{(n\times1)}+\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)}+\mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)}+\mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)}=\mathbf{0},\quad \frac{\partial L}{\partial \mathbf{v}_{(p\times1)}}=\mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)}-\mathbf{e}_{(p\times1)}=\mathbf{0}$$

$$\frac{\partial L}{\partial s_i}=u_i s_i'=0,\ i=0\ to\ m$$

Nonlinear indeterminate equations

$$\frac{\partial L}{\partial \mathbf{u}_{(n\times1)}}=\mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)}+\mathbf{s}'_{(m\times1)}-\mathbf{b}_{(m\times1)}=0$$

Linear indeterminate equations

Check whether the solution obtained from the linear indeterminate equations satisfies the nonlinear indeterminate equations and determine the solution.

Since these equations are linear in the variables $\mathbf{d},\mathbf{s}',\mathbf{u},\mathbf{v}$, this problem is a linear programming problem only having the equality constraints.

where $u_i, s_i' \geq 0;\ i=1\,to\ m$

Since the design variables $\mathbf{d}_{(n\times1)}$ are free in sign, we may decompose them as follows to use the Simplex method.

$$\mathbf{d}_{(n\times1)}=\mathbf{d}^+_{(n\times1)}-\mathbf{d}^-_{(n\times1)},\ (d_i^+\geq0,\ d_i^-\geq0; i=1\ to\ n)$$

Also, the Lagrange multipliers $\mathbf{v}_{(p\times1)}$ for the equality constraints are free in sign, we may decompose them as follows to use the Simplex method.

$$\mathbf{v}_{(p\times1)}=\mathbf{y}_{(p\times1)}-\mathbf{z}_{(p\times1)},\ (y_i\geq0,\ z_i\geq0;\ i=1\ \ to\ \ p)$$

7.1 Quadratic Programming(QP)
- Method 2: Simplex Method for Solving Quadratic Programming Problem

Kuhn-Tucker Necessary Condition: $\nabla L(\mathbf{d}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial \mathbf{d}_{(n\times1)}} = \mathbf{c}_{(n\times1)} + \mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times1)} + \mathbf{A}_{(n\times m)}\mathbf{u}_{(m\times1)} + \mathbf{N}_{(n\times p)}\mathbf{v}_{(p\times1)} = \mathbf{0},$$

$$\frac{\partial L}{\partial \mathbf{v}_{(p\times1)}} = \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times1)} - \mathbf{e}_{(p\times1)} = \mathbf{0}$$

$$\frac{\partial L}{\partial s_i} = u_i s_i' = 0,\ i = 0\ to\ m$$

Nonlinear indeterminate equations

$$\frac{\partial L}{\partial \mathbf{u}_{(n\times1)}} = \mathbf{A}^T_{(m\times n)}\mathbf{d}_{(n\times1)} + \mathbf{s}'_{(m\times1)} - \mathbf{b}_{(m\times1)} = 0$$

Linear indeterminate equations

Check whether the solution obtained from the linear indeterminate equations satisfies the nonlinear indeterminate equations and determine the solution.

Since these equations are linear in the variables $\mathbf{d}, \mathbf{s}', \mathbf{u}, \mathbf{v}$, this problem is a linear programming problem only having the equality constraints.

where $u_i, s_i' \geq 0;\ i = 1\, to\ m$

Because $\mathbf{d}$ and $\mathbf{v}$ are free in sign.

$$\mathbf{d}_{(n\times1)} = \mathbf{d}^+_{(n\times1)} - \mathbf{d}^-_{(n\times1)},\ (d_i^+ \geq 0,\ d_i^- \geq 0; i = 1\ to\ n)$$

$$\mathbf{v}_{(p\times1)} = \mathbf{y}_{(p\times1)} - \mathbf{z}_{(p\times1)},\ (y_i \geq 0,\ z_i \geq 0;\ i = 1\ to\ p)$$

Matrix From

$$\begin{bmatrix} \mathbf{H}_{(n\times n)} & -\mathbf{H}_{(n\times n)} & \mathbf{A}_{(n\times m)} & \mathbf{0}_{(n\times m)} & \mathbf{N}_{(n\times p)} & -\mathbf{N}_{(n\times p)} \\ \mathbf{A}^T{}_{(m\times n)} & -\mathbf{A}^T{}_{(m\times n)} & \mathbf{0}_{(m\times m)} & \mathbf{I}_{(m\times m)} & \mathbf{0}_{(m\times p)} & \mathbf{0}_{(m\times p)} \\ \mathbf{N}^T{}_{(p\times n)} & -\mathbf{N}^T{}_{(p\times n)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times p)} & \mathbf{0}_{(p\times p)} \end{bmatrix} \begin{bmatrix} \mathbf{d}^+_{(n\times1)} \\ \mathbf{d}^-_{(n\times1)} \\ \mathbf{u}_{(m\times1)} \\ \mathbf{s}'_{(m\times1)} \\ \mathbf{y}_{(p\times1)} \\ \mathbf{z}_{(p\times1)} \end{bmatrix} = \begin{bmatrix} -\mathbf{c}_{(n\times1)} \\ \mathbf{b}_{(m\times1)} \\ \mathbf{e}_{(p\times1)} \end{bmatrix}$$

Introduce the artificial variables, define the artificial objective function and solve the linear programming problem by using the Simplex method.

7.1 Quadratic Programming(QP)

Matrix From

Introduce the artificial variables, define the artificial objective function and solve the linear programming problem by using the Simplex method.

$$\begin{bmatrix} \mathbf{H}_{(n\times n)} & -\mathbf{H}_{(n\times n)} & \mathbf{A}_{(n\times m)} & \mathbf{0}_{(n\times m)} & \mathbf{N}_{(n\times p)} & -\mathbf{N}_{(n\times p)} \\ \mathbf{A}^T{}_{(m\times n)} & -\mathbf{A}^T{}_{(m\times n)} & \mathbf{0}_{(m\times m)} & \mathbf{I}_{(m\times m)} & \mathbf{0}_{(m\times p)} & \mathbf{0}_{(m\times p)} \\ \mathbf{N}^T{}_{(p\times n)} & -\mathbf{N}^T{}_{(p\times n)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times m)} & \mathbf{0}_{(p\times p)} & \mathbf{0}_{(p\times p)} \end{bmatrix} \begin{bmatrix} \mathbf{d}^+_{(n\times 1)} \\ \mathbf{d}^-_{(n\times 1)} \\ \mathbf{u}_{(m\times 1)} \\ \mathbf{s}'_{(m\times 1)} \\ \mathbf{y}_{(p\times 1)} \\ \mathbf{z}_{(p\times 1)} \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{n+m+p} \end{bmatrix} = \begin{bmatrix} -\mathbf{c}_{(n\times 1)} \\ \mathbf{b}_{(m\times 1)} \\ \mathbf{e}_{(p\times 1)} \end{bmatrix}$$

Artificial variables

How to define the artificial objective function

- Define an one equation by sum of all the equations from 1 row to n+m+p row.
- Define the sum of the all artificial variables($Y_1+Y_2+\ldots+Y_{n+m+p}$) as an objective function(w).

- Determine an initial basic feasible solution for the linear programming problem by using the Simplex method.

- Check whether the initial basic feasible solutions satisfy the following nonlinear indeterminate equations and determine that as a solution.

$$\frac{\partial L}{\partial s_i} = u_i s_i' = 0, \; i = 0 \; to \; m$$

7.1 Quadratic Programming(QP)
- Summary of Method 2 of Simplex Method for Solving Quadratic Programming Problem

Kuhn-Tucker Necessary Condition(Matrix form)

$$\mathbf{B}_{((n+m+p)\times(2n+2m+2p))}\mathbf{X}_{((2n+2m+2p)\times 1)} = \mathbf{D}_{((n+m+p)\times 1)}$$

Simplex Method for Solving Quadratic Programming Problem

1. The problem to solve the Kuhn-Tucker necessary condition is same with the problem having only the equality constraints(linear programming problem).

2. To solve the linear indeterminate equations, we introduce the artificial variables, define the artificial objective function, and determine the initial basic feasible solution by using the Simplex method.

$$\mathbf{B}_{((n+m+p)\times(2n+2m+2p))}\mathbf{X}_{((2n+2m+2p)\times 1)} + \mathbf{Y}_{((n+m+p)\times 1)} = \mathbf{D}_{((n+m+p)\times 1)}$$

If any of the elements in D is(are) negative, the corresponding equation must be multiplied by -1 to have a nonnegative element on the right sid

3. The artificial objective function is defined as follows.

$$w = \sum_{i=1}^{n+m+p} Y_i = \sum_{i=1}^{n+m+p} D_i - \sum_{j=1}^{2(n+m+p)} \sum_{i=1}^{n+m+p} B_{ij} X_j = w_0 + \sum_{j=1}^{2(n+m+p)} C_j X_j$$

where $\underline{C_j = -\sum_{i=1}^{n+m+p} B_{ij}}$, $\underline{w_0 = \sum_{i=1}^{n+m+p} D_i}$ Initial value of the artificial objective function

Add the elements of the j th column of the matrix B and change its sign.(Initial relative objective coefficient).

4. Solve the linear programming problem by using the Simplex and check whether the solution satisfies the following equation.

$u_i s_i' = 0;\ i = 1\ to\ m$: This equation is used to check whether the solution satisfies this equation.

7.1 Quadratic Programming(QP)
- Comparison between Method 1and 2

Optimization problem

Minimize $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$

Subject to $h_i(\mathbf{x}) = 0, \quad i = 1,\ldots,p$ **Equality constraint**

$g_i(\mathbf{x}) = 0, \quad i = 1,\ldots,m$ **Inequality constraint**

Definition of Lagrange function

$$L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{p} v_i h_i(\mathbf{x}) + \sum_{i=1}^{m} u_i (g_i(\mathbf{x}) + s_i^2)$$

v_i : Lagrange multiplier for the equality constraint(It is free in sign.)

u_i : Lagrange multiplier for the inequality constraint(Nonnegative)

s_i : Slack variable transforming an inequality constraint to an equality constraint

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{x}, \mathbf{v}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0, \qquad j = 1, \ \ldots \ , n$$

$$\frac{\partial L}{\partial v_i} = h_i(\mathbf{x}^*) = 0, \qquad i = 1, \ \ldots \ , p$$

$$\frac{\partial L}{\partial u_i} = g_i(\mathbf{x}^*) + s_i^{*2} = 0, \qquad i = 1, \ \ldots \ , \mathrm{m}$$

Linear indeterminate equations

$$\frac{\partial L}{\partial s_i} = u_i^* s_i^* = 0, \qquad i = 1, \ \ldots \ , m$$

Nonlinear indeterminate equations

$$u_i^* \geq 0, \quad i = 1, \ \ldots \ , m$$

Method 1.
- Find the solutions to satisfy the nonlinear indeterminate equations.
- Check whether the solutions satisfy the linear indeterminate equations and determine the solution of this problem.
- Human can find the solution of this problem easily by using this method.

Method 2.
- Find the solutions to satisfy the linear indeterminate equations by using the Simplex method.
- Check whether the solutions satisfy the nonlinear indeterminate equations and determine the solution of this problem.
- Since the algorithm of this method is more systematical, this method is useful for the computational approach.

Naval Architecture & Ocean Engineering

Ch.7 Constrained Nonlinear Optimization Method

7.2 Sequential Linear Programming

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

7.2 Sequential Linear Programming(SLP)

- Define the linear programming(LP) problem by linearizing the objective function and the constraints in the current design point.
- Compute the design change by solving the linear programming problem and obtain the improved design point.

$$\underbrace{\mathbf{x}^{(k+1)}}_{\text{Improved design point}} = \underbrace{\mathbf{x}^{(k)}}_{\text{Current design point}} + \underbrace{\mathbf{d}^{(k)}}_{\text{Design change obtained by solving the LP problem.}}$$

- This method is to find the optimal solution by solving the linear programming problem sequentially.

7.2 Sequential Linear Programming(SLP)
- [Example] Problem with Inequality Constraints (1)

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

$$\textit{Subject to} \quad g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$

$$g_2(\mathbf{x}) = -x_1 \le 0$$

$$g_3(\mathbf{x}) = -x_1 \le 0$$

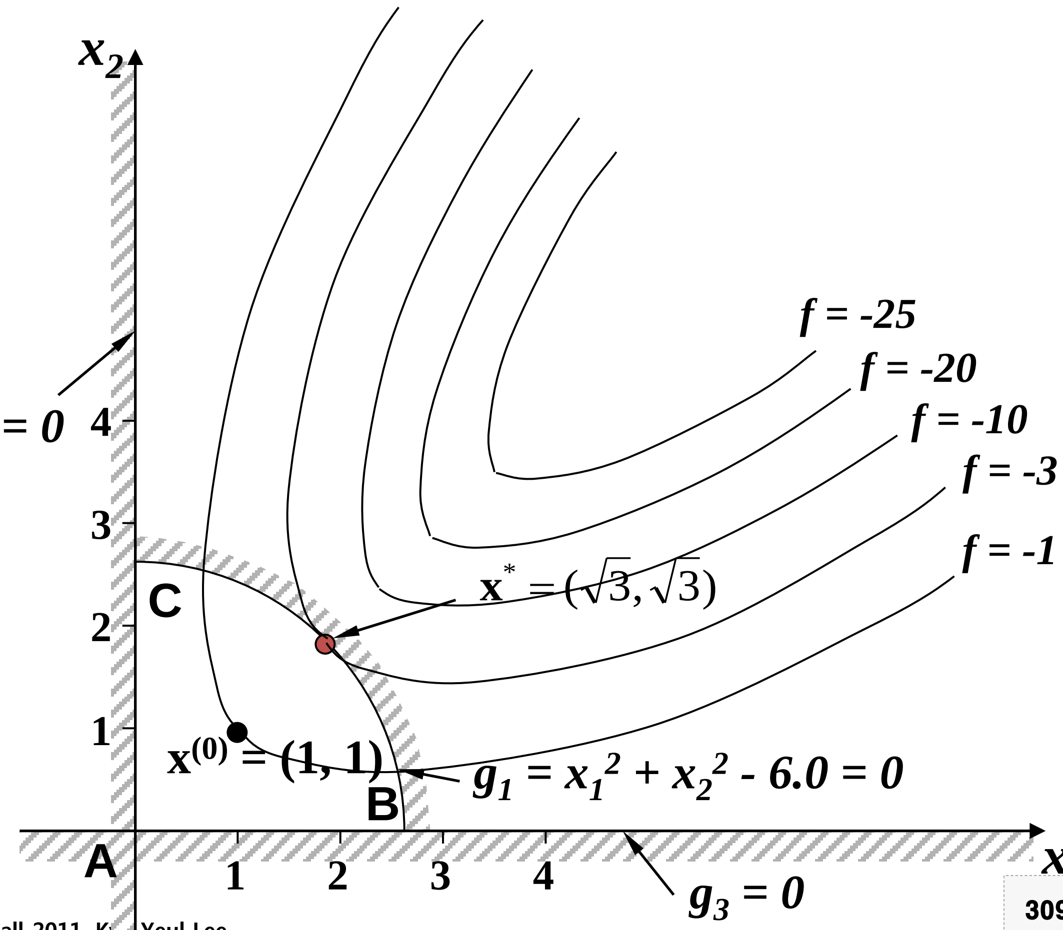
The starting design point: $\mathbf{x}^{(0)} = (1,1)$,

$$\varepsilon_1 = \varepsilon_2 = 0.001$$

Choose move limits such that a 15% design change is permissible.

The optimal solution:

$$\mathbf{x}^* = (\sqrt{3}, \sqrt{3}),\ f(\mathbf{x}^*) = -3$$



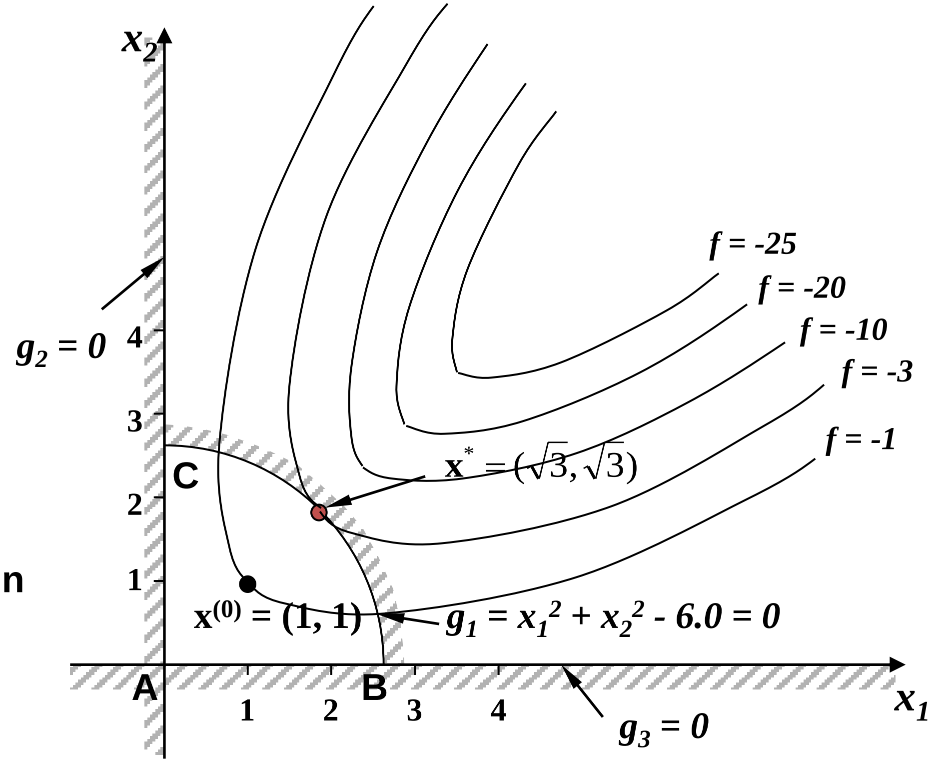
7.2 Sequential Linear Programming(SLP)
- [Example] Problem with Inequality Constraints (2)

$$\textit{Minimize} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

$$\textit{Subject to} \quad g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$

$$g_2(\mathbf{x}) = -x_1 \le 0$$

$$g_3(\mathbf{x}) = -x_1 \le 0$$



(1) Iteration 1(k = 0)

(i) Step 1

From the given point(starting point), the current design point is as follows.

$$\mathbf{x}^{(0)} = (1,1), \varepsilon_1 = \varepsilon_2 = 0.001$$

(ii) Step 2: Evaluate the objective and constraint function at the current design point.

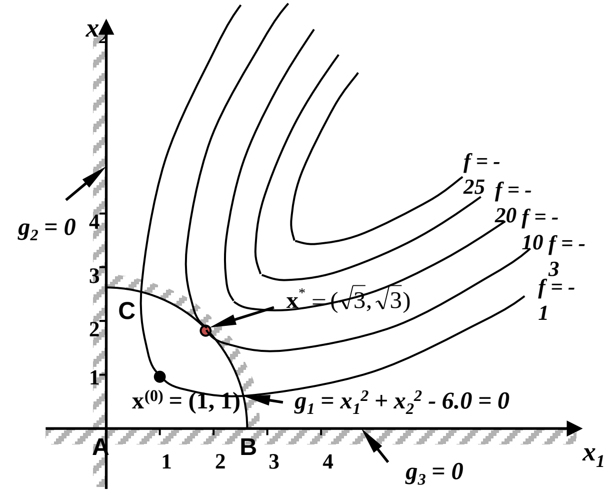
$$f(1,1) = -1$$

$g_1(1,1) = -\frac{2}{3} < 0$ ➡ **Constraint is satisfied.**

$g_2(1,1) = -1 < 0$ ➡ **Constraint is satisfied.**

$g_3(1,1) = -1 < 0$ ➡ **Constraint is satisfied.**

7.2 Sequential Linear Programming(SLP)
- [Example] Problem with Inequality Constraints (3)



$$\textbf{\textit{Minimize}} \quad f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$

$$\textbf{\textit{Subject to}} \quad g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$

$$g_2(\mathbf{x}) = -x_1 \le 0$$

$$g_3(\mathbf{x}) = -x_2 \le 0$$

(1) Iteration 1(k = 0) $\mathbf{x}^{(0)} = (1,1), \varepsilon_1 = \varepsilon_2 = 0.001$

(iii) Step 3: Define the LP problem(linearize the objective function).

$$\textbf{\textit{Minimize:}} \quad f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) \cong f(\mathbf{x}^{(0)}) + \nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}$$

The first-order(linear) Taylor series expansion of the objective function

$$\textbf{\textit{Minimize:}} \quad f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) - f(\mathbf{x}^{(0)}) \cong \nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$$

$$\Delta\mathbf{x}^{(0)} = \mathbf{d}^{(0)}, \nabla f^T = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}$$

$$\mathbf{x}^{(k)} = \begin{bmatrix} x_1^{(k)} & x_2^{(k)} \end{bmatrix}^T$$

$$\mathbf{d}^{(k)} = \begin{bmatrix} d_1^{(k)} & d_2^{(k)} \end{bmatrix}^T = \begin{bmatrix} \Delta x_1^{(k)} & \Delta x_2^{(k)} \end{bmatrix}^T$$

$$\textbf{\textit{Minimize:}} \quad f(\mathbf{x}^{(0)} + \mathbf{d}^{(0)}) - f(\mathbf{x}^{(0)}) \cong \begin{bmatrix} 2x_1 - 3x_2 & 2x_2 - 3x_1 \end{bmatrix}_{\mathbf{x}^{(0)}} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix}$$

$$\overline{f}(\mathbf{d}^{(0)}) \cong (2x_1^{(0)} - 3x_2^{(0)})d_1^{(0)} + (2x_2^{(0)} - 3x_1^{(0)})d_2^{(0)}$$ ← **Substitute** $\mathbf{x}^{(0)} = (1,1)$

$$\boxed{\overline{f}(\mathbf{d}^{(0)}) \cong -d_1^{(0)} - d_2^{(0)}}$$ ← **The linearized objective function**

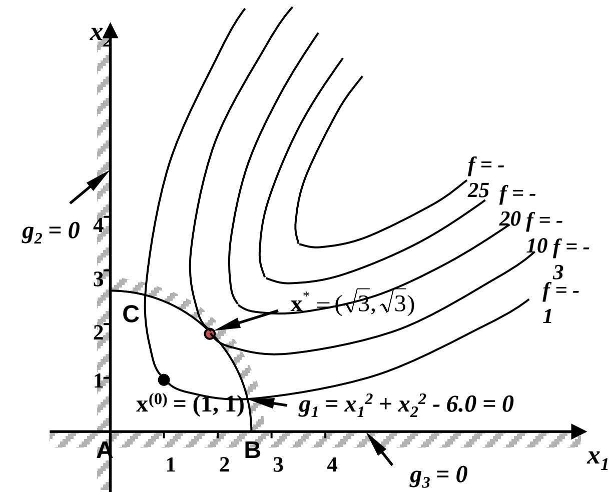
7.2 Sequential Linear Programming(SLP) - [Example] Problem with Inequality Constraints (4)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$



(1) Iteration 1(*k* = 0) $\mathbf{x}^{(0)} = (1,1), \varepsilon_1 = \varepsilon_2 = 0.001$

(iii) Step 3: Define the LP problem(linearize the constraints).

Subject to: $g_j(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) \Rightarrow g_j(\mathbf{x}^{(0)}) + \nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} \le 0;\ j = 1\ to\ m$

The first-order(linear) Taylor series expansion of the constraints

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$$

$$\nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} \le -g_j(\mathbf{x}^{(0)});\ j = 1\ to\ m$$

$$\Delta\mathbf{x}^{(0)} = \mathbf{d}^{(0)}, \nabla g_j^T = \begin{bmatrix} \frac{\partial g_j}{\partial x_1} & \frac{\partial g_j}{\partial x_2} \end{bmatrix}, \nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = \bar{g}_j(\Delta\mathbf{x}^{(0)}) = \bar{g}_j(\mathbf{d}^{(0)})$$

Subject to:

$$\bar{g}_1(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} \frac{1}{3}x_1^{(0)} & \frac{1}{3}x_2^{(0)} \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(\frac{1}{6}\left(x_1^{(0)}\right)^2 + \frac{1}{6}\left(x_2^{(0)}\right)^2 - 1.0\right)$$

$$\bar{g}_2(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} -x_1^{(0)} & 0 \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(-x_1^{(0)}\right)$$

$$\bar{g}_3(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} 0 & -x_2^{(0)} \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(-x_2^{(0)}\right)$$

Substitute $\mathbf{x}^{(0)} = (1,1)$

$$\bar{g}_1(\mathbf{d}^{(0)}) = \tfrac{1}{3}d_1^{(0)} + \tfrac{1}{3}d_2^{(0)} \le \frac{2}{3}$$

$$\bar{g}_2(\mathbf{d}^{(0)}) = -d_1^{(0)} \le 1$$

$$\bar{g}_3(\mathbf{d}^{(0)}) = -d_2^{(0)} \le 1$$

The linearized constraints

$$g_1(1,1) = -\tfrac{2}{3}$$

$$g_2(1,1) = -1$$

$$g_3(1,1) = -1$$

7.2 Sequential Linear Programming(SLP)
- [Example] Problem with Inequality Constraints (5)

(iv) Step 4: Solve LP problem for the design change($d^{(0)}$)

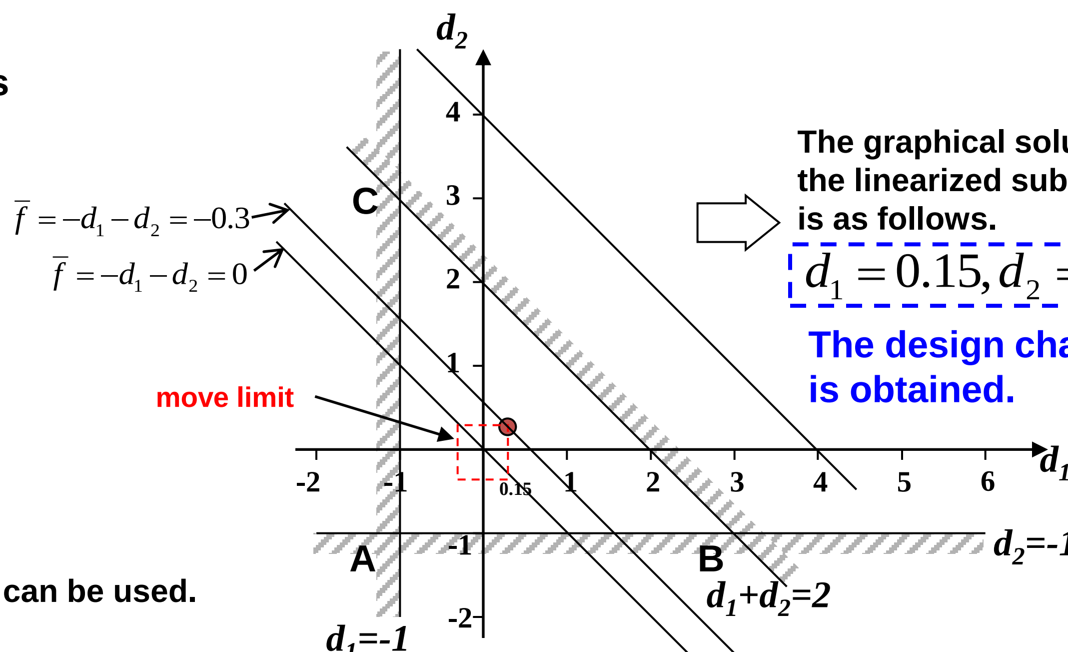
$$
\begin{aligned}
\textit{Minimize} \quad & \bar{f} = -d_1 - d_2 \\
\textit{Subject to} \quad & \tfrac{1}{3} d_1 + \tfrac{1}{3} d_2 \le \tfrac{2}{3} \\
& -d_1 \le 1 \\
& -d_2 \le 1 \\
& -0.15 \le d_1 \le 0.15 \\
& -0.15 \le d_2 \le 0.15
\end{aligned}
$$

Linearize the objective function and constraints.

$$
\begin{aligned}
\textit{Minimize} \quad & f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2 \\
\textit{Subject to} \quad & g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0 \\
& g_2(\mathbf{x}) = -x_1 \le 0 \\
& g_3(\mathbf{x}) = -x_1 \le 0
\end{aligned}
$$

$$
\begin{aligned}
& f(1,1) = -1, g_1(1,1) = -\tfrac{2}{3}, \\
& g_2(1,1) = -1, g_3(1,1) = -1 \\
& \nabla f = (-1,-1), \nabla g_1 = (\tfrac{1}{3}, \tfrac{1}{3}), \\
& \nabla g_2 = (-1,0), \nabla g_3 = (0,-1)
\end{aligned}
$$

Limits must be imposed on changes in design called move limit



The graphical solution for the linearized subproblem is as follows.

$$d_1 = 0.15, d_2 = 0.15$$

The design change is obtained.

To solve the problem, the Simplex method can be used.

7.2 Sequential Linear Programming(SLP)
- [Example] Problem with Inequality Constraints (6)

(v) Step 5: Check for convergence by using the obtained design change $\mathbf{d}^{(0)}$.

$$\mathbf{d}^{(0)} = (d_1, d_2) = (0.15, 0.15)$$

Since $\left\|\mathbf{d}^{(0)}\right\| = \sqrt{0.15^2 + 0.15^2} = 0.212 > \varepsilon_2 (= 0.001)$**, the criterion for convergence is not satisfied.**

(vi) Step 6: Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{d}^{(k)}$ **. Set** $k = k + 1$ **and go to Step 2**

$$\mathbf{x}^{(1)} = \mathbf{x}^{(1,1)} = \mathbf{x}^{(0)} + \mathbf{d}^{(0)} = (1,1) + (0.15, 0.15) = (1.15, 1.15)$$

$$k = k + 1 = 1$$

- **Step 1: Estimate a starting design point as $x^{(0)}$. Set k=0. Specify two small numbers, ε_1, ε_2(criterion for violating the constraints and convergence)**

- **Step 2: Evaluate objective and constraint function at current design point $x^{(k)}$. Also evaluate the objective and constraint function gradients at the current design point.**

- **Step 3: Select the proper move limits $\Delta x_{il}^{(k)}$ and $\Delta x_{iu}^{(k)}$ as some fraction of the current design point. Define the linear programming problem.**

$$\Delta x_{il}^{(k)} \le \Delta x_i^{(k)} \le \Delta x_{iu}^{(k)}$$

7.2 Sequential Linear Programming(SLP)
- Summary of Algorithm of SLP

- **Step 4: Solve the linear programming problem for $\mathbf{d}^{(k)}$ by using the Simplex method.**

- **Step 5: Check for convergence. If, $g_i \le \varepsilon_1 (i = 1 \text{ to } m)$, $|h_i| \le \varepsilon_1$ $(i = 1 \text{ to } p)$, and $\|\mathbf{d}^{(k)}\| \le \varepsilon_2$, then stop and the current design point $\mathbf{x}^{(k)}$ is the optimal solution. Otherwise, continue.**

- **Step 6: Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}$, Set $k = k+1$ and go to Step 2.**

7.2 Sequential Linear Programming(SLP)
- Summary of Algorithm of SLP

$$\textit{Minimize} \quad f(\mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}) \cong f(\mathbf{x}^{(k)}) + \nabla f^T(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)}$$

The first-order(linear) Taylor series expansion of the objective function

$$\textit{Subject to} \quad h_j(\mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}) \cong h_j(\mathbf{x}^{(k)}) + \nabla h_j^T(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} = 0;\ j = 1 \text{ to } p$$

The first-order(linear) Taylor series expansion of the equality constraints

$$g_j(\mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}) \cong g_j(\mathbf{x}^{(k)}) + \nabla g_j^T(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} \le 0;\ j = 1 \text{ to } m$$

The first-order(linear) Taylor series expansion of the inequality constraints

Define

$$\bar{f} = f(\mathbf{x} + \Delta\mathbf{x}) - f(\mathbf{x}),\ e_j = -h_j(\mathbf{x}),\ b_j = -g_j(\mathbf{x}),$$

$$c_i = \partial f(\mathbf{x})/\partial x_i,\ n_{ij} = \partial h_j(\mathbf{x})/\partial x_i,\ a_{ij} = \partial g_j(\mathbf{x})/\partial x_i,$$

$$d_i = \Delta x_i$$

$$\textit{Minimize} \quad \bar{f} = \sum_{i=1}^{n} c_i d_i$$

$$\textit{Subject to} \quad \sum_{i=1}^{n} n_{ij} d_i = e_j;\ j = 1 \text{ to } p$$

$$\sum_{i=1}^{n} a_{ij} d_i \le b_j;\ j = 1 \text{ to } m$$

where $d_{il} \le d_i \le d_{iu} \ (\Delta x_{il}^{(k)} \le \Delta x_i^{(k)} \le \Delta x_{iu}^{(k)})$

Matrix form

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T_{(1\times n)} \mathbf{d}_{(n\times 1)}$$ **: Linearized objective function**

$$\textit{Subject to} \quad \mathbf{N}^T_{(p\times n)} \mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$ **: Linearized equality constraint**

$$\mathbf{A}^T_{(m\times n)} \mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$ **: Linearized inequality constraint**

➡ **Linear Programming Problem**

➡ **It can be solved by using the Simplex method.**

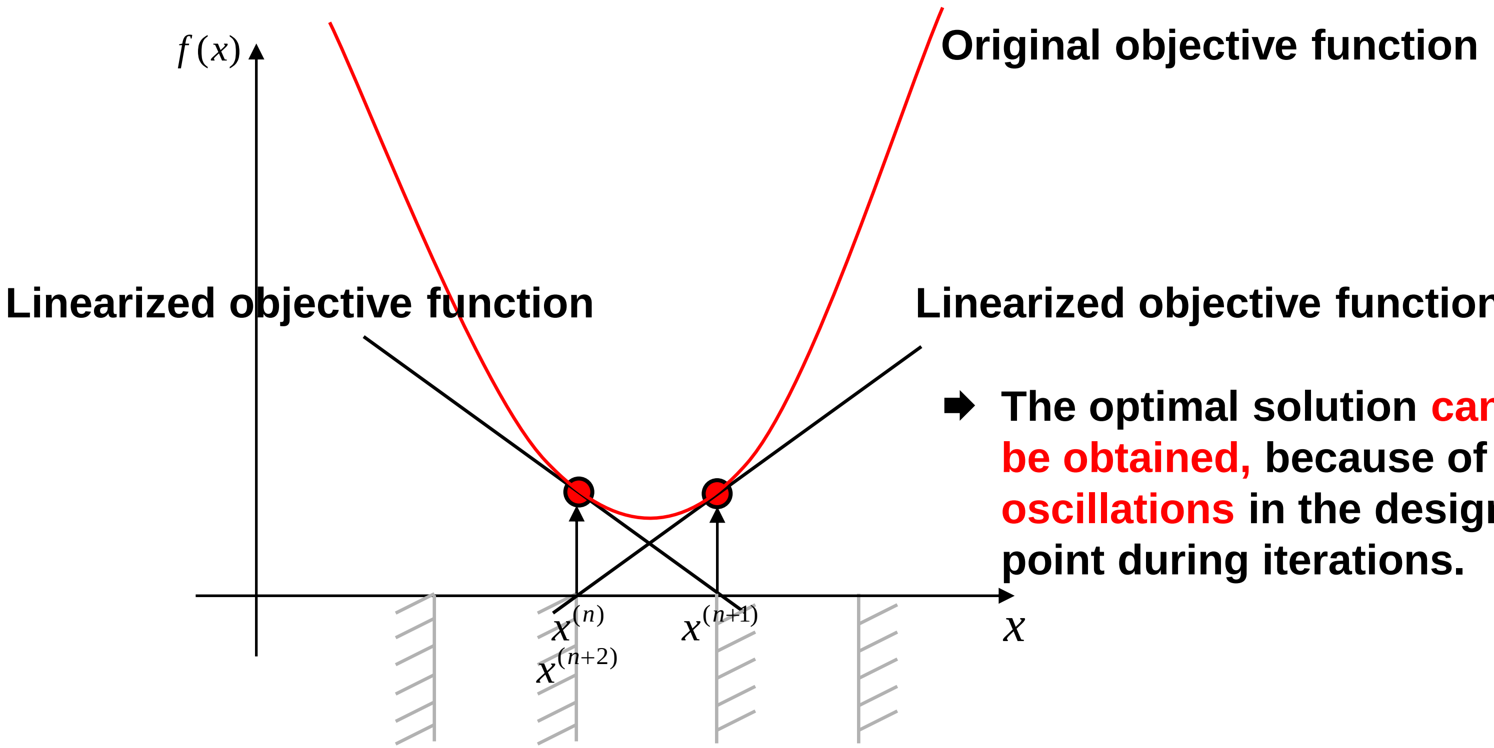
7.2 Sequential Linear Programming(SLP)
- Limitations of SLP Method

☑ The move limits of the design variables are defined by the user.

☑ If the move limits are too small, it take much time to find the optimal solution.

☑ If the move limits are too large, it can cause oscillations in the design point during iterations.

☑ Thus performance of the method depends heavily on selection of move limits



➡ The optimal solution cannot be obtained, because of the oscillations in the design point during iterations.

Ch.7 Constrained Nonlinear Optimization method

7.3 Sequential Quadratic Programming (SQP)

- Formulation of the Quadratic Programming Problem to Determine the Search Direction

$$\textit{Minimize} \quad f(\mathbf{x}+\Delta\mathbf{x}) \cong f(\mathbf{x}) + \nabla f^T(\mathbf{x})\Delta\mathbf{x} + 0.5\Delta\mathbf{x}^T\mathbf{H}\Delta\mathbf{x}$$

The second-order Taylor series expansion of the objective function

$$\textit{Subject to} \quad h_j(\mathbf{x}+\Delta\mathbf{x}) \cong h_j(\mathbf{x}) + \nabla h_j^T(\mathbf{x})\Delta\mathbf{x} = 0;\ j = 1\ to\ p$$

The first-order(linear) Taylor series expansion of the equality constraints

$$g_j(\mathbf{x}+\Delta\mathbf{x}) \cong g_j(\mathbf{x}) + \nabla g_j^T(\mathbf{x})\Delta\mathbf{x} \le 0;\ j = 1\ to\ m$$

The first-order(linear) Taylor series expansion of the inequality constraints

Assumption:
$$\bar{f} = f(\mathbf{x}+\Delta\mathbf{x}) - f(\mathbf{x}),\ e_j = -h_j(\mathbf{x}),\ b_j = -g_j(\mathbf{x}),$$
$$c_i = \partial f(\mathbf{x})/\partial x_i,\ n_{ij} = \partial h_j(\mathbf{x})/\partial x_i,\ a_{ij} = \partial g_j(\mathbf{x})/\partial x_i,$$
$$d_i = \Delta x_i$$

<u>**Matrix form**</u>

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$

: Quadratic objective function

$$\textit{Subject to} \quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$

: Linear equality constraints

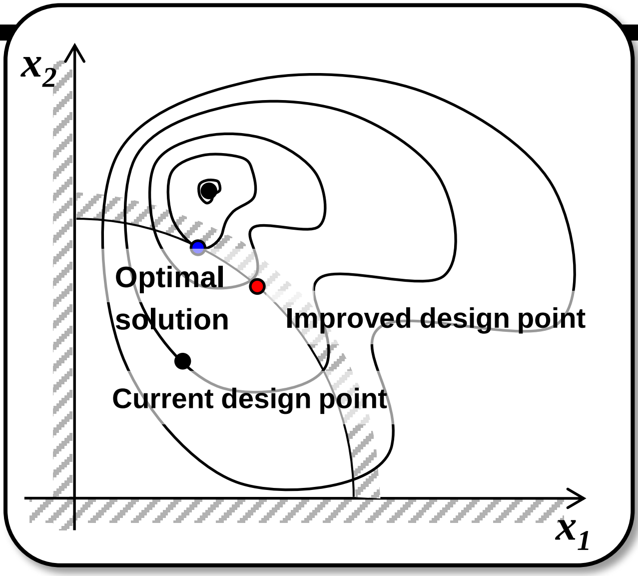
$$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$

: Linear inequality constraints

7.3 Sequential Quadratic Programming (SQP) - Algorithm of the SQP

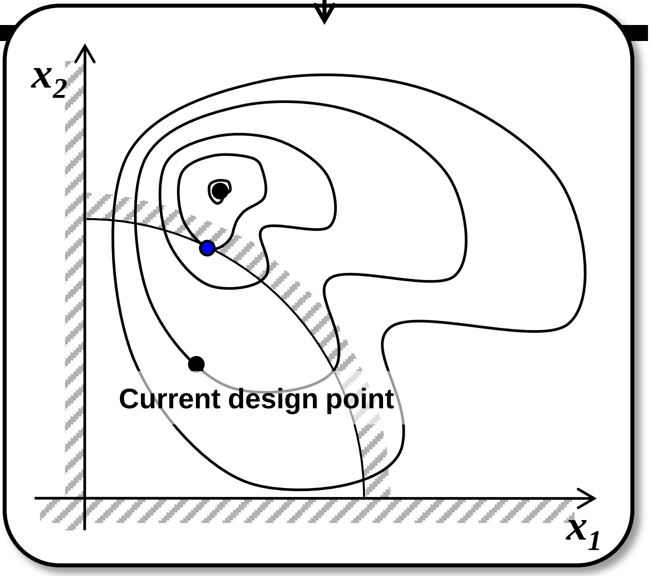
Objective function is approximated to the quadratic form

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form



Step 1

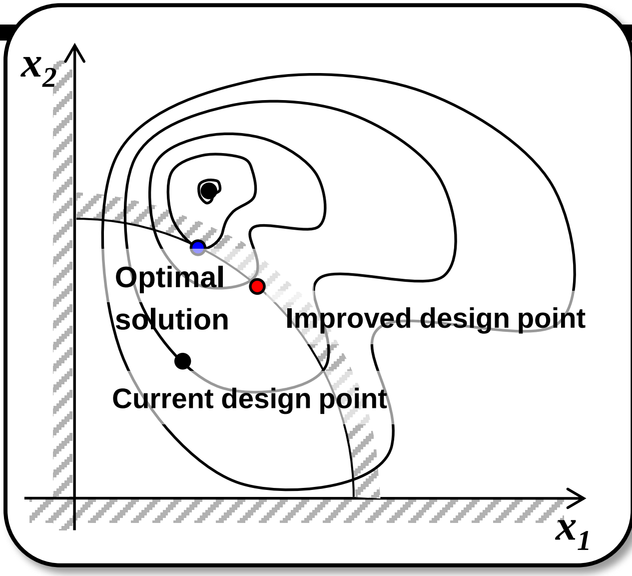
Define the quadratic programming problem at the current point.



7.3 Sequential Quadratic Programming (SQP) - Algorithm of the SQP

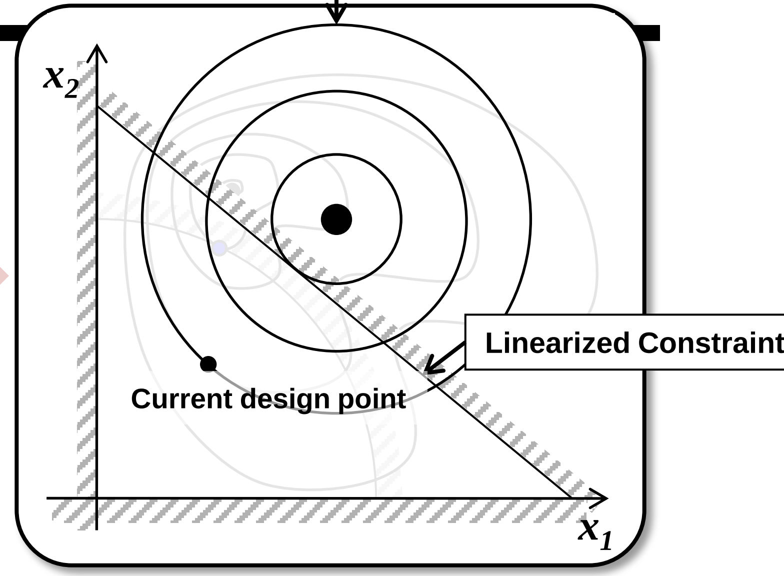
Objective function is approximated to the quadratic form

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form



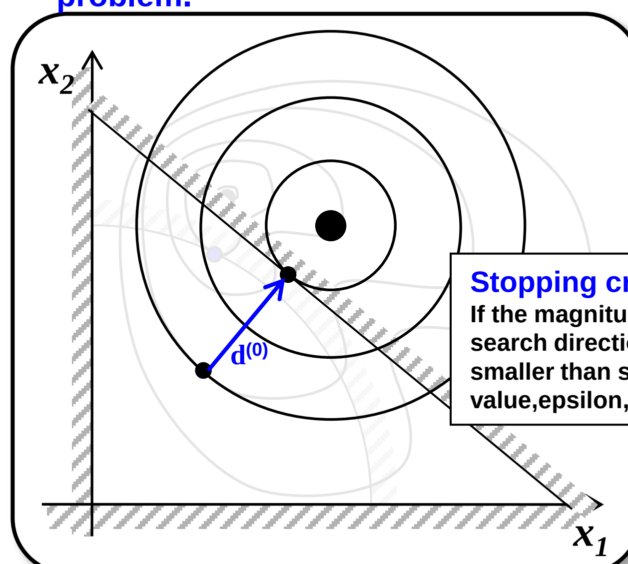
Step 1

Define the quadratic programming problem at the current point.



Step 2

Calculate the search direction($d^{(0)}$) by solving the quadratic programming problem.



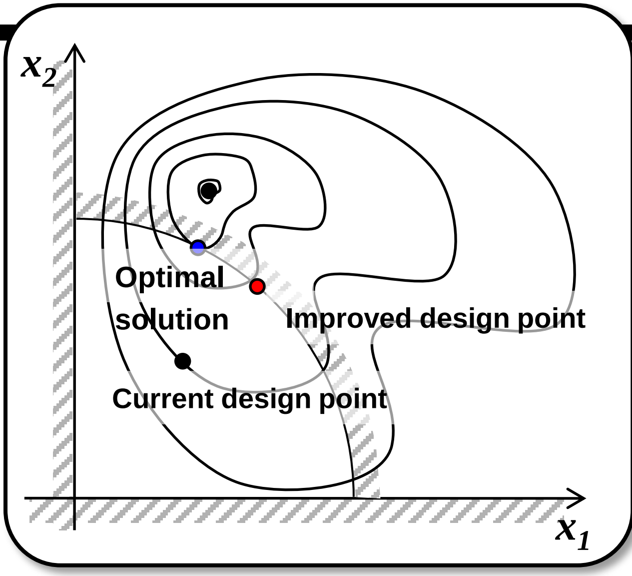
Stopping criteria

If the magnitude of the search direction $|d^{(0)}|$ is smaller than small value,epsilon, stop.

7.3 Sequential Quadratic Programming (SQP) - Algorithm of the SQP

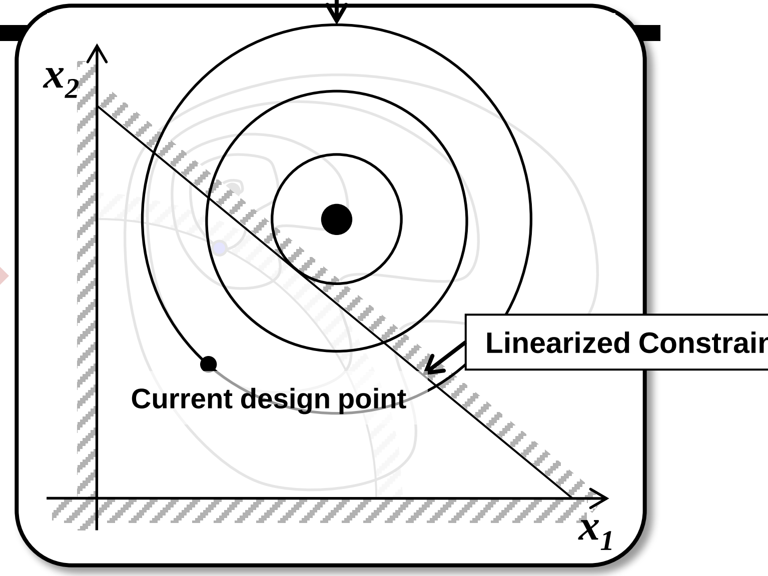
Objective function is approximated to the quadratic form

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form



Step 1

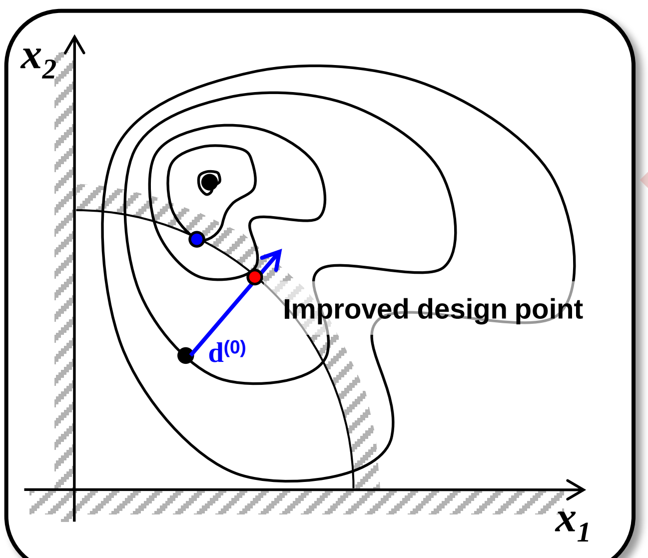
Define the quadratic programming problem at the current point.



Step 2

Calculate the **search direction($d^{(0)}$)** by solving **the quadratic programming problem.**

Go to the Step 1 at the improved design point.

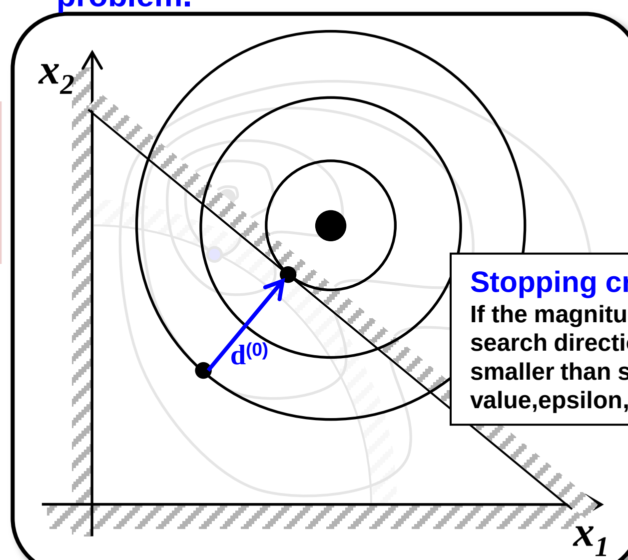


Step 3

After defining the penalty function, calculate the step size by using the **one dimensional search method.**

- **Penalty Function:**
Modified objective function by adding a penalty for possible constraint violations to the current value of the objective function(The method of transformation from the constrained optimal design problem to unconstrained optimal design problem)

- Example of the one dimensional search method: Golden section search method



Stopping criteria
If the magnitude of the search direction $|d^{(0)}|$ is smaller than small value,epsilon, stop.

☑ Sequential Quadratic Programming(SQP)

- ① After defining the quadratic programming problem about the objective function and constraints at the current design point, solve this problem and calculate the search direction $\mathbf{d}^{(k)}$.
- ② Define the penalty function by adding a penalty for possible constraint violations to the current value of the objective function and calculate the step size α_k to minimize the penalty function. For determination of the step size one dimensional search method, e.g., Golden section search method can be used. And determine the improved design point.
- ③ At the improved design point, go to ①
- The method is to find the optimal solution by solving the quadratic programming problem sequentially.

☑ CSD(Constrained Steepest Descent) method

- This method is a kind of the SQP method.
- When defining the quadratic programming problem, the Hessian matrix is assumed to be equal to the identity Matrix.
- This method uses the Pshenichny's penalty function.

If the Hessian matrix is equal to the Identity matrix , then the objective function is approximated as a centric circle form.

Define the QP problem

To find the search direction($\mathbf{d}^{(0)}$), define the QP problem at current design point.

Minimize: $f(\mathbf{x}^{(0)}+\Delta\mathbf{x}^{(0)}) \cong f(\mathbf{x}^{(0)})+\nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}+0.5\Delta\mathbf{x}^{(0)T}\mathbf{H}\Delta\mathbf{x}^{(0)}$

The second-order Taylor series expansion of the objective function

Minimize: $f(\mathbf{x}^{(0)}+\Delta\mathbf{x}^{(0)})-f(\mathbf{x}^{(0)}) \cong \nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}+0.5\Delta\mathbf{x}^{(0)T}\mathbf{H}\Delta\mathbf{x}^{(0)}$

$\Delta\mathbf{x}^{(0)}=\mathbf{d}^{(0)}, \nabla f^T=\begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}, \mathbf{H}=\mathbf{I}$ (In the CSD method, the Hessian matrix is assumed to be equal to the identity matrix.)

Minimize: $f(\mathbf{x}^{(0)}+\mathbf{d}^{(0)})-f(\mathbf{x}^{(0)}) \cong \begin{bmatrix}\frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}\end{bmatrix}_{\mathbf{x}^{(0)}}\begin{bmatrix}d_1^{(0)}\\ d_2^{(0)}\end{bmatrix}+0.5(d_1^{(0)2}+d_2^{(0)2})$

$$\bar{f}(\mathbf{d}^{(0)}) \cong \underbrace{\frac{\partial f(\mathbf{x}^{(0)})}{\partial x_1}}_{\text{constant}}d_1^{(0)}+\underbrace{\frac{\partial f(\mathbf{x}^{(0)})}{\partial x_2}}_{\text{constant}}d_2^{(0)}+0.5(d_1^{(0)2}+d_2^{(0)2})$$

$$\bar{f}(\mathbf{d}^{(0)}) \cong \underline{c_1 d_1^{(0)}+c_2 d_2^{(0)}+0.5(d_1^{(0)2}+d_2^{(0)2})}$$

It has the same form of the equation of circle.

Form of the equation of circle: $x_1^2+x_2^2+c_1x_1+c_2x_2+c_3=0$

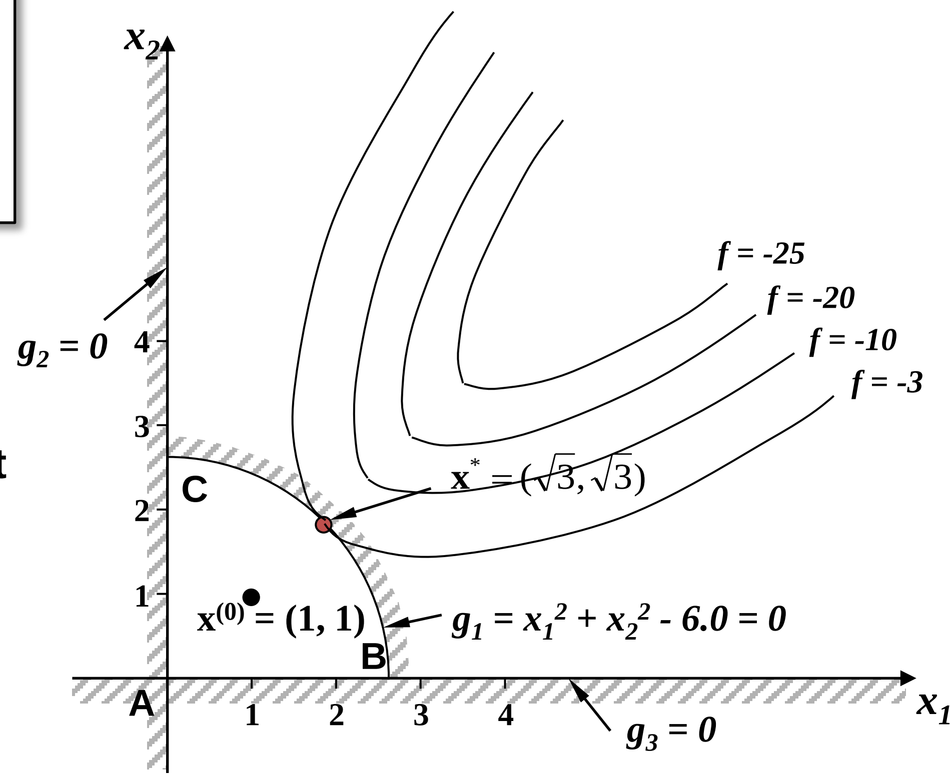
Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \dfrac{1}{6}x_1^2 + \dfrac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

Optimal solution: $\mathbf{x}^* = (\sqrt{3}, \sqrt{3}),\ f(\mathbf{x}^*) = -3$



Assume that the starting point is $\mathbf{x}^{(0)} = (1,1)$.

(1) Iteration 1(*k* = 0)

(i) Step 1: Evaluate the objective and constraint function at the current design point.

$f(1,1) = -1$

$g_1(1,1) = -\frac{2}{3} < 0$ ➡ **Constraint is satisfied.**

$g_2(1,1) = -1 < 0$ ➡ **Constraint is satisfied.**

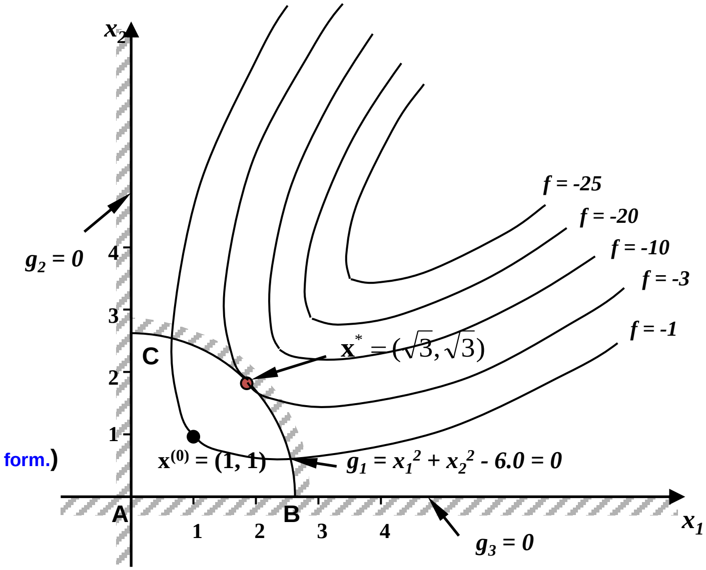
$g_3(1,1) = -1 < 0$ ➡ **Constraint is satisfied.**

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$



(1) Iteration 1(k = 0) $\mathbf{x}^{(0)} = (1,1)$

(ii) Step 2: Define the QP problem(The objective function is approximated to the quadratic form.)

$$\textit{Minimize: } f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) \cong f(\mathbf{x}^{(0)}) + \nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} + 0.5\Delta\mathbf{x}^{(0)T}\mathbf{H}\Delta\mathbf{x}^{(0)}$$

$$\textit{Minimize: } f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) - f(\mathbf{x}^{(0)}) \cong \nabla f^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} + 0.5\Delta\mathbf{x}^{(0)T}\mathbf{H}\Delta\mathbf{x}^{(0)}$$

$$\Delta\mathbf{x}^{(0)} = \mathbf{d}^{(0)}, \nabla f^T = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}, \ \mathbf{H} = \mathbf{I}$$

$$\textit{Minimize: } \quad f(\mathbf{x}^{(0)} + \mathbf{d}^{(0)}) - f(\mathbf{x}^{(0)}) \cong \begin{bmatrix} 2x_1 - 3x_2 & 2x_2 - 3x_1 \end{bmatrix}_{\mathbf{x}^{(0)}} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} + 0.5(d_1^{(0)2} + d_2^{(0)2})$$

$$\overline{f}(\mathbf{d}^{(0)}) \cong (2x_1^{(0)} - 3x_2^{(0)})d_1^{(0)} + (2x_2^{(0)} - 3x_1^{(0)})d_2^{(0)} + 0.5(d_1^{(0)2} + d_2^{(0)2})$$

$$\overline{f}(\mathbf{d}^{(0)}) \cong \underbrace{-d_1^{(0)} - d_2^{(0)}}_{\text{Objective function is approximated to the first order term}} + \underbrace{0.5(d_1^{(0)2} + d_2^{(0)2})}_{\text{Objective function is approximated to the second order term}}$$

Objective function is approximated to the first order term

Objective function is approximated to the second order term

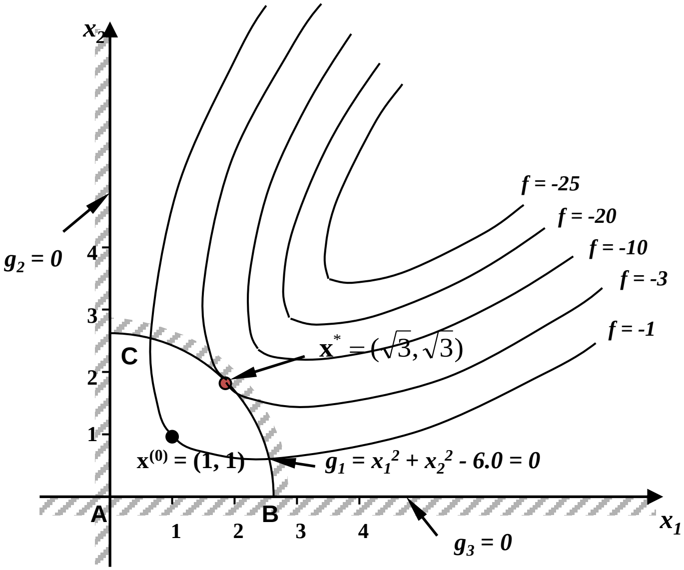
- Example of SQP – Iteration 1 (3)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$



(1) Iteration 1(k = 0) $\mathbf{x}^{(0)} = (1,1)$

Subject to: $g_j(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) \cong \underline{g_j(\mathbf{x}^{(0)}) + \nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}} \le 0;\ j = 1\ to\ m$

The first-order(linear) Taylor series expansion of constraint

$$\nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} \le -g_j(\mathbf{x}^{(0)});\ j = 1\ to\ m$$

$$\Delta\mathbf{x}^{(0)} = \mathbf{d}^{(0)}, \nabla g_j^T = \begin{bmatrix} \frac{\partial g_j}{\partial x_1} & \frac{\partial g_j}{\partial x_2} \end{bmatrix}, \nabla g_j^T(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = \bar{g}_j(\Delta\mathbf{x}^{(0)}) = \bar{g}_j(\mathbf{d}^{(0)})$$

Subject to:

$$\bar{g}_1(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} \frac{1}{3}x_1^{(0)} & \frac{1}{3}x_2^{(0)} \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(\frac{1}{6}\left(x_1^{(0)}\right)^2 + \frac{1}{6}\left(x_2^{(0)}\right)^2 - 1.0 \right)$$

$$\bar{g}_2(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} -x_1^{(0)} & 0 \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(-x_1^{(0)}\right)$$

$$\bar{g}_3(\mathbf{d}^{(0)}) \Rightarrow \begin{bmatrix} 0 & -x_2^{(0)} \end{bmatrix} \begin{bmatrix} d_1^{(0)} \\ d_2^{(0)} \end{bmatrix} \le -\left(-x_2^{(0)}\right)$$

The constraints are linearized

Substitute $\mathbf{x}^{(0)} = (1,1)$

$$\bar{g}_1(\mathbf{d}^{(0)}) = \tfrac{1}{3}d_1^{(0)} + \tfrac{1}{3}d_2^{(0)} \le \frac{2}{3}$$

$$\bar{g}_2(\mathbf{d}^{(0)}) = -d_1^{(0)} \le 1$$

$$\bar{g}_3(\mathbf{d}^{(0)}) = -d_2^{(0)} \le 1$$

(iii) Step 3: Solve the QP problem to determine the search direction($\mathbf{d}^{(0)}$)

Constrained Optimal Design Problem (Original problem)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \dfrac{1}{6}x_1^2 + \dfrac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

①

$f(1,1) = -1, g_1(1,1) = -\frac{2}{3},$
$g_2(1,1) = -1, g_3(1,1) = -1$
$\nabla f = (-1,-1), \nabla g_1 = (\frac{1}{3},\frac{1}{3}),$
$\nabla g_2 = (-1,0), \nabla g_3 = (0,-1)$

Quadratic Programming Problem

Minimize $\bar{f} = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $\frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3}$

$-d_1 \le 1$

$-d_2 \le 1$

$x_1^{(0)} + d_1 = x_1, \quad x_2^{(0)} + d_2 = x_2$

Substitute $x_1^{(0)} = 1, \quad x_2^{(0)} = 1$

$1 + d_1 = x_1, \quad 1 + d_2 = x_2$

②

Lagrange function

$$L = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2) + u_1[\tfrac{1}{3}(d_1 + d_2 - 2) + s_1^2] + u_2(-d_1 - 1 + s_2^2) + u_3(-d_2 - 1 + s_3^2)$$

③

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{d}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1^2 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2^2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3^2 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u \ge 0,\ i = 1, 2, 3$$

The optimal direction is

$$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0, 0, 0),$$

$$\mathbf{s}^{(0)} = (s_1, s_2, s_3) = (0, 1.414, 1.414),$$

$$\boxed{\mathbf{d}^{(0)} = (d_1, d_2) = (1,1)}$$

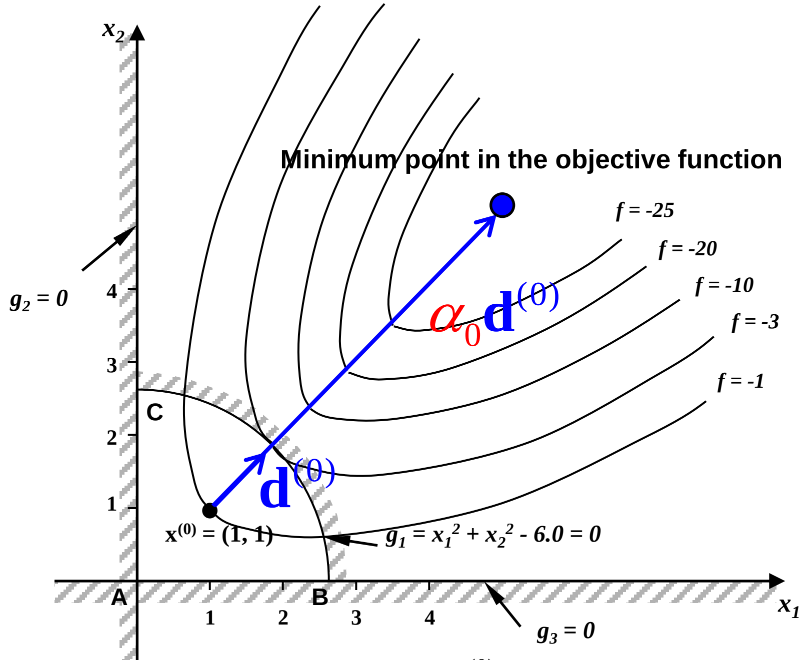
The search direction is determined.

*** The search direction also can be determined by using the Simplex method.**

$$\boxed{\mathbf{d}^{(0)} = (d_1, d_2) = (1,1)}$$

The search direction is determined.

(iv) Step 4: After the search direction($\mathbf{d}^{(0)}$) is determined, calculate the step size.



$$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0),$$

$$\mathbf{d}^{(0)} = (d_1, d_2) = (1,1)$$

Step size minimizing the value of the objective function along the search direction

$$\underline{\mathbf{x}^{(1)}} = \underline{\mathbf{x}^{(0)}} + \underline{\alpha_0 \mathbf{d}^{(0)}}$$

Improved design point

Current design point

Search direction obtained from the QP problem

Find α_k : *Minimize* $f\left(\mathbf{x}^{(1)}\right) = f\left(\mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}\right) = f(\alpha_0)$

Find

Objective function to be given

Given

The improved design point can be found along the search direction by minimizing the step size of the value of the objective function. However, it may violates the constraints, when without considering the original constraints.

Therefore, a penalty function, which considers the constraints, should be constructed by adding the penalty for possible constraint violations to the current value of the objective function.

By property of the nature, the objective function is decreased when the constraints is violated, we can find the improved design point of minimizing the penalty function while the constraints are satisfied.

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

Penalty function(Pshenichny's descent function, $\Phi(\mathbf{x}^{(k)})$)

By adding a penalty for possible constraint violations to the current value of the objective function, the constrained optimal design problem is transformed to the unconstrained optimal design problem

$$\Phi(\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}) + R_k \cdot V(\mathbf{x}^{(k)})$$

where,

k: iteration number how many times the QP problem is defined approximately

$f(\mathbf{x}^{(k)})$: current(k[th] iteration) value of the objective function

$V(\mathbf{x}^{(k)})$ is either the maximum constraint violation among all the constraints or zero.

$V(\mathbf{x}^{(k)})$ is nonnegative. If all the constraints are satisfied, the value of the $V(\mathbf{x}^{(k)})$ is zero.

$$V(\mathbf{x}^{(k)}) = \max\{0; |h_1|, |h_2|, \cdots, |h_p|; g_1, g_2, \cdots, g_m\}$$

where,

h_p: value of the equality constraint function at the design point $\mathbf{x}^{(k)}$

g_p: value of the inequality constraint function at the design point $\mathbf{x}^{(k)}$

R_k is a positive number called the penalty parameter

$$R_k = \max\{R_0, \ r_k\}$$

The initially value of R_k is specified by the user: R_0

r_k : Summation of all the Lagrange multipliers

$$r_k = \sum_{i=1}^{p} |v_i^{(k)}| + \sum_{i=1}^{m} u_i^{(k)}$$

$v_i^{(k)}$:Lagrange multipliers for the equality constraints(free in sign)

$u_i^{(k)}$:Lagrange multiplier for the inequality constraints(nonnegative)

7.3 Sequential Quadratic Programming (SQP)

- Penalty function : Pshenichny's Descent Function(2)

Penalty function(Pshenichny's descent function, $\Phi(\mathbf{x}^{(k)})$)

By adding a penalty for possible constraint violations to the current value of the objective function($f(\mathbf{x})$), the constrained optimal design problem is transformed to the unconstrained optimal design problem

$$\Phi(\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}) + R_k \cdot V(\mathbf{x}^{(k)})$$

(k is the iteration number how many times the QP problem is defined approximately.)

$V(\mathbf{x}^{(k)})$ is either the maximum constraint violation among all the constraints or zero.

$V(\mathbf{x}^{(k)})$ is nonnegative. If all the constraints are satisfied, the value of the $V(\mathbf{x}^{(k)})$ is zero.

$$V(\mathbf{x}^{(k)}) = \max\{0; |h_1|, |h_2|, \cdots, |h_p|; g_1, g_2, \cdots, g_m\}$$

h_p: value of the equality constraint function at the design point $\mathbf{x}^{(k)}$

g_p: value of the inequality constraint function at the design point $\mathbf{x}^{(k)}$

R_k is a positive number called the penalty parameter (initially specified by the user).

$$R_k = \max\left\{R_0,\ \underline{r_k}\left(= \sum_{i=1}^{p}\left|v_i^{(k)}\right| + \sum_{i=1}^{m} u_i^{(k)}\right)\right\}$$

$v_i^{(k)}$:Lagrange multipliers for the equality constraints(free in sign)

$u_i^{(k)}$:Lagrange multiplier for the inequality constraints(nonnegative)

: Summation of all the Lagrange multipliers

(v) Step 5: Calculate the penalty parameter R_k (In this example, the initial penalty parameter is assumed as R_0=10.)

$$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0) \text{ and } r_k = \sum_{i=1}^{p}\left|v_i^{(k)}\right| + \sum_{i=1}^{m} u_i^{(k)} \qquad r_0 = \sum_{i=1}^{m} u_i^{(0)} = 0$$

Since this problem does not have the equality constraints, we do not consider the v_i.

Therefore, $R_0 = \max\{R_0, r_0\} = \max\{10, 0\} = 10$

$$\Phi(\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}) + R_k \cdot V(\mathbf{x}^{(k)})$$

$$= x_1^2 + x_2^2 - 3x_1x_2 + 10 \cdot V(\mathbf{x}^{(k)}),\ V(\mathbf{x}^{(k)}) = \max\{0, g_1(\mathbf{x}^{(k)}), g_2(\mathbf{x}^{(k)}), g_3(\mathbf{x}^{(k)})\},\ (k=0)$$

$g_1(\mathbf{x}^{(k)}) = \frac{1}{6}(x_1^{(k)})^2 + \frac{1}{6}(x_2^{(k)})^2 - 1.0$

$g_2(\mathbf{x}^{(k)}) = -x_1^{(k)}$

$g_3(\mathbf{x}^{(k)}) = -x_2^{(k)}$

7.3 Sequential Quadratic Programming (SQP)
- Penalty function (3)

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

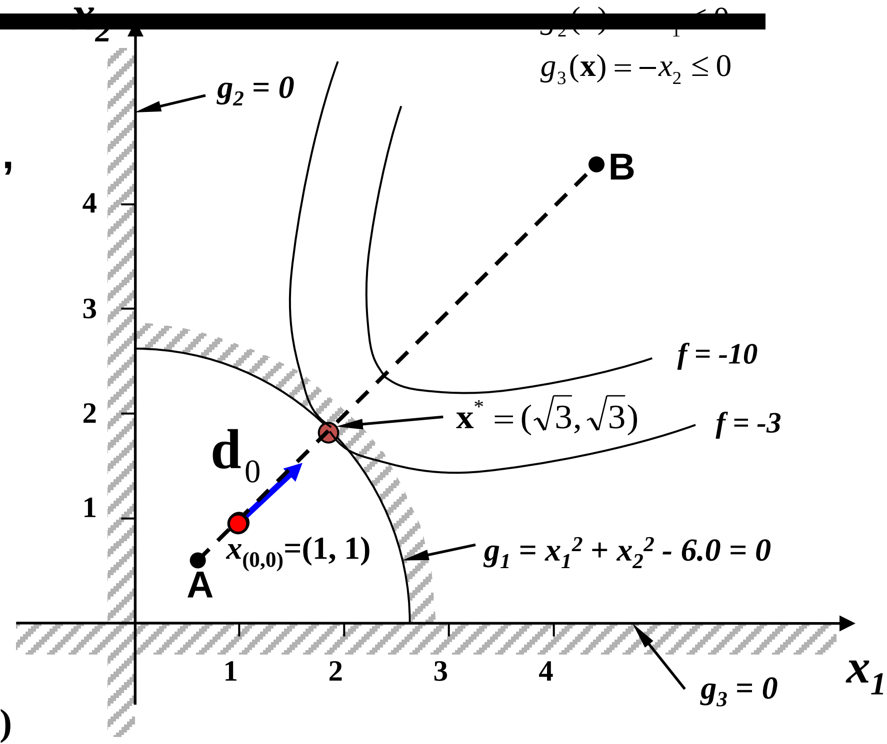
$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

(vi) Step 6:

By using the one dimensional search method, e.g., Golden section search method, calculate the step size to minimize the penalty function along the search direction($\mathbf{d}^{(0)}$), and determine the improved design point.



$$\Phi(\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}) + R_k \cdot V(\mathbf{x}^{(k)})$$
$$= x_1^2 + x_2^2 - 3x_1x_2 + 10 \cdot V(\mathbf{x}^{(k)})$$

$$V(\mathbf{x}^{(k)}) = \max\{0, g_1(\mathbf{x}^{(k)}), g_2(\mathbf{x}^{(k)}), g_3(\mathbf{x}^{(k)})\} \quad ,(k=0)$$

After the k-th search direction is found, one dimensional search for step size is started.

$$\mathbf{x}^{(k,j)} = \underline{\mathbf{x}^{(k)}} + \alpha_{(k,j)}\underline{\mathbf{d}^{(k)}}$$

The iteration number k does not change during the one dimensional search .

$$\Phi\left(\mathbf{x}^{(k,j)}\right) = f\left(\mathbf{x}^{(k,j)}\right) + \underline{R_k} \cdot V\left(\mathbf{x}^{(k,j)}\right), \quad V(\mathbf{x}^{(k,j)}) = \max\{0, g_1(\mathbf{x}^{(k,j)}), g_2(\mathbf{x}^{(k,j)}), g_3(\mathbf{x}^{(k,j)})\}$$

The iteration number k does not change during the one dimensional search method

After completing the one dimensional search, k is changed to k+1:

$\mathbf{x}^{(k,j)}$ **is changed to** $\mathbf{x}^{(k+1)}$

7.3 Sequential Quadratic Programming (SQP) - Determination of the Step Size by Using the Golden Section Search Method (1)

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

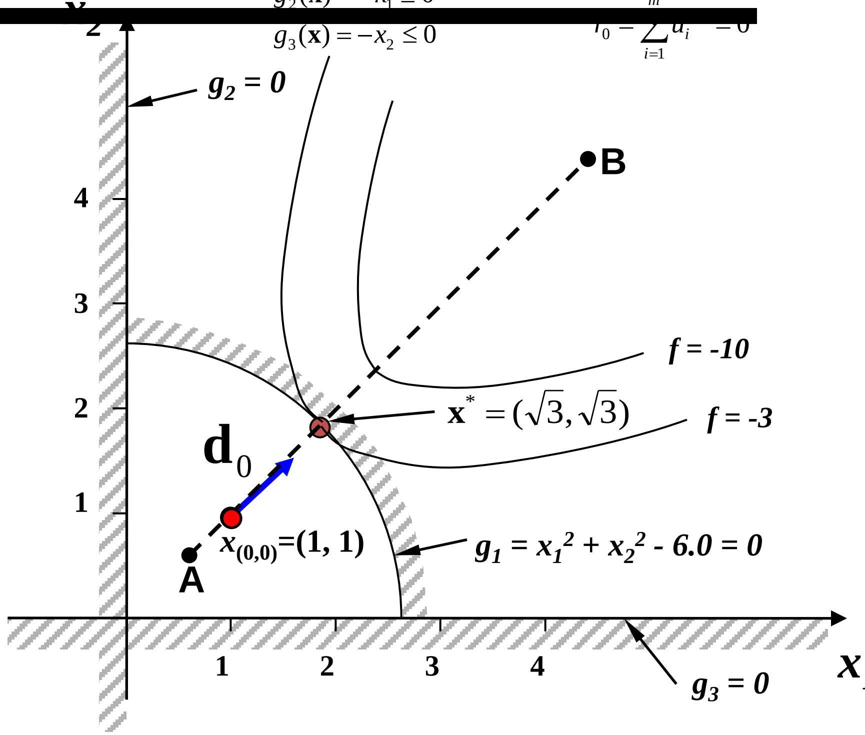
$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

$R_0 = \max\{R_0, r_0\} = \max\{10, 0\} = 10$

$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0)$

$r_0 = \sum_{i=1}^{m} u_i = 0$



(vi) Step 6:

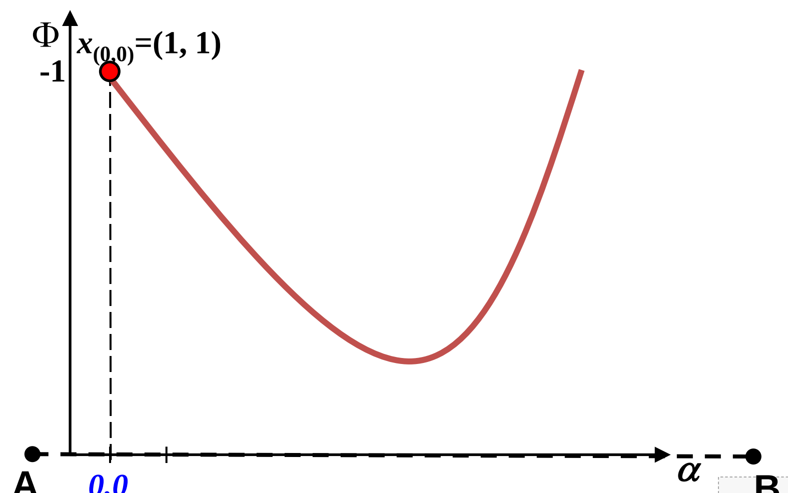
Search direction: $\mathbf{d}_0 = (1,1), \quad k = 0,\ j = 0$

When $\alpha_{(0,j)} = 0.0$

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + \alpha_{(0,j)} \cdot d^{(0)} = (1,1) + 0 \cdot (1,1) = (1,1)$$

$$\Phi\left(\mathbf{x}^{(0,j)}\right) = f\left(\mathbf{x}^{(0,j)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -1 + 10 \times 0 = -1$$

$$\textit{where, } V(\mathbf{x}^{(0,j)}) = \max\{0, g_1(\mathbf{x}^{(0,j)}), g_2(\mathbf{x}^{(0,j)}), g_3(\mathbf{x}^{(0,j)})\}$$
$$= \max\{0, -\tfrac{2}{3}, -1, -1\} = 0$$



7.3 Sequential Quadratic Programming (SQP) - Determination of the Step Size by Using the Golden Section Search Method (2)

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

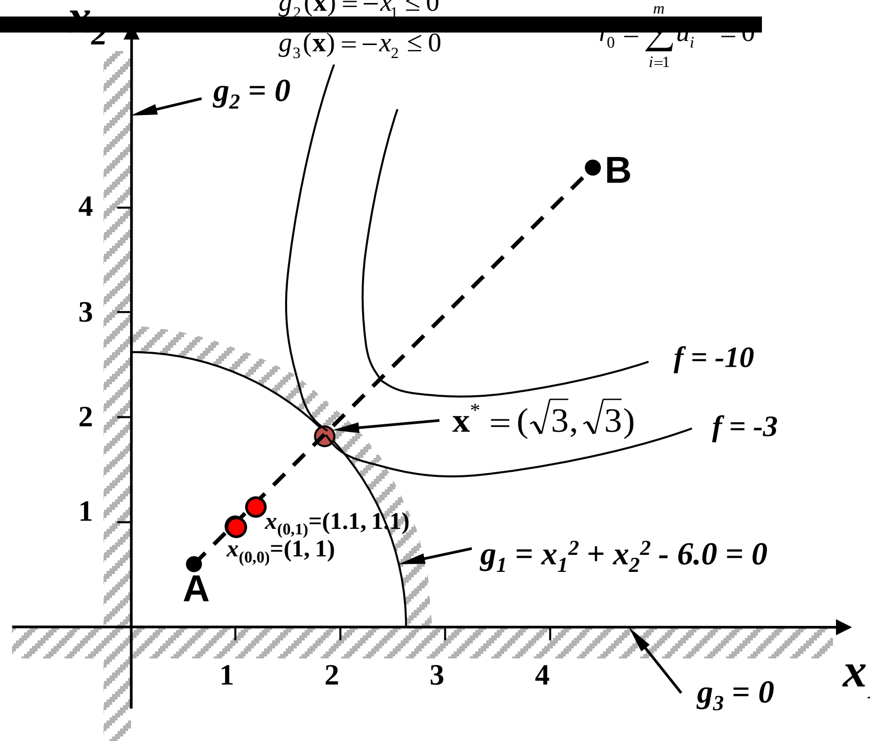
$g_3(\mathbf{x}) = -x_2 \le 0$

$R_0 = \max\{R_0, r_0\} = \max\{10, 0\} = 10$

$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0)$

$r_0 = \sum_{i=1}^{m} u_i = 0$

(vi) Step 6:



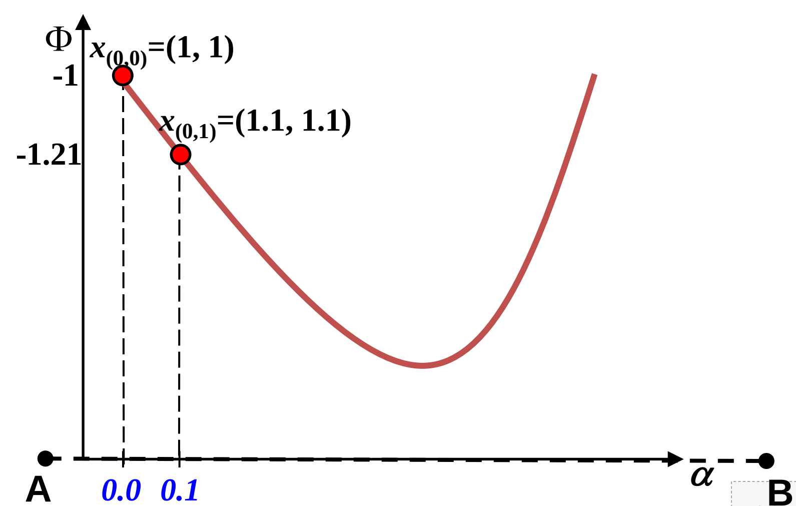
Search direction: $\mathbf{d}_0 = (1,1), \quad k = 0, \; j = 1$

Assume $\alpha_{(0,j)} = 0.1$*,

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + \alpha_{(0,j)} \cdot d^{(0)} = (1,1) + 0.1 \cdot (1,1) = (1.1, 1.1)$$

$$\Phi\left(\mathbf{x}^{(0,j)}\right) = f\left(\mathbf{x}^{(0,j)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -1.21 + 10 \times 0 = -1.21$$

$$where,\ V(\mathbf{x}^{(0,j)}) = \max\{0, g_1(\mathbf{x}^{(0,j)}), g_2(\mathbf{x}^{(0,j)}), g_3(\mathbf{x}^{(0,j)})\}$$
$$= \max\{0, -0.57, -1.1, -1.1\} = 0$$



*** The initial value of $\alpha_{(k,j)}$ (0.1) is defined by user, we can also define that as another value(ex. 0.5).**

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

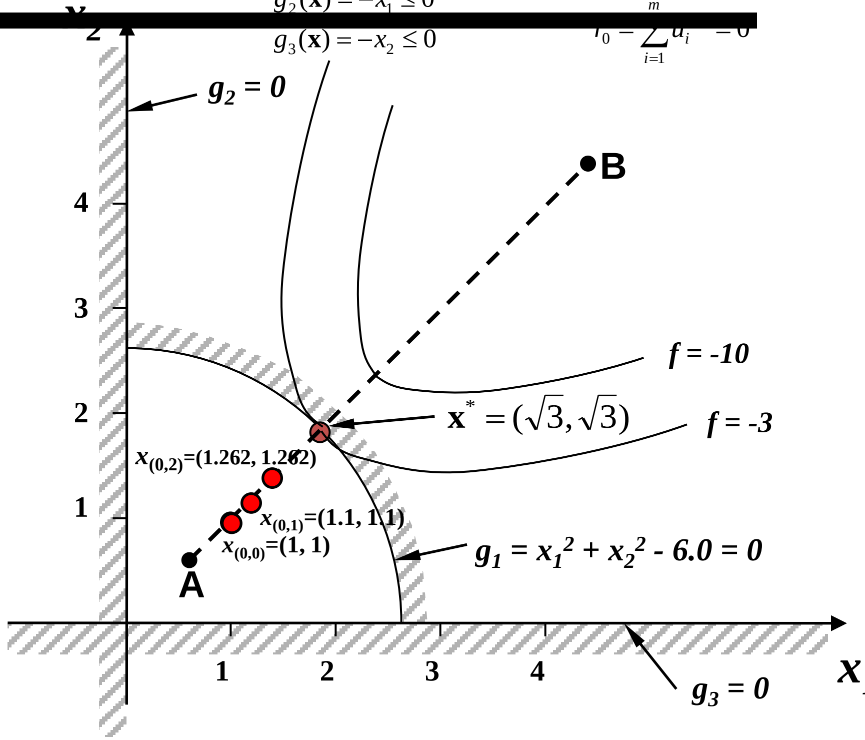
$g_3(\mathbf{x}) = -x_2 \le 0$

$R_0 = \max\{R_0, r_0\}$
$= \max\{10, 0\} = 10$

$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0)$

$r_0 = \sum_{i=1}^{m} u_i = 0$

(vi) Step 6:



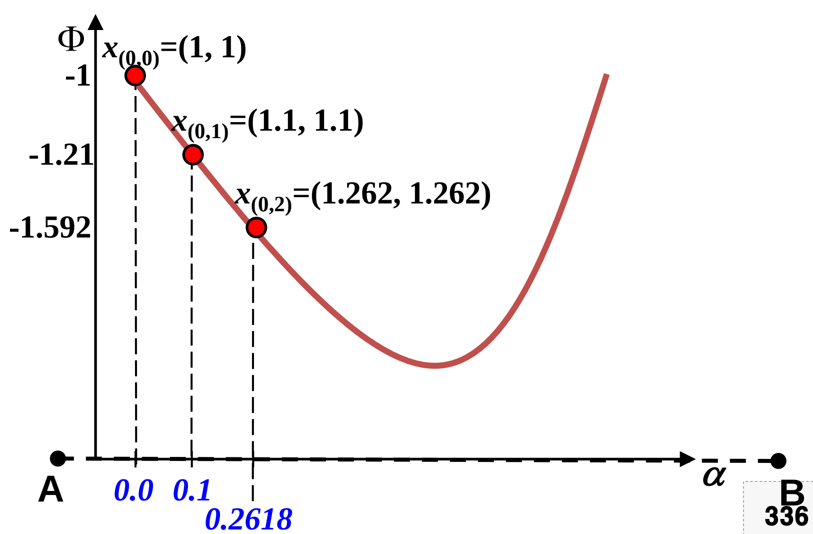
Search direction: $\mathbf{d}_0 = (1,1), \quad k = 0, \ j = 2$

When $\alpha_{(0,j)} = 0.1 + 1.618(0.1) = 0.2618$

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + \alpha_{(0,j)} \cdot d^{(0)} = (1,1) + 0.262 \cdot (1,1) = (1.262, 1.262)$$

$$\Phi\left(\mathbf{x}^{(0,j)}\right) = f\left(\mathbf{x}^{(0,j)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -1.592 + 10 \times 0 = -1.592$$

$$where,\ V(\mathbf{x}^{(0,2)}) = \max\{0, g_1(\mathbf{x}^{(0,2)}), g_2(\mathbf{x}^{(0,2)}), g_3(\mathbf{x}^{(0,2)})\}$$
$$= \max\{0, -0.469, -1.262, -1.262\} = 0$$



$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

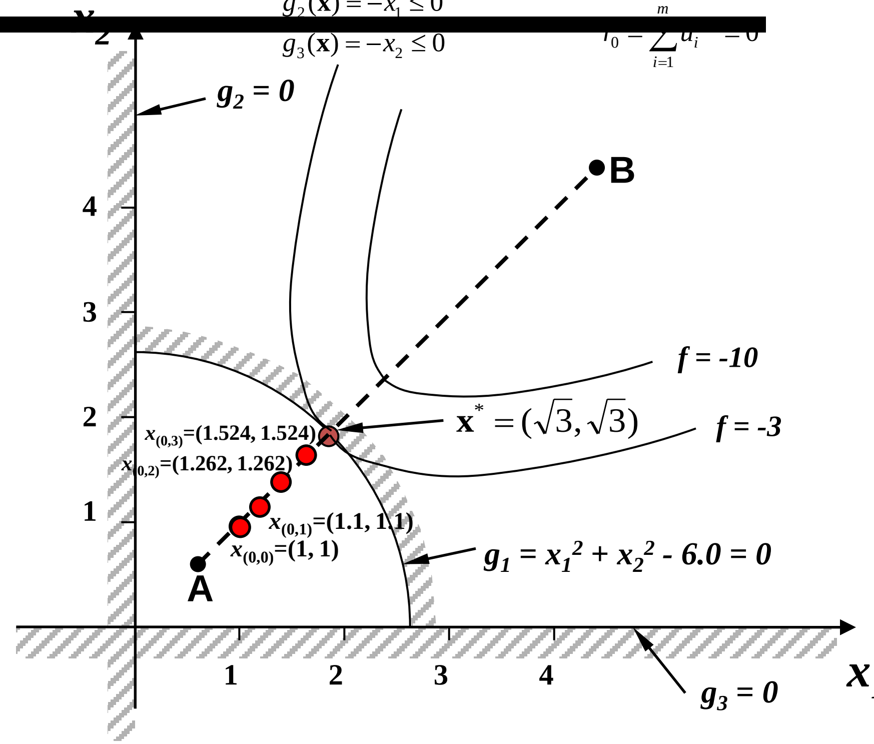
$g_3(\mathbf{x}) = -x_2 \le 0$

$R_0 = \max\{R_0, r_0\} = \max\{10, 0\} = 10$

$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0)$

$r_0 = \sum_{i=1}^{m} u_i = 0$

(vi) Step 6:



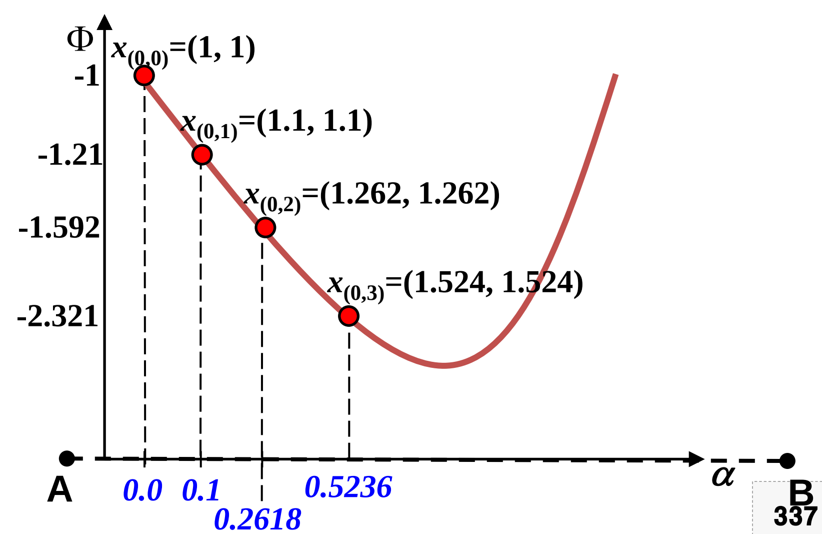
Search direction: $\mathbf{d}_0 = (1,1), \quad k = 0, \ j = 3$

When $\alpha_{(0,j)} = 0.1 + 1.618(0.1) + 1.618^2(0.1) = 0.5236$

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + \alpha_{(0,j)} \cdot d^{(0)} = (1,1) + 0.524 \cdot (1,1) = (1.524, 1.524)$$

$$\Phi\left(\mathbf{x}^{(0,j)}\right) = f\left(\mathbf{x}^{(0,j)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -2.321 + 10 \times 0 = -2.321$$

$$where,\ V(\mathbf{x}^{(0,j)}) = \max\{0, g_1(\mathbf{x}^{(0,j)}), g_2(\mathbf{x}^{(0,j)}), g_3(\mathbf{x}^{(0,j)})\}$$
$$= \max\{0, -0.226, -1.524, -1.524\} = 0$$

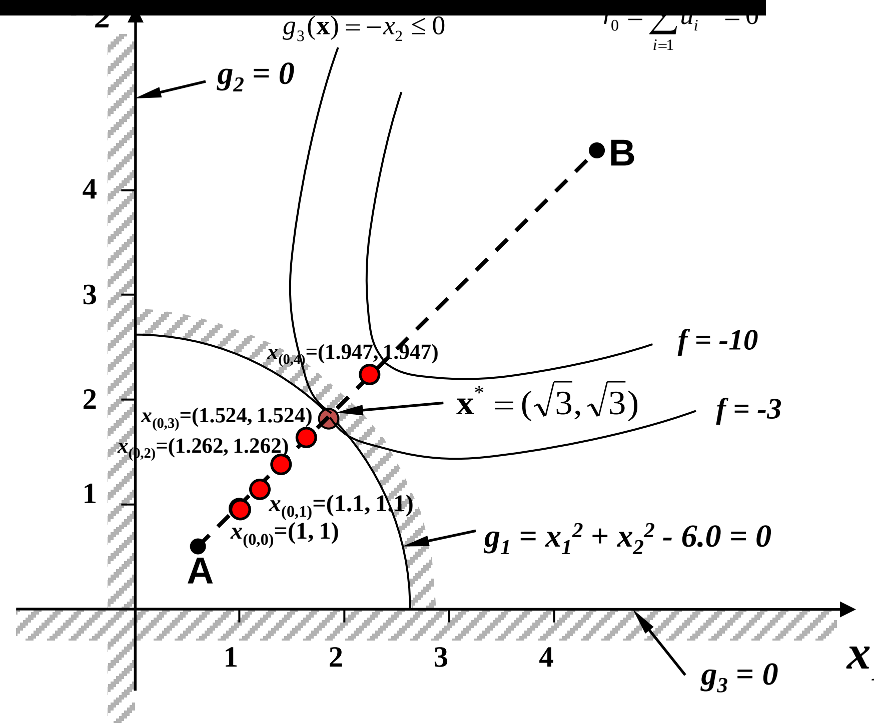


7.3 Sequential Quadratic Programming (SQP) - Determination of the Step Size by Using the Golden Section Search Method (5)

$$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$$
$$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$
$$g_2(\mathbf{x}) = -x_1 \le 0$$
$$g_3(\mathbf{x}) = -x_2 \le 0$$

$$R_0 = \max\{R_0, r_0\} = \max\{10, 0\} = 10$$
$$\mathbf{u}^{(0)} = (u_1, u_2, u_3) = (0,0,0)$$
$$r_0 = \sum_{i=1}^{m} u_i = 0$$

(vi) Step 6:



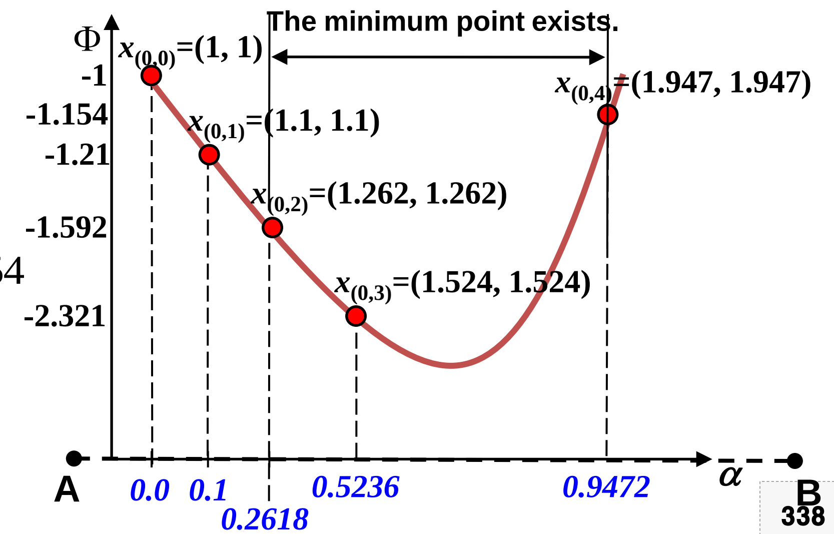
Search direction: $\mathbf{d}_0 = (1,1), \quad k = 0,\ j = 4$

When $\alpha_{(0,j)} = 0.1 + 1.618(0.1) + 1.618^2(0.1) + 1.618^3(0.1) = 0.9472$

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + \alpha_{(0,j)} \cdot d^{(0)} = (1,1) + 0.947 \cdot (1,1) = (1.947, 1.947)$$

$$\Phi\left(\mathbf{x}^{(0,j)}\right) = f\left(\mathbf{x}^{(0,j)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -3.792 + 10 \times 0.2638 = -1.154$$

$$where,\ V(\mathbf{x}^{(0,4)}) = \max\{0, g_1(\mathbf{x}^{(0,4)}), g_2(\mathbf{x}^{(0,4)}), g_3(\mathbf{x}^{(0,4)})\}$$
$$= \max\{0, 0.2638, -1.947, -1.947\} = 0.2638$$



$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

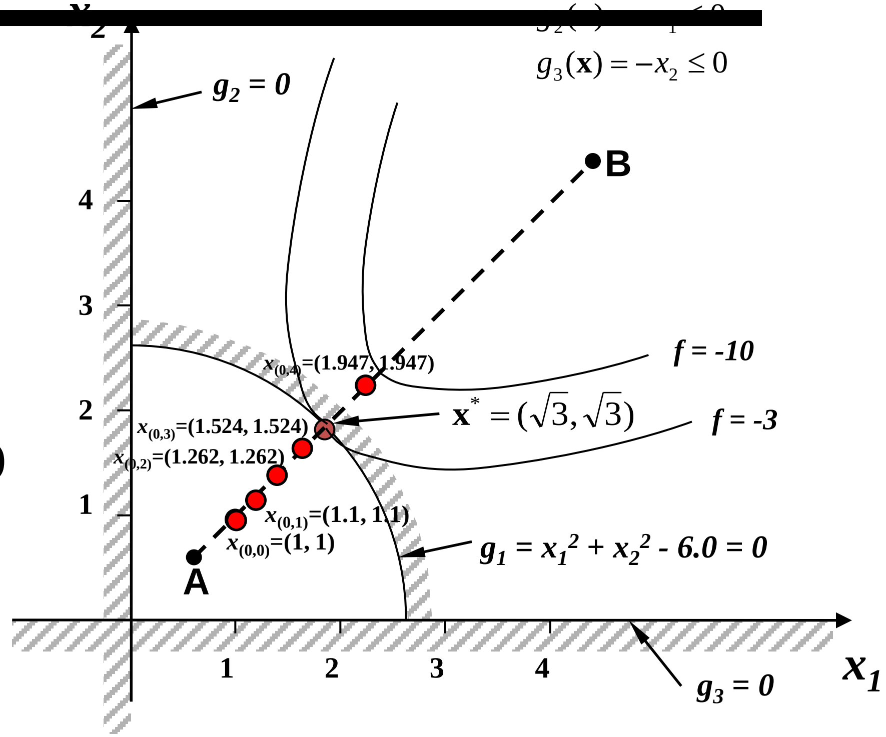
$g_3(\mathbf{x}) = -x_2 \le 0$

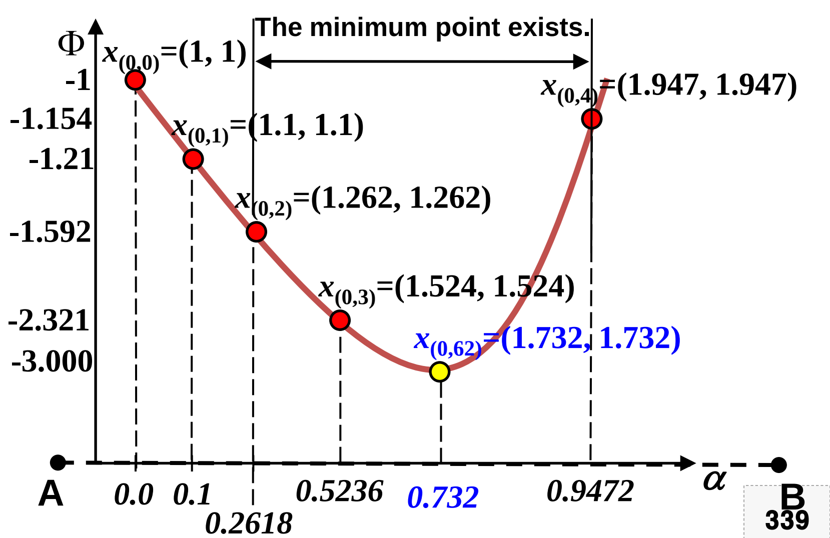
(vi) Step 6:

The value of the $\alpha_0 = 0.732$ **is found which minimizes the value of the penalty function in the interval between** $\mathbf{x}^{(0,2)}$ **and** $\mathbf{x}^{(0,4)}$**.**

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \cdot \mathbf{d}^{(0)} = (1,1) + 0.732 \cdot (1,1) = (1.732, 1.732)$$

$$f\left(\mathbf{x}^{(1)}\right) = f\left(1.732, 1.732\right) = -3$$





Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

(2) Iteration 2(k = 1)

(ii)Step 2: Calculate maximum constraint violation among all the constraints

From the previous stage,

$$\mathbf{x}^{(1)} = (1.732, 1.732)$$

$$f\left(\mathbf{x}^{(1)}\right) = f\left(1.732, 1.732\right) = -2.999824$$

$$g_1\left(\mathbf{x}^{(1)}\right) = g_1\left(1.732, 1.732\right) = -5.866 \times 10^{-5}$$

➡ **Constraint is satisfied.**

$$g_2\left(\mathbf{x}^{(1)}\right) = -1.732$$

➡ **Constraint is satisfied.**

$$g_3\left(\mathbf{x}^{(1)}\right) = -1.732$$

➡ **Constraint is satisfied.**

$$V_1 = V\left(\mathbf{x}^{(1)}\right) = \max\{0; -5.866 \times 10^{-5}, -1.732, -1.732\} = 0$$

And,

$$\nabla f(\mathbf{x}^{(1)}) = (2x_1 - 3x_2, 2x_2 - 3x_1) = (-1.732, -1.732)$$

$$\nabla g_1(\mathbf{x}^{(1)}) = (\tfrac{1}{3}x_1, \tfrac{1}{3}x_2) = (0.577, 0.577), \nabla g_2 = (-1, 0), \nabla g_3 = (0, -1)$$

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

(iii) Step 3: Solve the QP problem to determine the search direction($d^{(0)}$)

Constrained Optimal Design Problem (Original problem)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_1 \le 0$

① ⇒

Quadratic Programming Problem

Minimize $\bar{f} = (-1.732d_1 - 1.732d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $0.577d_1 + 0.577d_2 \le 5.866\times10^{-5}$

$-d_1 \le 1.732$

$-d_2 \le 1.732$

where,
$d_1 = x_1 - 1.732,$
$d_2 = x_2 - 1.732$

$f(1.732,1.732) = -3, \nabla f = (-1.732,-1.732)$

$g_1(1.732,1.732) = -5.866\times10^{-5}, \nabla g_1 = (0.577,0.577)$

$g_2(1.732,1.732) = -1.732, \nabla g_2 = (-1,0)$

$g_3(1.732,1.732) = -1.732, \nabla g_3 = (0,-1)$

②

Lagrange function

$$L = (-1.732d_1 - 1.732d_2) + 0.5(d_1^2 + d_2^2) + u_1[0.577(d_1 + d_2) - 5.866\times10^{-5} + s_1^2] + u_2(-d_1 - 1.732 + s_2^2) + u_3(-d_2 - 1.732 + s_3^2)$$

③ ⇒

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{d},\mathbf{u},\mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial d_1} = -1.732 + d_1 + 0.577u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1.732 + d_2 + 0.577u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = 0.577\left(d_1 + d_2\right) - 5.866\times10^{-6} + s_1^2 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1.732 + s_2^2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1.732 + s_3^2 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u \ge 0,\ i = 1,2,3$$

The optimal solution is

$$\mathbf{d}^{(1)} = (d_1, d_2) = (5.081\times10^{-5},\ 5.081\times10^{-5})$$

$$\mathbf{u}^{(1)} = (u_1, u_2, u_3) = (3, 0, 0)$$

$$\mathbf{s}^{(1)} = (s_1, s_2, s_3) = (0, 1.316, 1.316)$$

* The search direction also can be determined by using the Simplex method.

(iv) Step 4: Check for the following stopping criteria.

$$\mathbf{d}^{(1)} = (d_1, d_2) = (5.081\times10^{-5}, \quad 5.081\times10^{-5})$$

$$\left\|\mathbf{d}^{(1)}\right\| = \sqrt{\left(5.081\times10^{-5}\right)^2 + \left(5.081\times10^{-5}\right)^2} = 7.186\times10^{-5} < \varepsilon_2(=0.001)$$

The stopping criteria is satisfied.

(iv) Step 5: Stop

The optimal solution: $\mathbf{x}^* = (\sqrt{3}, \sqrt{3}),\ f(\mathbf{x}^*) = -3$

The Lagrange multiplier:

$$\mathbf{u}^* = (3, 0, 0),\ \mathbf{s}^* = (0, 1.316, 1.316)$$

Optimization Problem

Minimize $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$

Subject to $h_i(\mathbf{x}) = 0, \quad i = 1,\ldots,p$ **Equality constraints**

$g_i(\mathbf{x}) = 0, \quad i = 1,\ldots,m$ **Inequality constraints**

Pshenichny's descent function:

the penalty function is constructed by adding a penalty for possible constraint violations to the current value of the objective function

$$\Phi(\mathbf{x}^{(k)}) = f(\mathbf{x}^{(k)}) + R_k \cdot V(\mathbf{x}^{(k)})$$

(k is the iteration number how many times the QP problem is defined.)

$V(\mathbf{x}^{(k)})$ **is either the maximum constraint violation among all the constraints or zero.**

$V(\mathbf{x}^{(k)})$ **is nonnegative. If all the constraints are satisfied, the value of the** $V(\mathbf{x}^{(k)})$ **is zero.**

$$V(\mathbf{x}^{(k)}) = \max\{0; |h_1|, |h_2|, \cdots, |h_p|; g_1, g_2, \cdots, g_m\}$$

➡ **If all the constraints are satisfied, the value of the** $V(\mathbf{x}^{(k)})$ **is zero.**

R_k **is a positive number called the penalty parameter(initially specified by the user).**

$$R_k = \max\left\{R_0,\ \underline{r_k}\left(= \sum_{i=1}^{p}\left|v_i^{(k)}\right| + \sum_{i=1}^{m} u_i^{(k)}\right)\right\}$$

: Summation of the all Lagrange multipliers

The improved design point is determined as follows.

$$\underbrace{\mathbf{x}^{(k+1)}}_{\text{Improved design point}} = \underbrace{\mathbf{x}^{(k)}}_{\text{Current design point}} + \alpha_k \cdot \mathbf{d}^{(k)}$$

α_k: **Step size calculated by one dimensional search method(ex. Golden section search method)**

$\mathbf{d}^{(k)}$: **Search direction obtained from the QP problem**

7.3 Sequential Quadratic Programming (SQP)

- **Solution of the Quadratic Programming Problem to Determine the Search Direction by using the Simplex Method**

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab. http://asdal.snu.ac.kr

- Determine the Search Direction by using the Simplex Method

[Iteration 1] (1)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Solve the QP problem to determine the search direction($d^{(0)}$)

Constrained Optimal Design Problem (Original problem)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

①

$f(1,1) = -1,\ g_1(1,1) = -\frac{2}{3},$
$g_2(1,1) = -1,\ g_3(1,1) = -1$
$\nabla f = (-1,-1), \nabla g_1 = (\frac{1}{3},\frac{1}{3}),$
$\nabla g_2 = (-1,0), \nabla g_3 = (0,-1)$

Quadratic Programming Problem

Minimize $\bar{f} = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $\frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3}$

$-d_1 \le 1$

$-d_2 \le 1$

where $d_1 = x_1 - 1,\ d_2 = x_2 - 1$

②

Lagrange function

$$L = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2) + u_1[\tfrac{1}{3}(d_1 + d_2 - 2) + s_1^2] + u_2(-d_1 - 1 + s_2^2) + u_3(-d_2 - 1 + s_3^2)$$

③

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

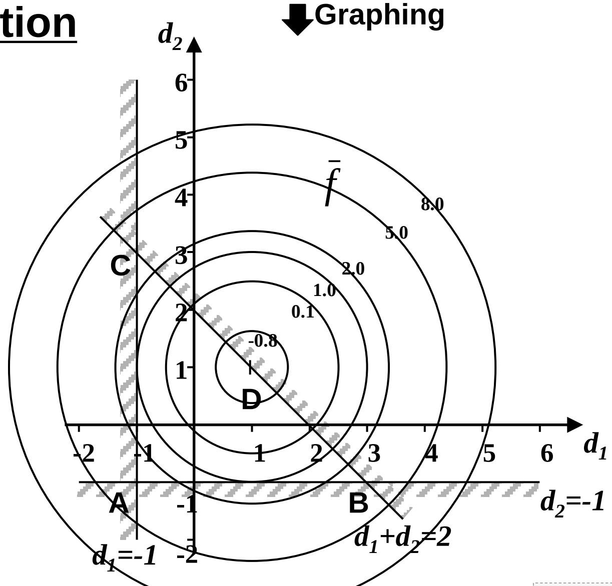
$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1^2 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2^2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3^2 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u_i \ge 0,\ i = 1,2,3$$

Graphing



- Determine the Search Direction by using the Simplex Method [Iteration 1] (2)

Quadratic Programming Problem

$$\textit{Minimize}\ \ \bar{f} = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2)$$

$$\textit{Subject to}\ \tfrac{1}{3}d_1 + \tfrac{1}{3}d_2 \le \tfrac{2}{3}$$

$$-d_1 \le 1$$

$$-d_2 \le 1$$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1^2 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2^2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3^2 = 0$$

$$\frac{\partial L}{\partial s_i} = \underline{u_i s_i = 0},\ u_i \ge 0,\ i = 1,2,3 \ \Rightarrow\ u_i s_i^2 = 0,\ u_i \ge 0,\ i = 1,2,3$$

Multiply the both side of equations by s_i

Replace s_i^2 with s_i' $\xrightarrow{\hspace{2em}}$ $s_i^2 = s_i' \ge 0$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1' = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2' = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3' = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i' = 0$$

$$u_i, s_i' \ge 0,\ i = 1,2,3$$

Represent s_i' to s_i for the convenience $\xrightarrow{\hspace{2em}}$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0$$

$$u_i, s_i \ge 0,\ i = 1,2,3$$

- Determine the Search Direction by using the Simplex Method [Iteration 1] (3)

Quadratic Programming Problem

Minimize $\bar{f} = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $\frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3}$

$-d_1 \le 1$

$-d_2 \le 1$

Matrix form

Minimize $\bar{f} = \mathbf{c}^T{}_{(1\times2)}\mathbf{d}_{(2\times1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times2)}\mathbf{H}_{(2\times2)}\mathbf{d}_{(2\times1)}$

Subject to $\mathbf{A}^T{}_{(3\times2)}\mathbf{d}_{(2\times1)} \le \mathbf{b}_{(3\times1)}$

Assume that $\mathbf{H}_{(2\times2)}$ is equal to $\mathbf{I}_{(2\times2)}$.

where $\mathbf{d}_{(2\times1)} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \mathbf{c}_{(2\times1)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{H}_{(2\times2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$

$\mathbf{A}_{(2\times3)} = \begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{1}{3} & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)} = \begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1 + d_1 + \tfrac{1}{3}u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1 + d_2 + \tfrac{1}{3}u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = \tfrac{1}{3}\left(d_1 + d_2 - 2\right) + s_1 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1 + s_2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1 + s_3 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0$$

$$u_i, s_i \ge 0,\ i = 1, 2, 3$$

How can we express the Kuhn-Tucker necessary condition in a matrix form($\mathbf{d}, \mathbf{c}, \mathbf{H}, \mathbf{A}, \mathbf{b}$)?

[Iteration 1](4)

$$\mathbf{d}_{(2\times1)}=\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \mathbf{c}_{(2\times1)}=\begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{H}_{(2\times2)}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_{(2\times3)}=\begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{1}{3} & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)}=\begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1}=-1+d_1+\tfrac{1}{3}u_1-u_2=0$$

$$\frac{\partial L}{\partial d_2}=-1+d_2+\tfrac{1}{3}u_1-u_3=0$$

$$\frac{\partial L}{\partial u_1}=\tfrac{1}{3}\left(d_1+d_2-2\right)+s_1=0$$

$$\frac{\partial L}{\partial u_2}=-d_1-1+s_2=0$$

$$\frac{\partial L}{\partial u_3}=-d_2-1+s_3=0$$

$$\frac{\partial L}{\partial s_i}=u_i s_i=0$$

$$u_i, s_i \ge 0,\ i=1,2,3$$

$$\begin{bmatrix} -1+d_1+\frac{1}{3}u_1-u_2 \\ -1+d_2+\frac{1}{3}u_1-u_3 \end{bmatrix}=\begin{bmatrix} -1 \\ -1 \end{bmatrix}+\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}+\begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{1}{3} & 0 & -1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$=\mathbf{c}_{(2\times1)}+\mathbf{H}_{(2\times2)}\mathbf{d}_{(2\times1)}+\mathbf{A}_{(2\times3)}\mathbf{u}_{(3\times1)}=\mathbf{0}$$

$$\begin{bmatrix} \frac{1}{3}(d_1+d_2-2)+s_1 \\ -d_1-1+s_2 \\ -d_2-1+s_3 \end{bmatrix}=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}+\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}-\begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$

$$=\mathbf{A}^T{}_{(3\times2)}\mathbf{d}_{(2\times1)}+\mathbf{s}_{(3\times1)}-\mathbf{b}_{(3\times1)}=\mathbf{0}$$

Since the design variables $\mathbf{d}_{(n\times1)}$ are free in sign, we may decompose them as follows for using the Simplex method.

$$\mathbf{d}_{(2\times1)}=\mathbf{d}^+_{(2\times1)}-\mathbf{d}^-_{(2\times1)}$$

Matrix form

$$\underbrace{\begin{bmatrix} \mathbf{H}_{(2\times2)} & -\mathbf{H}_{(2\times2)} & \mathbf{A}_{(2\times3)} & \mathbf{0}_{(2\times3)} \\ \mathbf{A}^T{}_{(3\times2)} & -\mathbf{A}^T{}_{(3\times2)} & \mathbf{0}_{(3\times3)} & \mathbf{I}_{(3\times3)} \end{bmatrix}}_{=\mathbf{B}_{(5\times10)}}\underbrace{\begin{bmatrix} \mathbf{d}^+_{(2\times1)} \\ \mathbf{d}^-_{(2\times1)} \\ \mathbf{u}_{(3\times1)} \\ \mathbf{s}_{(3\times1)} \end{bmatrix}}_{=\mathbf{X}_{(10\times1)}}=\underbrace{\begin{bmatrix} -\mathbf{c}_{(2\times1)} \\ \mathbf{b}_{(3\times1)} \end{bmatrix}}_{=\mathbf{D}_{(5\times1)}}$$

- Determine the Search Direction by using the Simplex Method [Iteration 1] (5)

Kuhn-Tucker necessary condition : $\nabla L(\mathbf{d}^+, \mathbf{d}^-, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\underbrace{\left[\begin{array}{c:c:c:c} \mathbf{H}_{(2\times2)} & -\mathbf{H}_{(2\times2)} & \mathbf{A}_{(2\times3)} & \mathbf{0}_{(2\times3)} \\ \hdashline \mathbf{A}^T{}_{(3\times2)} & -\mathbf{A}^T{}_{(3\times2)} & \mathbf{0}_{(3\times3)} & \mathbf{I}_{(3\times3)} \end{array}\right]}_{=\mathbf{B}_{(5\times10)}} \underbrace{\begin{bmatrix} \mathbf{d}^+_{(2\times1)} \\ \mathbf{d}^-_{(2\times1)} \\ \mathbf{u}_{(3\times1)} \\ \mathbf{s}_{(3\times1)} \end{bmatrix}}_{=\mathbf{X}_{(10\times1)}} = \underbrace{\begin{bmatrix} -\mathbf{c}_{(2\times1)} \\ \mathbf{b}_{(3\times1)} \end{bmatrix}}_{=\mathbf{D}_{(5\times1)}}$$

where $\mathbf{d}^+_{(2\times1)} = \begin{bmatrix} d_1^+ \\ d_2^+ \end{bmatrix}, \mathbf{d}^-_{(2\times1)} = \begin{bmatrix} d_1^- \\ d_2^- \end{bmatrix}, \mathbf{c}_{(2\times1)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{H}_{(2\times2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A}_{(2\times3)} = \begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{1}{3} & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)} = \begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$

$$\mathbf{B}_{(5\times10)} = \left[\begin{array}{cc:cc:ccc:ccc} 1 & 0 & -1 & 0 & \frac{1}{3} & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{3} & 0 & -1 & 0 & 0 & 0 \\ \hdashline \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\mathbf{X}^T{}_{(1\times10)} = \begin{bmatrix} d_1^+ & d_2^+ & d_1^- & d_2^- & u_1 & u_2 & u_3 & s_1 & s_2 & s_3 \end{bmatrix}, \mathbf{D}^T{}_{(1\times5)} = \begin{bmatrix} 1 & 1 & \frac{2}{3} & 1 & 1 \end{bmatrix}$$

Kuhn-Tucker necessary condition(matrix form)

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)} = \mathbf{D}_{(5\times1)}$$

$$\begin{bmatrix} 1 & 0 & -1 & 0 & \frac{1}{3} & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{3} & 0 & -1 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1^+ \\ d_2^+ \\ d_1^- \\ d_2^- \\ u_1 \\ u_2 \\ u_3 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$

We want to find.

➡ **This problem is to find X in the linear programming problem only having the equality constraints.**

➡ $u_i s_i = 0; i = 1\, to\, 3$ **: Check whether the solution obtained from the linear indeterminate equation satisfies the nonlinear indeterminate equation and determine the solution.**

Simplex method to solve the quadratic programming problem

1. The problem to solve the Kuhn-Tucker necessary condition is the same with the problem having only the equality constraints(linear programming problem).
2. To solve the linear indeterminate equation, we introduce the artificial variables, define the artificial objective function and determine the initial basic feasible solution by using the Simplex method.

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)} + \underbrace{\mathbf{Y}_{(5\times1)}}_{\text{Artificial variables}} = \mathbf{D}_{(5\times1)}$$

3. The artificial objective function is defined as follows.

$$w = \sum_{i=1}^{5} Y_i = \sum_{i=1}^{5} D_i - \sum_{j=1}^{10}\sum_{i=1}^{5} B_{ij} X_j = w_0 + \sum_{j=1}^{10} C_j X_j$$

where $C_j = -\sum_{i=1}^{5} B_{ij}$: Add the elements of the j th column of the matrix B and change the its sign.(Initial relative objective coefficient).

$$w_0 = \sum_{i=1}^{5} D_i = 1 + 1 + \tfrac{2}{3} + 1 + 1 = \tfrac{14}{3}$$

: Initial value of the artificial objective function (summation of the all elements of the matrix D)

4. Solve the linear programming problem by using the Simplex and check whether the solution satisfies the following equation.

$u_i s_i = 0; i = 1\ to\ 3$: Check whether the solution obtained from the linear indeterminate equation satisfies the nonlinear indeterminate equation and determine the solution.

7.3 Sequential Quadratic Programming (SQP)

- Determine the Search Direction by using the Simplex Method

[Iteration 1](8)

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)}+\underbrace{\mathbf{Y}_{(5\times1)}}_{\text{Artificial variables}}=\mathbf{D}_{(5\times1)} \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 & -1 & 0 & \frac{1}{3} & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{3} & 0 & -1 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1^+(=X_1) \\ d_2^+(=X_2) \\ d_1^-(=X_3) \\ d_2^-(=X_4) \\ u_1(=X_5) \\ u_2(=X_6) \\ u_3(=X_7) \\ s_1(=X_8) \\ s_2(=X_9) \\ s_3(=X_{10}) \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$

Define the artificial objective function for using the Simplex method

Sum the all column(1~5): $\frac{1}{3}X_1+\frac{1}{3}X_2-\frac{1}{3}X_3-\frac{1}{3}X_4+\frac{2}{3}X_5-X_6-X_7+X_8+X_9+X_{10}+\underbrace{Y_1+Y_2+Y_3+Y_4+Y_5}_{w}=\frac{14}{3}$

Replace the summation of the all artificial to w and rearrange: $-\frac{1}{3}X_1-\frac{1}{3}X_2+\frac{1}{3}X_3+\frac{1}{3}X_4-\frac{2}{3}X_5+X_6+X_7-X_8-X_9-X_{10}=w-\frac{14}{3}$

1

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
Y1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
Y2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
Y3	1/3	1/3	-1/3	-1/3	0	0	0	1	0	0	0	0	1	0	0	2/3	**2/3**
Y4	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	-
Y5	0	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	1	-
A. Obj.	-1/3	-1/3	1/3	1/3	-2/3	1	1	**-1**	-1	-1	0	0	0	0	0	w-14/3	-

Artificial objective function

Sum all the elements of the row and change the its sign (ex. 1 row: -(1+0+1/3-1+0)=-1/3)

7.3 Sequential Quadratic Programming (SQP)

- Determine the Search Direction by using the Simplex Method [Iteration 1] (9)

2

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
Y1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
Y2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
X8	1/3	1/3	-1/3	-1/3	0	0	0	1	0	0	0	0	1	0	0	2/3	-
Y4	-1	0	1	0	0	0	0	0	**1**	0	0	0	0	1	0	1	**1**
Y5	0	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	1	-
A. Obj.	0	0	0	0	-2/3	1	1	0	**-1**	-1	0	0	1	0	0	w-4	-

3

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
Y1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
Y2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
X8	1/3	1/3	-1/3	-1/3	0	0	0	1	0	0	0	0	1	0	0	2/3	-
X9	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	-
Y5	0	-1	0	1	0	0	0	0	0	**1**	0	0	0	0	1	1	**1**
A. Obj.	-1	0	1	0	-2/3	1	1	0	0	**-1**	0	0	1	1	0	w-3	-

4

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
Y1	**1**	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	**1**
Y2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
X8	1/3	1/3	-1/3	-1/3	0	0	0	1	0	0	0	0	1	0	0	2/3	2
X9	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1	-
X10	0	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	1	-
A. Obj.	**-1**	-1	1	1	-2/3	1	1	0	0	0	0	0	1	1	1	w-2	-

7.3 Sequential Quadratic Programming (SQP)

- Determine the Search Direction by using the Simplex Method [Iteration 1] (10)

5

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
Y2	0	**1**	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	**1**
X8	0	1/3	0	-1/3	-1/9	1/3	0	1	0	0	-1/3	0	1	0	0	1/3	1
X9	0	0	0	0	1/3	-1	0	0	1	0	1	0	0	1	0	2	-
X10	0	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	1	-
A. Obj.	0	**-1**	0	1	-1/3	0	1	0	0	0	1	0	1	1	1	w-1	-

6

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
X2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
X8	0	0	0	0	-2/9	1/3	1/3	1	0	0	-1/3	-1/3	1	0	0	0	-
X9	0	0	0	0	1/3	-1	0	0	1	0	1	0	0	1	0	2	-
X10	0	0	0	0	1/3	0	-1	0	0	1	0	1	0	0	1	2	-
A. Obj.	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	w-0	-

Since the value of the objective function becomes zero, the initial basic feasible solution is obtained.

6

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X1	1	0	-1	0	1/3	-1	0	0	0	0	1	0	0	0	0	1	-
X2	0	1	0	-1	1/3	0	-1	0	0	0	0	1	0	0	0	1	-
X8	0	0	0	0	-2/9	1/3	1/3	1	0	0	-1/3	-1/3	1	0	0	0	-
X9	0	0	0	0	1/3	-1	0	0	1	0	1	0	0	1	0	2	-
X10	0	0	0	0	1/3	0	-1	0	0	1	0	1	0	0	1	2	-
A. Obj.	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	w-0	-

Since the value of the objective function becomes zero, the initial basic feasible solution is obtained.

$$\mathbf{X}^T{}_{(1\times 10)} = \begin{bmatrix} d_1^+ & d_2^+ & d_1^- & d_2^- & u_1 & u_2 & u_3 & s_1 & s_2 & s_3 \end{bmatrix}$$

Basic solution:

$$X_1 = 1, \quad X_2 = 1, \quad X_8 = 0, \quad X_9 = 2, \quad X_{10} = 2$$

Nonbasic solution:

$$X_3 = X_4 = X_5 = X_6 = X_7 = 0$$

This solution satisfies the nonlinear indeterminate equation($X_i X_{3+i} = 0; i = 5\ to\ 7, X_i \geq 0; i = 1\ to\ 10$ **)**

So, the optimal solution is $d_1 = d_2 = 1, u_1 = u_2 = u_3 = 0, s_1 = 0, s_2 = s_3 = 2.$

Why are the values of $\mathbf{u_1}$ and $\mathbf{s_1}$ zero at the same time?

➡ In the Pivot step, if the smallest(i.e., the most negative) coefficient of the artificial objective function or the smallest positive ratio“bi/ai” appears more than one time, the initial basic feasible solution can be changed by depending on the selection of the pivot element in the pivot procedure.

➡ We have to find and check the solution until the nonlinear indeterminate equation(u_i*s_i=0) is satisfied.

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

The optimal solution in this problem is $d_1 = d_2 = 1, u_1 = u_2 = u_3 = 0, s_1 = 0, s_2 = s_3 = 2$ **.**

Why are the values of $\mathbf{u_1}$ **and** $\mathbf{s_1}$ **are zero at the same time?**

Quadratic Programming Problem

Minimize $\bar{f} = (-d_1 - d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $\frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3}$

$-d_1 \le 1$

$-d_2 \le 1$

This example is graphical displayed as the right side.

$\longrightarrow s_1 = 0$

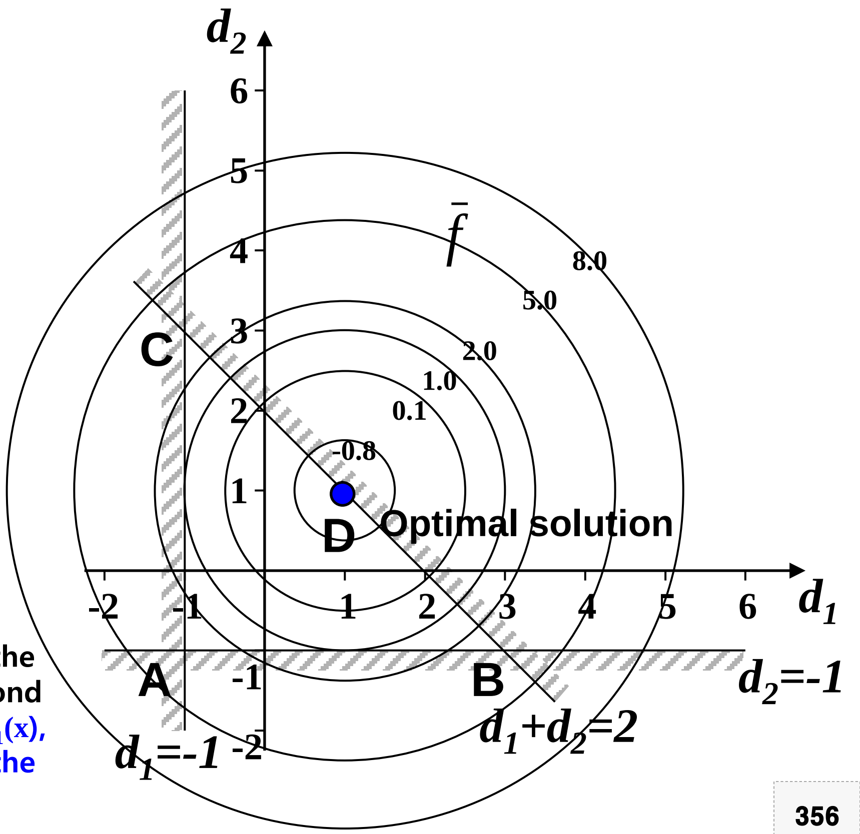
The optimal solution is on the linearized constraint($g_1(\mathbf{x})$, $d_1+d_2=2$).

$\longrightarrow u_2 = u_3 = 0$

The optimal solution is not in the region satisfying the inequality constraint.

$\longrightarrow u_1 = 0$

The optimal solution is on the inequality constraint($g_1(\mathbf{x})$) and is equal to the optimal solution of the objective function to be approximated to the second order. Therefore, although we do not consider the inequality constraint $g_1(\mathbf{x})$, the optimal solution of QP problem is not changed.($g_1(\mathbf{x})$ does not affect the optimal solution of this problem.)



[Iteration 2](1)

Quadratic programming problem
- Objective function: quadratic form
- Constraint: linear form

Solve the QP problem to determine the search direction($d^{(0)}$)

Constrained Optimal Design Problem (Original problem)

Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_1 \le 0$

①

$f(1.732,1.732) = -3, \nabla f = (-1.732, -1.732)$

$g_1(1.732,1.732) = -5.866\times10^{-5}, \nabla g_1 = (0.577, 0.577)$

$g_2(1.732,1.732) = -1.732, \nabla g_2 = (-1,0)$

$g_3(1.732,1.732) = -1.732, \nabla g_3 = (0,-1)$

Quadratic Programming Problem

Minimize $\bar{f} = (-1.732d_1 - 1.732d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $0.577d_1 + 0.577d_2 \le 5.866\times10^{-5}$

$-d_1 \le 1.732$

$-d_2 \le 1.732$

where

$d_1 = x_1 - 1.732,$

$d_2 = x_2 - 1.732$

②

Lagrange function

$$L = (-1.732d_1 - 1.732d_2) + 0.5(d_1^2 + d_2^2)$$
$$+ u_1[0.577(d_1 + d_2) - 5.866\times10^{-5} + s_1^2]$$
$$+ u_2(-d_1 - 1.732 + s_2^2)$$
$$+ u_3(-d_2 - 1.732 + s_3^2)$$

③

Kuhn-Tucker necessary condition: $\nabla L(\mathbf{d}, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\frac{\partial L}{\partial d_1} = -1.732 + d_1 + 0.577u_1 - u_2 = 0$$
$$\frac{\partial L}{\partial d_2} = -1.732 + d_2 + 0.577u_1 - u_3 = 0$$
$$\frac{\partial L}{\partial u_1} = 0.577\left(d_1 + d_2\right) - 5.866\times10^{-6} + s_1^2 = 0$$
$$\frac{\partial L}{\partial u_2} = -d_1 - 1.732 + s_2^2 = 0$$
$$\frac{\partial L}{\partial u_3} = -d_2 - 1.732 + s_3^2 = 0$$
$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u \ge 0,\ i = 1,2,3$$

1. Multiply s_i and the both side and replace s_i^2 with s_i'
2. Represent s_i' to s_i for the convenience

[Iteration 2] (2)

Quadratic Programming Problem

Minimize $\bar{f} = (-1.732d_1 - 1.732d_2) + 0.5(d_1^2 + d_2^2)$

Subject to $0.577d_1 + 0.577d_2 \le 0$

$-d_1 \le 1.732$

$-d_2 \le 1.732$

Matrix form

Minimize $\bar{f} = \mathbf{c}^T{}_{(1\times2)}\mathbf{d}_{(2\times1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times2)}\mathbf{H}_{(2\times2)}\mathbf{d}_{(2\times1)}$

Subject to $\mathbf{A}^T{}_{(3\times2)}\mathbf{d}_{(2\times1)} \le \mathbf{b}_{(3\times1)}$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1.732 + d_1 + 0.577u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1.732 + d_2 + 0.577u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = 0.577\left(d_1 + d_2\right) - 5.866\times10^{-6} + s_1 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1.732 + s_2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1.732 + s_3 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u_i, s_i \ge 0,\ i = 1,2,3$$

Assume that $\mathbf{H}_{(2\times2)}$ **is equal to** $\mathbf{I}_{(2\times2)}$**.**

where

$$\mathbf{d}_{(2\times1)} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \mathbf{c}_{(2\times1)} = \begin{bmatrix} -1.732 \\ -1.732 \end{bmatrix}, \mathbf{H}_{(2\times2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_{(2\times3)} = \begin{bmatrix} 0.577 & -1 & 0 \\ 0.577 & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)} = \begin{bmatrix} 0 \\ 1.732 \\ 1.732 \end{bmatrix}$$

[Iteration 2] (3)

$$\mathbf{d}_{(2\times1)} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \mathbf{c}_{(2\times1)} = \begin{bmatrix} -1.732 \\ -1.732 \end{bmatrix}, \mathbf{H}_{(2\times2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_{(2\times3)} = \begin{bmatrix} 0.577 & -1 & 0 \\ 0.577 & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)} = \begin{bmatrix} 0 \\ 1.732 \\ 1.732 \end{bmatrix}$$

Kuhn-Tucker necessary condition

$$\frac{\partial L}{\partial d_1} = -1.732 + d_1 + 0.577u_1 - u_2 = 0$$

$$\frac{\partial L}{\partial d_2} = -1.732 + d_2 + 0.577u_1 - u_3 = 0$$

$$\frac{\partial L}{\partial u_1} = 0.577(d_1 + d_2) - 5.866\times10^{-6} + s_1 = 0$$

$$\frac{\partial L}{\partial u_2} = -d_1 - 1.732 + s_2 = 0$$

$$\frac{\partial L}{\partial u_3} = -d_2 - 1.732 + s_3 = 0$$

$$\frac{\partial L}{\partial s_i} = u_i s_i = 0,\ u_i, s_i \ge 0,\ i = 1,2,3$$

$$\begin{bmatrix} -1.732 + d_1 + 0.577u_1 - u_2 \\ -1.732 + d_2 + 0.577u_1 - u_3 \end{bmatrix} = \begin{bmatrix} -1.732 \\ -1.732 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} 0.577 & -1 & 0 \\ 0.577 & 0 & -1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$= \mathbf{c}_{(2\times1)} + \mathbf{H}_{(2\times2)}\mathbf{d}_{(2\times1)} + \mathbf{A}_{(2\times3)}\mathbf{u}_{(3\times1)} = \mathbf{0}$$

$$\begin{bmatrix} -0.577d_1 - 0.577d_2 + s_1 \\ -d_1 - 1.732 + s_2 \\ -d_2 - 1.732 + s_3 \end{bmatrix} = \begin{bmatrix} 0.577 & 0.577 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} - \begin{bmatrix} 0 \\ 1.732 \\ 1.732 \end{bmatrix}$$

$$= \mathbf{A}^T{}_{(3\times2)}\mathbf{d}_{(2\times1)} + \mathbf{s}_{(3\times1)} - \mathbf{b}_{(3\times1)} = \mathbf{0}$$

Matrix form

Since the design variables $\mathbf{d}_{(n\times1)}$ are free in sign, we may decompose them as follows for using the Simplex method.

$$\mathbf{d}_{(2\times1)} = \mathbf{d}^+_{(2\times1)} - \mathbf{d}^-_{(2\times1)}$$

$$\underbrace{\begin{bmatrix} \mathbf{H}_{(2\times2)} & -\mathbf{H}_{(2\times2)} & \mathbf{A}_{(2\times3)} & \mathbf{0}_{(2\times3)} \\ \mathbf{A}^T{}_{(3\times2)} & -\mathbf{A}^T{}_{(3\times2)} & \mathbf{0}_{(3\times3)} & \mathbf{I}_{(3\times3)} \end{bmatrix}}_{=\mathbf{B}_{(5\times10)}} \underbrace{\begin{bmatrix} \mathbf{d}^+_{(2\times1)} \\ \mathbf{d}^-_{(2\times1)} \\ \mathbf{u}_{(3\times1)} \\ \mathbf{s}_{(3\times1)} \end{bmatrix}}_{=\mathbf{X}_{(10\times1)}} = \underbrace{\begin{bmatrix} -\mathbf{c}_{(2\times1)} \\ \mathbf{b}_{(3\times1)} \end{bmatrix}}_{=\mathbf{D}_{(5\times1)}}$$

- Determine the Search Direction by using the Simplex Method [Iteration 2] (4)

Kuhn-Tucker necessary condition : $\nabla L(\mathbf{d}^+, \mathbf{d}^-, \mathbf{u}, \mathbf{s}) = \mathbf{0}$

$$\underbrace{\left[\begin{array}{c:c:c:c} \mathbf{H}_{(2\times2)} & -\mathbf{H}_{(2\times2)} & \mathbf{A}_{(2\times3)} & \mathbf{0}_{(2\times3)} \\ \hdashline \mathbf{A}^T{}_{(3\times2)} & -\mathbf{A}^T{}_{(3\times2)} & \mathbf{0}_{(3\times3)} & \mathbf{I}_{(3\times3)} \end{array}\right]}_{=\mathbf{B}_{(5\times10)}} \underbrace{\begin{bmatrix} \mathbf{d}^+_{(2\times1)} \\ \mathbf{d}^-_{(2\times1)} \\ \mathbf{u}_{(3\times1)} \\ \mathbf{s}_{(3\times1)} \end{bmatrix}}_{=\mathbf{X}_{(10\times1)}} = \underbrace{\begin{bmatrix} -\mathbf{c}_{(2\times1)} \\ \mathbf{b}_{(3\times1)} \end{bmatrix}}_{=\mathbf{D}_{(5\times1)}}$$

where $\mathbf{d}^+_{(2\times1)} = \begin{bmatrix} d_1^+ \\ d_2^+ \end{bmatrix}, \mathbf{d}^-_{(2\times1)} = \begin{bmatrix} d_1^- \\ d_2^- \end{bmatrix}, \mathbf{c}_{(2\times1)} = \begin{bmatrix} -1.732 \\ -1.732 \end{bmatrix}, \mathbf{H}_{(2\times2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{A}_{(2\times3)} = \begin{bmatrix} 0.577 & -1 & 0 \\ 0.577 & 0 & -1 \end{bmatrix}, \mathbf{b}_{(3\times1)} = \begin{bmatrix} 0 \\ 1.732 \\ 1.732 \end{bmatrix}$

$$\mathbf{B}_{(5\times10)} = \left[\begin{array}{cc:cc:ccc:ccc} 1 & 0 & -1 & 0 & 0.577 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0.577 & 0 & -1 & 0 & 0 & 0 \\ \hdashline 0.577 & 0.577 & -0.577 & -0.577 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\mathbf{X}^T{}_{(1\times10)} = \begin{bmatrix} d_1^+ & d_2^+ & d_1^- & d_2^- & u_1 & u_2 & u_3 & s_1 & s_2 & s_3 \end{bmatrix}, \mathbf{D}^T{}_{(1\times5)} = \begin{bmatrix} 1.732 & 1.732 & 0 & 1.732 & 1.732 \end{bmatrix}$$

- Determine the Search Direction by using the Simplex Method [Iteration 2] (5)

Kuhn-Tucker necessary condition(matrix form)

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)} = \mathbf{D}_{(5\times1)}$$

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 0.577 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0.577 & 0 & -1 & 0 & 0 & 0 \\ 0.577 & 0.577 & -0.577 & -0.577 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1^+ \\ d_2^+ \\ d_1^- \\ d_2^- \\ u_1 \\ u_2 \\ u_3 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1.732 \\ 1.732 \\ 0 \\ 1.732 \\ 1.732 \end{bmatrix}$$

We want to find.

➡ **This problem is to find X in the linear programming problem only having the equality constraints.**

➡ $u_i s_i = 0; i = 1\, to\, 3$ **: Check whether the solution obtained from the linear indeterminate equation satisfies the nonlinear indeterminate equation and determine the solution.**

Simplex method to solve the quadratic programming problem

1. The problem to solve the Kuhn-Tucker necessary condition is the same with the problem having only the equality constraints(linear programming problem).
2. To solve the linear indeterminate equation, we introduce the artificial variables, define the artificial objective function and determine the initial basic feasible solution by using the Simplex method.

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)} + \underline{\mathbf{Y}_{(5\times1)}} = \mathbf{D}_{(5\times1)}$$

Artificial variables

3. The artificial objective function is defined as follows.

$$w = \sum_{i=1}^{5} Y_i = \sum_{i=1}^{5} D_i - \sum_{j=1}^{10}\sum_{i=1}^{5} B_{ij}X_j = w_0 + \sum_{j=1}^{10} C_j X_j$$

where $C_j = -\sum_{i=1}^{5} B_{ij}$: Add the elements of the j th column of the matrix B and change the its sign.(Initial relative objective coefficient).

$$w_0 = \sum_{i=1}^{5} D_i = 1 + 1 + \tfrac{2}{3} + 1 + 1 = \tfrac{14}{3}$$: Initial value of the artificial objective function (summation of the all elements of the matrix D)

4. Solve the linear programming problem by using the Simplex and check whether the solution satisfies the following equation.

$u_i s_i = 0; i = 1\,to\,3$: Check whether the solution obtained from the linear indeterminate equation satisfies the nonlinear indeterminate equation and determine the solution.

- Determine the Search Direction by using the Simplex Method

[Iteration 2] (7)

$$\mathbf{B}_{(5\times10)}\mathbf{X}_{(10\times1)}+\underbrace{\mathbf{Y}_{(5\times1)}}_{\text{Artificial variables}}=\mathbf{D}_{(5\times1)} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & 0 & 0.577 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0.577 & 0 & -1 & 0 & 0 & 0 \\ 0.577 & 0.577 & -0.577 & -0.577 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} d_1^+ \\ d_2^+ \\ d_1^- \\ d_2^- \\ u_1 \\ u_2 \\ u_3 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix}=\begin{bmatrix} 1.732 \\ 1.732 \\ 0 \\ 1.732 \\ 1.732 \end{bmatrix}$$

Define the artificial objective function for using the Simplex method

Sum the all column(1~5): $0.577X_1+0.577X_2-0.577X_3-0.577X_4+1.154X_5-X_6-X_7+X_8+X_9+X_{10}+\underbrace{Y_1+Y_2+Y_3+Y_4+Y_5}_{w}=6.928$

Replace the summation of the all artificial to w and rearrange: $-0.577X_1-0.577X_2+0.577X_3+0.577X_4-1.154X_5+X_6+X_7-X_8-X_9-X_{10}=w-6.928$

1

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
Y1	1	0	-1	0	0.577	-1	0	0	0	0	1	0	0	0	0	1.732	3
Y2	0	1	0	-1	0.577	0	-1	0	0	0	0	1	0	0	0	1.732	3
Y3	0.577	0.577	-0.577	-0.577	0	0	0	1	0	0	0	0	1	0	0	0	-
Y4	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	1.732	-
Y5	0	-1	0	1	0	0	0	0	0	1	0	0	0	0	1	1.732	-
A. Obj.	-0.577	-0.577	0.577	0.577	-1.154	1	1	-1	-1	-1	0	0	0	0	0	w-6.928	-

Artificial objective function

Sum all the elements of the row and change the its sign (ex. 1 row: -(1+0+1/3-1+0)=-1/3)

7.3 Sequential Quadratic Programming (SQP)

- Determine the Search Direction by using the Simplex Method [Iteration 2] (8)

2	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	1.732	0.000	-1.732	0.000	1.000	-1.732	0.000	0.000	0.000	0.000	1.732	0.000	0.000	0.000	0.000	3.000	-1.732
Y2	-1.000	1.000	1.000	-1.000	0.000	1.000	-1.000	0.000	0.000	0.000	-1.000	1.000	0.000	0.000	0.000	0.000	0.000
Y3	0.577	0.577	-0.577	-0.577	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
Y4	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	1.732	1.732
Y5	0.000	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.732	-
A. Obj.	1.423	-0.577	-1.423	0.577	0.000	-1.000	1.000	-1.000	-1.000	-1.000	2.000	0.000	0.000	0.000	0.000	w-3.464	

3	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	0.000	1.732	0.000	-1.732	1.000	0.000	-1.732	0.000	0.000	0.000	0.000	1.732	0.000	0.000	0.000	3.000	-
X3	-1.000	1.000	1.000	-1.000	0.000	1.000	-1.000	0.000	0.000	0.000	-1.000	1.000	0.000	0.000	0.000	0.000	-
Y3	0.000	1.155	0.000	-1.155	0.000	0.577	-0.577	1.000	0.000	0.000	-0.577	0.577	1.000	0.000	0.000	0.000	-
Y4	0.000	-1.000	0.000	1.000	0.000	-1.000	1.000	0.000	1.000	0.000	1.000	-1.000	0.000	1.000	0.000	1.732	-
Y5	0.000	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.732	1.732
A. Obj.	0.000	0.845	0.000	-0.845	0.000	0.423	-0.423	-1.000	-1.000	-1.000	0.577	1.423	0.000	0.000	0.000	w-3.464	

4	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	0.000	1.732	0.000	-1.732	1.000	0.000	-1.732	0.000	0.000	0.000	0.000	1.732	0.000	0.000	0.000	3.000	-
X3	-1.000	1.000	1.000	-1.000	0.000	1.000	-1.000	0.000	0.000	0.000	-1.000	1.000	0.000	0.000	0.000	0.000	-
Y3	0.000	1.155	0.000	-1.155	0.000	0.577	-0.577	1.000	0.000	0.000	-0.577	0.577	1.000	0.000	0.000	0.000	-
Y4	0.000	-1.000	0.000	1.000	0.000	-1.000	1.000	0.000	1.000	0.000	1.000	-1.000	0.000	1.000	0.000	1.732	1.732
X10	0.000	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.732	-
A. Obj.	0.000	-0.155	0.000	0.155	0.000	0.423	-0.423	-1.000	-1.000	0.000	0.577	1.423	0.000	0.000	1.000	w-1.732	

7.3 Sequential Quadratic Programming (SQP)

- Determine the Search Direction by using the Simplex Method [Iteration 2] (9)

5	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	0.000	1.732	0.000	-1.732	1.000	0.000	-1.732	0.000	0.000	0.000	0.000	1.732	0.000	0.000	0.000	3.000	1.732
X3	-1.000	1.000	1.000	-1.000	0.000	1.000	-1.000	0.000	0.000	0.000	-1.000	1.000	0.000	0.000	0.000	0.000	0.000
Y3	0.000	1.155	0.000	-1.155	0.000	0.577	-0.577	1.000	0.000	0.000	-0.577	0.577	1.000	0.000	0.000	0.000	0.000
X9	0.000	-1.000	0.000	1.000	0.000	-1.000	1.000	0.000	1.000	0.000	1.000	-1.000	0.000	1.000	0.000	1.732	-1.732
X10	0.000	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	1.732	-1.732
A. Obj.	0.000	-1.155	0.000	1.155	0.000	-0.577	0.577	-1.000	0.000	0.000	1.577	0.423	0.000	1.000	1.000	w-0.000	

6	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	1.732	0.000	-1.732	0.000	1.000	-1.732	0.000	0.000	0.000	0.000	1.732	0.000	0.000	0.000	0.000	3.000	1.732
X2	-1.000	1.000	1.000	-1.000	0.000	1.000	-1.000	0.000	0.000	0.000	-1.000	1.000	0.000	0.000	0.000	0.000	0.000
Y3	1.155	0.000	-1.155	0.000	0.000	-0.577	0.577	1.000	0.000	0.000	0.577	-0.577	1.000	0.000	0.000	0.000	0.000
X9	-1.000	0.000	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	1.732	-1.732
X10	-1.000	0.000	1.000	0.000	0.000	1.000	-1.000	0.000	0.000	1.000	-1.000	1.000	0.000	0.000	1.000	1.732	-1.732
A. Obj.	-1.155	0.000	1.155	0.000	0.000	0.577	-0.577	-1.000	0.000	0.000	0.423	1.577	0.000	1.000	1.000	w-0.000	

7	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	0.000	0.000	0.000	0.000	1.000	-0.866	-0.866	-1.500	0.000	0.000	0.866	0.866	-1.500	0.000	0.000	3.000	
X2	0.000	1.000	0.000	-1.000	0.000	0.500	-0.500	0.866	0.000	0.000	-0.500	0.500	0.866	0.000	0.000	0.000	
X1	1.000	0.000	-1.000	0.000	0.000	-0.500	0.500	0.866	0.000	0.000	0.500	-0.500	0.866	0.000	0.000	0.000	
X9	0.000	0.000	0.000	0.000	0.000	-0.500	0.500	0.866	1.000	0.000	0.500	-0.500	0.866	1.000	0.000	1.732	
X10	0.000	0.000	0.000	0.000	0.000	0.500	-0.500	0.866	0.000	1.000	-0.500	0.500	0.866	0.000	1.000	1.732	
A. Obj.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	w-0.000	

7.3 Sequential Quadratic Programming (SQP)
- Determine the Search Direction by using the Simplex Method [Iteration 2](10)

7	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Y1	Y2	Y3	Y4	Y5	bi	bi/ai
X5	0.000	0.000	0.000	0.000	1.000	-0.866	-0.866	-1.500	0.000	0.000	0.866	0.866	-1.500	0.000	0.000	3.000	
X2	0.000	1.000	0.000	-1.000	0.000	0.500	-0.500	0.866	0.000	0.000	-0.500	0.500	0.866	0.000	0.000	0.000	
X1	1.000	0.000	-1.000	0.000	0.000	-0.500	0.500	0.866	0.000	0.000	0.500	-0.500	0.866	0.000	0.000	0.000	
X9	0.000	0.000	0.000	0.000	0.000	-0.500	0.500	0.866	1.000	0.000	0.500	-0.500	0.866	1.000	0.000	1.732	
X10	0.000	0.000	0.000	0.000	0.000	0.500	-0.500	0.866	0.000	1.000	-0.500	0.500	0.866	0.000	1.000	1.732	
A. Obj.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	w-0.000	

$$\mathbf{X}^T{}_{(1\times 10)} = \begin{bmatrix} d_1^+ & d_2^+ & d_1^- & d_2^- & u_1 & u_2 & u_3 & s_1 & s_2 & s_3 \end{bmatrix}$$

Basic solution:

$$X_5 = 3, \quad X_2 = 0, \quad X_1 = 0, \quad X_9 = 1.732, \quad X_{10} = 1.732$$

Nonbasic solution:

$$X_3 = X_4 = X_6 = X_7 = X_8 = 0$$

This solution satisfy the nonlinear indeterminate equation($X_i X_{3+i} = 0; i = 5\ to\ 7, X_i \geq 0; i = 1\ to\ 10$).

So, the optimal solution is $d_1 = d_2 = 0,\ u_1 = 3,\ u_2 = u_3 = 0,\ s_1 = 0,\ s_2 = s_3 = 1.732$.

➡ **In the Pivot step, if the smallest(i.e., the most negative) coefficient of the artificial objective function or the smallest positive ratio"bi/ai" appears more than one time, the initial basic feasible solution can be changed by depending on the selection of the pivot element in the pivot procedure.**

➡ **We have to find and check the solution until the nonlinear indeterminate equation(u_i*s_i=0) is satisfied.**

Naval Architecture & Ocean Engineering

7.3 Sequential Quadratic Programming (SQP)

- Summary of the Sequential Quadratic Programming

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

- [Summary] Formulation of the Quadratic Programming Problem to Determine the Search Direction

$$\textit{Minimize} \quad f(\mathbf{x}+\Delta\mathbf{x}) \cong f(\mathbf{x}) + \nabla f^T(\mathbf{x})\Delta\mathbf{x} + 0.5\Delta\mathbf{x}^T\mathbf{H}\Delta\mathbf{x}$$

The second-order Taylor series expansion of the objective function

$$\textit{Subject to} \quad h_j(\mathbf{x}+\Delta\mathbf{x}) \cong h_j(\mathbf{x}) + \nabla h_j^T(\mathbf{x})\Delta\mathbf{x} = 0;\ j = 1\ to\ p$$

The first-order(linear) Taylor series expansion of the equality constraints

$$g_j(\mathbf{x}+\Delta\mathbf{x}) \cong g_j(\mathbf{x}) + \nabla g_j^T(\mathbf{x})\Delta\mathbf{x} \le 0;\ j = 1\ to\ m$$

The first-order(linear) Taylor series expansion of the inequality constraints

Assumption:

$$\bar{f} = f(\mathbf{x}+\Delta\mathbf{x}) - f(\mathbf{x}),\ e_j = -h_j(\mathbf{x}),\ b_j = -g_j(\mathbf{x}),$$

$$c_i = \partial f(\mathbf{x})/\partial x_i,\ n_{ij} = \partial h_j(\mathbf{x})/\partial x_i,\ a_{ij} = \partial g_j(\mathbf{x})/\partial x_i,$$

$$d_i = \Delta x_i$$

Matrix form

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$

: Quadratic objective function

$$\textit{Subject to} \quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$

: Linear equality constraints

$$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$

: Linear inequality constraints

- [Summary] Determination of the Step Size by using the Golden Section Search Method

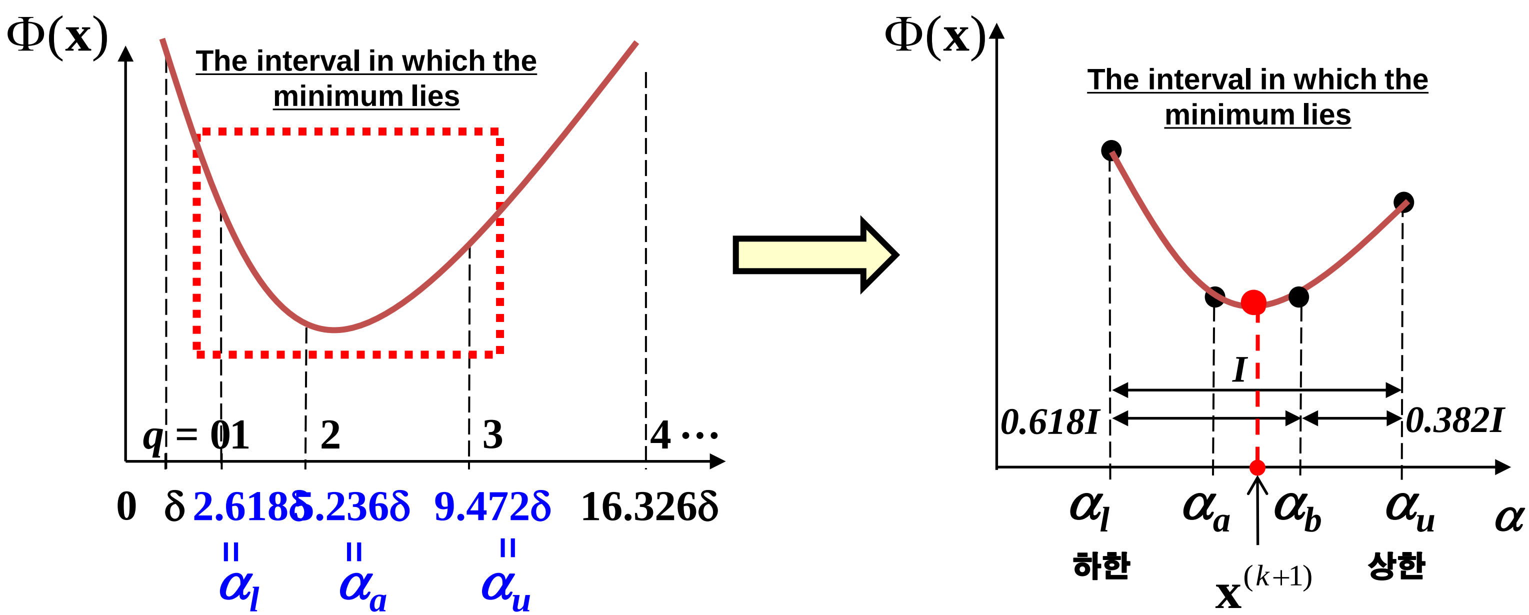
Trial design point for which the descent condition is checked

$$\mathbf{x}^{(k,j)} = \mathbf{x}^{(k)} + \alpha_{(k,j)}\mathbf{d}^{(k)}$$ ➡ How can we determine the value of the $\alpha_{(k,j)}$ to find the improved design point?

Find the improved design point which minimizes the descent function more than the current point by changing $\alpha_{(k,j)}$. (One dimensional search method, such as the Golden section search method, can be used.)

Determination of the improved design point $\mathbf{x}^{(k+1)}$ by using the one dimensional search method such as the Golden section search method($\mathbf{x}^{(k,j)}$ is changed to $\mathbf{x}^{(k+1)}$.)

After finding the interval in which the minimum lies, find the minimum point, x, by reducing the interval(Golden section search method)



- [Summary] Formulation of the Quadratic Programming Problem

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{H}_{(n\times n)}\mathbf{d}_{(n\times 1)}$$

$$\textit{Subject to} \quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$

$$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$

$$\Big\downarrow \ \textbf{Assumption: } \mathbf{H}_{(n\times n)} = \mathbf{I}_{(n\times n)}$$

$$\textit{Minimize} \quad \bar{f} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{I}_{(n\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{c}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)} + \frac{1}{2}\mathbf{d}^T{}_{(1\times n)}\mathbf{d}_{(n\times 1)}$$

$$\textit{Subject to} \quad \mathbf{N}^T{}_{(p\times n)}\mathbf{d}_{(n\times 1)} = \mathbf{e}_{(p\times 1)}$$

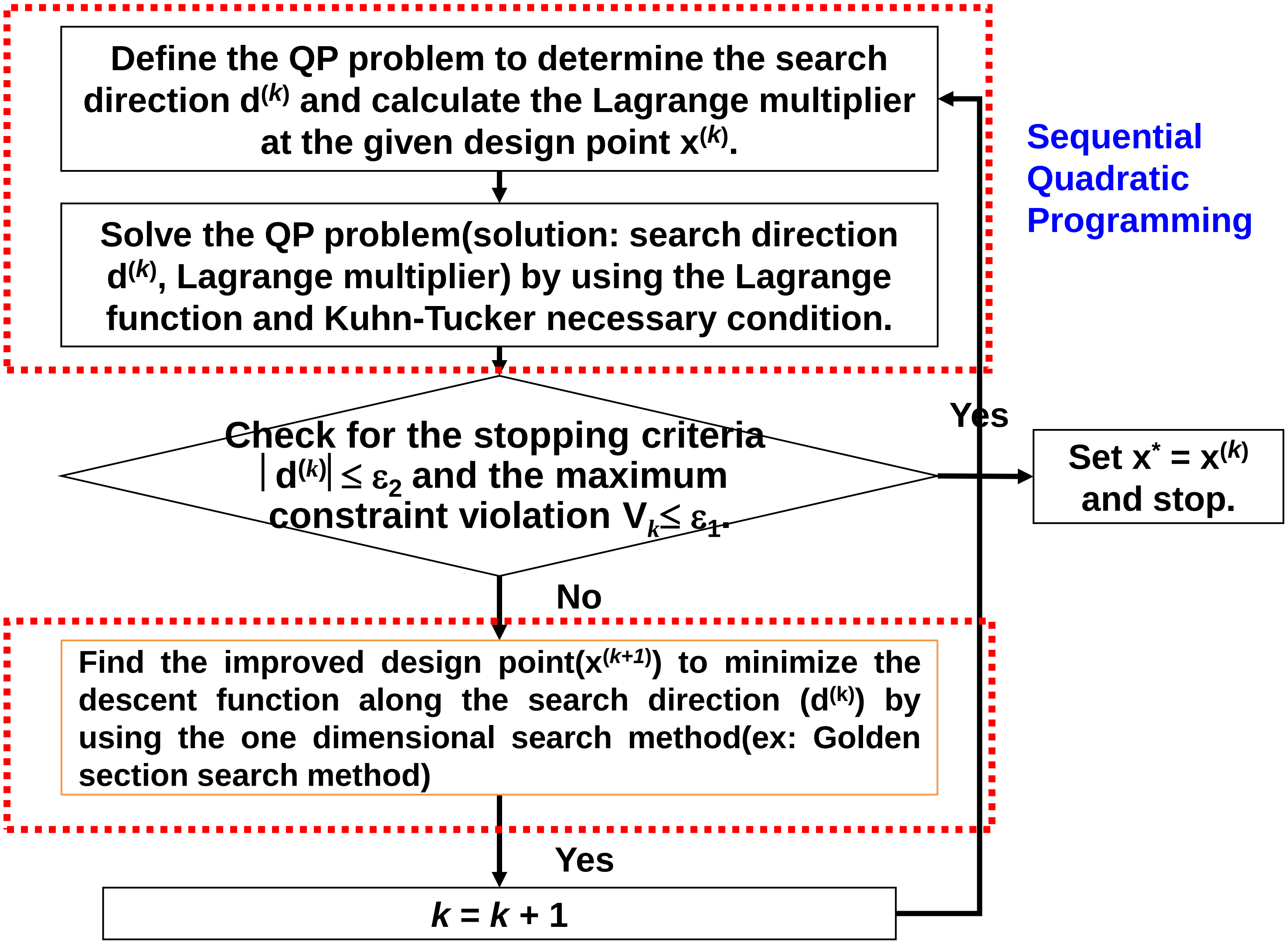
$$\mathbf{A}^T{}_{(m\times n)}\mathbf{d}_{(n\times 1)} \le \mathbf{b}_{(m\times 1)}$$

➡ Since $\mathbf{H}_{(n\times n)} = \mathbf{I}_{(n\times n)}$, the objective function is a quadratic form.

➡ All constraints are linear.

➡ This problem is called the convex programming problem and any local optimum solution is also a global optimum solution.

7.3 Sequential Quadratic Programming (SQP)
- Flow Diagram of the SQP Algorithm



☑ **Step 1: Set k=0. Estimate the initial value for the design variables as $\mathbf{x}^{(0)}$. Select an appropriate initial value for the penalty parameter R_0, and two small number ε_1, ε_2 that define the permissible constraint violation and convergence parameter values, respectively.**

☑ **Step 2: At $\mathbf{x}^{(k)}$ compute the objective and constraint functions and their gradient. Calculate the maximum constraint violation V_k.**

☑ **Step 3: Using the objective and constraints function values and their gradients, define the QP problem. Solve the QP problem to obtain the search direction $\mathbf{d}^{(k)}(= \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)})$ and Lagrange multiplier $\mathbf{v}^{(k)}$, $\mathbf{u}^{(k)}$.**

7.3 Sequential Quadratic Programming (SQP)
- Summary of the SQP Algorithm (2)

☑ **Step 4: Check for the stopping criteria $|\mathbf{d}^{(k)}| \le \varepsilon_2$ and the maximum constraint violation $V_k \le \varepsilon_1$. If these criteria are satisfied then stop. Otherwise continue.**

☑ **Step 5: Calculate the sum r_k of the Lagrange multiplier. Set $R = \max\{R_k, r_k\}$.**

☑ **Step 6: Set $\mathbf{x}^{(k,j)} = \mathbf{x}^{(k)} + \alpha_{(k,j)}\mathbf{d}^{(k)}$ where $\alpha = \alpha_{(k,j)}$ is a proper step size. As for the unconstrained problems, the step size can be obtained by minimizing the descent function along the search direction $\mathbf{d}^{(k)}$. The one dimensional search method, such as the Golden section search, can be used to determine a step size.**
(If the one dimensional search method is end, the current design point $\mathbf{x}^{(k,j)}$ is changed to $\mathbf{x}^{(k+1)}$.)

☑ **Step 7: Save the current penalty parameter as R_k = R. Update the iteration counter as $k = k+1$ and go to Step 2.**

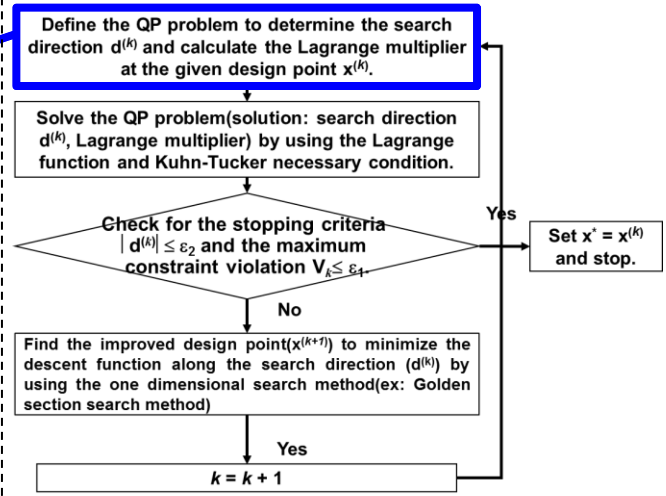
7.3 Sequential Quadratic Programming (SQP)

- Effect of the Starting Point in the SQP Algorithm

The starting point can affect performance of the algorithm.

For example, at some points, the Quadratic Programming problem defined to determine the search direction may not have any solution.
This need not mean that the original problem is infeasible.
The original problem may be highly nonlinear, so that the linearized constraints may be inconsistent giving infeasible Quadratic Programming problem.

This situation can be handled by either temporarily deleting the inconsistent constraints or starting from another point.



Sequential Quadratic Programming(SQP)

7.3 Sequential Quadratic Programming (SQP)

- Use of the Descent Condition for SQP Instead of the Golden Section Search Method

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$
$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$
$g_2(\mathbf{x}) = -x_1 \le 0$
$g_3(\mathbf{x}) = -x_2 \le 0$

(vi) Step 6: By using the one dimensional search method(ex. Descent Condition method) calculate the step size to minimize the descent function along the search direction($\mathbf{d}^{(0)}$) and determine the improved design point.

$$\Phi(\mathbf{x}^{(k,j)}) = f(\mathbf{x}^{(k,j)}) + R_k \cdot V(\mathbf{x}^{(k,j)})$$

$$= x_1^2 + x_2^2 - 3x_1x_2 + 10 \cdot V(\mathbf{x}^{(k,j)})$$

$$V(\mathbf{x}^{(k,j)}) = \max\{0, g_1(\mathbf{x}^{(k,j)}), g_2(\mathbf{x}^{(k,j)}), g_3(\mathbf{x}^{(k,j)})\}, \ (k=0)$$

$$\mathbf{x}^{(k,j)} = \mathbf{x}^{(k)} + t_{(k,j)}\mathbf{d}^{(k)}$$

$\mathbf{x}^{(k,j)}$ ⇨ **k iteration of CSD algorithm**
j iteration of one dimensional search method

$$\mathbf{d}_0 = (1,1) \quad \mathbf{x}^{(0,0)} = (1,1),$$

$$\Phi\left(\mathbf{x}^{(1,j)}\right) = \Phi\left(\mathbf{x}^{(0)} + t_{(0,j)}\mathbf{d}^{(0)}\right) = \boxed{\Phi\left(t_{(0,j)}\right)}$$

$$\boxed{\Phi\left(\mathbf{x}^{(0,0)}\right) - t_{(0,j)}\beta_k} \ \textbf{where, } \beta_k = \gamma\left\|\mathbf{d}^{(k)}\right\|^2, \ (\gamma = 0.5, \text{Defined by user})$$

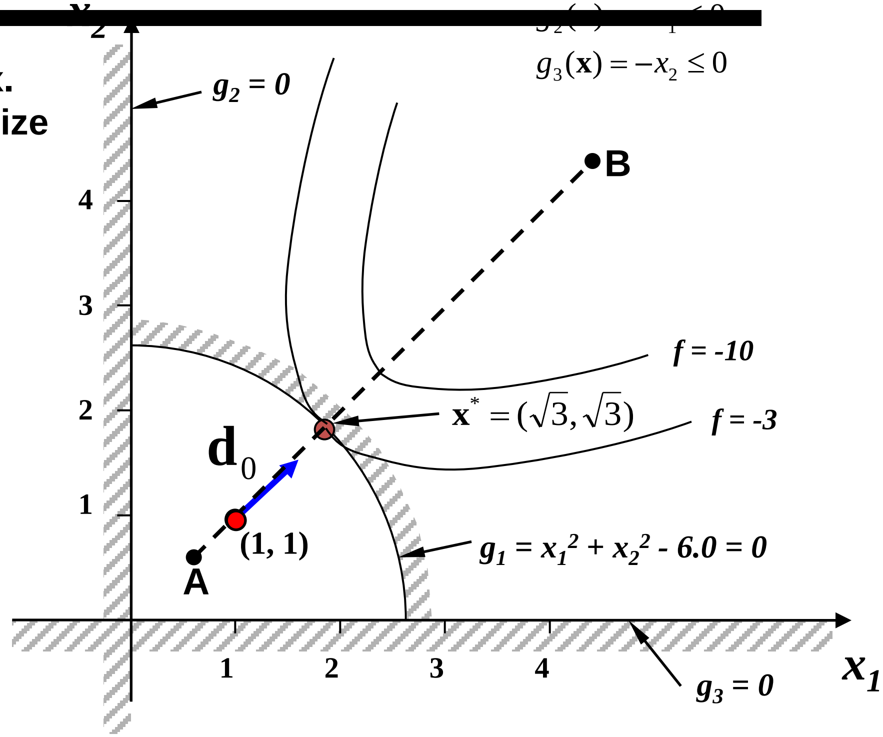
By reducing the value of t from 1 to a half, find the point to satisfy the following equation.

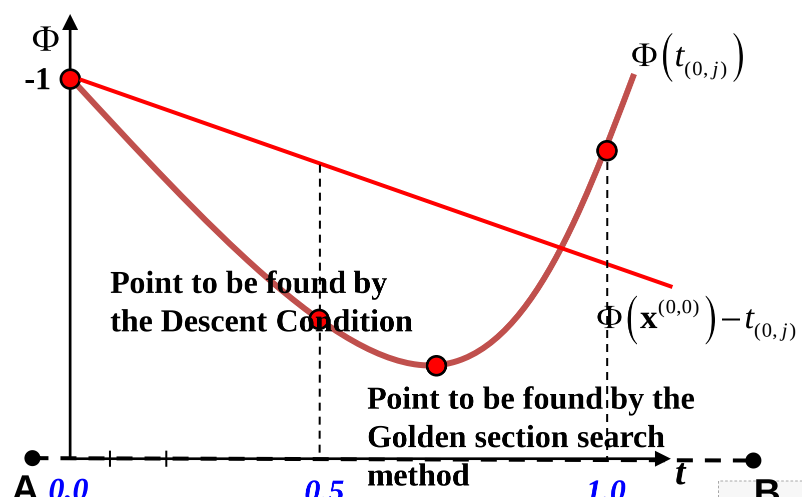
$$\boxed{\Phi\left(t_{(0,j)}\right)} \le \boxed{\Phi\left(\mathbf{x}^{(0,0)}\right) - t_{(0,j)}\beta_k}$$

$$\Phi\left(t_{(0,j)}\right) \le -1 - t_{(0,j)} \quad where,\ \beta_k = \gamma\left\|\mathbf{d}^{(k)}\right\|^2 = 0.5(1^2 + 1^2) = 1$$

$$\Phi\left(\mathbf{x}^{(0,0)}\right) = f\left(\mathbf{x}^{(0,0)}\right) + R_0 \cdot V\left(\mathbf{x}^{(0,0)}\right) = -1 + 10 \times 0 = -1$$

$$V(\mathbf{x}^{(0,0)}) = \max\{0, -\tfrac{2}{3}, -1, -1\} = 0$$





Use of the Descent Condition for SQP Instead of the Golden Section Search Method (2)

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

(vi) Step 6: By using the one dimensional search method(ex. Descent Condition method) calculate the step size to minimize the descent function along the search direction($\mathbf{d}^{(0)}$) and determine the improved design point.

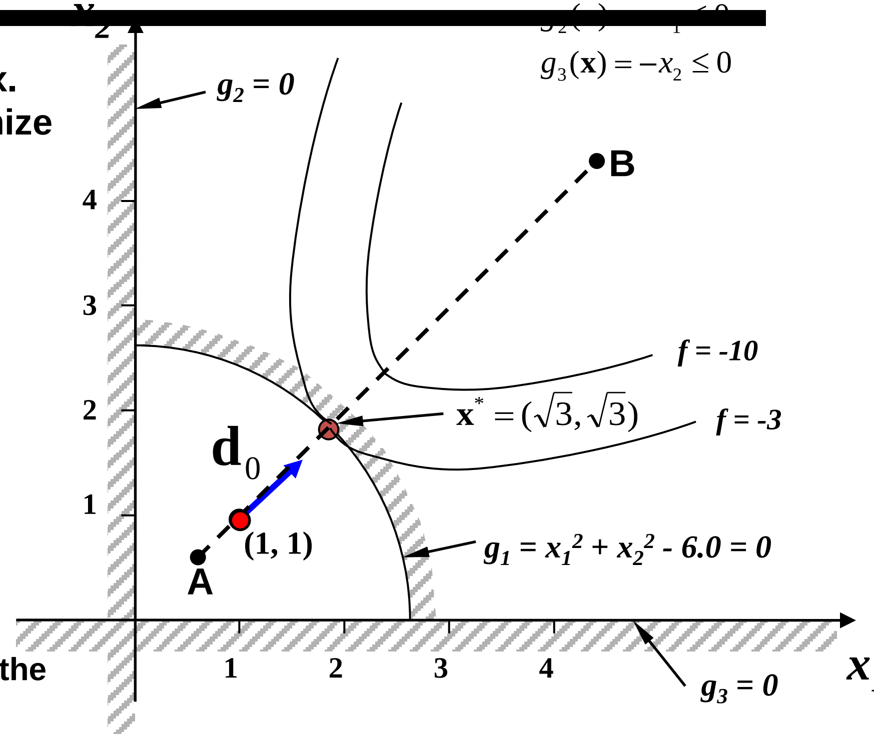
$$\Phi(\mathbf{x}^{(k,j)}) = f(\mathbf{x}^{(k,j)}) + R_k \cdot V(\mathbf{x}^{(k,j)})$$

$$= x_1^2 + x_2^2 - 3x_1x_2 + 10 \cdot V(\mathbf{x}^{(k,j)})$$

$$V(\mathbf{x}^{(k,j)}) = \max\{0, g_1(\mathbf{x}^{(k,j)}), g_2(\mathbf{x}^{(k,j)}), g_3(\mathbf{x}^{(k,j)})\},\ (k=0)$$

$$\mathbf{x}^{(k,j)} = \mathbf{x}^{(k)} + t_{(k,j)}\mathbf{d}^{(k)}$$

$\mathbf{x}^{(k,j)}$ ⇨ **k iteration of CSD algorithm**
j iteration of one dimensional search method



By reducing the value of t from 1 to a half, find the point to satisfy the following equation.

$$\Phi\left(t_{(0,j)}\right) \le -1 - t_{(0,j)} \qquad k = 0,\ j = 0$$

When $t_{(0,j)} = 1$

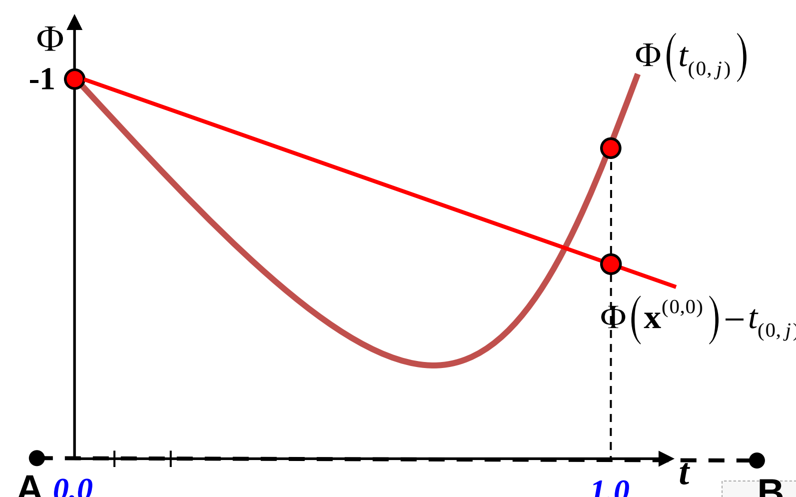
$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + t_{(0,j)} \cdot d^{(0)} = (1,1) + 1 \cdot (1,1) = (2,2)$$

$$\Phi\left(t_{(0,j)}\right) = f\left(2,2\right) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -4 + 10 \times 0.333 = -0.667$$

$$where,\ V(\mathbf{x}^{(0,j)}) = \max\{0, \tfrac{1}{3}, -2, -2\} = 0.333$$

$$-1 - t_{(0,j)} = -1 - 1 = -2$$

If $\Phi\left(t_{(0,j)}\right) \le -1 - t_{(0,j)}$ **is not satisfied, t is reduced to 0.5.**



$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$
$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$
$g_2(\mathbf{x}) = -x_1 \le 0$
$g_3(\mathbf{x}) = -x_2 \le 0$

(vi) Step 6: By using the one dimensional search method(ex. Descent Condition method) calculate the step size to minimize the descent function along the search direction($\mathbf{d}^{(0)}$) and determine the improved design point.

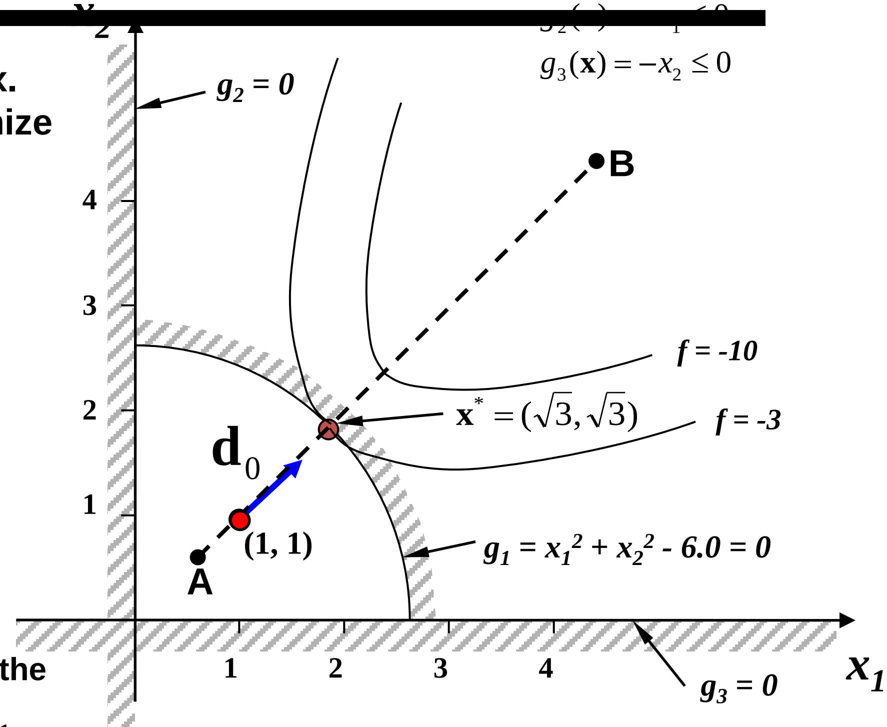
$$\Phi(\mathbf{x}^{(k,j)}) = f(\mathbf{x}^{(k,j)}) + R_k \cdot V(\mathbf{x}^{(k,j)})$$

$$= x_1^2 + x_2^2 - 3x_1x_2 + 10 \cdot V(\mathbf{x}^{(k,j)})$$

$$V(\mathbf{x}^{(k,j)}) = \max\{0, g_1(\mathbf{x}^{(k,j)}), g_2(\mathbf{x}^{(k,j)}), g_3(\mathbf{x}^{(k,j)})\}, \ (k=0)$$

$$\mathbf{x}^{(k,j)} = \mathbf{x}^{(k)} + t_{(k,j)}\mathbf{d}^{(k)}$$

$\mathbf{x}^{(k,j)}$ ➡ **k iteration of CSD algorithm**
j iteration of one dimensional search method



By reducing the value of t from 1 to a half, find the point to satisfy the following equation.

$$-2.25 = \Phi\left(t_{(0,j)}\right) \le -1 - t_{(0,j)} = -1.5 \qquad k=0,\ j=1$$

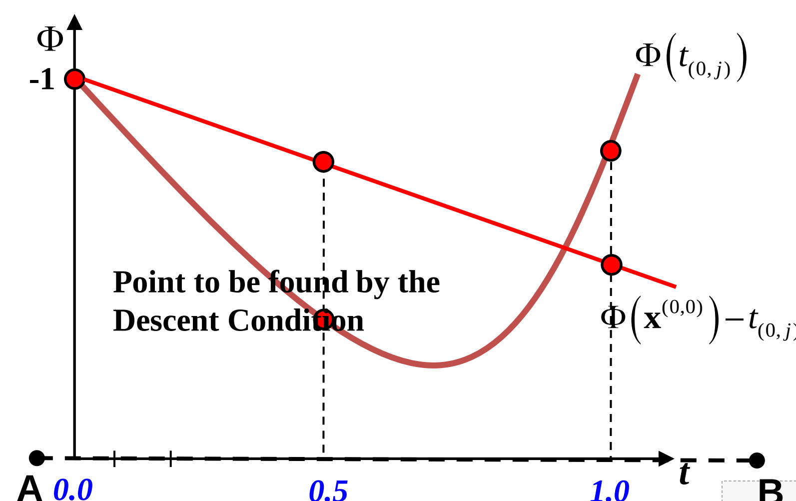
When $t_{(0,j)} = 0.5$

$$\mathbf{x}^{(0,j)} = \mathbf{x}^{(0)} + t_{(0,j)} \cdot d^{(0)} = (1,1) + 0.5 \cdot (1,1) = (1.5, 1.5)$$

$$\Phi\left(t_{(0,j)}\right) = f(1.5, 1.5) + R_0 \cdot V\left(\mathbf{x}^{(0,j)}\right) = -2.25 + 10 \times 0 = -2.25$$

$$where,\ V(\mathbf{x}^{(0,j)}) = \max\{0, -\tfrac{2}{8}, -1.5, -1.5\} = 0$$

$$-1 - t_{(0,j)} = -1 - 0.5 = -1.5$$



Since $\Phi\left(t_{(0,j)}\right) \le -1 - t_{(0,j)}$ **is satisfied, (1.5, 1.5) is the next design point.**

$f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

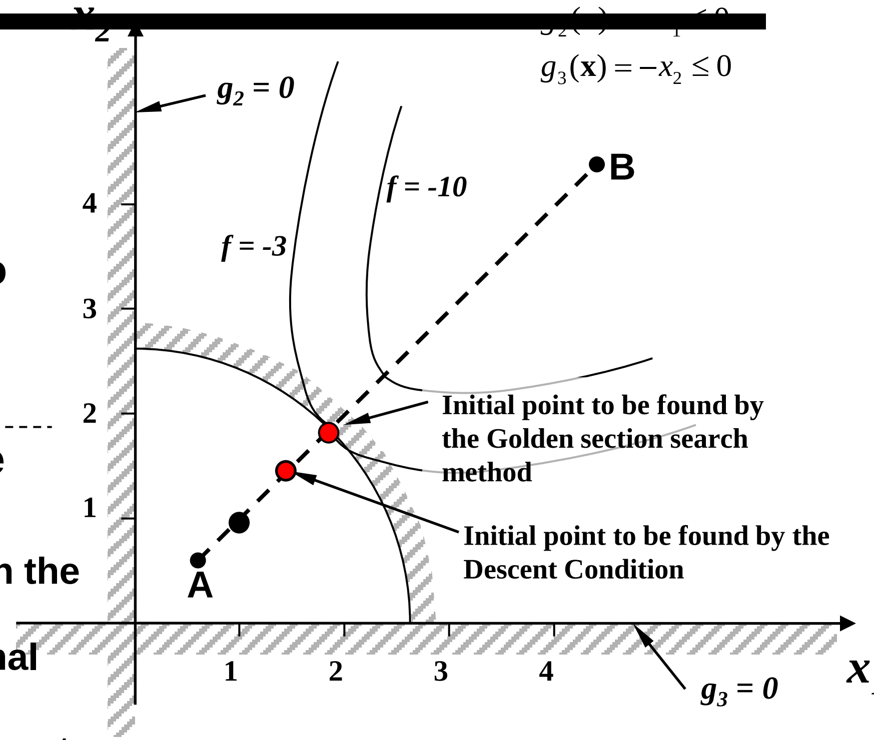
$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_2 \le 0$

The step size obtained by Descent Condition is different from the step size obtained by Golden section search method.

Since the improved design points obtained by two method are different, the number of iteration of defining the QP problem is changed.



If we use the Golden section search method in the right example,

- The number of iteration of the one dimensional search in the first iteration of CSD is 62.

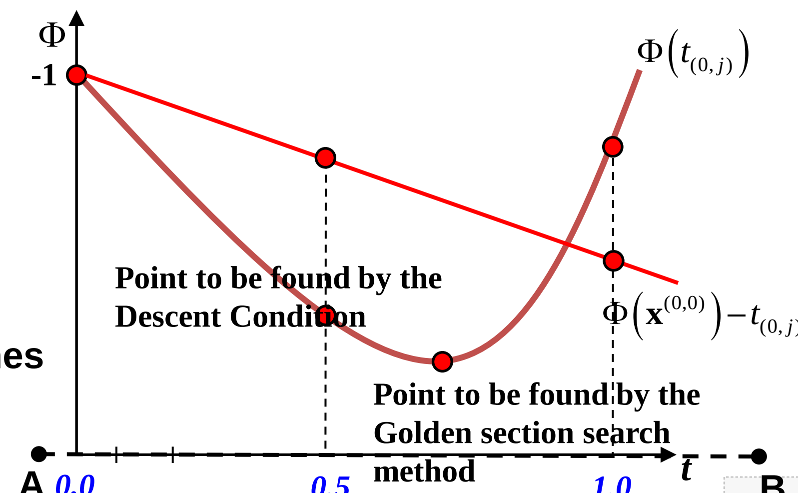
- By defining the QP problem two times, we find the optimal design point.

+ The step size obtained by one dimensional search direction is exact size respectively.

If we use the Descent condition in the right example,

- The number of iteration of the one dimensional search in the first iteration of CSD is 1.

- Since the step size obtained by one dimensional search direction is not exact size, the QP problem is defined in 20 times to find the optimal solution

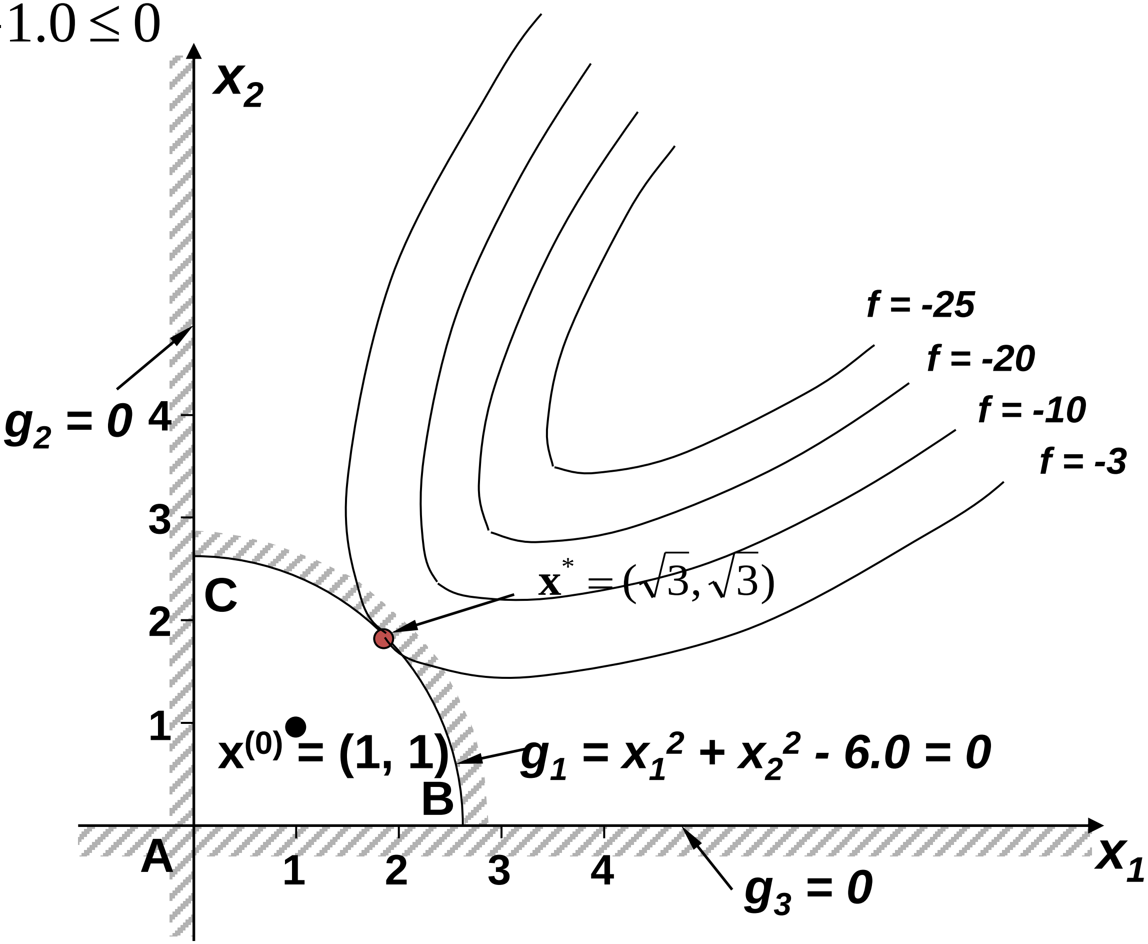


Minimize $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to $g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$

$g_2(\mathbf{x}) = -x_1 \le 0$

$g_3(\mathbf{x}) = -x_1 \le 0$



Optimal Solution:

$$\mathbf{x}^* = (\sqrt{3}, \sqrt{3}),\ f(\mathbf{x}^*) = -3$$

Comparison between the Golden Section Search Method and Descent Condition Method (1)

Minimize: $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$

Subject to:

$$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$$

$$g_2(\mathbf{x}) = -x_1 \le 0$$

$$g_3(\mathbf{x}) = -x_1 \le 0$$

Solution: $\mathbf{x} = (\sqrt{3},\ \sqrt{3}),\ \ f(\mathbf{x}) = -3$

Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(1, 1)	Descent Condition	r = 0.0	19	19	(1.732, 1.732)	-3.0
		r = 0.1	19	19	(1.732, 1.732)	-3.0
		r = 0.5	19	19	(1.732, 1.732)	-3.0
		r = 0.9	19	19	(1.732, 1.732)	-3.0
	Golden section search method		1	62	(1.732, 1.732)	-3.0
(0.1, 0.1)	Descent Condition	r = 0.0	35	85	(1.732, 1.732)	-3.0
		r = 0.1	36	52	(1.732, 1.732)	-3.0
		r = 0.5	29	44	(1.732, 1.732)	-3.0
		r = 0.9	44	124	(1.732, 1.732)	-3.0
	Golden section search method		1	38	(1.732, 1.732)	-3.0
(1.5, 1.5)	Descent Condition	r = 0.0	18	18	(1.732, 1.732)	-3.0
		r = 0.1	18	18	(1.732, 1.732)	-3.0
		r = 0.5	18	18	(1.732, 1.732)	-3.0
		r = 0.9	18	18	(1.732, 1.732)	-3.0
	Golden section search method		2	68	(1.732, 1.732)	-3.0

Comparison between the Golden Section Search Method and Descent Condition Method (2)

Minimize: $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

Solution: $\mathbf{x} = (-1.0,\ 1.5),\ \ f(\mathbf{x}) = -1.25$

Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(0, 0)	Descent Condition	r = 0.0	39	59	(-1.0, 1.5)	-1.25
		r = 0.1	38	58	(-1.0, 1.5)	-1.25
		r = 0.5	41	67	(-1.0, 1.5)	-1.25
		r = 0.9	60	127	(-1.0, 1.5)	-1.25
	Golden section search method		17	329	(-1.0, 1.5)	-1.25
(1, 1)	Descent Condition	r = 0.0	40	63	(-1.0, 1.5)	-1.25
		r = 0.1	40	63	(-1.0, 1.5)	-1.25
		r = 0.5	40	66	(-1.0, 1.5)	-1.25
		r = 0.9	72	194	(-1.0, 1.5)	-1.25
	Golden section search method		17	282	(-1.0, 1.5)	-1.25
(-1, 2)	Descent Condition	r = 0.0	35	55	(-1.0, 1.5)	-1.25
		r = 0.1	35	55	(-1.0, 1.5)	-1.25
		r = 0.5	37	61	(-1.0, 1.5)	-1.25
		r = 0.9	66	177	(-1.0, 1.5)	-1.25
	Golden section search method		18	299	(-1.0, 1.5)	-1.25

Minimize

$$f(x_1, x_2) = -\left[25 - (x_1 - 5)^2 - (x_2 - 5)^2\right]$$

Subject to

$$g_1(x_1, x_2) = -32 + 4x_1 + {x_2}^2 \le 0$$

$$g_2(x_1, x_2) = -x_1 \le 0$$

$$g_3(x_1, x_2) = x_1 \le 10$$

$$g_4(x_1, x_2) = -x_2 \le 0$$

$$g_5(x_1, x_2) = x_2 \le 10$$

Solution

$$x_1^* = 4.374, x_2^* = 3.808, f(x_1^*, x_2^*) = -4.815$$

Comparison between the Golden Section Search Method and Descent Condition Method (3)

Minimize: $f(x_1, x_2) = -\left[25 - (x_1 - 5)^2 - (x_2 - 5)^2\right]$

Subject to: $g_1(x_1, x_2) = -32 + 4x_1 + x_2^{\ 2} \le 0$

$g_2(x_1, x_2) = -x_1 \le 0$

$g_3(x_1, x_2) = x_1 \le 10$

$g_4(x_1, x_2) = -x_2 \le 0$

$g_5(x_1, x_2) = x_2 \le 10$

Solution: $\mathbf{x} = (4.374,\ 3.808),\ \ f(\mathbf{x}) = -4.815$

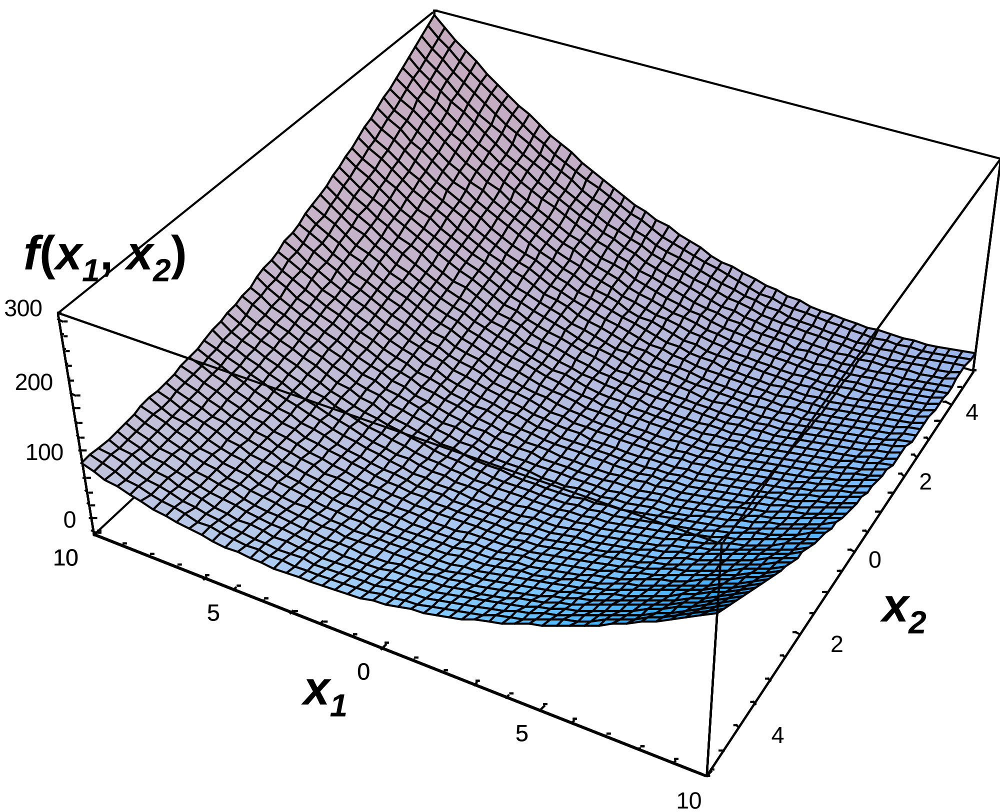
Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(0, 0)	Descent Condition	r = 0.0	22	23	(4.374, 3.808)	-23.188
		r = 0.1	22	23	(4.374, 3.808)	-23.188
		r = 0.5	22	23	(4.374, 3.808)	-23.188
		r = 0.9	22	24	(4.374, 3.808)	-23.188
	Golden section search method		590	13,509	(4.374, 3.808)	-23.188
(7, 1)	Descent Condition	r = 0.0	15	22	(4.374, 3.808)	-23.188
		r = 0.1	15	22	(4.374, 3.808)	-23.188
		r = 0.5	15	22	(4.374, 3.808)	-23.188
		r = 0.9	24	45	(4.374, 3.808)	-23.188
	Golden section search method		1143	26,804	(4.374, 3.808)	-23.188
(-3, -10)	Descent Condition	r = 0.0	19	35	(4.374, 3.808)	-23.188
		r = 0.1	19	35	(4.374, 3.808)	-23.188
		r = 0.5	19	35	(4.374, 3.808)	-23.188
		r = 0.9	28	61	(4.374, 3.808)	-23.188
	Golden section search method		884	20,005	(4.374, 3.808)	-23.188

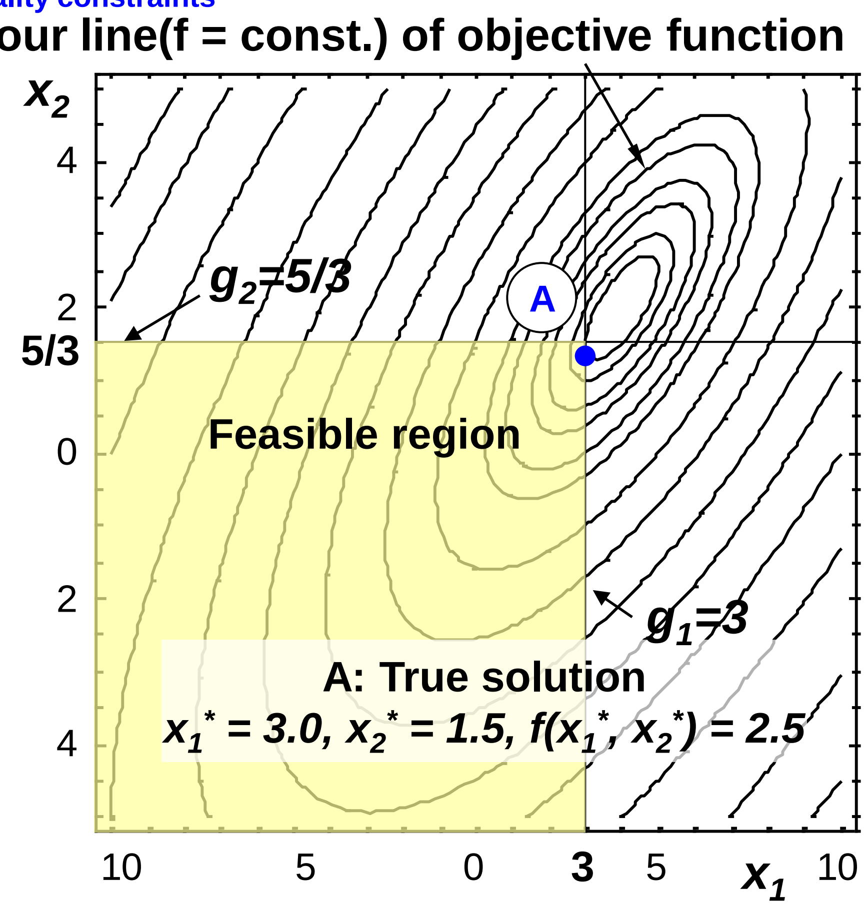
Find $x_1(=B/T), x_2(=1/C_B)$

Minimize $f(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 + 10$

Subject to $g_1(x_1, x_2) = x_1 - 3 \le 0$ ➡ **Optimization problem having two unknown variables and two inequality constraints**

$g_2(x_1, x_2) = x_2 - 5/3 \le 0$





Comparison between the Golden Section Search Method and Descent Condition Method (4)

Minimize: $f(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 + 10$

Subject to: $g_1(x_1, x_2) = x_1 - 3 \le 0$

$g_2(x_1, x_2) = x_2 - 5/3 \le 0$

Solution: $\mathbf{x} = (3.0,\ 1.5),\ \ f(\mathbf{x}) = 2.5$

Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(0, 0)	Descent Condition	r = 0.0	22	24	(3.0, 1.5)	2.5
		r = 0.1	22	24	(3.0, 1.5)	2.5
		r = 0.5	22	26	(3.0, 1.5)	2.5
		r = 0.9	24	33	(3.0, 1.5)	2.5
	Golden section search method		13	203	(3.0, 1.5)	2.5
(2, 1)	Descent Condition	r = 0.0	19	20	(3.0, 1.5)	2.5
		r = 0.1	19	20	(3.0, 1.5)	2.5
		r = 0.5	19	20	(3.0, 1.5)	2.5
		r = 0.9	19	20	(3.0, 1.5)	2.5
	Golden section search method		4	89	(3.0, 1.5)	2.5
(-3, -5)	Descent Condition	r = 0.0	26	52	(3.0, 1.5)	2.5
		r = 0.1	25	28	(3.0, 1.5)	2.5
		r = 0.5	25	28	(3.0, 1.5)	2.5
		r = 0.9	25	30	(3.0, 1.5)	2.5
	Golden section search method		9	255	(3.0, 1.5)	2.5

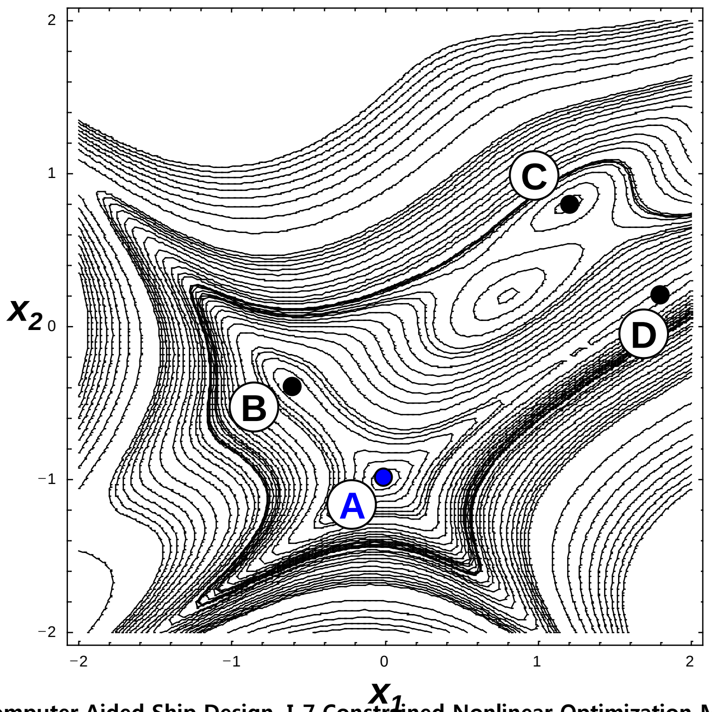
Goldstein-Price Function

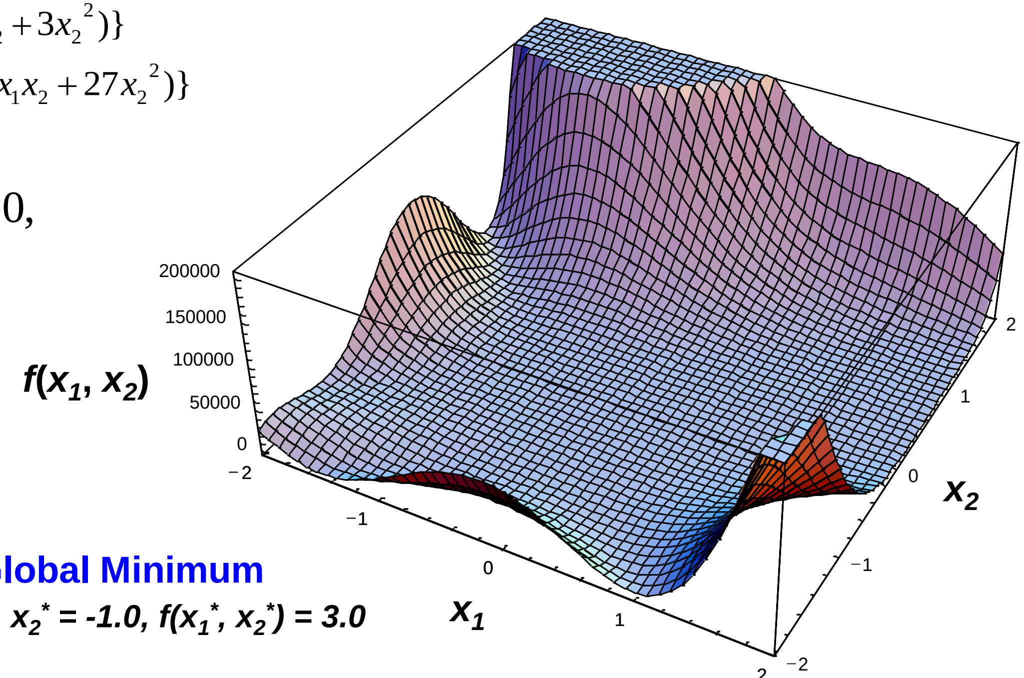
Minimize

$$f(x_1,x_2)=\{1+(x_1+x_2+1)^2\cdot(19-14x_1+3x_1^{\ 2}-14x_2+6x_1x_2+3x_2^{\ 2})\}$$
$$\cdot\{30+(2x_1-3x_2)^2\cdot(18-32x_1+12x_1^{\ 2}+48x_2-36x_1x_2+27x_2^{\ 2})\}$$

Subject to

$$g_1(x_1,x_2)=-2-x_1\le 0,\ g_2(x_1,x_2)=-2-x_2\le 0,$$
$$g_3(x_1,x_2)=x_1-2\le 0,\ g_4(x_1,x_2)=x_2-2\le 0$$





A : Global Minimum

$x_1^* = 0.0,\ x_2^* = -1.0,\ f(x_1^*, x_2^*) = 3.0$

B : Local Minimum

$x_1^* = -0.6,\ x_2^* = -0.4,\ f(x_1^*, x_2^*) = 30.0$

C : Local Minimum

$x_1^* = 1.2,\ x_2^* = 0.8,\ f(x_1^*, x_2^*) = 840.0$

D : Local Minimum

$x_1^* = 1.8,\ x_2^* = 0.2,\ f(x_1^*, x_2^*) = 84.0$

Comparison between the Golden Section Search Method and Descent Condition Method (5)

Minimize:

$$f(x_1,x_2)=\{1+(x_1+x_2+1)^2 \times(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)\} \times\{30+(2x_1-3x_2)^2 \times(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2)\}$$

Subject to:

$$g_1(x_1,x_2)=-2-x_1\le 0,\ g_2(x_1,x_2)=-2-x_2\le 0,$$
$$g_3(x_1,x_2)=x_1-2\le 0,\ g_4(x_1,x_2)=x_2-2\le 0$$

In this example, since there are some local minimum design points, the optimal solution to be obtained is changed depending on the initial design point. So, the calculating the optimal solutions by assuming the initial design point in many times and comparing the results are needed.

Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(0, 0)	Descent Condition	r = 0.0	30	302	(-0.6, -0.4)	30.0
		r = 0.1	26	258	(-0.6, -0.4)	30.0
		r = 0.5	21	208	(-0.6, -0.4)	30.0
		r = 0.9	62	739	(-0.6, -0.4)	30.0
	Golden section search method		15	467	(-0.6, -0.4)	30.0
(2, 3)	Descent Condition	r = 0.0	77	605	(0.0, -1.0)	3.0
		r = 0.1	31	194	(0.0, -1.0)	3.0
		r = 0.5	28	172	(0.0, -1.0)	3.0
		r = 0.9	56	523	(0.0, -1.0)	3.0
	Golden section search method		13	417	(0.0, -1.0)	3.0
(-5, -5)	Descent Condition	r = 0.0	70	545	(0.0, -1.0)	3.0
		r = 0.1	24	135	(0.0, -1.0)	3.0
		r = 0.5	24	136	(0.0, -1.0)	3.0
		r = 0.9	51	459	(0.0, -1.0)	3.0
	Golden section search method		17	497	(0.0, -1.0)	3.0

Rastrigin's Function

Minimize

$$f(x_1, x_2) = 20 + x_1^2 - 10\cos(2\pi \cdot x_1) + x_2^2 - 10\cos(2\pi \cdot x_2)$$

Subject to

$$g_1(x_1, x_2) = -5.12 - x_1 \le 0$$

$$g_2(x_1, x_2) = -5.12 - x_2 \le 0$$

$$g_3(x_1, x_2) = x_1 - 5.12 \le 0$$

$$g_4(x_1, x_2) = x_2 - 5.12 \le 0$$

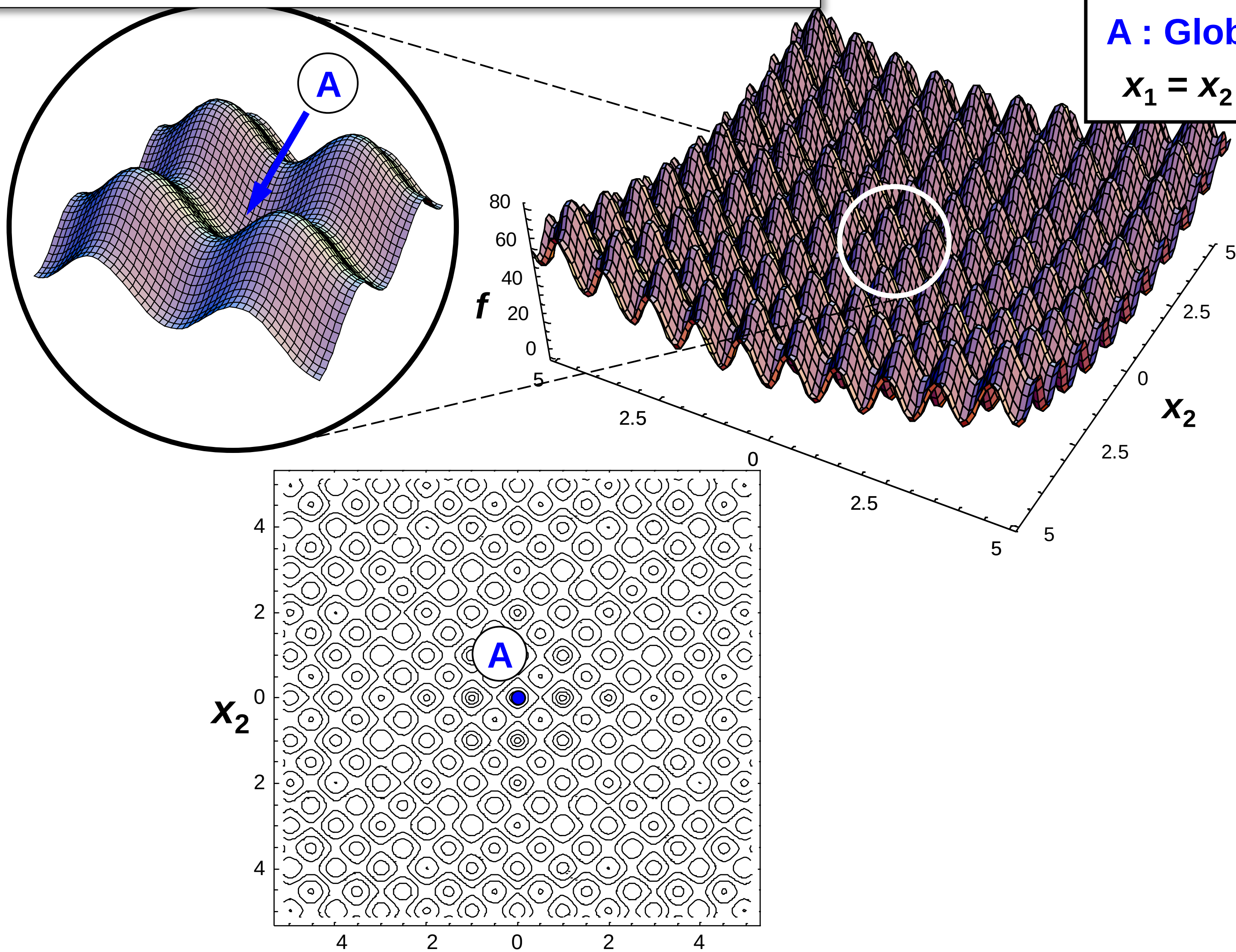
Solution

$$x_1^* = 0.0, x_2^* = 0.0, f(x_1^*, x_2^*) = 0.0$$

Global and Local minimum point of the Rastrigins's Function

A : Global Optimum

$x_1 = x_2 = 0.0, f = 0.0$



Comparison between the Golden Section Search Method and Descent Condition Method (6)

***Minimize*:**

$$f(x_1, x_2) = 20 + x_1^2 - 10\cos(2\pi \cdot x_1) + x_2^2 - 10\cos(2\pi \cdot x_2)$$

***Subject to*:**

$$g_1(x_1, x_2) = -5.12 - x_1 \le 0$$

$$g_2(x_1, x_2) = -5.12 - x_2 \le 0$$

$$g_3(x_1, x_2) = x_1 - 5.12 \le 0$$

$$g_4(x_1, x_2) = x_2 - 5.12 \le 0$$

In this example, since there are some local minimum design points, the optimal solution to be obtained is changed depending on the initial design point. So, the calculating the optimal solutions by assuming the initial design point in many times and comparing the results are needed.

Initial vale	Method		Iteration of defining the QP problem	Iteration of one dimensional search method	Local Optimum Point	Optimum Value
(0.1, 0.1)	Descent Condition	r = 0.0	18	147	(0.0, 0.0)	0.0
		r = 0.1	18	147	(0.0, 0.0)	0.0
		r = 0.5	9	82	(0.0, 0.0)	0.0
		r = 0.9	39	427	(0.0, 0.0)	0.0
	Golden section search method		1	47	(0.0, 0.0)	0.0
(2.1, 2.1)	Descent Condition	r = 0.0	16	134	(1.990, 1.990)	7.960
		r = 0.1	16	134	(1.990, 1.990)	7.960
		r = 0.5	7	69	(1.990, 1.990)	7.960
		r = 0.9	32	358	(1.990, 1.990)	7.960
	Golden section search method		1	45	(1.990, 1.990)	7.960
(-2.1, -3)	Descent Condition	r = 0.0	18	144	(-1.990, -2.985)	12.934
		r = 0.1	18	144	(-1.990, -2.985)	12.934
		r = 0.5	9	82	(-1.990, -2.985)	12.934
		r = 0.9	36	395	(-1.990, -2.985)	12.934
	Golden section search method		7	229	(-1.990, -2.985)	12.934

Step of calculation	Descent Condition method	Golden Section Search method
Iteration number of defining the QP problem	Many	Little
Iteration number of one dimensional search method	Little	Many

☑ Comparison between the Golden Section Search Method and Descent Condition Method

- **When we use the one dimensional search method, we have to calculate the value of the objective function and constraints repetitively.**
- **If it takes much time to calculate the value of the objective function and constraints, the Descent condition method is more useful.**

Computer Aided Ship Design

Part I. Optimization Method

Ch.8 Determination of the Optimum Main Dimensions of a Ship by using an Optimization Method

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.8 Determination of the Optimum Main Dimensions of a Ship by using an Optimization Method

8.1 Owner's Requirements

8.1 Owner's Requirements

☑ **Owner's Requirements**

- **Ship's Type**
- **Deadweight(DWT)**

- **Cargo Hold Capacity(V_{CH})**
  - Cargo Capacity: Cargo Hold Volume / Containers in Hold & on Deck / Car Deck Area.
  - Water Ballast Capacity.

- **Service Speed (V_s)**
  - Service Speed at Draft with Sea Margin, Engine Power & RPM.

- **Dimensional Limitations : Panama canal, Suez canal, Strait of Malacca, St. Lawrence Seaway, Port limitations.**

- **Maximum Draft(T_{max})**

- **Daily Fuel Oil Consumption(DFOC) : Related with ship's economy.**

- **Special Requirements**
  - Ice Class, Air Draft, Bow/Stern Thruster, Special Rudder, Twin Skeg.

- **Delivery Day**
  - Delivery day, with (   )$ delay penalty per day.
  - Abt. 21 months from contract.

- **The Price of a ship**
  - Material & Equipment Cost + Construction Cost + Additional Cost + Margin.

Ch.8 Determination of the Optimum Main Dimensions of a Ship by using an Optimization Method

8.2 Design Model for the Determination of the Optimum Main Dimensions(L,B,D,T,C_B)

This section presents the summary of the design Model for the Determination of the Optimum Main Dimensions. For the detailed description of the design model, please refer to "OCW, 2012 Innovative Ship Design"

Seoul National Univ.
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Design Model for the Determination of the Optimum Main Dimensions(L,B,D,T,C_B)

Find(Design variables) L, B, D, C_B T_d

length breadth depth block coefficient design draft

Given(Owner's requirement) DWT, V_{H_req}, $T_s(=T_{\max})$, V

deadweight Required cargo hold capacity Scantling Draft (maximum) ship speed

Physical constraint

→ **Hydrostatic equilibrium(Weight equation)** – Equality constraint

$$L \cdot B \cdot T_d \cdot C_B \cdot \rho_{sw} \cdot C_\alpha = DWT_{given} + LWT(L, B, D, C_B)$$

$$= DWT_{given} + C_s \cdot L^{1.6}(B+D) + C_o \cdot L \cdot B$$

$$+ C_{power} \cdot (L \cdot B \cdot T_d \cdot C_B)^{2/3} \cdot V^3 \quad \cdots (2.3)$$

Economical constraints(Owner's requirements)

→ **Required cargo hold capacity(Volume equation)** - Equality constraint

$$V_{H\_req} = C_H \cdot L \cdot B \cdot D \cdots (3.1)$$

- DFOC(Daily Fuel Oil Consumption)
  : It is related with the resistance and propulsion.

- Delivery date
  : It is related with the shipbuilding process.

Regulatory constraint

→ **Freeboard regulation(1966 ICLL)** - Inequality constraint

$$D \ge T_s + C_{FB} \cdot D \cdots (4)$$

Objective Function(Criteria to determine the proper main dimensions)

$$Building\ \ Cost = C_{PS} \cdot C_s \cdot L^{1.6}(B+D) + C_{PO} \cdot C_o \cdot L \cdot B + C_{PM} \cdot C_{power} \cdot (L \cdot B \cdot T_d \cdot C_B)^{2/3} \cdot V^3$$

Naval Architecture & Ocean Engineering

Computer Aided Ship Design

Part I. Optimization Method

Ch.9 Determination of Optimal Operating Conditions for the Liquefaction Cycle of the LNG FPSO

September, 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Ch.9 Determination of Optimal Operating Conditions for the Liquefaction Cycle of the LNG FPSO

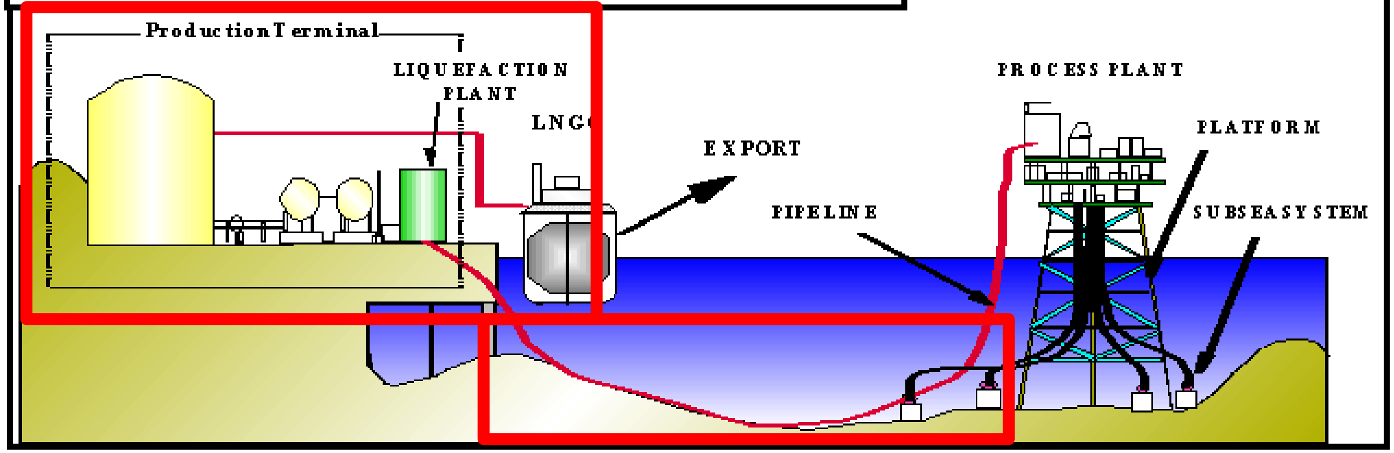
Introduction

9.1. What is the Liquefaction Cycle of a LNG FPSO?

1. What is the Liquefaction Cycle of a LNG FPSO?

Concept of a LNG FPSO(Floating Production Storage Offloading)

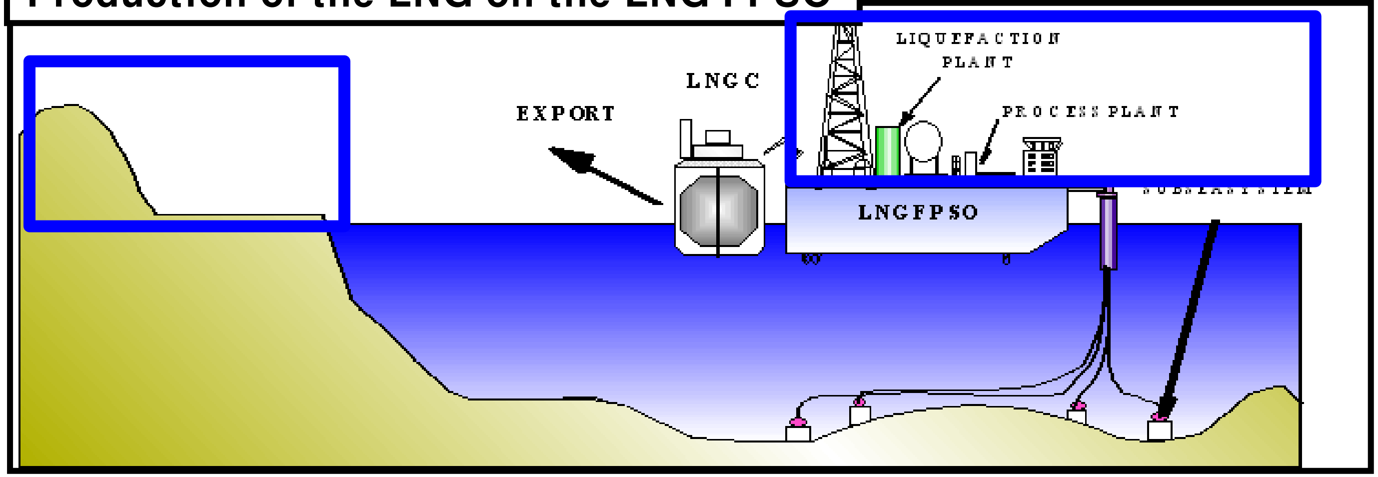
Production of the LNG by onshore facility



Natural gas on the offshore production site is transported by the pipe line to the onshore LNG plant where the NG is liquefied to the LNG.

The LNG FPSO is a floating vessel having the production facility, storage tanks, offloading system for the LNG, and turret system.

Production of the LNG on the LNG FPSO



The natural gas is liquefied direct on the LNG FPSO.
→ Onshore LNG plant and transport pipeline are not needed

[Article] Shell decides to move forward with groundbreaking LNG FPSO

Shell, the world's largest oil company, is now ready to start construction of what will be the world's first LNG FPSO, in a ship yard, Samsung heavy industry, in South Korea.

LNG FPSO cools down the temperature of the natural gas(NG) from 27° C to -162° C to shrink in volume by 600 times.

The liquefaction process system for LNG is most important system of the LNG FPSO.

The World's First LNG FPSO

Reference) [Article]Yonhapnews, SHELL DECIDES TO MOVE FORWARD WITH GROUNDBREAKING FLOATING LNG, 2011. 5. 20

1) MTPA: Million Ton per Annual

1. What is a LNG(Liquefied Natural Gas) FPSO(Floating Production Storage Off-loading) ?

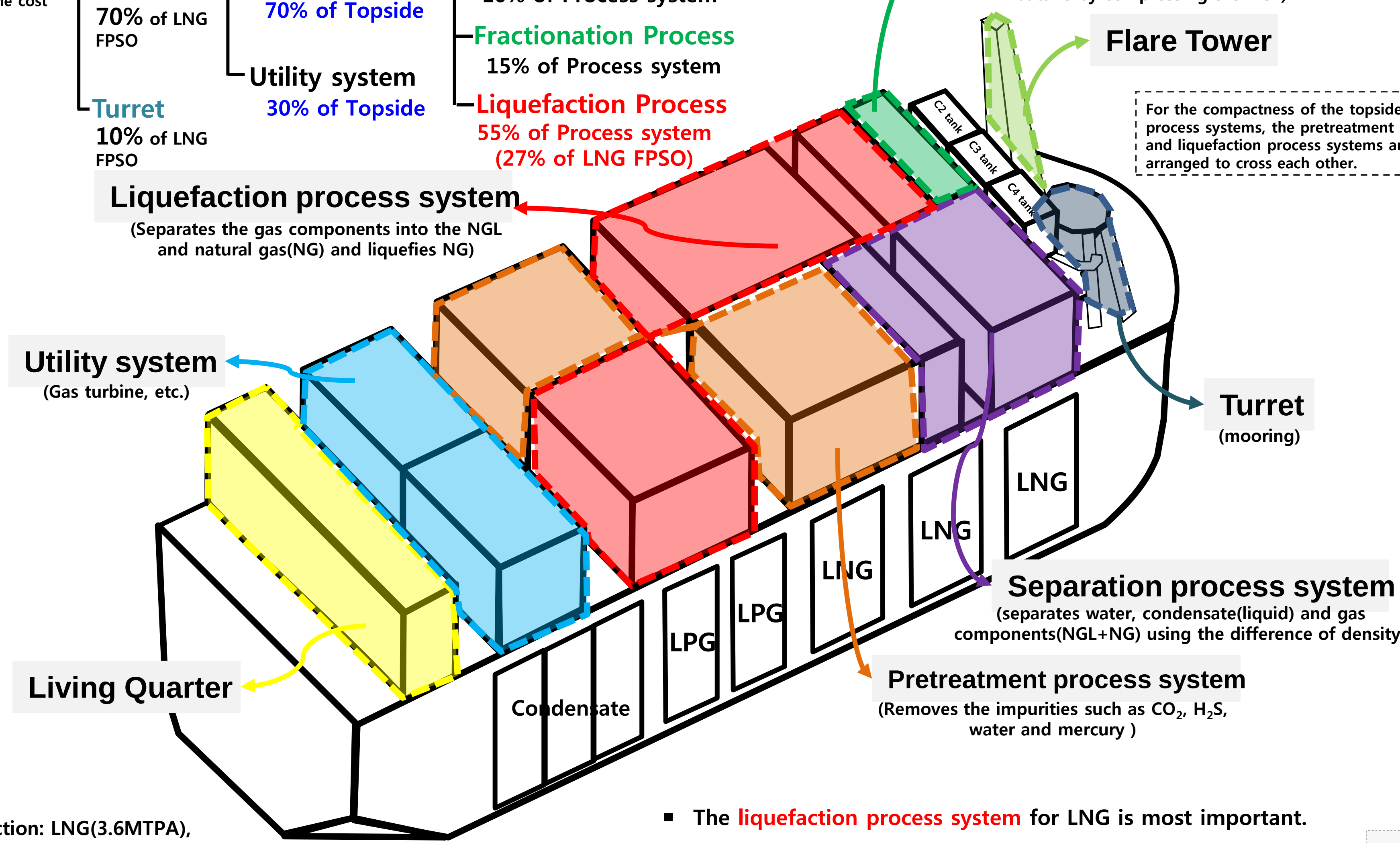
Configuration of a LNG FPSO

- **NG(Natural Gas):** Main component is methane(CH_4).
- **LPG(Liquefied Petroleum Gas):** Main components are propane(C_3H_8) and butane(C_4H_{10}).
- **NGL(Natural Gas Liquids):** Main components are ethane(C_2H_6), propane(C_3H_8) and butane(C_4H_{10}). It exists in the gas phase at 1 atm and 20°C. That is, **NGL = LPG + ethane(C_2H_6)**
- **Condensate:** Main components are Pentane(C_5H_{12}) and Hexane(C_6H_{14}). It exists in the liquid phase at 1 atm and 20°C, so called oil.

LNG FPSO

(Numbers under the parts mean the cost distribution)

- **Hull**
  20% of LNG FPSO
- **Topside**
  70% of LNG FPSO
  - **Process system**
    70% of Topside
    - **Separation process**
      20% of Process system
    - **Pretreatment Process**
      10% of Process system
    - **Fractionation Process**
      15% of Process system
    - **Liquefaction Process**
      55% of Process system
      (27% of LNG FPSO)
  - **Utility system**
    30% of Topside
- **Turret**
  10% of LNG FPSO



- Production: LNG(3.6MTPA),

*MTPA: million ton per annual

- The liquefaction process system for LNG is most important.

Reference) [Article]Yonhapnews, SHELL DECIDES TO MOVE FORWARD WITH GROUNDBREAKING FLOATING LNG, 2011. 5. 20

1. What is the Liquefaction Cycle of the LNG FPSO?

Topside Process Systems of LNG FPSO (1/3)

- **NG(Natural Gas):** Main component is methane(CH_4).
- **LPG(Liquefied Petroleum Gas):** Main components are propane(C_3H_8) and butane(C_4H_{10}).
- **NGL(Natural Gas Liquids):** Main components are ethane(C_2H_6), propane(C_3H_8) and butane(C_4H_{10}). It exists in the **gas** phase at 1 atm and 20°C. That is, **NGL = LPG + ethane(C_2H_6)**
- **Condensate:** Main components are Pentane(C_5H_{12}) and Hexane(C_6H_{14}). It exists in the liquid phase at 1 atm and 20°C, so called **oil**.

: condensate + NGL+ NG
: condensate(C5~C6)
: NGL + NG

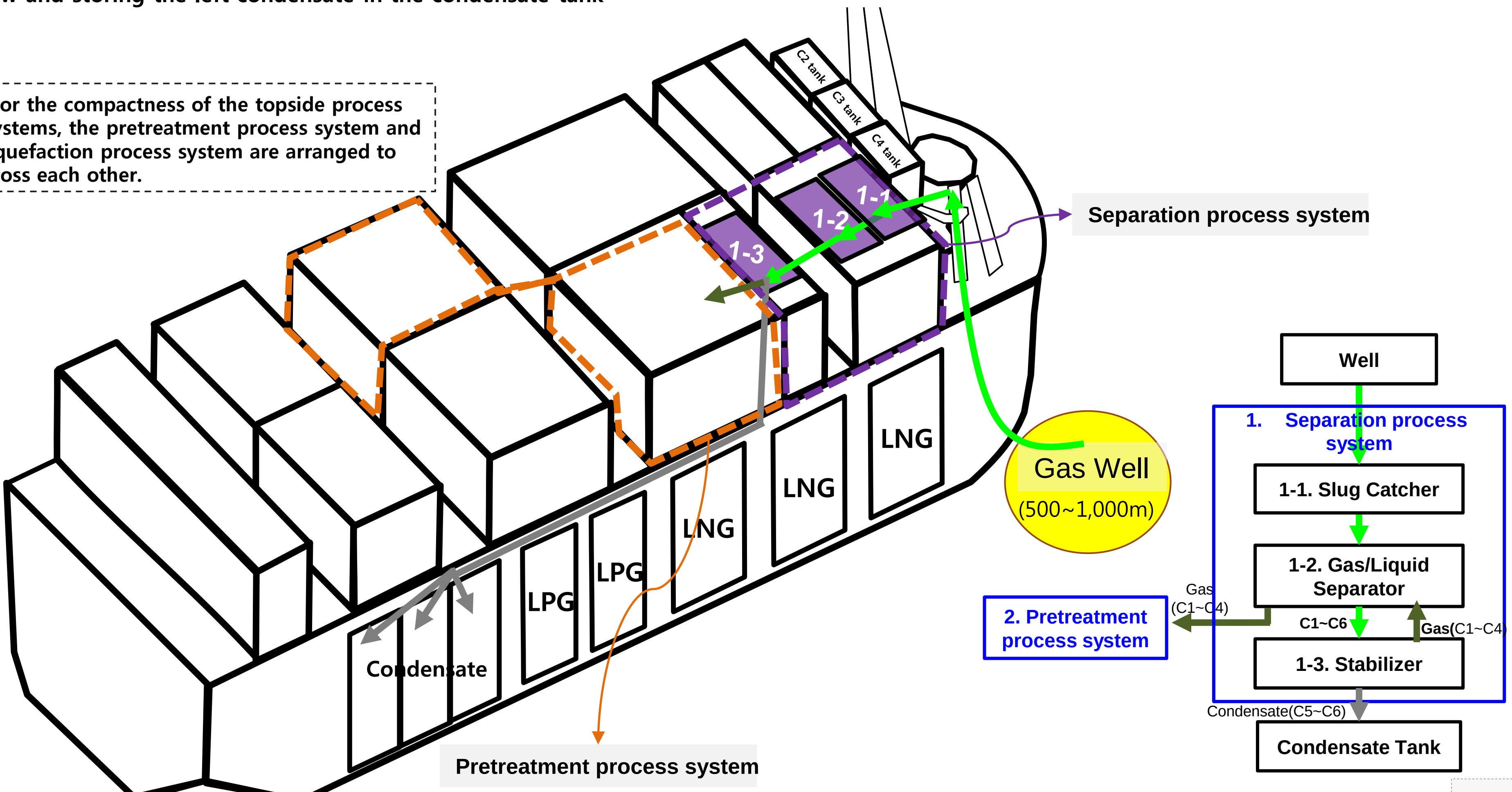
1. Separation process system

1-1. Slug Catcher: Stabilizing Slug Flow from gas well

1-2. Gas/Liquid Separator: Separating water, condensate(liquid) and gas components(NGL+NG) using the difference of density

1-3. Stabilizer: Since a part of gas components is not separated from the condensate completely in the Gas/Liquid separator, separating the gas components from the condensate again, returning it to the gas flow and storing the left condensate in the condensate tank

For the compactness of the topside process systems, the pretreatment process system and liquefaction process system are arranged to cross each other.



1. What is the Liquefaction Cycle of the LNG FPSO?

Topside Process Systems of LNG FPSO (2/3)

- **NG(Natural Gas):** Main component is methane(CH4).
- **LPG(Liquefied Petroleum Gas):** Main components are propane(C3H8) and butane(C4H10).
- **NGL(Natural Gas Liquids):** Main components are ethane(C2H6), propane(C3H8) and butane(C4H10). It exists in the gas phase at 1 atm and 20°C. That is, **NGL = LPG + ethane(C2H6)**
- **Condensate:** Main components are Pentane(C5H12) and Hexane(C6H14). It exists in the liquid phase at 1 atm and 20°C, so called **oil.**

2. Pretreatment process system

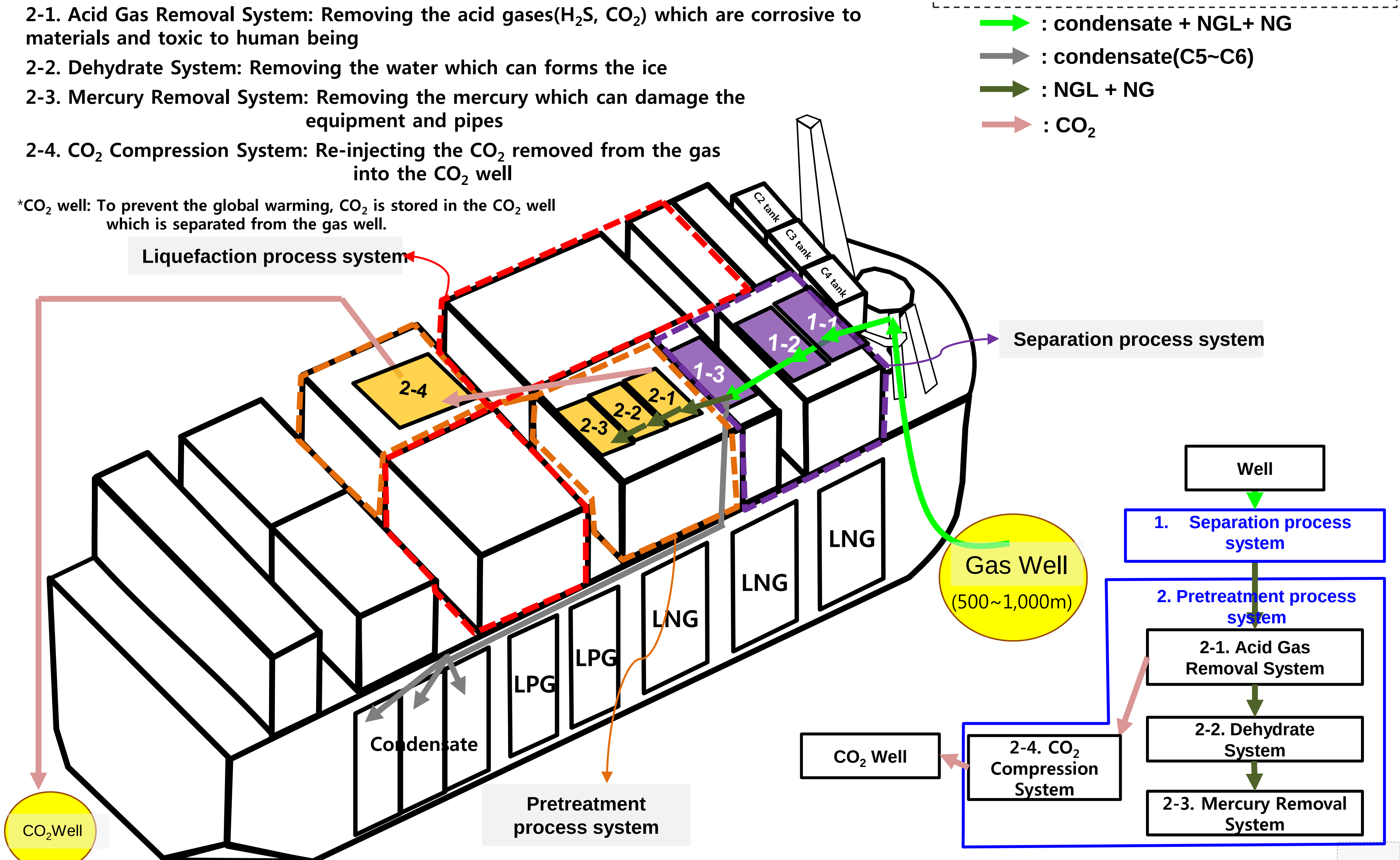
2-1. Acid Gas Removal System: Removing the acid gases(H_2S, CO_2) which are corrosive to materials and toxic to human being

2-2. Dehydrate System: Removing the water which can forms the ice

2-3. Mercury Removal System: Removing the mercury which can damage the equipment and pipes

2-4. CO_2 Compression System: Re-injecting the CO_2 removed from the gas into the CO_2 well

*CO_2 well: To prevent the global warming, CO_2 is stored in the CO_2 well which is separated from the gas well.



1. What is the Liquefaction Cycle of the LNG FPSO?

Topside Process Systems of LNG FPSO (3/3)

- **NG(Natural Gas):** Main component is methane(CH_4).
- **LPG(Liquefied Petroleum Gas):** Main components are propane(C_3H_8) and butane(C_4H_{10}).
- **NGL(Natural Gas Liquids):** Main components are ethane(C_2H_6), propane(C_3H_8) and butane(C_4H_{10}). It exists in the gas phase at 1 atm and 20°C. That is, **NGL = LPG + ethane(C_2H_6)**
- **Condensate:** Main components are Pentane(C_5H_{12}) and Hexane(C_6H_{14}). It exists in the liquid phase at 1 atm and 20°C, so called **oil.**

3. Fractionation process system(분류)

3.1 Ethane Distillation System: Separating NGL into the ethane and LPG and using it as the refrigerant in the liquefaction system. Some of LPG is stored in the LPG tank and the rest goes to Propane Distillation system

3.2 Propane Distillation System: Separating LPG into the propane(C3) and butane(C4) by compressing them which are used as the refrigerant in the liquefaction system

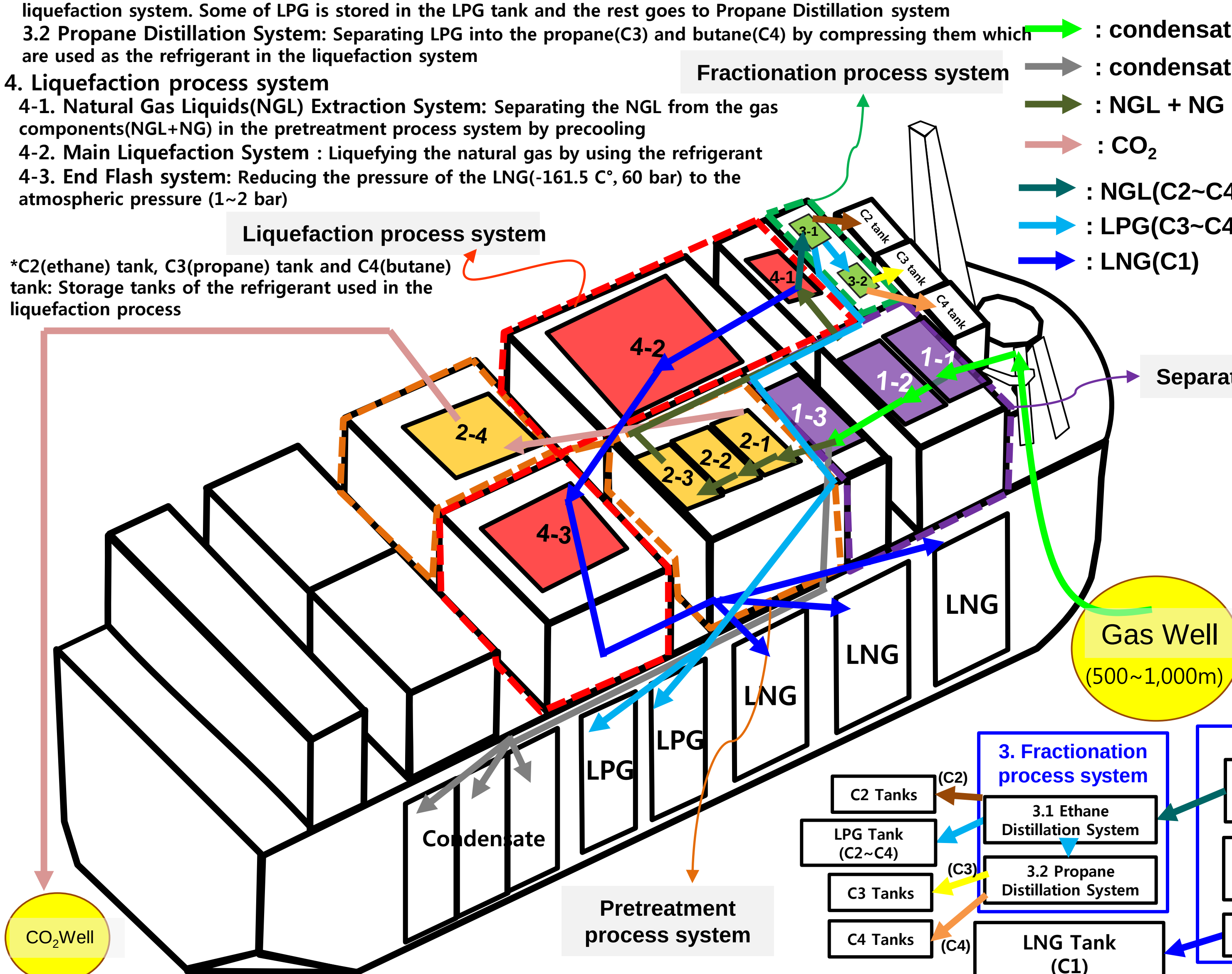
4. Liquefaction process system

4-1. Natural Gas Liquids(NGL) Extraction System: Separating the NGL from the gas components(NGL+NG) in the pretreatment process system by precooling

4-2. Main Liquefaction System : Liquefying the natural gas by using the refrigerant

4-3. End Flash system: Reducing the pressure of the LNG(-161.5 C°, 60 bar) to the atmospheric pressure (1~2 bar)

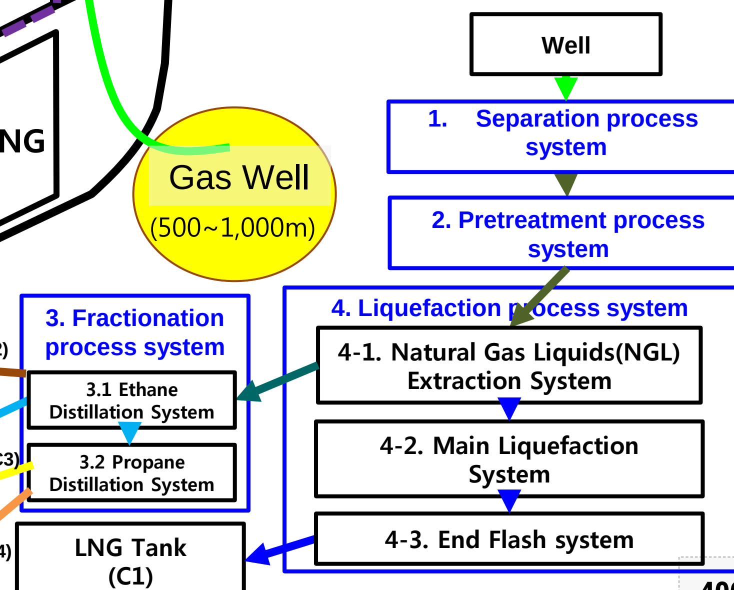
*C2(ethane) tank, C3(propane) tank and C4(butane) tank: Storage tanks of the refrigerant used in the liquefaction process



- Natural gas composition:

Components	Natural gas(mol %)
Nitrogen	6.0
Methane	83.2
Ethane	7.1
Propane	2.25
i-Butane	0.40
n-Butane	0.60
i-Pentane	0.12
n-Pentane	0.33

: NG composition for maximizing the amount of methane and satisfying the required the heating value of NG



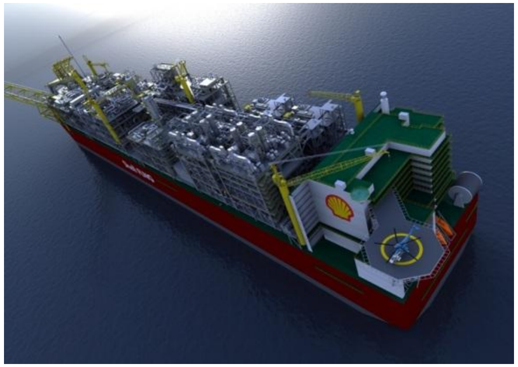
1. What is the Liquefaction Cycle of the LNG FPSO?

Major Considerations for the Selection of the Liquefaction Cycle for Offshore Application

<Liquefaction process system>

+

<Exploration and Production of the Natural Gas>



<LNG FPSO>

➢ Reliability

- All major oil companies required that liquefaction cycles shall have reliability based on the results from previous onshore projects.
- Dual Mixed Refrigerant(DMR) cycle was verified from the SAKHALIN onshore liquefaction cycle in 2005.

➢ SAFETY

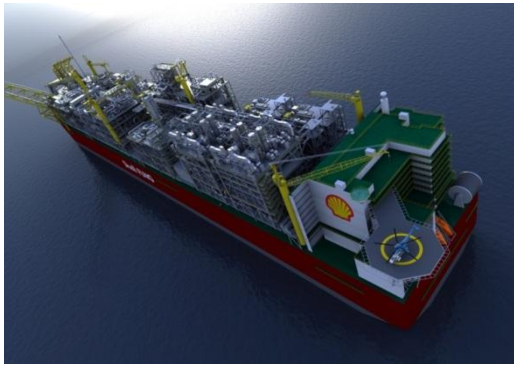
- Safety studies : HAZard and Operability(HAZOP), HAZard Identification (HAZID), Failure Modes and Effects Analysis(FMEA), Fault Tree Analysis(FTA), Event Tree Analysis (ETA), CFD Exhausts Dispersion Study – Helideck Study Report, Dropped Object Study , Explosion Risk Analysis, Failure, etc.

Major Considerations for the Selection of the Liquefaction Cycle for Offshore Application

<Liquefaction process system>

+

<Exploration and Production of the Natural Gas>



<LNG FPSO>

➢ Ship Motion Effect

- If the LNG FPSO is inclined more than 1.5 degrees, the capacity of LNG production can be reduced by 10%.
- Therefore, the liquefaction cycle in the LNG FPSO has to be designed by considering compactness, mechanical damping devices, internal turret system, and dynamic positioning system.

➢ COMPACTNESS

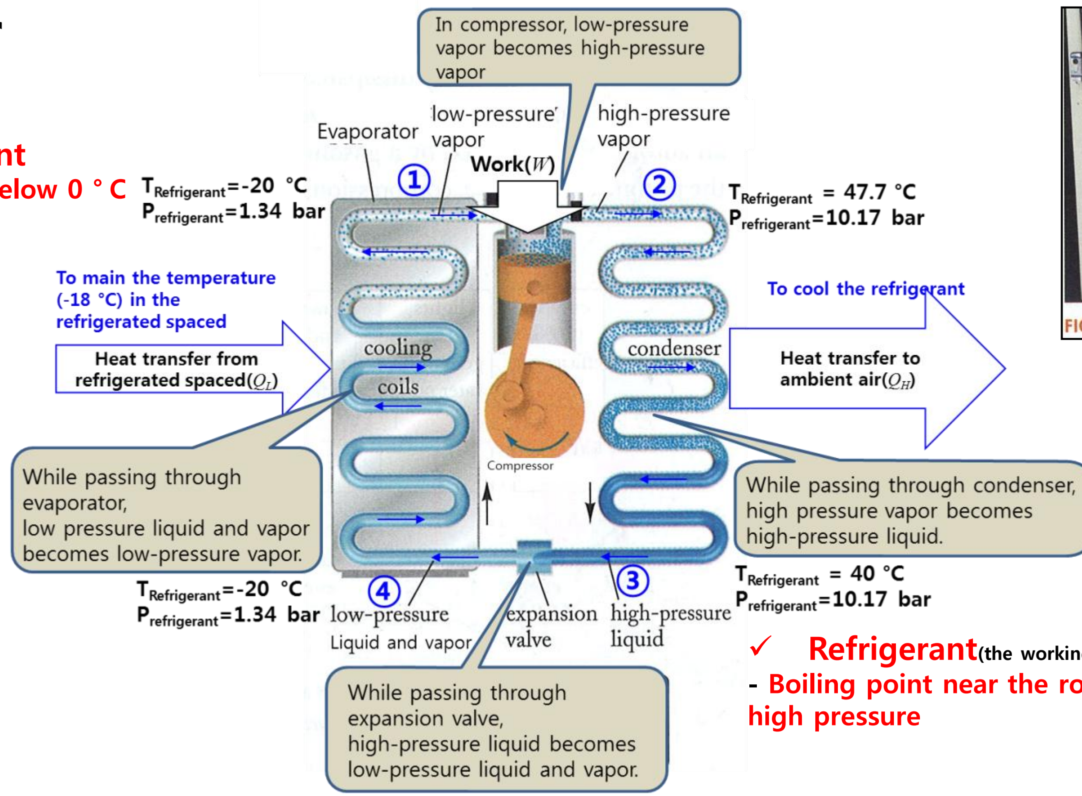
- Available area for the liquefaction cycle of offshore application is smaller than that of onshore plant.
- By determining the optimal operating conditions and doing the optimal synthesis of the liquefaction cycle, the required power for the compressors can be reduced which will result in the reduction of the compressor size and the flow rate of the refrigerant. Thus, the overall sizes of the liquefaction cycle including the pipe diameter, equipment and instrument can be reduced.
- Therefore, the compactness can be achieved by optimization studies such as determination of the optimal operating condition or optimal synthesis of the liquefaction cycle.

9.2. PROCESS OF THE REFRIGERATOR

Introduction to the Cooling System for Refrigerator (1/2)

- **Refrigerator**

✓ **Refrigerant**
-Boiling point below 0 ° C at low pressure



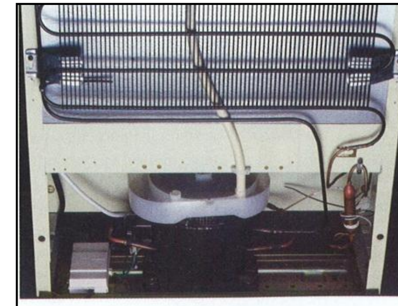


FIGURE 21.19 Condenser coils of a refrigerator.

✓ **Refrigerant**(the working substance in the refrigerator)
- Boiling point near the room temperature at high pressure

✓ **Refrigerator** :heat engine that operates backward to extract heat from a low-temperature reservoir and transfer it to a high-temperature reservoir. Because the natural tendency of heat is to flow a hot region to a cold one, energy must be provided to a refrigerator to reverse the flow, and this energy adds to the heat exhausted by the refrigerator.

✓ Refrigeration system process
- ①→② : The compressor, usually driven by an electric motor brings the refrigerant to a high pressure, which raises its temperature as well.
- ②→③ : The hot refrigerant passes through the condenser corresponded with the sea water cooler in the liquefaction cycle, an array of thin tubes that give off heat from the refrigerant to the atmosphere. The condenser is on the back of most house hold refrigerators. As it cools, the refrigerant becomes a liquid under high pressure.
- ③→④ : The liquid refrigerant goes into the expansion valve, from which it emerges at a lower pressure and temperature.
- ④→① : In the evaporator corresponded with the heat exchanger in the liquefaction cycle, the cool liquid refrigerant absorbs heat from the storage chamber and vaporizes. Along in the evaporator, the refrigerant vapor absorbs more heat and becomes warmer. The warm vapor then goes back to the compressor to start another cycle

9.2. Process of the Refrigerator

9.2.1 Equation of state

Equation of State for an Ideal Gas

Equation of state

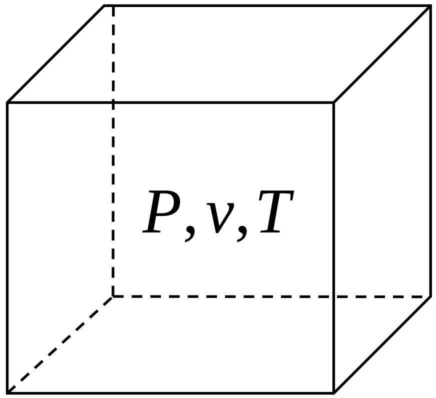
: Any equation that relates the pressure(P), temperature(T) and specific volume(v) of a substance

Ideal gas state: The condition that

1) the volume of the molecules is negligible compared with the total volume of the gas
2) the force that binds the molecules to each other is zero.

Equation of state for an ideal gas

$$P \cdot v = R \cdot T$$



P : pressure

T : temperature

v : specific volume(the volume that the molecules can move = the volume of the box)

R : gas constant

Cubic Equations of State for Liquids and Vapors

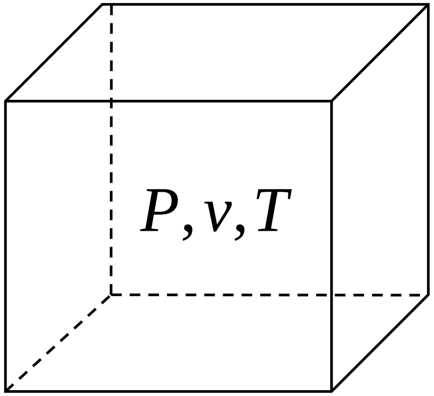
Ideal gas state: The condition that

1) the volume of the molecules is negligible compared with the total volume of the gas
2) the force that binds the molecules to each other is zero.

P : pressure

T : temperature

v : specific volume(the volume that the molecules can move = the volume of the box)



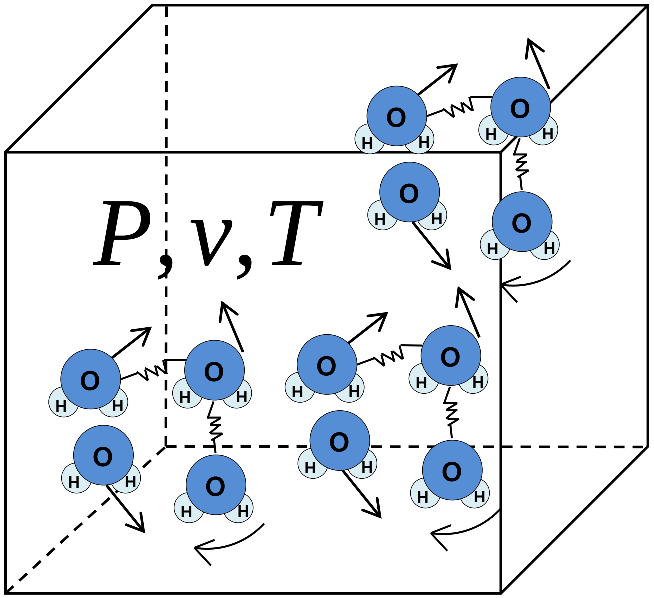
In case of the liquids and vapors,

1) the volume of the molecules is not negligible compared with the total volume of the gas

→ The specific volume(v) has to be decreased.

2) the force that binds the molecules to each other is not zero.

→ The pressure(P) has to be modified.



Note: Summary of the History of the Cubic Equations of State for Liquids and Vapors

*EoS: Equation of State

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

Cubic equations of state

(1) Ideal gas EoS* (1802)

(1) $Pv = RT$

(2) van der Waals EoS(1873)

(2) $\left(P + \frac{a}{v^2}\right)(v-b) = RT$

$a = \frac{24}{64} \cdot \frac{R^2 \cdot T_c^{\,2}}{P_c},\ b = \frac{1}{8} \cdot \frac{R \cdot T_c}{P_c}$

(3) Redlich-Kwong EoS(1949)

(3) $\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$

$a(T) = \frac{0.42748 \cdot (T/T_c)^{-1/2} \cdot R^2 \cdot T_c^{\,2}}{P_c},\ b = \frac{0.08664 \cdot R \cdot T_c}{P_c}$

(4) Soave-Redlich-Kwong EoS(1972)

(4) $\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$

$a(T) = \frac{0.42748 \cdot \alpha_{SRK}(T/T_c;\omega) \cdot R^2 \cdot T_c^{\,2}}{P_c}$

$\alpha_{SRK}(T/T_c;\omega) =$

$\left[1 + (0.480 + 1.574 \cdot \omega - 0.176 \cdot \omega^2) \cdot (1 - (T/T_c)^{1/2}\right]^2$

$b = \frac{0.08664 \cdot R \cdot T_c}{P_c}$

(5) Peng-Robinson EoS(1976)

(5) $\left(P + \frac{a(T)}{(v + (1-\sqrt{2}) \cdot b) \cdot (v + (1+\sqrt{2}) \cdot b)}\right)(v-b) = RT$

$a(T) = \frac{0.45724 \cdot \alpha_{PR}(T/T_c;\omega) \cdot R^2 \cdot T_c^{\,2}}{P_c}$

$\alpha_{PR}(T/T_c;\omega) =$

$\left[1 + (0.37464 + 1.54226 \cdot \omega - 0.26992 \cdot \omega^2) \cdot (1 - (T/T_c)^{1/2}\right]^2$

$b = \frac{0.07780 \cdot R \cdot T_c}{P_c}$

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]

T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor

R: gas constant(=8.314[m3Pa/(mol·K)]

Note: History of the Cubic Equations of State for Liquids and Vapors(1)

*EoS: Equation of State

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

(1) Ideal gas EoS* (1802)

(1) $Pv = RT$

(2) van der Waals EoS(1873)

(2) $\left(P + \frac{a}{v^2}\right)(v - b) = RT$

$a = \frac{24}{64} \cdot \frac{R^2 \cdot T_c^2}{P_c}, \; b = \frac{1}{8} \cdot \frac{R \cdot T_c}{P_c}$

(1) Ideal gas equation → (2) van der Waals(vdW) Eos

① Considering the attractive forces between molecules
: The pressure depends on both the frequency of collisions with the walls and the force of each collision. Because both the frequency and the force of the collisions are reduced by the attractive forces, the pressure(P) is reduced in proportion to the square of the concentration(a/v^2, a is a positive constant characteristic of each gas).

② Considering the volume of the molecules
: The volume that the molecules can move(molar volume, v) is decreased by the volume of the molecules(b)

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]

T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor

R: gas constant(=8.314[m3Pa/(mol·K)]

*EoS: Equation of State

Note: History of the Cubic Equations of State for Liquids and Vapors(2)

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

(1) Ideal gas EoS* (1802)

(1) $$Pv = RT$$

(2) van der Waals EoS(1873)

(2) $$\left(P + \frac{a}{v^2}\right)(v-b) = RT$$

$$a = \frac{24}{64} \cdot \frac{R^2 \cdot T_c^{\,2}}{P_c},\ b = \frac{1}{8} \cdot \frac{R \cdot T_c}{P_c}$$

(3) Redlich-Kwong EoS(1949)

(3) $$\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.42748 \cdot (T/T_c)^{-1/2} \cdot R^2 \cdot T_c^{\,2}}{P_c},\ b = \frac{0.08664 \cdot R \cdot T_c}{P_c}$$

(2) van der Waals(vdW) EoS → (3) Redlich-Kwong(RK) EoS

① Modify the pressure reduction due to the attractive forces

: The fact that the pressure reduction depends on the temperature(T)(inverse proportional to the $T^{1/2}$) and taking $v(v+b)$ instead of v^2 to calculate the pressure reduction is more accurate is proved by the experiment.

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor
R: gas constant(=8.314[m3Pa/(mol·K)]

Note: History of the Cubic Equations of State for Liquids and Vapors(3)

*EoS: Equation of State

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

(1) Ideal gas EoS* (1802)

(1) $Pv = RT$

(2) van der Waals EoS(1873)

(2) $\left(P + \frac{a}{v^2}\right)(v-b) = RT$

$a = \frac{24}{64} \cdot \frac{R^2 \cdot T_c^{\,2}}{P_c}, \; b = \frac{1}{8} \cdot \frac{R \cdot T_c}{P_c}$

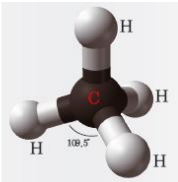
(3) Redlich-Kwong EoS(1949)

(3) $\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$

$a(T) = \frac{0.42748 \cdot (T/T_c)^{-1/2} \cdot R^2 \cdot T_c^{\,2}}{P_c}, \; b = \frac{0.08664 \cdot R \cdot T_c}{P_c}$

This equations are exact for the simple fluid such as argon and methane.

- The force between the molecules is acting on the center of that.
- The shape of the molecules is sphere



<methane>

(4) Soave-Redlich-Kwong EoS(1972)

(4) $\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$

$a(T) = \frac{0.42748 \cdot \alpha_{SRK}(T/T_c;\omega) \cdot R^2 \cdot T_c^{\,2}}{P_c}$

$\alpha_{SRK}(T/T_c;\omega) =$

$\left[1 + (0.480 + 1.574 \cdot \omega - 0.176 \cdot \omega^2) \cdot (1 - (T/T_c)^{1/2})\right]^2$

$b = \frac{0.08664 \cdot R \cdot T_c}{P_c}$

(3) Redlich-Kwong(RK) EoS → (4) Soave-Redlich-Kwong EoS

① Modify the pressure reduction due to the attractive forces

: The pressure reduction depending on the temperature(T) is modified by introducing the acentric factor(ω) for general fluid including the simple fluid.

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor
R: gas constant(=8.314[m3Pa/(mol·K)]

*EoS: Equation of State

Note: History of the Cubic Equations of State for Liquids and Vapors(4)

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

(1) Ideal gas EoS* (1802)

$$\text{(1)} \quad Pv = RT$$

(2) van der Waals EoS(1873)

$$\text{(2)} \quad \left(P + \frac{a}{v^2}\right)(v-b) = RT$$

$$a = \frac{24}{64}\cdot\frac{R^2\cdot T_c^{\,2}}{P_c},\ b = \frac{1}{8}\cdot\frac{R\cdot T_c}{P_c}$$

(3) Redlich-Kwong EoS(1949)

$$\text{(3)} \quad \left(P + \frac{a(T)}{v\cdot(v+b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.42748\cdot(T/T_c)^{-1/2}\cdot R^2\cdot T_c^{\,2}}{P_c},\ b = \frac{0.08664\cdot R\cdot T_c}{P_c}$$

(4) Soave-Redlich-Kwong EoS(1972)

$$\text{(4)} \quad \left(P + \frac{a(T)}{v\cdot(v+b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.42748\cdot\alpha_{SRK}(T/T_c;\omega)\cdot R^2\cdot T_c^{\,2}}{P_c}$$

$$\alpha_{SRK}(T/T_c;\omega) =$$

$$\left[1+(0.480+1.574\cdot\omega-0.176\cdot\omega^2)\cdot(1-(T/T_c)^{1/2}\right]^2$$

$$b = \frac{0.08664\cdot R\cdot T_c}{P_c}$$

(5) Peng-Robinson EoS(1976)

$$\text{(5)} \quad \left(P + \frac{a(T)}{(v+(1-\sqrt{2})\cdot b)\cdot(v+(1+\sqrt{2})\cdot b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.45724\cdot\alpha_{PR}(T/T_c;\omega)\cdot R^2\cdot T_c^{\,2}}{P_c}$$

$$\alpha_{PR}(T/T_c;\omega) =$$

$$\left[1+(0.37464+1.54226\cdot\omega-0.26992\cdot\omega^2)\cdot(1-(T/T_c)^{1/2}\right]^2$$

$$b = \frac{0.07780\cdot R\cdot T_c}{P_c}$$

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor
R: gas constant(=8.314[m3Pa/(mol·K)]

(4) Soave-Redlich-Kwong EoS → (5) Peng-Robinson EoS

① **Modify the pressure reduction due to the attractive forces**
: The pressure reduction depending on the molar volume(v) is modified by $(v+(1-\sqrt{2})\cdot b)\cdot(v+(1+\sqrt{2})\cdot b)$ **instead of $v(v+b)$.**

Note: General Form of the Cubic Equations of State for Liquids and Vapors

The van der Waals(vdW), Redlich-Kwong(RK), Soave-Redlich-Kwong(SRK) and Peng-Robinson(PR) equation of state are represented as the following cubic equations form.

$$\left(P+\frac{a(T)}{(v+\varepsilon b)(v+\sigma b)}\right)(v-b)=RT \quad \Rightarrow \quad P=\frac{RT}{v-b}-\frac{a(T)}{(v+\varepsilon b)(v+\sigma b)}$$

$$a(T)=\Psi\frac{\alpha(T_r)R^2{T_c}^2}{P_c}, \quad b=\Omega\frac{RT_c}{P_c}$$

EoS	$\alpha(T_r)$	σ	ε	Ω	Ψ	Z_c
vdW(1873)	1	0	0	1/8	24/64	3/8
RK(1949)	$T_r^{-0.5}$	1	0	0.08664	0.42748	1/3
SRK(1972)	$\alpha_{SPK}(T_r, \omega)$	1	0	0.08664	0.42748	1/3
PR(1976)	$\alpha_{SR}(T_r, \omega)$	$1+\sqrt{2}$	$1-\sqrt{2}$	0.07780	0.45724	0.30740
$\alpha_{SPK}(T_r;\omega)=\left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-T_r^{\,0.5}\right)\right]^2$						
$\alpha_{PR}(T_r;\omega)=\left[1+(0.37464+1.54226\omega-0.26992\omega^2)\left(1-T_r^{\,0.5}\right)\right]^2$						

$$T_r=T/T_c$$

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor

R: gas constant(=8.314[m3Pa/(mol·K)]

Note: Cubic Equations of State for Liquids and Vapors
- Compressible Factor for the Ideal Gas

The compressible factor(Z):

$$Z \equiv \frac{P \cdot v}{(P \cdot v)^{ig}} = \frac{P \cdot v}{R \cdot T}$$

R : gas constant(=8.314 [$m^3Pa/((mol \cdot K)$])
P : pressure[Pa]
T : temperature[K]
v : molar volume[m^3/mol]

- If $P \to 0$, $v \to 0$. So, the volume of the molecules is not negligible compared with the total volume of the gas and the force that binds the molecules to each other is not zero(Ideal gas state).

$$P \cdot v \to R \cdot T$$

$$Z = 1$$

The compressible factor(Z):

$$Z \equiv \frac{P \cdot v}{(P \cdot v)^{ig}} = \frac{P \cdot v}{R \cdot T}$$

R : gas constant(=8.314 [$m^3 Pa/((mol \cdot K)$])

P : pressure[Pa]

T : temperature[K]

V : molar volume[m^3/mol]

- **By using the cubic equations of state for the liquids and vapors, we can obtain the compressible factor(Z) of the liquids and vapors.**

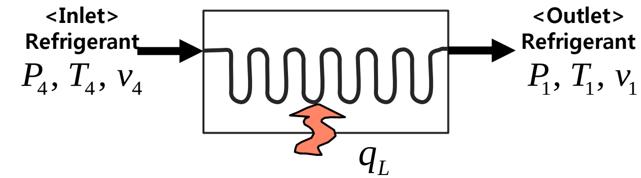
$$P = \frac{R \cdot T}{v - b} - \frac{a(T)}{(v + \varepsilon \cdot b)(v + \sigma \cdot b)}$$

$$\Big\Downarrow \times \frac{V}{R \cdot T}$$

$$\frac{P \cdot v}{R \cdot T} = \frac{v}{v - b} - \frac{v}{R \cdot T} \cdot \frac{a(T)}{(v + \varepsilon \cdot b)(v + \sigma \cdot b)}$$

$$\Big\Downarrow Z \equiv \frac{P \cdot V}{R \cdot T},$$

$$\therefore Z = \frac{v}{v - b} - \frac{v}{R \cdot T} \cdot \frac{a(T)}{(v + \varepsilon \cdot b)(v + \sigma \cdot b)}$$



Equations of state: Any equation that relates the pressure(P), temperature(T) and specific volume(V) of a substance.

$$P_1 v_1 = RT_1$$

$$P_4 v_4 = RT_4$$

[Equation of state for an ideal gas]

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

Example) Soave, Redlich, Kwong(SRK) equation

$$v = \frac{RT}{P} + b - \frac{a(T)}{P}\frac{v-b}{(v-\varepsilon b)(v-\sigma b)}$$

$v_1 = \frac{RT_1}{P} + b - \frac{a(T_1)}{P_1}\frac{v_1-b}{(v_1-\varepsilon b)(v_1-\sigma b)}$
$v_{4,v} = \frac{RT_4}{P} + b - \frac{a(T_4)}{P_4}\frac{v_{4,v}-b}{(v_{4,v}-\varepsilon b)(v_{4,v}-\sigma b)}$
$v_{4,l} = \frac{RT_4}{P} + b - \frac{a(T_4)}{P_4}\frac{v_{4,l}-b}{(v_{4,l}-\varepsilon b)(v_{4,l}-\sigma b)}$

[Given]

$$\dot{Q}_L = 20[kW]$$

[Find]

Operating Conditions [20]:

P_i, T_i, v_i, T_s, v_s, $v_{4,l}$, $v_{4,v}$, v_f, w, M, q_L, q_H (i=1,2,3,4)

$$\text{Minimize}(\dot{m}\cdot w)$$

Refrigerant: Ammonia
P: pressure [bar]
T: temperature [K]
v: specific volume [m³/kg]
T_s: temperature of the refrigerant in the compressor at isentropic process [K]
v_s: specific volume of the refrigerant in the compressor at isentropic process [m³/kg]
v_f: vapor fraction
q_H: specific heat transfer from the refrigerant to the atmosphere [kJ/kg]
q_L: specific heat transfer from the refrigerated space to the refrigerant [kJ/kg]
M: mass flow rate of the refrigerant [kg/s]
$\dot{Q}_L$: heat transfer from the refrigerated space to the refrigerant [kW]

s: specific entropy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ for SRK equation

R: gas constant (=8.314 Jmol⁻¹K⁻¹)
P_c: critical pressure of the refrigerant
T_c: critical temperature of the refrigerant

$$b = \Omega \frac{RT_C}{P_C}$$

$\Omega = 0.08664$ for SRK equation
$\varepsilon = 0$ for SRK equation
$\sigma = 1$ for SRK equation

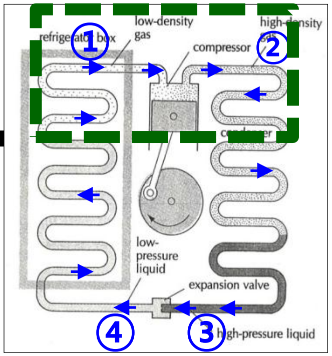
Example) Ammonia:

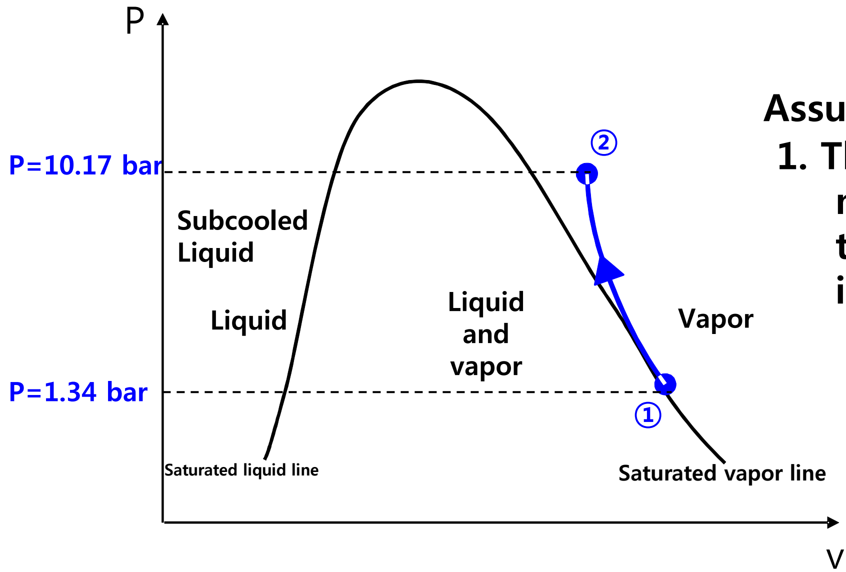
ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

9.2. Process of the Refrigerator

9.2.2 Compression

Pressure(P)-Specific Volume(v) Diagram

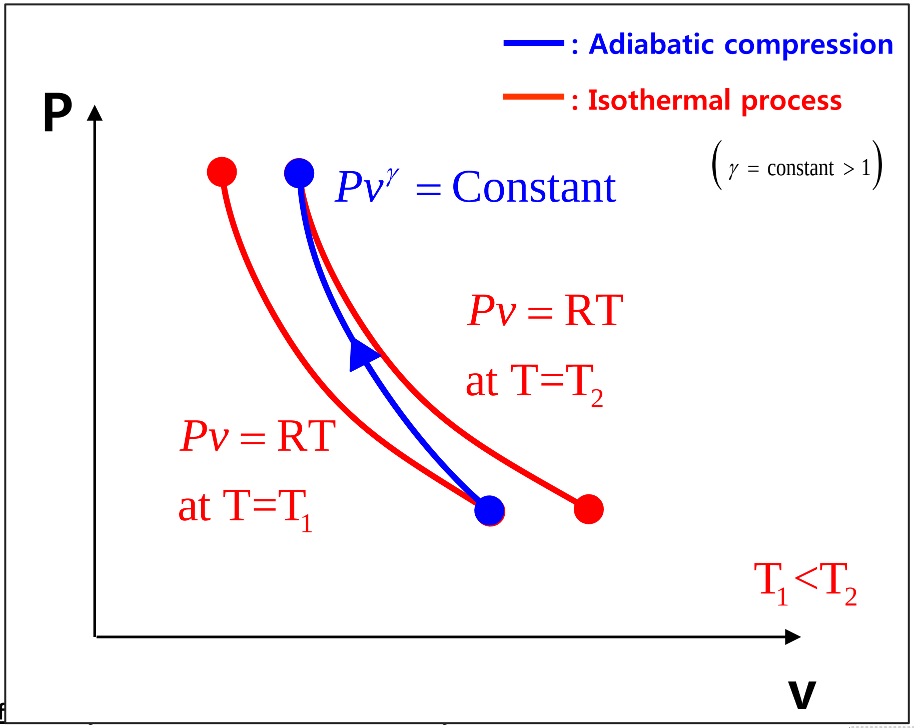




Assumption:

1. There is not sufficient time to transfer much heat from the refrigerant and the compressor is typically well insulated. "**Adiabatic process**"

* Adiabatic process: Process for which there is no heat transfer between system and it surroundings.



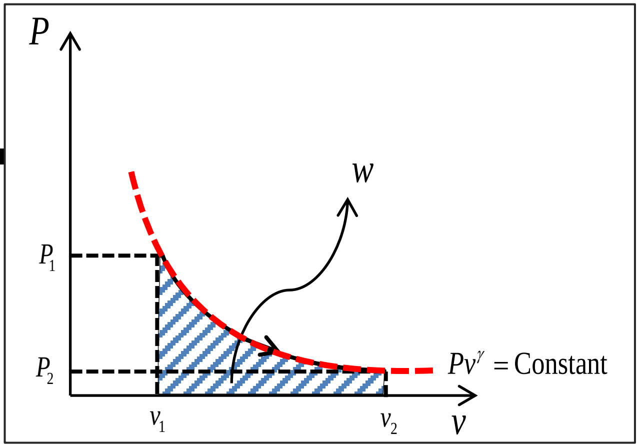
Purpose, **Assumption**, **Result**

		Process	Temperature	Pressure
Compressor	1→2	Adiabatic compression	↑	↑

Why is the temperature of the refrigerant raised after the compressor?

Because of the adiabatic compression, the temperature of the refrigerant is increased form T_1 to T_2.

Note: P-V Graph of the Adiabatic Process (1/3)



u: specific internal energy
q: heat transfer from the surrounding to the system
w: work acting toward the system
C_V^{ig} : constant volume heat-capacity for ideal gas
T: temperature
P: pressure
v: specific volume

- According to the first law for the closed system,

$$\Delta u = q + w$$

1) In the adiabatic process, $q = 0$.

2) In the first law, the sign of w acting toward the system is positive.

$$\Delta u = -w$$

$$du = -dw$$

1) If the state of the substance in the system is an ideal gas state,

$$u = u^{ig} , \; du^{ig} = C_V^{ig} dT$$

2) $dw = P \cdot dv$

$$C_V^{ig} dT = -P \cdot dv$$

1) Equation of state for ideal gas $P \cdot v = R \cdot T \rightarrow T = \dfrac{P \cdot v}{R}$ ⇨ $dT = \dfrac{v}{R} dP + \dfrac{P}{R} dv$

(R : gas constant)

$$C_V^{ig} \left(\frac{v}{R} dP + \frac{P}{R} dv \right) = -P \cdot dv$$

Note: P-V Graph of the Adiabatic Process (2/3)

u: specific internal energy
q: heat transfer from the surrounding to the system
w: work acting toward the system
C_V^{ig} : constant volume heat-capacity for ideal gas
T: temperature
P: pressure
v: specific volume
R: gas constant

$$C_V^{ig} \cdot \left(\frac{v}{R} \cdot dP + \frac{P}{R} \cdot dv \right) = -P \cdot dv$$

$$C_V^{ig} \cdot \frac{v}{R} \cdot dP = -P \cdot dv - C_V^{ig} \cdot \frac{P}{R} \cdot dv$$

$$C_V^{ig} \cdot \frac{v}{R} \cdot dP = -\left(1 + \frac{C_V^{ig}}{R} \right) \cdot P \cdot dv$$

$$\left(\frac{C_V^{ig}}{R} \right) v \cdot dP = -\left(1 + \frac{C_V^{ig}}{R} \right) \cdot P \cdot dv$$

$$\left(\frac{C_V^{ig}}{R} \right) \cdot \frac{dP}{P} = -\left(1 + \frac{C_V^{ig}}{R} \right) \cdot \frac{dv}{v}$$

$$\frac{dP}{P} = -\left(\frac{R}{C_V^{ig}} + 1 \right) \cdot \frac{dv}{v}$$

$$\int \frac{dP}{P} = \int -\left(\frac{R}{C_V^{ig}} + 1 \right) \cdot \frac{dv}{v}$$

If the C_V^{ig} is the constant,

$$\int \frac{dP}{P} = -\left(\frac{R}{C_V^{ig}} + 1 \right) \int \cdot \frac{dv}{v}$$

$$\ln \frac{P_2}{P_1} = -\left(\frac{R}{C_V^{ig}} + 1 \right) \ln \frac{v_2}{v_1}$$

$$\frac{P_2}{P_1} = \left(\frac{v_2}{v_1} \right)^{-\left(\frac{R}{C_V^{ig}} + 1 \right)}$$

Note: P-V Graph of the Adiabatic Process (3/3)

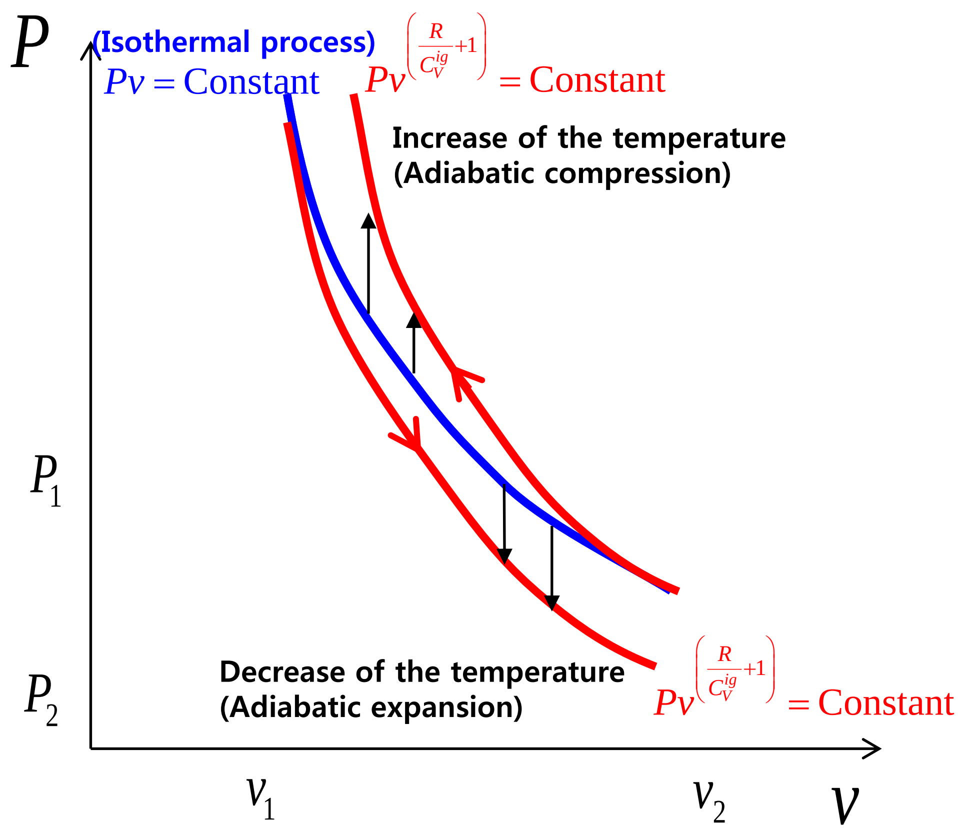
u: specific internal energy
q: heat transfer from the surrounding to the system
w: work acting toward the system
C_V^{ig} : constant volume heat-capacity for ideal gas
T: temperature
P: pressure
v: specific volume
R: gas constant

$$\frac{P_2}{P_1} = \left(\frac{v_2}{v_1}\right)^{-\left(\frac{R}{C_V^{ig}}+1\right)}$$

$$\frac{P_2}{P_1} = \left(\frac{v_1}{v_2}\right)^{\left(\frac{R}{C_V^{ig}}+1\right)}$$

$$P_2 \cdot v_2^{\left(\frac{R}{C_V^{ig}}+1\right)} = P_1 \cdot v_1^{\left(\frac{R}{C_V^{ig}}+1\right)} = \text{constant}$$

$$\left(\frac{R}{C_V^{ig}}+1 = \gamma(\text{constant}) > 1\right)$$



Cf) P-V Graph of the Isothermal Process

$$P \cdot v = R \cdot T$$
$$= \text{Constant}$$

Note: Specific Enthalpy(h)-(1/2)

u: specific internal energy
P: pressure
v: specific volume

1. Definition:

$$h = \left[Internal\ \ energy\ \ of\ \ the\ \ flow \right] + \left[Flow\ \ work \right]$$
$$= u + P \cdot v$$

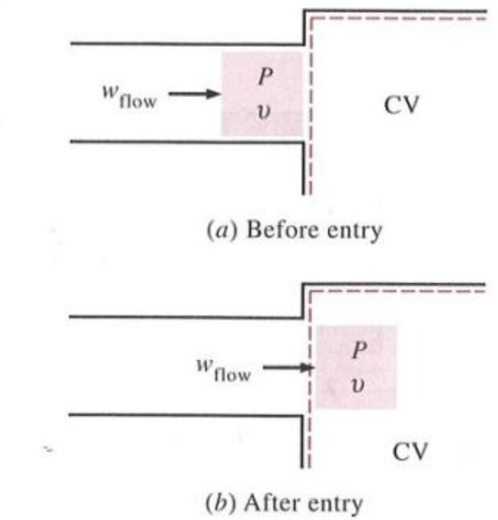
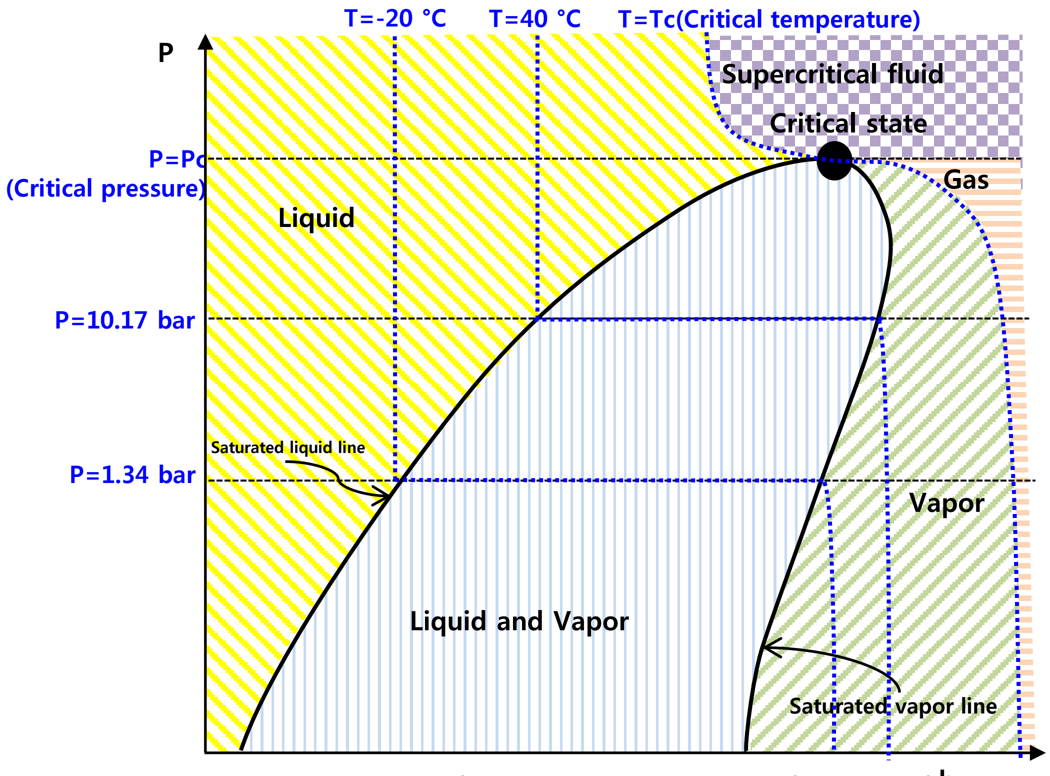


FIGURE 4-12
Flow work is the energy needed to push a fluid into or out of a control volume, and it is equal to Pv.

2. Pressure(P)-Specific Enthalpy(h) diagram



* Vapor: Vapor can be condensed either by compression at constant temperature or by cooling at constant pressure.

* Gas: The vapor phase of a substance is customarily called a gas when it is above the critical temperature. Gas cannot be condensed by compression at constant temperature.

* Supercritical fluid: A single phase at and above the critical temperature and pressure

Note: Specific Enthalpy(h)-(2/2)

$$h = u + P \cdot v$$

3. Calculation of the specific enthalpy(h) for a pure substance

Many tables of thermodynamics properties does not give values for internal energy. To allow calculation of enthalpy from the pressure, specific volume and temperature, the following equation is derived by using the definition($h=u+Pv$), equation of state and experiment.

$$h = h^{IG} + h^{R}$$

h^{ig}: **Ideal gas value of the enthalpy**
h^{R}: **Residual enthalpy(correction of the ideal gas state values to the real gas values)**

$$h^{IG} = h^{IG}(T) = a + b \cdot T + c \cdot T^2 + d \cdot T^3 + e \cdot T^4 + f \cdot T^5$$

where

a, b, c, d, e **and** f: **constants characteristic of the particular substance**
T: **temperature**

$$h^{R} = h^{R}(P, v, T) = RT(Z-1) + \frac{T\left(\frac{da}{dT}\right) - a}{b} \frac{1}{(\sigma - \varepsilon)} \ln\left[\frac{Z + \sigma \cdot \beta}{Z + \varepsilon \cdot \beta}\right]$$

where

$$a = \psi \frac{\alpha(T/T_c, \omega) R^2 T_c^2}{P_c}$$

Equation of state	$\alpha(Tr, \omega)$	σ	ε	Ω	ψ
SRK	$\alpha(T_r;\omega) = \left[1 + (0.480 + 1.574\omega - 0.176\omega^2)\left(1 - (T/T_C)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

P: **pressure**
v: **specific volume**
P_c: **critical pressure of the substance**
T_c: **critical temperature of the substance**
Z: **compressible factor**

$$Z = \frac{v}{v-b} - \frac{v}{R \cdot T} \cdot \frac{a(T)}{(v + \varepsilon \cdot b)(v + \sigma \cdot b)} \qquad \beta = \frac{bP}{RT} \qquad b = \Omega \frac{RT_c}{P_c}$$

• Calculation of the specific enthalpy(h) for a pure substance

Many tables of thermodynamics properties does not give values for internal energy. To allow calculation of enthalpy from the pressure, specific volume and temperature, the following equation is derived by using the definition($h=u+Pv$), equation of state and experiment.

P[Pa]
v[m^3]
T[K]

$$h=h^{IG}+h^{R}$$

h^{IG}: Ideal gas value of the specific enthalpy
h^{R}: Residual specific enthalpy(correction of the ideal gas state values to the real gas values)

$$h^{IG}=h^{IG}(T)=a+b\cdot T+c\cdot T^2+d\cdot T^3+e\cdot T^4+f\cdot T^5$$

where
a, b, c, d, e and f: constants characteristic of the particular substance
h^{ig}[J/g], T: temperature[K]

Example) Ammonia:
a = -1.8514, b = 1.9937, c = -5.3266x10^{-4}
d = 2.0615x10^{-6}, e = -1.3386x10^{-9},
f = 3.0533x10^{-13}

$$h^{R}=h^{R}(P,v,T)=RT(Z-1)+\frac{T\left(\frac{da(T)}{dT}\right)-a(T)}{b}\frac{1}{(\sigma-\varepsilon)}\ln\left[\frac{Z+\sigma\cdot\beta}{Z+\varepsilon\cdot\beta}\right]$$

R: Gas constant (=8.314 J/(mol*K)
= 8.314m^{3}Pa(mol*K))

* 1 bar = 100kPa

where

$$a(T)=\psi\frac{\alpha(T_r,\omega)R^2T_c^2}{P_c} \qquad T_r=\frac{T}{T_c} \qquad Z=\frac{v}{v-b}-\frac{v}{R\cdot T}\cdot\frac{a(T)}{(v+\varepsilon\cdot b)(v+\sigma\cdot b)} \qquad \beta=\frac{bP}{RT} \qquad b=\Omega\frac{RT_c}{P_c}$$

1) The values of parameters a, σ, ε, Ω, Ψ are depending on the type of the cubic equation of state. For example, the value of the parameters for the Soave-Redlich-Kwong(SRK) equation of state are given in the following table.

$\alpha(T_r, \omega)$	σ	ε	Ω	ψ
$\alpha(T_r;\omega)=\left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-(T_r)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

2) The values of ω, critical pressure(Pc), and temperature(Tc) are depending on the substance.

Example) Ammonia:
ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

3) Since the unit of h^{IG} is J/g and h^{R} is J/mol, h^{R} is devided by molar mass(M, g/mol).

$$h^{R}[J/g]=\frac{h^{R}[J/mol]}{M[g/mol]}$$

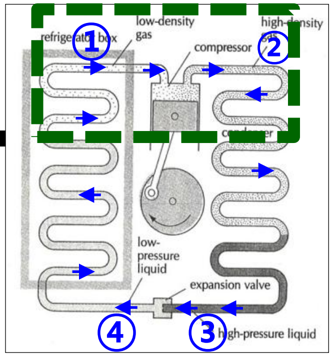
Example) Ammonia
$M_{Ammonia}$ = 17.031 (g/mol)

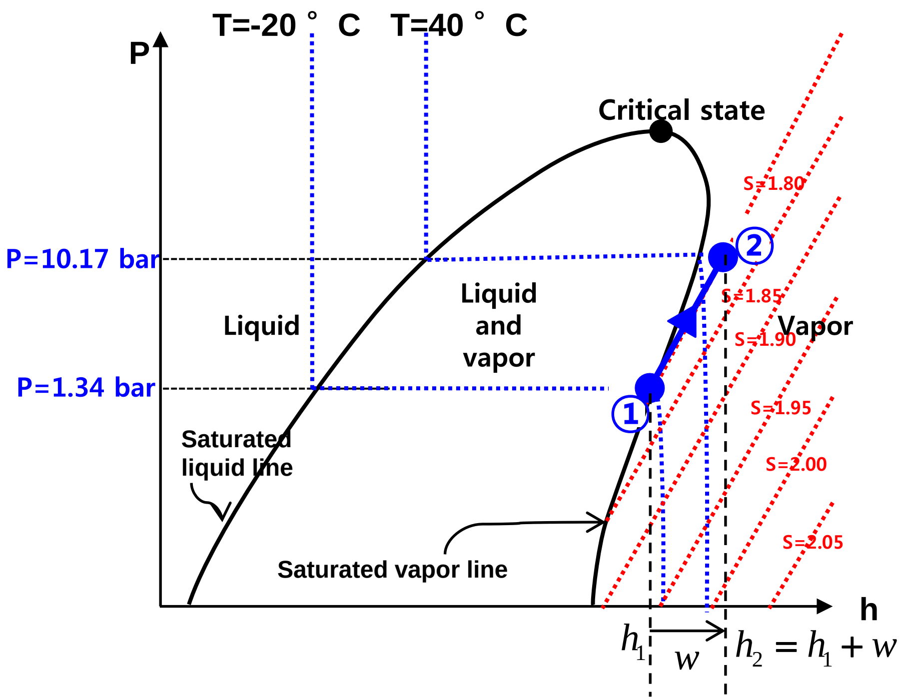
4) Central difference approximation

$$\left(\frac{da}{dT}\right)=\frac{a(T+e)-a(T-e)}{2\cdot e},\ (e=10^{-6})$$

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

Pressure(P)-Specific Enthalpy(h) Diagram





·······:Isentropic curve

Assumption:

1. Adiabatic process
2. The process of the compressor is "reversible".

→Since this process is "Adiabatic and reversible", the quality of energy "entropy" is not changed "**Isentropic process**".

- **Enthalpy**

$h = [\text{Internal energy of the flow}] + [\text{Flow work}]$
$= u + P \cdot v$

h : specific enthalpy
u : internal energy=$u(T,P)$
P : pressure
T : temperature

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↑	↑	+W	↑	0

According to the first law of thermodynamics (The total quantity of energy is constant)

$$\begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant entering}\\ \text{the compressor}\end{bmatrix} + \begin{bmatrix}\text{Total entering energy}\\ \text{as Work in the compressor}\end{bmatrix} = \begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant leaving}\\ \text{the compressor}\end{bmatrix}$$

$$h_1(P_1, v_1, T_1) + w = h_2(P_2, v_2, T_2)$$

Note: Specific Entropy(s)-(1/2)

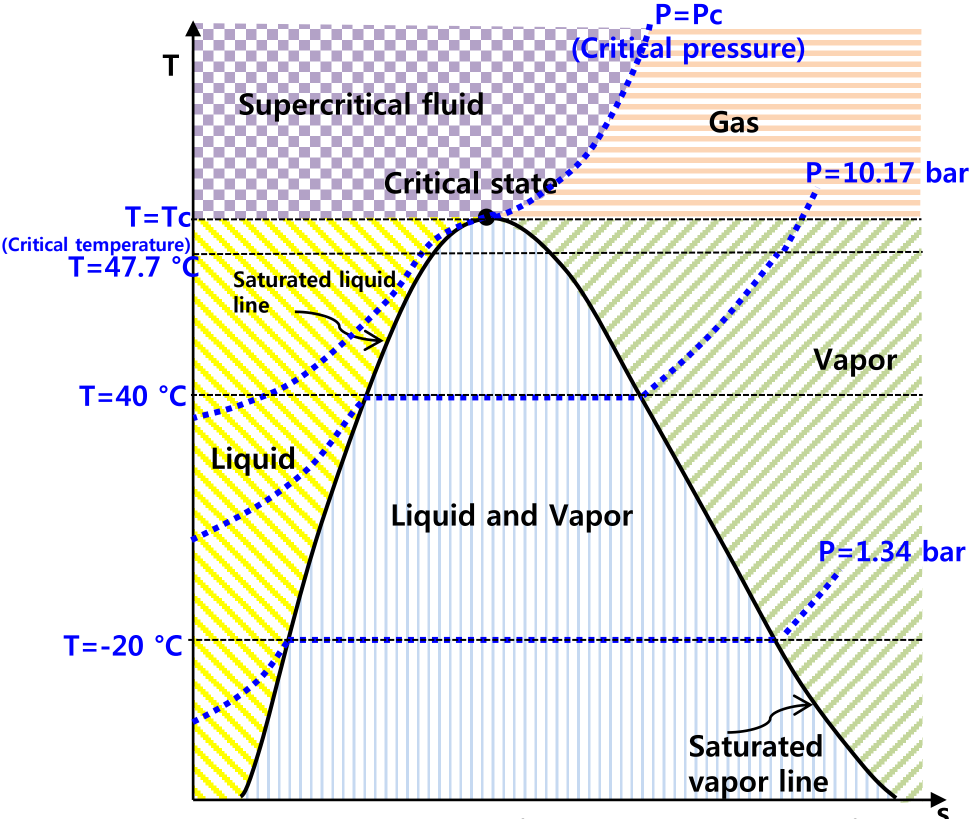
q: specific heat transfer
T: Temperature

The second law of the thermodynamics : Actual processes occur in the direction of decreasing quality of energy, "Entropy".

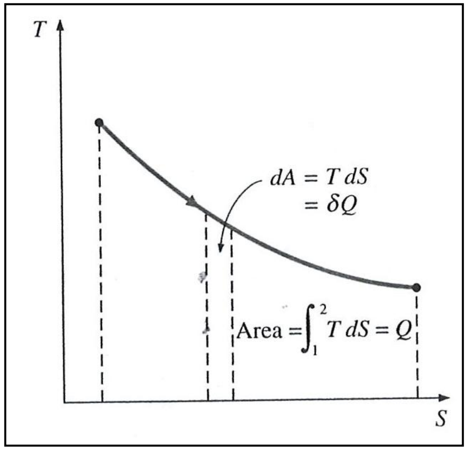
1. Definition: The quality of energy

$$ds = \frac{dq}{T}$$

2. Temperature(T)-Specific Entropy(s) diagram



Entropy can be viewed as a measure of molecular disorder, or molecular randomness.
As a system becomes more disordered, the positions of the molecules become less predictable and the entropy increases.



On the T-S diagram, the area under the process curve represents the heat transfer of the process.

3. Calculation of the specific entropy(s) for a pure substance

To allow calculation of entropy from the pressure, specific volume and temperature, the following equation is derived by using the definition($ds=dq/T$), equation of state and experiment.

$$s = s^{IG} + s^{R}$$

s^{ig}: **Ideal gas value of the entropy**

s^{R}: **Residual entropy(correction of the ideal gas state values to the real gas values)**

$$s^{ig} = s^{ig}\left(P,T\right) = R\int_{T_0}^{T} \frac{C_{P,j}^{ig}}{R}\frac{dT}{T} - \ln\frac{P}{P_0}$$

where

$\frac{C_P^{ig}(T)}{R} = A + B\cdot T + C\cdot T^2 + D\cdot T^{-2}$: **heat capacity of the particular substance(A, B, C, and D are constants characteristic of the particular substance)**

T: **temperature,** R: **gas constant,** P: **pressure,** v: **specific volume**

$$s^{R} = s^{R}\left(P,v,T\right) = R\ln(Z-\beta) + \frac{\left(\frac{da}{dT}\right)}{b}\frac{1}{(\sigma-\varepsilon)}\ln\left[\frac{Z+\sigma\cdot\beta}{Z+\varepsilon\cdot\beta}\right]$$

P_c: **critical pressure of the substance**

T_c: **critical temperature of the substance**

Z: **compressible factor**

where

$$a = \psi\frac{\alpha\left(T/T_c,\omega\right)R^2T_c^2}{P_c}$$

Equation of state	$\alpha(Tr, \omega)$	σ	ε	Ω	ψ
SRK	$\alpha(T_r;\omega) = \left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-\left(T/T_C\right)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

$$Z = \frac{v}{v-b} - \frac{v}{R\cdot T}\cdot\frac{a(T)}{(v+\varepsilon\cdot b)(v+\sigma\cdot b)} \qquad b = \Omega\frac{RT_c}{P_c} \qquad \beta = \frac{bP}{RT}$$

$$ds = \frac{dq}{T}$$

• Calculation of the specific entropy(s) for a pure substance

To allow calculation of entropy from the pressure, specific volume and temperature, the following equation is derived by using the definition($ds=dq/T$), equation of state and experiment.

P[Pa]
v[m^3]
T[K]

$$s = s^{IG} + s^{R}$$

s^{ig}: Ideal gas value of the entropy
s^{R}: Residual entropy(correction of the ideal gas state values to the real gas values)

$$S^{IG} = g + b\cdot\ln(T) + 2\cdot c\cdot T + \frac{3}{2}\cdot d\cdot T^2 + \frac{4}{3}\cdot e\cdot T^3 + \frac{5}{4}\cdot f\cdot T^4$$

where

a, b, c, d, e and f: coefficients of the ideal gas Enthalpy equation
s^{ig}[J/(g·K)], T: temperature[K]
g : Entropy coefficient (i.e. the Entropy of the ideal gas at T=0 K) = 1.00

Example) Ammonia:
a = **-1.8514**, b = **1.9937**, c = **-5.3266x10^{-4}**
d = **2.0615x10^{-6}**, e = **-1.3386x10^{-9}**,
f = **3.0533x10^{-13}**

$$s^R = s^R(P,v,T) = R\ln(Z-\beta) - R\ln\left(\frac{P}{P_0}\right) + \frac{\left(\frac{\partial a}{\partial T}\right)}{b}\frac{1}{(\sigma-\varepsilon)}\ln\left[\frac{Z+\sigma\cdot\beta}{Z+\varepsilon\cdot\beta}\right]$$

P_c: critical pressure of the substance
T_c: critical temperature of the substance
Z: compressible factor

where

$$a(T) = \psi\frac{\alpha(T_r,\omega)R^2T_c^2}{P_c} \qquad T_r = \frac{T}{T_c} \qquad Z = \frac{v}{v-b} - \frac{v}{R\cdot T}\cdot\frac{a(T)}{(v+\varepsilon\cdot b)(v+\sigma\cdot b)} \qquad \beta = \frac{bP}{RT} \qquad b = \Omega\frac{RT_c}{P_c}$$

1) The values of parameters $a, \sigma, \varepsilon, \Omega, \Psi$ are depending on the type of the cubic equation of state. For example, the value of the parameters for the Soave-Redlich-Kwong(SRK) equation of state are given in the following table.

$\alpha(T_r, \omega)$	σ	ε	Ω	ψ
$\alpha(T_r;\omega) = \left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-(T_r)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

2) The values of ω, critical pressure(Pc), and temperature(Tc) are depending on the substance.

Example) Ammonia:
ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

* 1 bar = 100kPa

3) Since the unit of s^{IG} is J/(g·K) and h^R is J/(mol·K), h^R is devided by molar mass(M, g/mol).

$$s^R[\mathrm{J/(g\cdot K)}] = \frac{s^R[\mathrm{J/(mol\cdot K)}]}{M[g/mol]}$$

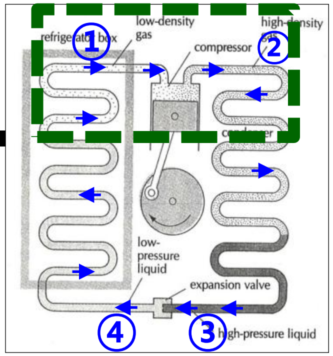
Example) Ammonia
$M_{Ammonia}$ = 17.031 (g/mol)

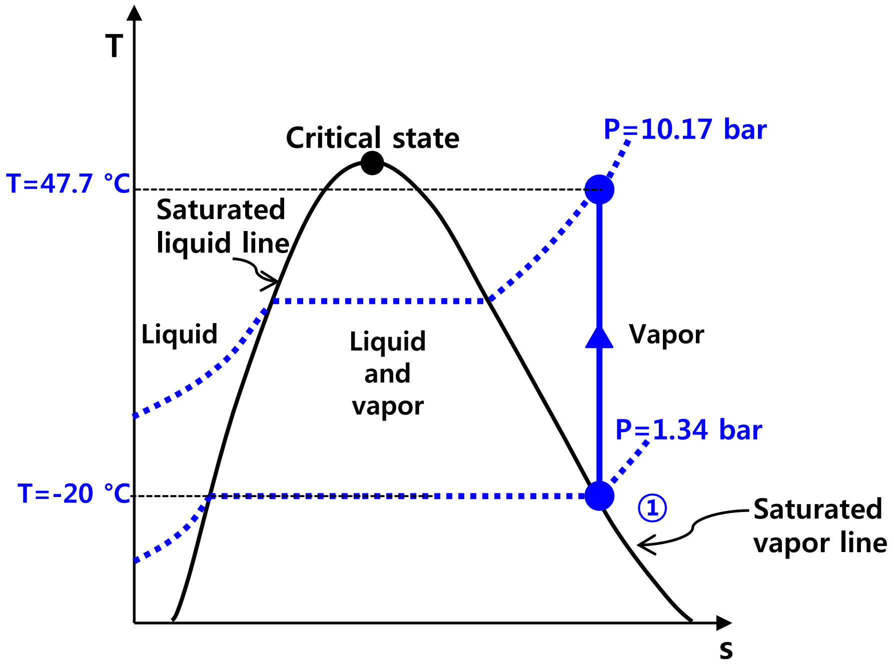
4) Central difference approximation

$$\left(\frac{da}{dT}\right) = \frac{a(T+e)-a(T-e)}{2\cdot e}, \left(e=10^{-6}\right)$$

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

Temperature(T)-Specific Entropy(s) Diagram





Assumption:
1. Adiabatic process
2. The process of the compressor is "reversible".
→Since, this process is "Adiabatic and reversible", the quality of energy "entropy" is not changed "**Isentropic process**".

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↑	↑	+W	↑	0

According to the assumption

$$\begin{bmatrix}\text{Specific entropy of the} \\ \text{refrigerant entering} \\ \text{the compressor}\end{bmatrix} = \begin{bmatrix}\text{Specific entropy of the} \\ \text{refrigerant leaving} \\ \text{the compressor}\end{bmatrix}$$

$$s_1\left(P_1, v_1, T_1\right) = s_2\left(P_2, v_2, T_2\right)$$

- Entropy
: The quality of energy

$$s = s^{IG} + s^{R}$$

$$= \int_{T_0}^{T} \frac{C_P(T)}{T} dT - \ln\frac{P}{P_0} + s^R(P, v, T)$$

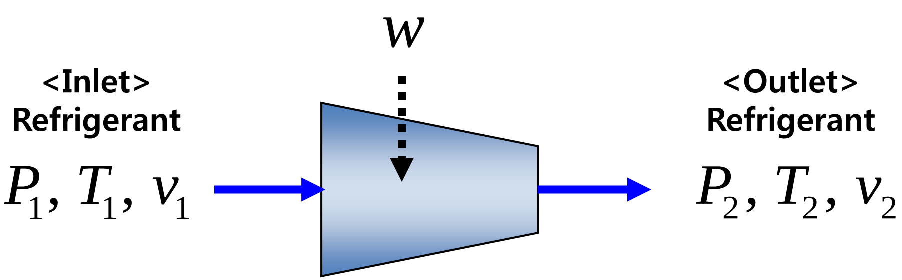
$C_p^{ig}(T)$: heat capacity at constant pressure

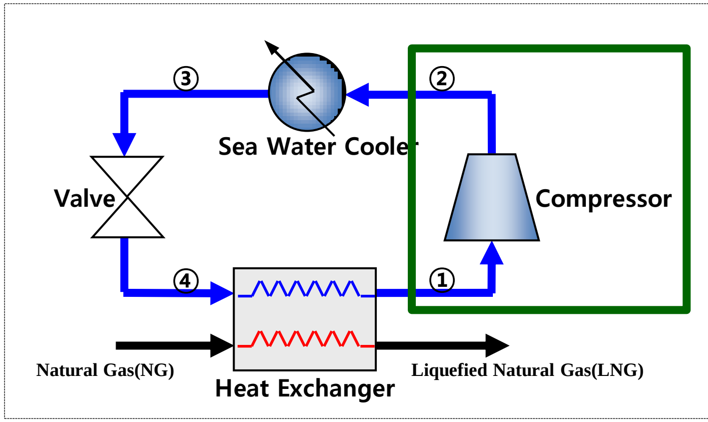
$$\frac{C_P^{ig}(T)}{R} = A + B\cdot T + C\cdot T^2 + D\cdot T^{-2}$$

R: gas constant(=8.314 Jmol^{-1}K^{-1})
T: temperature
SR(P,T): residual entropy

Mathematical Model of the Compressor (1/2)

- **Compressor**: brings the vapor refrigerant to a high pressure, which raises its temperature as well





1. Design variables(Operating Conditions): $P_1, T_1, v_1, P_2, T_2, v_2, w$

2. Assumption:
   1) There is not sufficient time to transfer much heat from the refrigerant*. "Adiabatic process"
   2) The process of the compressor is "reversible".
      →Since, this process is "Adiabatic and reversible", the quality of energy "entropy" is not changed.

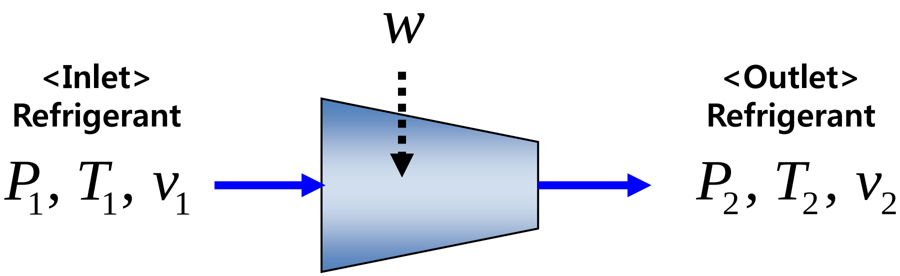
3. Equality constraints
   1) The first law of the thermodynamics(Energy conservation)

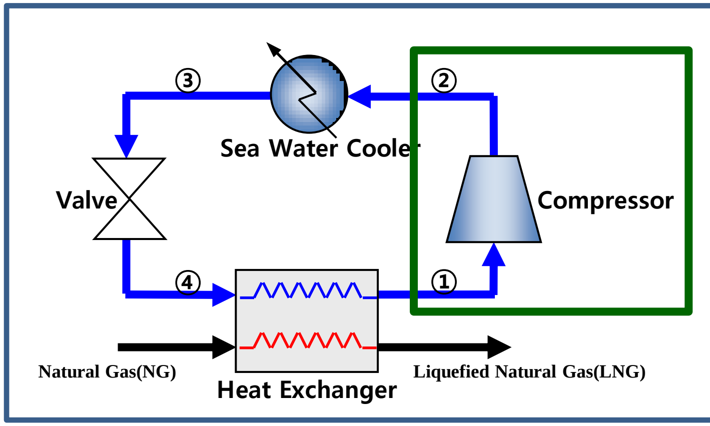
$$\underbrace{h_1(P_1, v_1, T_1)}_{\text{Energy of the refrigerant at the inlet}} + \underbrace{w}_{\text{Work input to the compressor per mass}} = \underbrace{h_2(P_2, v_2, T_2)}_{\text{Energy of the refrigerant at the outlet}}$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy

Mathematical Model of the Compressor (2/2)

- **Compressor**: brings the vapor refrigerant to a high pressure, which raises its temperature as well





3. Equality constraints

2) The second law of the thermodynamics
(For the adiabatic and reversible process, the quality of energy(entropy) is not changed.)

$$\underbrace{s_1(P_1, v_1, T_1)}_{\text{Quality of energy of the refrigerant at inlet}} = \underbrace{s_2(P_2, v_2, T_2)}_{\text{Quality of energy of the refrigerant at outlet}}$$

3) Equations of state(Soave, Redlich, Kwong(SRK) equation)

$$v_1 = \frac{RT_1}{P} + b - \frac{a(T_1)}{P_1} \frac{v_1 - b}{(v_1 - \varepsilon b)(v_1 - \sigma b)}$$

$$v_2 = \frac{RT_2}{P} + b - \frac{a(T_2)}{P_2} \frac{v_2 - b}{(v_2 - \varepsilon b)(v_2 - \sigma b)}$$

Equation of state
: Any equation that relates the pressure(P), temperature(T) and specific volume(V) of a substance.

Example) Equation of state for an ideal gas

$$Pv = RT$$

T: temperature
P: pressure
v: specific volume
s: specific entropy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ for SRK equation

R: gas constant (=8.314 Jmol[-1]K[-1])
P_c: critical pressure of the refrigerant
T_c: critical temperature of the refrigerant

$$b = \Omega \frac{RT_C}{P_C}$$

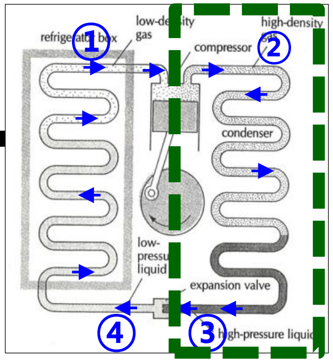
$\Omega = 0.08664$ for SRK equation
$\varepsilon = 0$ for SRK equation
$\sigma = 1$ for SRK equation

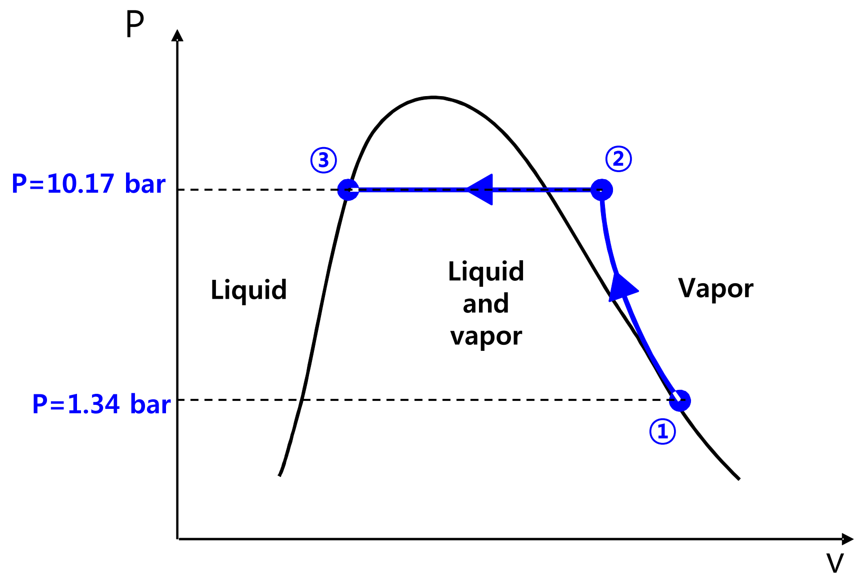
9.2. Process of the Refrigerator

9.2.3 CONDENSATION

2.3 Process of the Refrigerator - Condensation

Pressure(P)-Specific Volume(v) Diagram





Assumption:

There is no pressure drop of the refrigerant through the condenser. "Isobaric process"

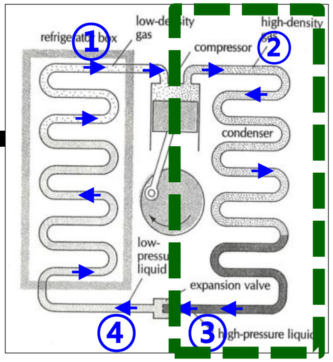
Purpose, Assumption, Result

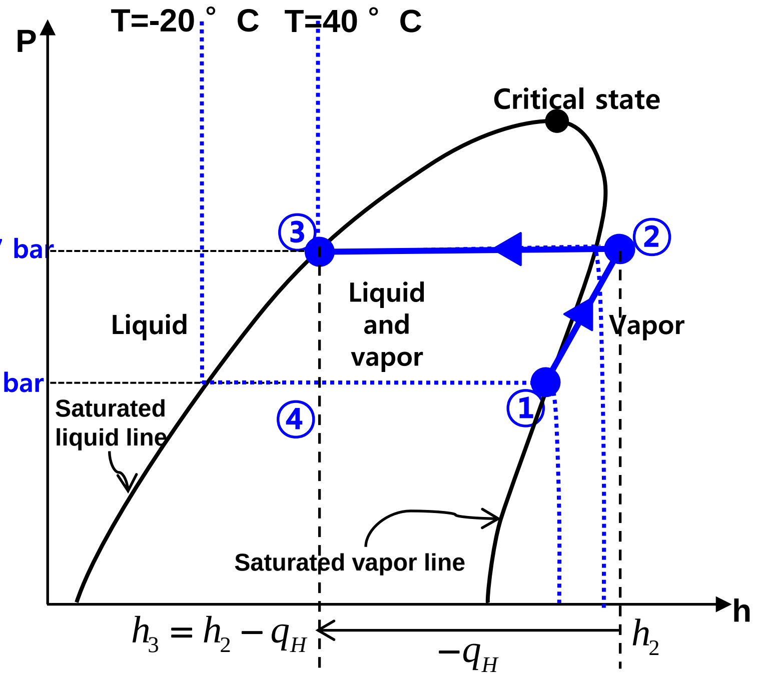
		Process	Temperature	Pressure
Condenser	2→3	Isobaric heat reinjection	⬇	0

Why do we assume the isobaric process in the condenser?

1. **Carnot cycle is the most efficient and ideal refrigeration cycle and condenses the refrigerant isothermally(Isothermal process).**
2. **However, the isothermal heat transfer from the refrigerant in a single phase is not easy to accomplish in practice.**
3. **Since maintaining a constant pressure in the condenser means maintain ng a constant temperature when the refrigerant is in a two-phase(liquid and vapor), the isothermal process is replaced by the isobaric process in the condenser**

Pressure(P)-Specific Enthalpy(h) Diagram





Assumption:

There is no pressure drop of the refrigerant through the condenser. "Isobaric process"

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↓	0	$-Q_H$	↓	↓

According to the first law of thermodynamics

$$\begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant entering}\\ \text{the condenser}\end{bmatrix} - \begin{bmatrix}\text{Total leaving energy}\\ \text{as Heat in the condenser}\end{bmatrix} = \begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant leaving}\\ \text{the condenser}\end{bmatrix}$$

$$h_2\left(P_2, v_2, T_2\right) - q_H = h_3\left(P_3, v_3, T_3\right)$$

- **Enthalpy**

$$h = [\textit{Internal energy of the flow}] + [\textit{Flow work}]$$
$$= u + P \cdot v$$

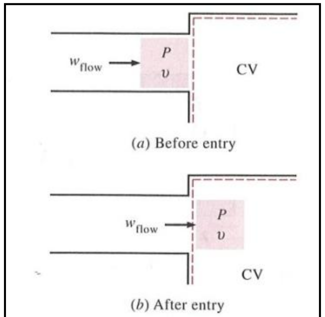


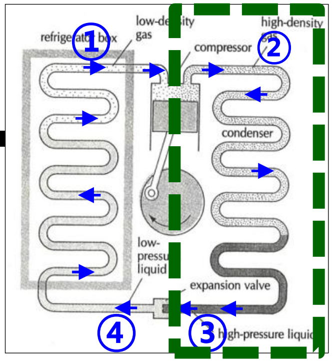
FIGURE 4-12
Flow work is the energy needed to push a fluid into or out of a control volume, and it is equal to Pv.

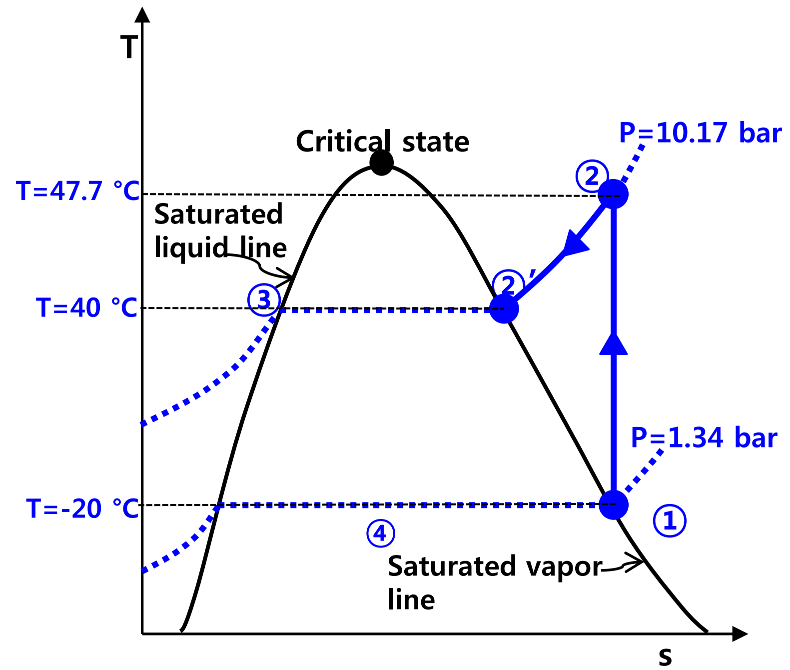
h : specific enthalpy
u : internal energy=$u(T,P)$
P : pressure
T : temperature

- **The first law of the thermodynamics:**
The total quantity of energy is constant.

2.3 Process of the Refrigerator - Condensation

Temperature(T)-Specific Entropy(s) Diagram (1/2)





Assumption:

There is no pressure drop of the refrigerant through the condenser. “Isobaric process”

- Entropy
  : The quality of energy

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↓	0	$-Q_H$	↓	↓

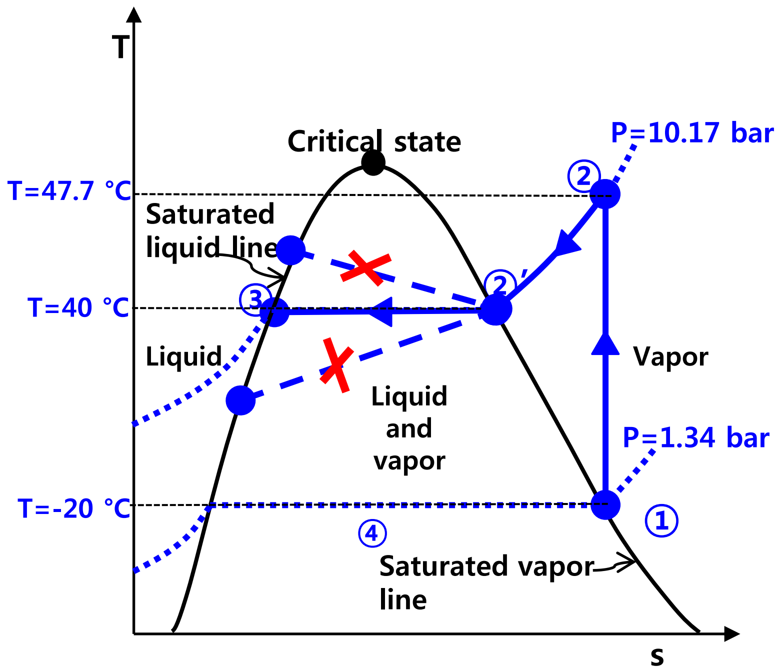
2 → 2′ Decrease of the temperature of the refrigerant:

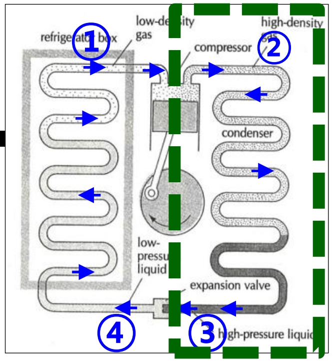
Because the heat of refrigerant is taken off to the atmosphere.

2 → 2′ Decrease of the entropy of the refrigerant:

Entropy can be viewed as a measure of molecular disorder, or molecular randomness.
The molecular disorder of the substance is decreased when the temperature of that is decreased.
Therefore, since the temperature of the refrigerant is decreased, the entropy of that is decreased.

Temperature(T)-Specific Entropy(s) Diagram (2/2)





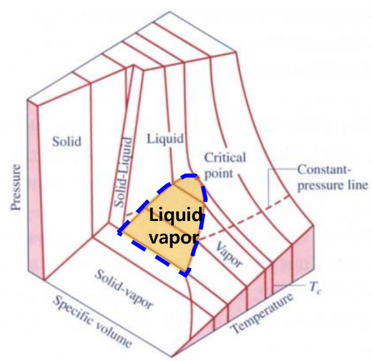
Assumption:

There is no pressure drop of the refrigerant through the condenser. "Isobaric process"

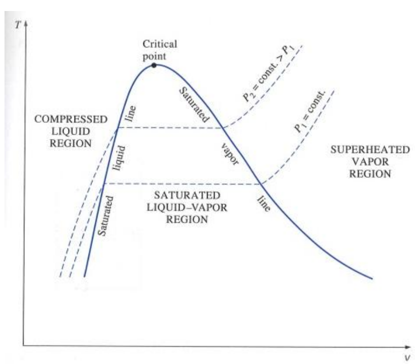
- Entropy
  : The quality of energy

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↓	0	$-Q_H$	↓	↓



P-v-T surface



v-T Diagram

2 ' → 3 Constant temperature of the refrigerant:

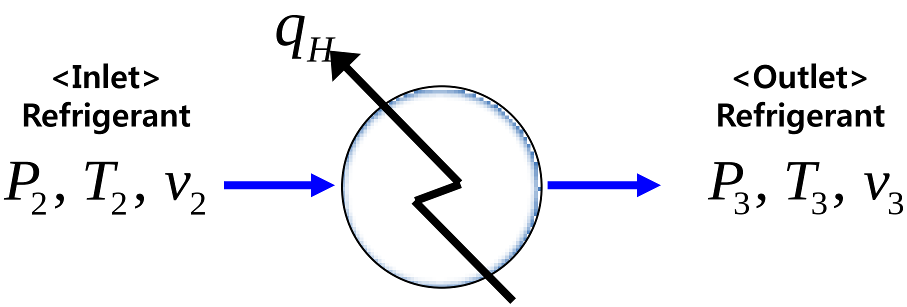
The temperature remains constant during the entire phase-change process if the pressure is held constant.

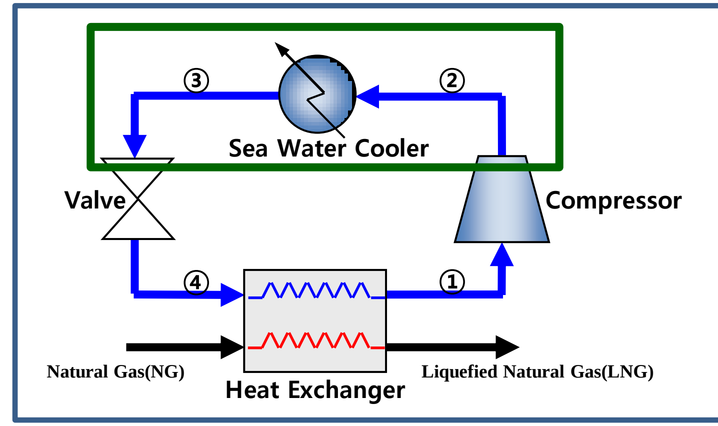
2 ' → 3 Decrease of the entropy of the refrigerant:

Entropy can be viewed as a measure of molecular disorder, or molecular randomness.
The molecular of the substance in the vapor phase is more disordered than that in liquid phase.
Therefore, since the liquid part of the refrigerant increases, the entropy of that is decreased.

Mathematical Model of the Sea Water Cooler

- **Sea Water(SW) Cooler**: takes off the heat from the hot vapor refrigerant to the sea water





1. Design variables(Operating Conditions): $P_2, T_2, v_2, P_3, T_3, v_3$

2. Assumption:
- There is no pressure drop of the refrigerant through the sea water cooler. "Isobaric process"

3. Equality constraints

1) The first law of the thermodynamics(Energy conservation)

$$\underbrace{h_2(P_2, v_2, T_2)}_{\text{Energy of the refrigerant at the inlet}} - q_H = \underbrace{h_3(P_3, v_3, T_3)}_{\text{Energy of the refrigerant at the outlet}}$$

q_H: Specific heat transfer from the refrigerant to sea water(Given)

2) Isobaric process

$$P_2 = P_3$$

3) Equations of state(Soave, Redlich, Kwong(SRK) equation)

$$v_3 = \frac{RT_3}{P} + b - \frac{a(T_3)}{P_3} \frac{v_3 - b}{(v_3 - \varepsilon b)(v_3 - \sigma b)}$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ for SRK equation

R: gas constant (=8.314 Jmol^{-1}K^{-1})
P_c: critical pressure of the refrigerant
T_c: critical temperature of the refrigerant

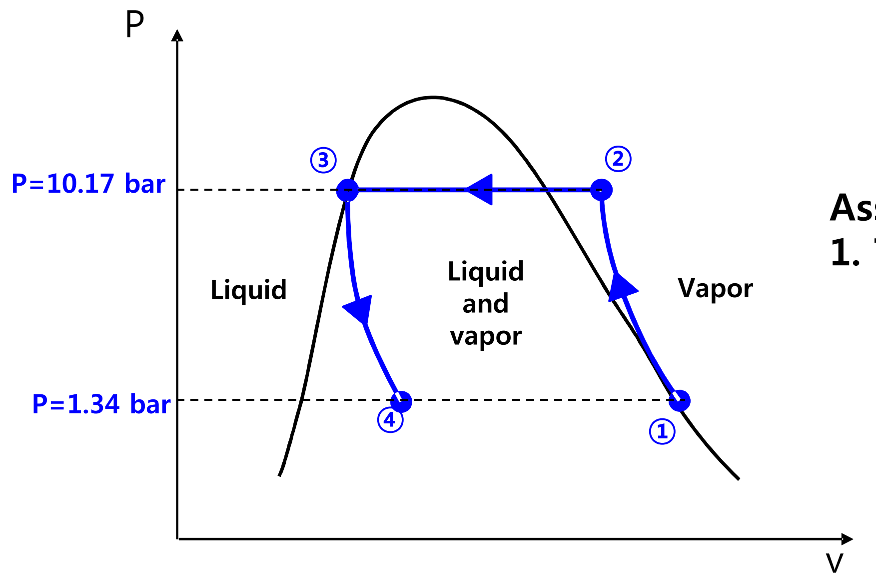
$$b = \Omega \frac{RT_C}{P_C}$$

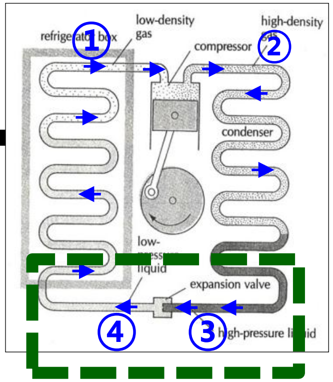
$\Omega = 0.08664$ for SRK equation
$\varepsilon = 0$ for SRK equation
$\sigma = 1$ for SRK equation

9.2. Process of the Refrigerator

9.2.4 Expansion

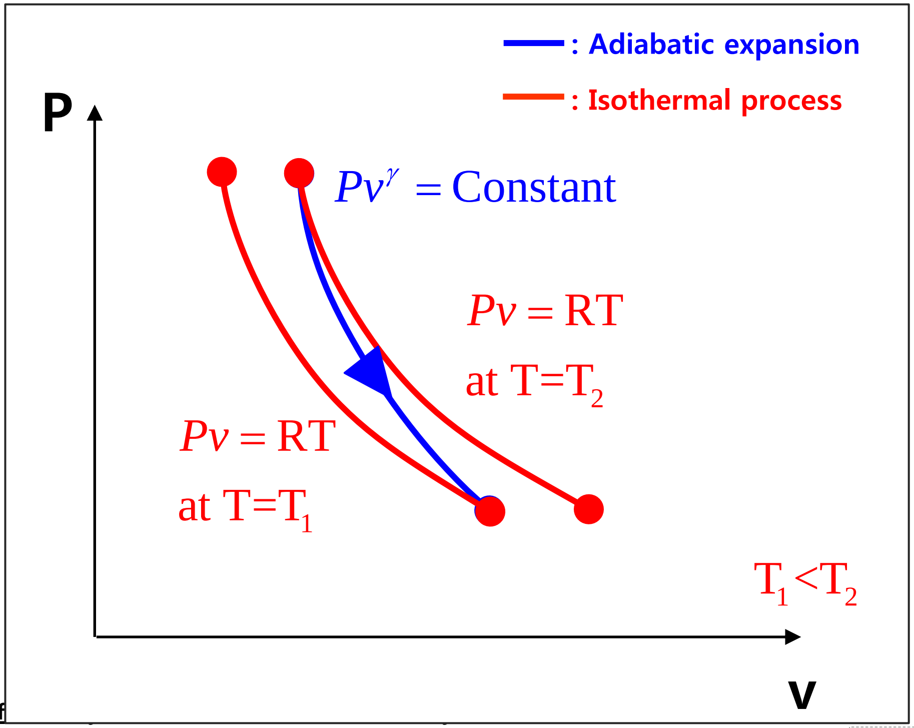
Pressure(P)-Specific Volume(v) Diagram





Assumption:

1. There is not sufficient time to transfer much heat from the refrigerant. "Adiabatic process"



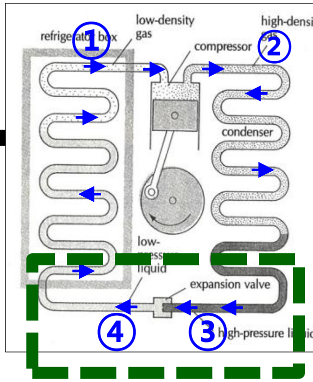
Purpose, Assumption, Result

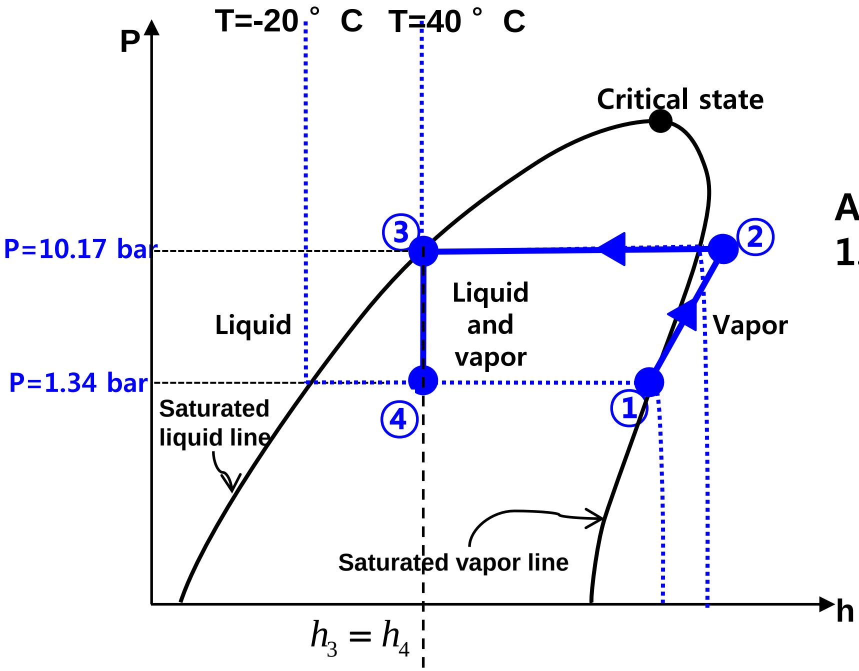
		Process	Temperature	Pressure
Expansion valve	3→4	Adiabatic expansion	↓	↓

Why is the temperature of the refrigerant decreased after the expansion valve?

Because of the adiabatic expansion, the temperature of the refrigerant is decreased from T_2 to T_1.

Pressure(P)-Specific Enthalpy(h) Diagram





Assumption:

1. There is not sufficient time to transfer much heat from the refrigerant. "Adiabatic process"

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↓	↓	0	0	↑

According to the first law of thermodynamics

$$\begin{bmatrix} \text{Total energy of the} \\ \text{refrigerant entering} \\ \text{the valve} \end{bmatrix} = \begin{bmatrix} \text{Total energy of the} \\ \text{refrigerant leaving} \\ \text{the valve} \end{bmatrix}$$

$$h_3\left(P_3, v_3, T_3\right) = h_4\left(P_4, v_4, T_4\right)$$

- Enthalpy

$$h = \left[\textit{Internal energy of the flow}\right] + \left[\textit{Flow work}\right]$$
$$= u + P \cdot v$$

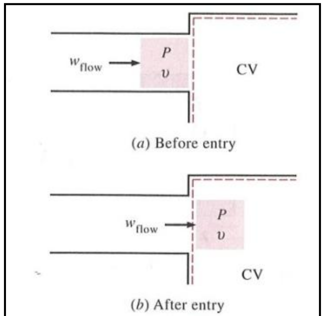
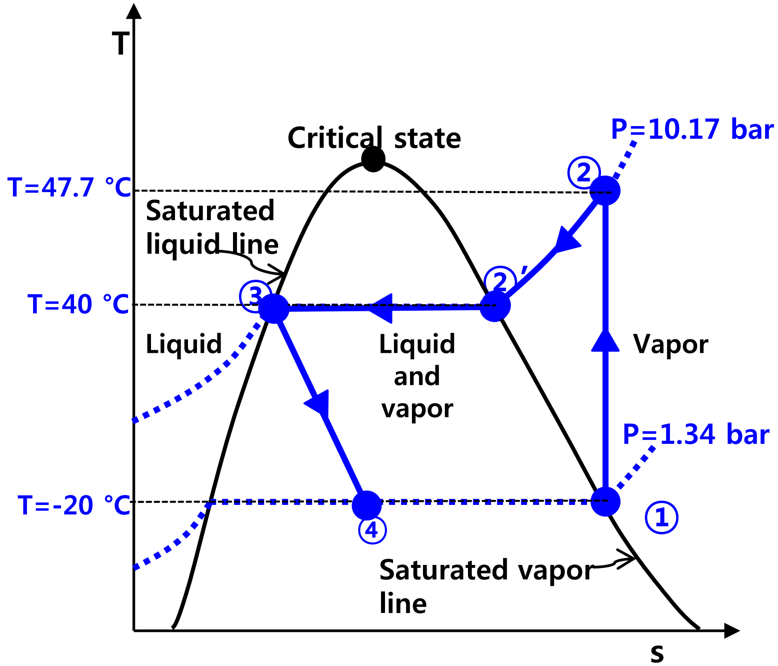


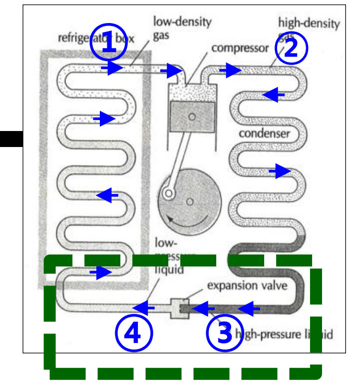
FIGURE 4-12
Flow work is the energy needed to push a fluid into or out of a control volume, and it is equal to Pv.

h : specific enthalpy
u : internal energy=$u(T,P)$
P : pressure
T : temperature

- The first law of the thermodynamics:
The total quantity of energy is constant.

Temperature(T)-Specific Entropy(s) Diagram





Natural Phenomena

By restricting the flow of the refrigerant, the pressure of the refrigerant is decreased. "Irreversible process"

→ Increase the specific entropy of the refrigerant

Assumption:

1. There is not sufficient time to transfer much heat from the refrigerant. "Adiabatic process"

- Entropy
  : The quality of energy

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↓	↓	0	0	↑

3 → 4 <u>Decrease of the temperature of the refrigerant</u>:

When the pressure of the refrigerant is decreased, the boiling temperature of that is also decreased. Since the boiling temperature is decreased, a part of the liquid refrigerant is evaporated by absorbing the heat from itself. Therefore, the temperature of the refrigerant is decreased.

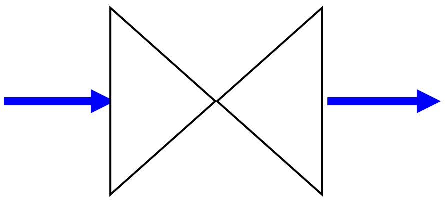
3 → 4 <u>Increase of the entropy of the refrigerant</u>:

<u>Entropy</u> can be viewed as a <u>measure of molecular disorder, or molecular randomness.</u>
The molecular disorder of the substance is decreased when the temperature of that is decreased.
The molecular of the substance in the vapor phase is more disordered than that in liquid phase.
<u>Since the increase of the entropy caused by the phase-change is larger than the decrease of that caused by the decrease of the temperature, the entropy of the refrigerant is increased.</u>

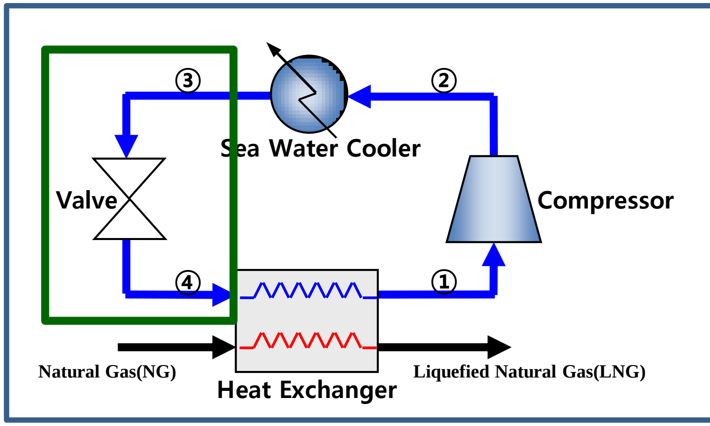
Mathematical Model of the Valve

- **Valve**: decreases the pressure of the liquid refrigerant, which decreases its temperature as well

<Inlet>
Refrigerant
P_3, T_3, v_3



<Outlet>
Refrigerant
P_4, T_4, v_4



1. Design variables(Operating Conditions): P_3, T_3, v_3, P_4, T_4, v_4

2. Assumption:
   1) There is not sufficient time to transfer much heat from the refrigerant. "Adiabatic process"
   2) By restricting the flow of the refrigerant, the pressure of the refrigerant is decreased. "Irreversible process"

3. Equality constraints
   1) The first law of the thermodynamics(Energy conservation)

$$\underbrace{h_3\left(P_3, v_3, T_3\right)}_{\text{Energy of the refrigerant at the inlet}} = \underbrace{h_4\left(P_4, v_4, T_4\right)}_{\text{Energy of the refrigerant at the outlet}}$$

   2) Equations of state(Soave, Redlich, Kwong(SRK) equation)

$$v_4 = \frac{RT_4}{P} + b - \frac{a(T_4)}{P_4}\frac{v_4 - b}{(v_4 - \varepsilon b)(v_4 - \sigma b)}$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ *for SRK equation*

R: gas constant (=8.314 $Jmol^{-1}K^{-1}$)
P_c: critical pressure of the refrigerant
T_c: critical temperature of the refrigerant

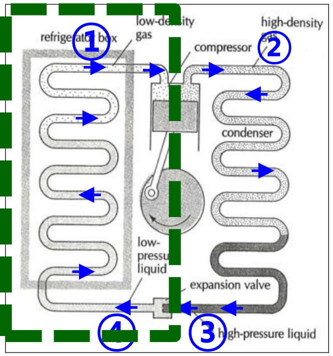
$$b = \Omega \frac{RT_C}{P_C}$$

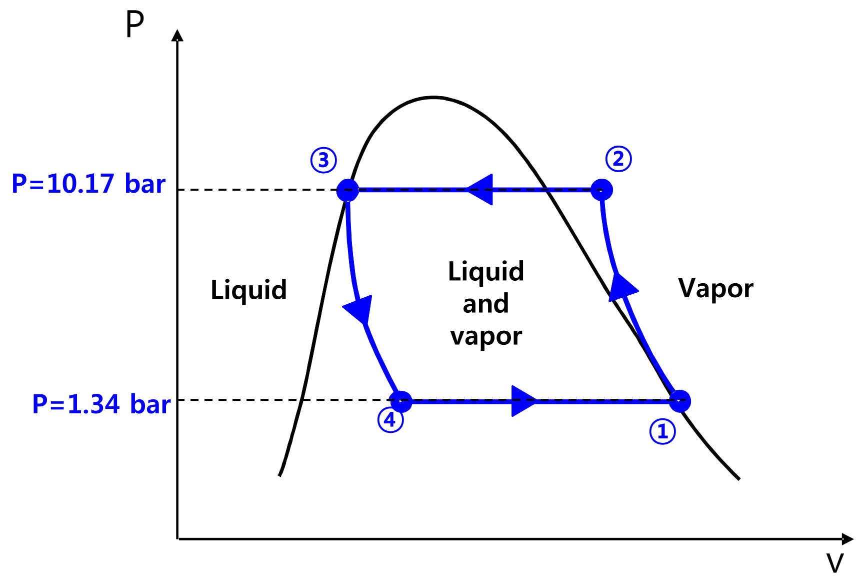
$\Omega = 0.08664$ *for SRK equation*
$\varepsilon = 0$ *for SRK equation*
$\sigma = 1$ *for SRK equation*

9.2. Process of the Refrigerator

9.2.5 Evaporation

Pressure(P)-Specific Volume(v) Diagram





Assumption:

There is no pressure drop of the refrigerant through the evaporator. "Isobaric process"

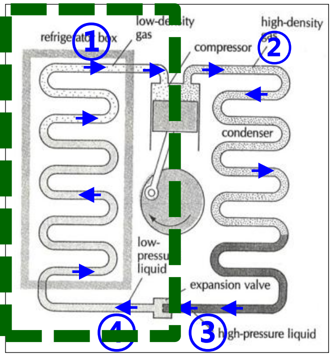
Purpose, Assumption, Result

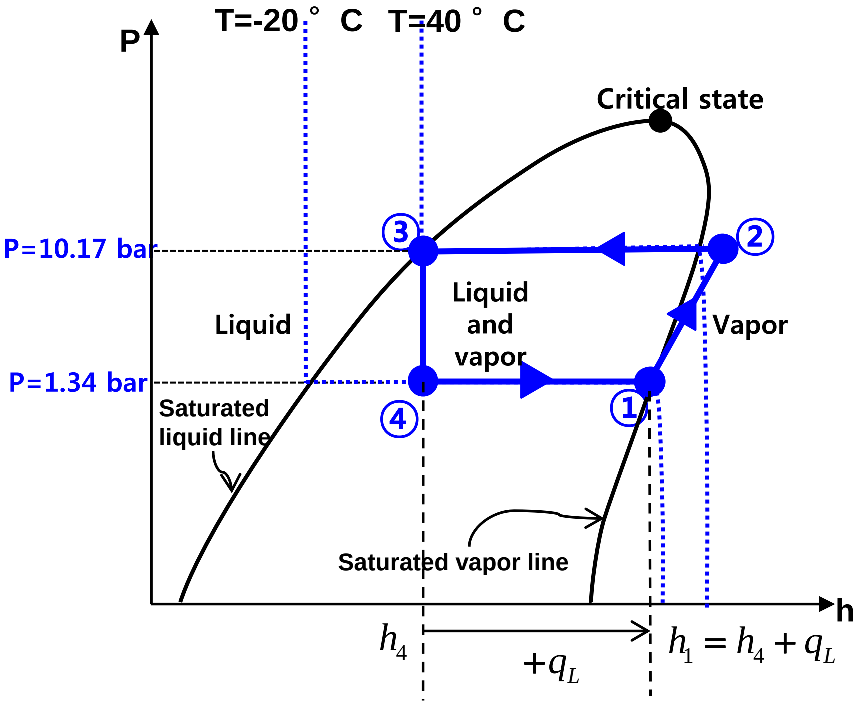
		Process	Temperature	Pressure
Evaporator	4→1	Isobaric heat absorption	↑	0

Why do we assume the isobaric process in the condenser?

1. Carnot cycle is the most efficient and ideal refrigeration cycle and condenses the refrigerant isothermally(Isothermal process).
2. However, the isothermal heat transfer from the refrigerant in a single phase is not easy to accomplish in practice.
3. Since maintaining a constant pressure in the condenser fixes the temperature when the refrigerant is in a two-phase(liquid and vapor), the isothermal process is replaced by the isobaric process in the condenser

Pressure(P)-Specific Enthalpy(h) Diagram





Assumption:

There is no pressure drop of the refrigerant through the evaporator. "Isobaric process"

Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↑	0	$+Q_L$	↑	↑

According to the first law of thermodynamics

$$\begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant entering}\\ \text{the evaporator}\end{bmatrix} + \begin{bmatrix}\text{Total leaving energy}\\ \text{as Heat in the evaporator}\end{bmatrix} = \begin{bmatrix}\text{Total energy of the}\\ \text{refrigerant leaving}\\ \text{the evaporator}\end{bmatrix}$$

$$h_4\left(P_4, v_4, T_4\right) + q_L = h_1\left(P_1, v_1, T_1\right)$$

- **Enthalpy**

$$h = [\textit{Internal energy of the flow}] + [\textit{Flow work}]$$
$$= u + P \cdot v$$

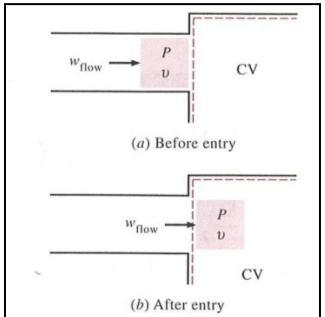


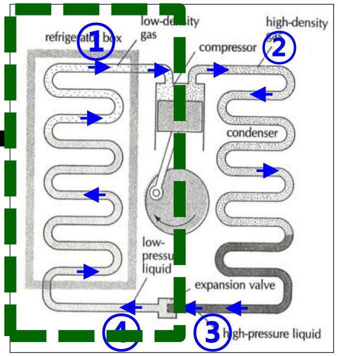
FIGURE 4-12
Flow work is the energy needed to push a fluid into or out of a control volume, and it is equal to Pv.

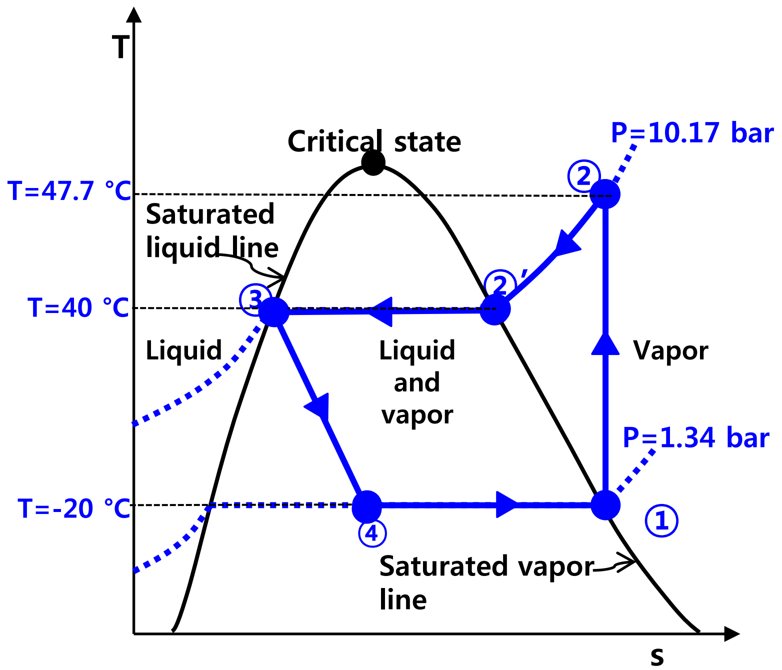
h : specific enthalpy
u : internal energy=$u(T,P)$
P : pressure
T : temperature

- **The first law of the thermodynamics: The total quantity of energy is constant.**

2.5 Process of the Refrigerator - Evaporation

Temperature(T)-Specific Entropy(s) Diagram

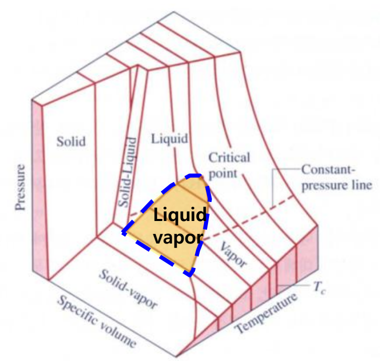


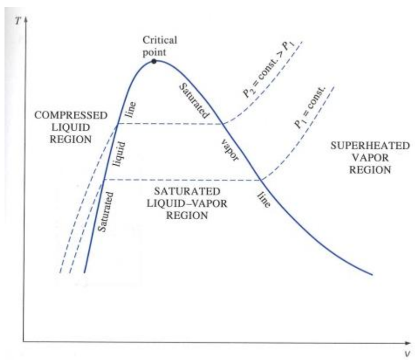


Assumption:

There is no pressure drop of the refrigerant through the condenser. "Isobaric process"

- Entropy
  : The quality of energy





Purpose, Assumption, Result

Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
↑	0	$+Q_L$	↑	↑

4 → 1 <u>Constant temperature of the refrigerant</u>:

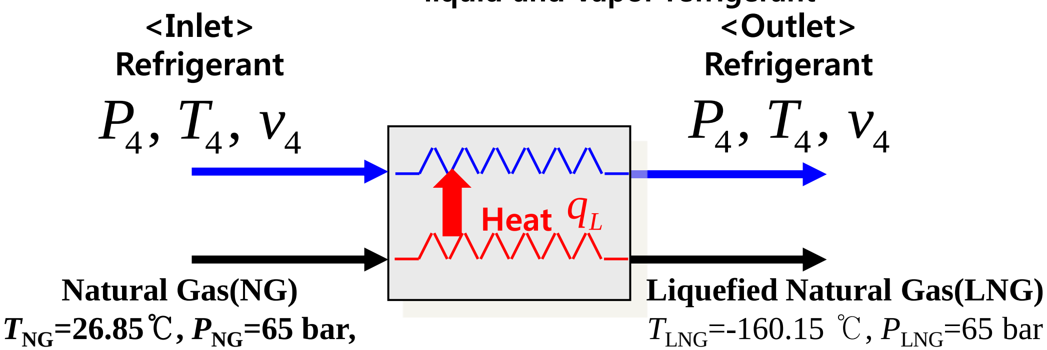
The temperature remains constant during the entire phase-change process if the pressure is held constant.

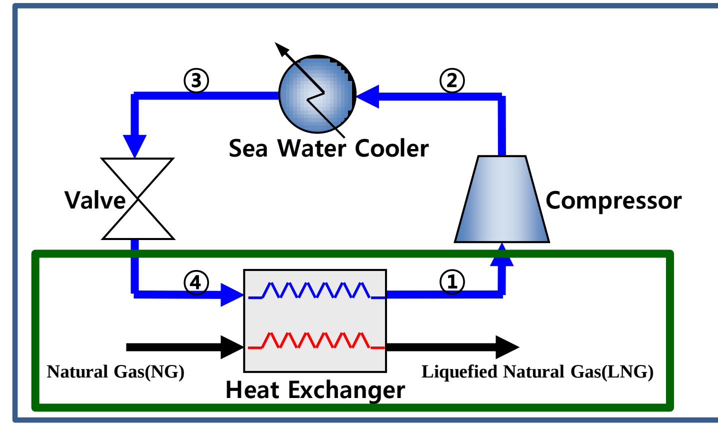
4 → 1 <u>Increase of the entropy of the refrigerant</u>:

Entropy can be viewed as a measure of molecular disorder, or molecular randomness.
The molecular of the substance in the vapor phase is more disordered than that in liquid phase.
<u>Therefore, since the vapor part of the refrigerant increases, the entropy of that is also increased</u>.

Mathematical Model of the Heat Exchanger (1/2)

- **Heat Exchanger**: takes off the heat from the natural gas to cool liquid and vapor refrigerant





1. Design variables(Operating Conditions): $P_1, T_1, v_1, P_4, T_4, v_4$

2. Assumption:
- There is no pressure drop of the refrigerant through the heat exchanger. "Isobaric process"

T: temperature
P: pressure
v: specific volume
h: specific enthalpy

3. Equality constraints

1) The first law of the thermodynamics(Energy conservation)

$$\underbrace{h_4\left(P_4, v_4, T_4\right)}_{\text{Energy of the refrigerant at the inlet}} + q_L = \underbrace{h_1\left(P_1, v_1, T_1\right)}_{\text{Energy of the refrigerant at the outlet}}$$

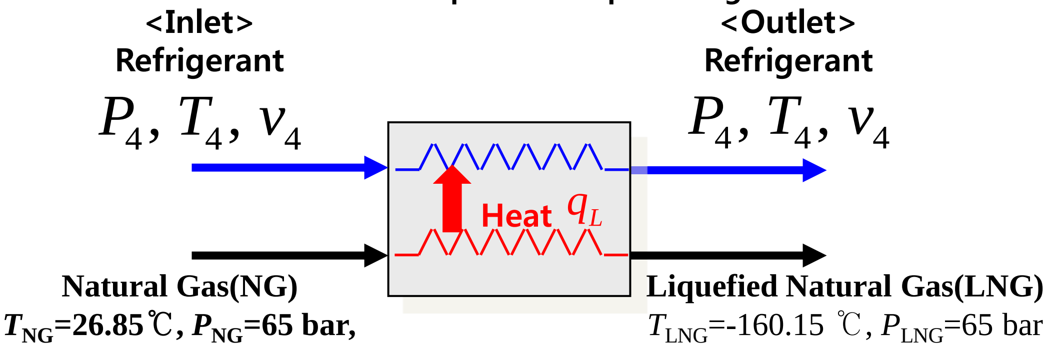
2) Isobaric process

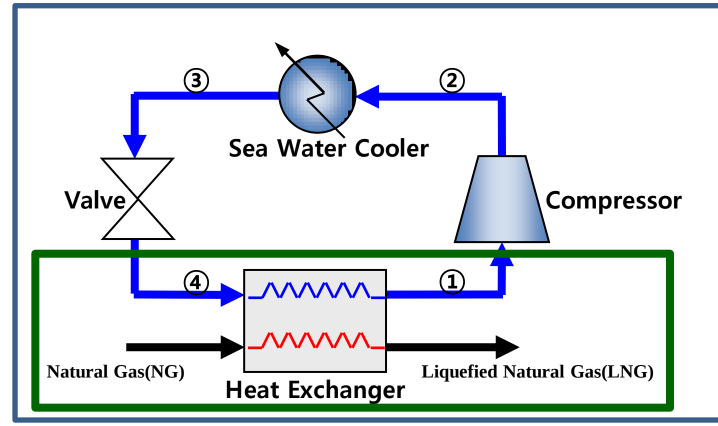
$$P_4 = P_1$$

q_L: Heat transfer for the liquefaction of the natural gas(Given)

Mathematical Model of the Heat Exchanger (2/2)

- **Heat Exchanger**: takes off the heat from the natural gas to cool liquid and vapor refrigerant





To produce the $\dot{m}_{NG}$ MTPA(Million ton per annual) LNG, the refrigerant has to take off the heat Q_L from NG.

$$Q_L = \dot{m}_{NG} \cdot q_L$$

$\dot{m}_{NG}$: Mass flow rate of the natural gas(Given, usually 3.6 MTPA)

q_L : Specific heat transfer for the liquefaction of the natural gas(Given)

$\dot{m}_{ref}$: Mass flow rate of the refrigerant

1. Design variables(Operating Conditions): P_1, T_1, v_1, P_4, T_4, v_4, $\dot{m}_{ref}$

3. Equality constraints

1) The first law of the thermodynamics(Energy conservation)

Energy of the refrigerant at the inlet

Energy of the refrigerant at the outlet

2) Isobaric process

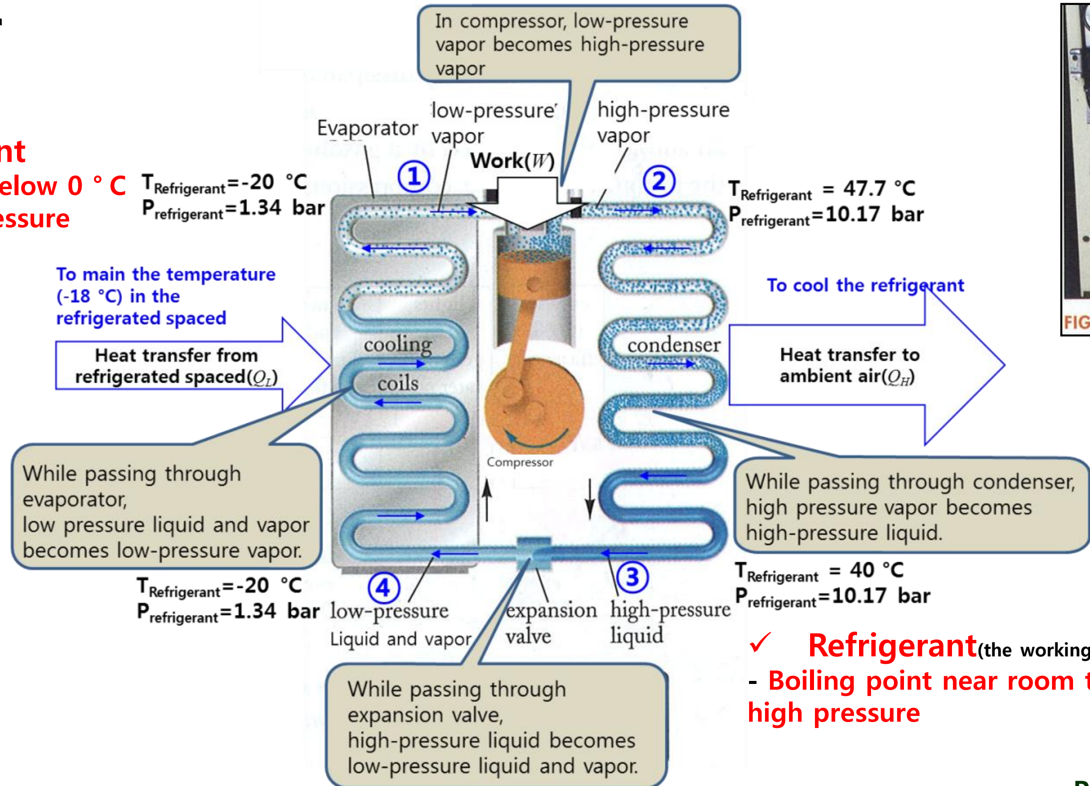
$$P_4 = P_1$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy

Introduction to the Cooling System for Refrigerator (2/2)

- **Refrigerator**

✓ **Refrigerant**
-Boiling point below 0 ° C when at low pressure



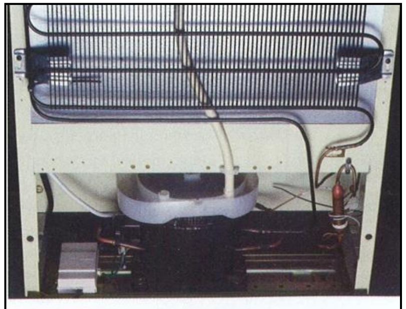


FIGURE 21.19 Condenser coils of a refrigerator.

✓ **Refrigerant**(the working substance in the refrigerator)
- Boiling point near room temperature when at high pressure

Purpose, Assumption, Result

		Process	Temperature	Pressure	Work or heat transfer	Enthalpy	Entropy
Compressor	1→2	Adiabatic compression	↑	↑	+W	↑	0
Condenser	2→3	Isobaric heat reinjection	↓	0	$-Q_H$	↓	↓
Expansion valve	3→4	Adiabatic expansion	↓	↓	0	0	↑
Evaporator	4→1	Isobaric heat absorption	↑	0	$+Q_L$	↑	↑

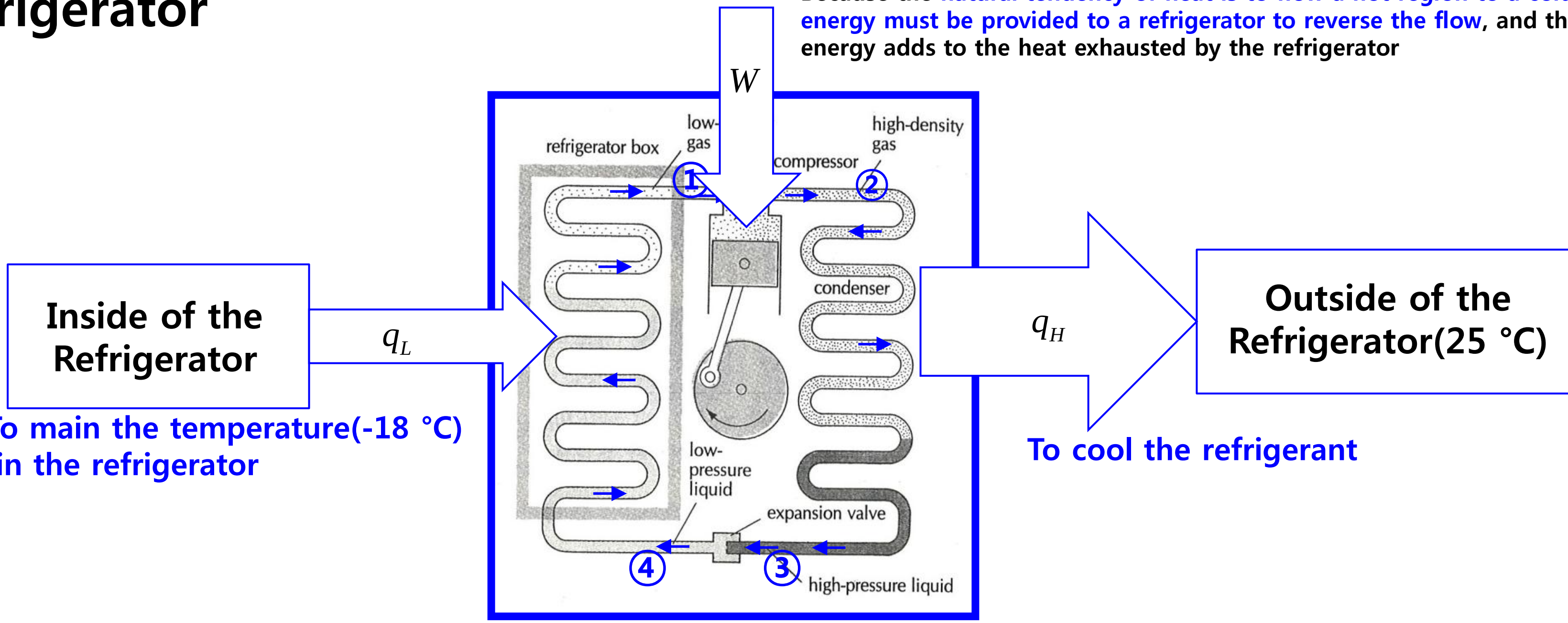
9.2. Process of the Refrigerator

9.2.6 OPERATING CONDITION

2.6 Efficiency of a Refrigerator(CP: coefficient of performance)

- **Refrigerator**

Because the natural tendency of heat is to flow a hot region to a cold one, energy must be provided to a refrigerator to reverse the flow, and this energy adds to the heat exhausted by the refrigerator



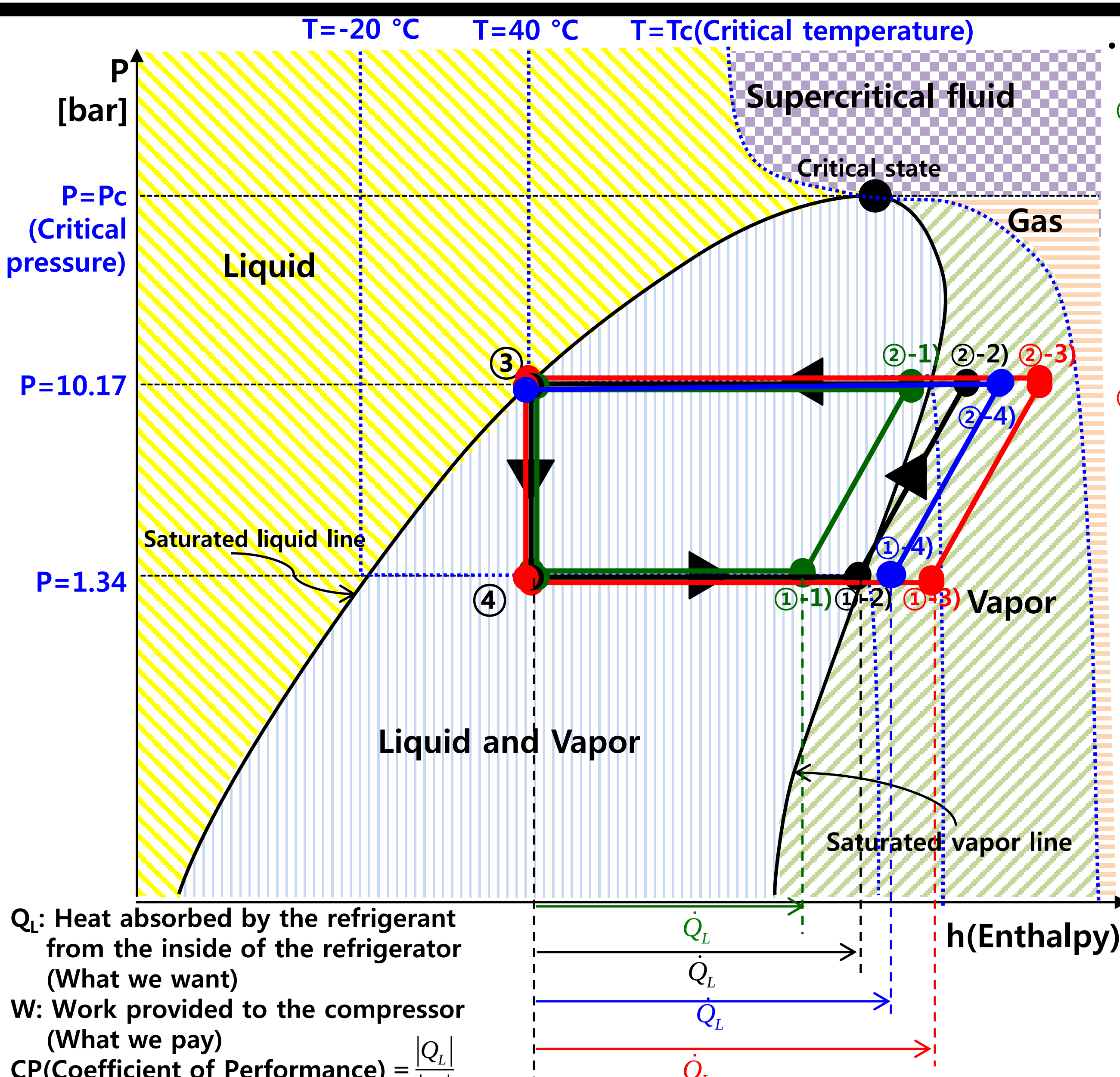
What is the efficiency of a refrigerator(CP: coefficient of performance)?

$$CP = \frac{\text{What we want}}{\text{What we pay for}}$$

$$= \frac{|Q_L|}{|W|}$$

To increase the efficiency of a refrigerator, when Q_L is given, we have to determine the operating conditions such as pressure, temperature, specific volume and flow rate for decreasing the work provided to the compressor.

2.6 Effect of the Operating Condition to the Refrigerator – Position of the Point ① (Superheating)



Q_L: Heat absorbed by the refrigerant from the inside of the refrigerator (What we want)

W: Work provided to the compressor (What we pay)

CP(Coefficient of Performance) = $\frac{|Q_L|}{|W|}$

- State of the refrigerant entering the compressor

①-1): Liquid and Vapor → Bad

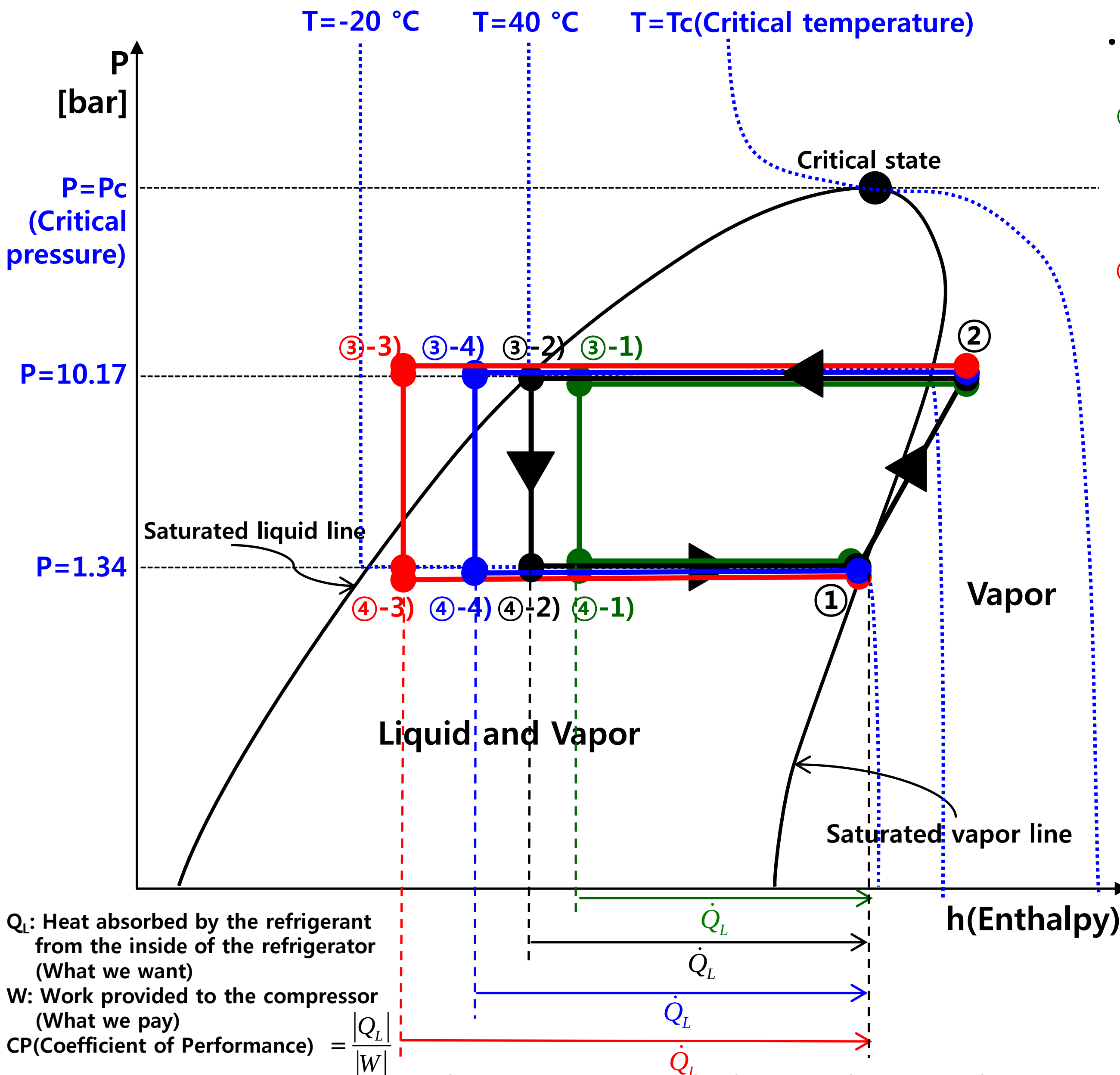
1. Since Q_L of ①-1) is smaller than that of ①-2), CP is decreased comparing with ①-2).
2. Since refrigeration compressors are designed as vapor pumps, if any amount of liquid is allowed to enter the compressor, serious mechanical damage to the compressor may result.

①-3), ①-4): Vapor state

1. For the superheated cycles ①-3) and ①-4), a greater quantity of heat must be dissipated at the condenser than for the cycle ①-2). → Increase of the size of the condenser
2. The work provided to the compressor for superheated cycle is slightly greater than that for the cycle ①-2).
3. For the cycle ①-4), since the increase in Q_L is greater than the increase in W, the CP is higher than that of the cycle ①-2). → Good
4. For the cycle ①-3), since the increase in Q_L is smaller than the increase in W, the CP is lower than that of the cycle ①-2). → Bad

Therefore, we have to determine the proper temperature of the refrigerant entering the compressor.

2.6 Effect of the Operating Condition to the Refrigerator – Position of the Point ③ (Subcooling)



- State of the refrigerant entering the valve

③-1): Liquid and Vapor → Bad

1. Since Q_L of ①-1) is smaller than that of ①-2), CP is decreased comparing with ①-2).

③ -3), ③ -4): Liquid state

1. For the subcooled cycles ③-3) and ③ -4), the increase of QL is accomplished without increasing the energy input to the compressor.
   → Increase of CP→ Good

2. However, to subcool the refrigerant in condenser, the additional equipment is needed to cool the refrigerant. → Bad

Therefore, we have to determine the proper temperature of the refrigerant entering the valve considering the Q_L and cost of the additional equipment caused by the subcooling.

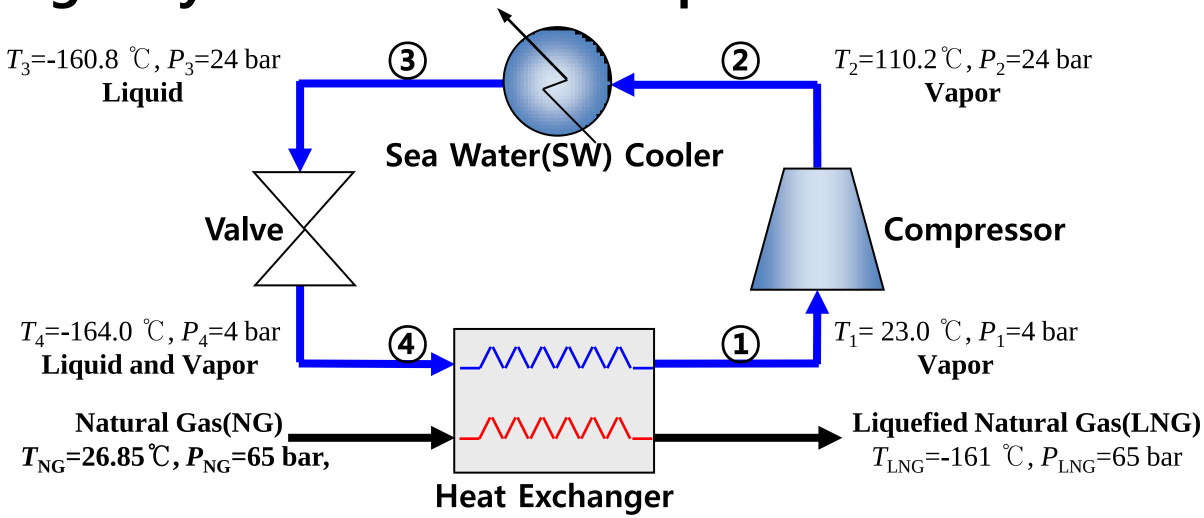
Q_L: Heat absorbed by the refrigerant from the inside of the refrigerator (What we want)
W: Work provided to the compressor (What we pay)
CP(Coefficient of Performance) $= \frac{|Q_L|}{|W|}$

2. Process of the Refrigerator

2.7 Mathematical Model of the Single Liquefaction Cycle

2.7 Mathematical Model of the Single Liquefaction Cycle (1/2)

- **Single Cycle for the LNG Liquefaction Process**



T: temperature
P: pressure
v: specific volume

$\dot{m}_{ref}$: Mass flow rate of the refrigerant

1. **Design Variables(Operating Condition, 14):** P_i, T_i, v_i, w, $\dot{m}_{ref}$ (i=1,2,3,4)

2. **Equality Constraint(11)**
   - Compressor(4)
   - Sea Water Cooler(3)
   - Valve(2)
   - Heat Exchanger(2)

The first law of the thermodynamics (Conservation of energy, Enthalpy)
The second law of the thermodynamics(Actual processes occur in the direction of decreasing quality of energy, Entropy)
Equation of state(Example: Soave, Redlich, Kwong(SRK) equation)

→ **Number of the design variables is larger than the number of the equality constraints.** → **Indeterminate Equation!**

3. **Objective Function:** Minimize the compressor powe $\mathrm{Min}(\dot{m}_{ref} \cdot w)$

→ **Optimization Problem!**

2.7 Mathematical Model of the Single Liquefaction Cycle (2/2)

1. Design Variables(Operating Condition, 14): P_i, T_i, v_i, w, $\dot{m}_{ref}$ $(i=1,2,3,4)$

2. Equality Constraint(11)

1) Compressor(4)

$h_1(P_1,v_1,T_1)+w=h_2(P_2,v_2,T_2)$ **[The first law of the thermodynamics]**

$s_1(P_1,v_1,T_1)=s_2(P_2,v_2,T_2)$ **[The second law of the thermodynamics]**

$$v_1=\frac{RT_1}{P}+b-\frac{a(T_1)}{P_1}\frac{v_1-b}{(v_1-\varepsilon b)(v_1-\sigma b)}$$ **[Equation of state]**

$$v_2=\frac{RT_2}{P}+b-\frac{a(T_2)}{P_2}\frac{v_2-b}{(v_2-\varepsilon b)(v_2-\sigma b)}$$ **[Equation of state]**

2) Sea Water Cooler(3)

$h_2(P_2,v_2,T_2)-q_H=h_3(P_3,v_3,T_3)$ **[The first law of the thermodynamics]**

$P_2=P_3$ **[Isobaric process]**

$$v_3=\frac{RT_3}{P}+b-\frac{a(T_3)}{P_3}\frac{v_3-b}{(v_3-\varepsilon b)(v_3-\sigma b)}$$ **[Equation of state]**

3) Valve(2)

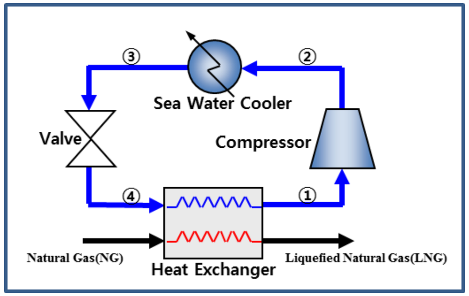
$h_3(P_3,v_3,T_3)=h_4(P_4,v_4,T_4)$ **[The first law of the thermodynamics]**

$$v_4=\frac{RT_4}{P}+b-\frac{a(T_4)}{P_4}\frac{v_4-b}{(v_4-\varepsilon b)(v_4-\sigma b)}$$ **[Equation of state]**

4) Heat Exchanger(2)

$\dot{m}_{ref}\cdot h_4(P_4,v_4,T_4)+\dot{m}_{NG}\cdot q_L=\dot{m}_{ref}\cdot h_1(P_1,v_1,T_1)$
[The first law of the thermodynamics]

$P_4=P_1$ **[Isobaric process]**



T: temperature, h: specific enthalpy, s: specific entropy,
P: pressure
v: specific volume
w: work provided to the compressor per mass
q_H: Specific heat transfer from the refrigerant to sea water(Given)
q_L: Heat transfer for the liquefaction of the natural gas(Given)
$\dot{m}_{NG}$: Mass flow rate of the natural gas(Given, usually 3.6 MTPA)
$\dot{m}_{ref}$: Mass flow rate of the refrigerant

$$a(T)=\psi\frac{\alpha(T_r)R^2T_C^2}{P_C}$$
$$b=\Omega\frac{RT_C}{P_C}$$

R: gas constant (=8.314 Jmol^{-1}K^{-1})
P_c: critical pressure of the refrigerant
T_c: critical temperature of the refrigerant

$\psi=0.42748,\ \Omega=0.08664,\ \varepsilon=0$ and $\sigma=1$ *for SRK equation*

3. Objective function(f)

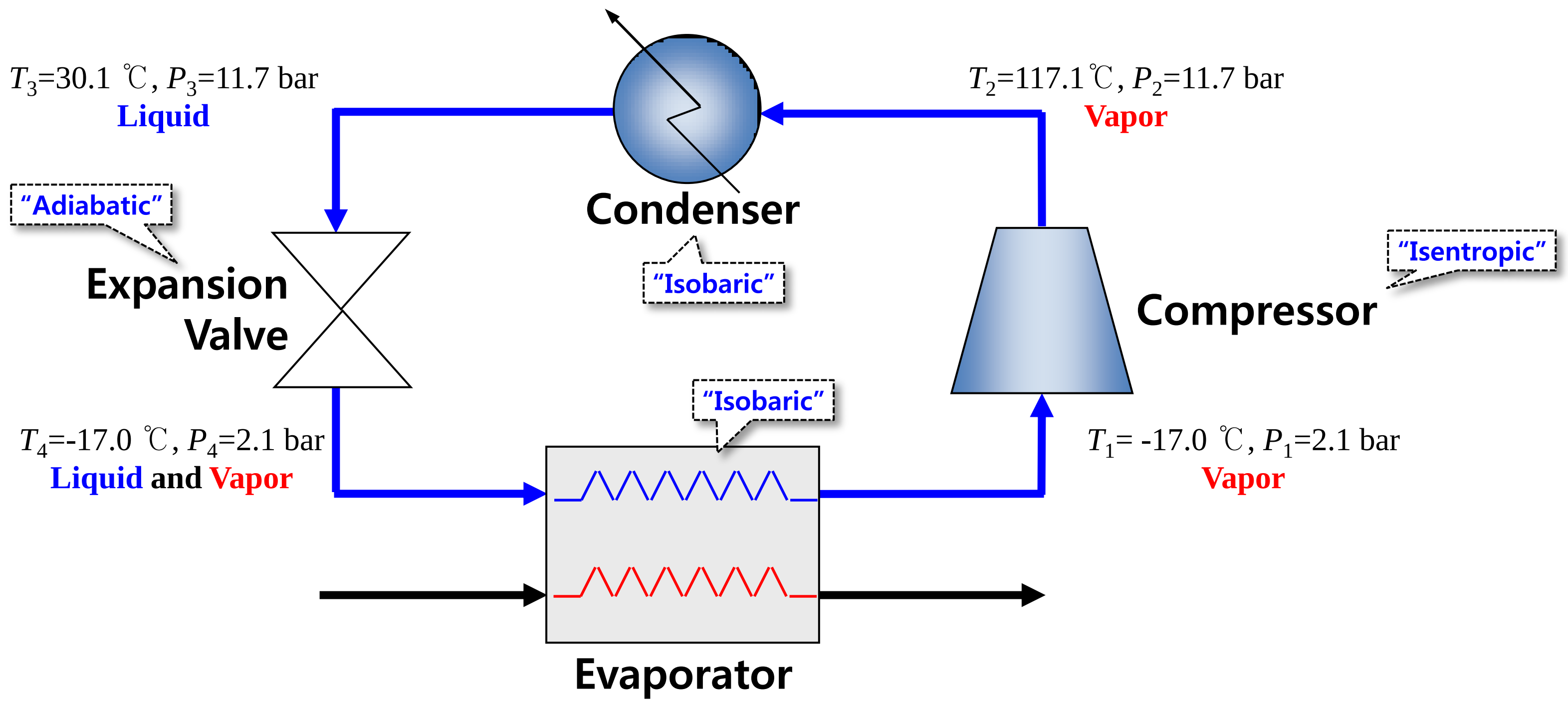
$$f=\dot{m}_{ref}\cdot w$$

9.3. Concept of Optimal Synthesis of Liquefaction Cycle

What is "Optimal Synthesis(Design)" of Liquefaction Cycle of a LNG FPSO?
- Synthesis: Combination of Equipment

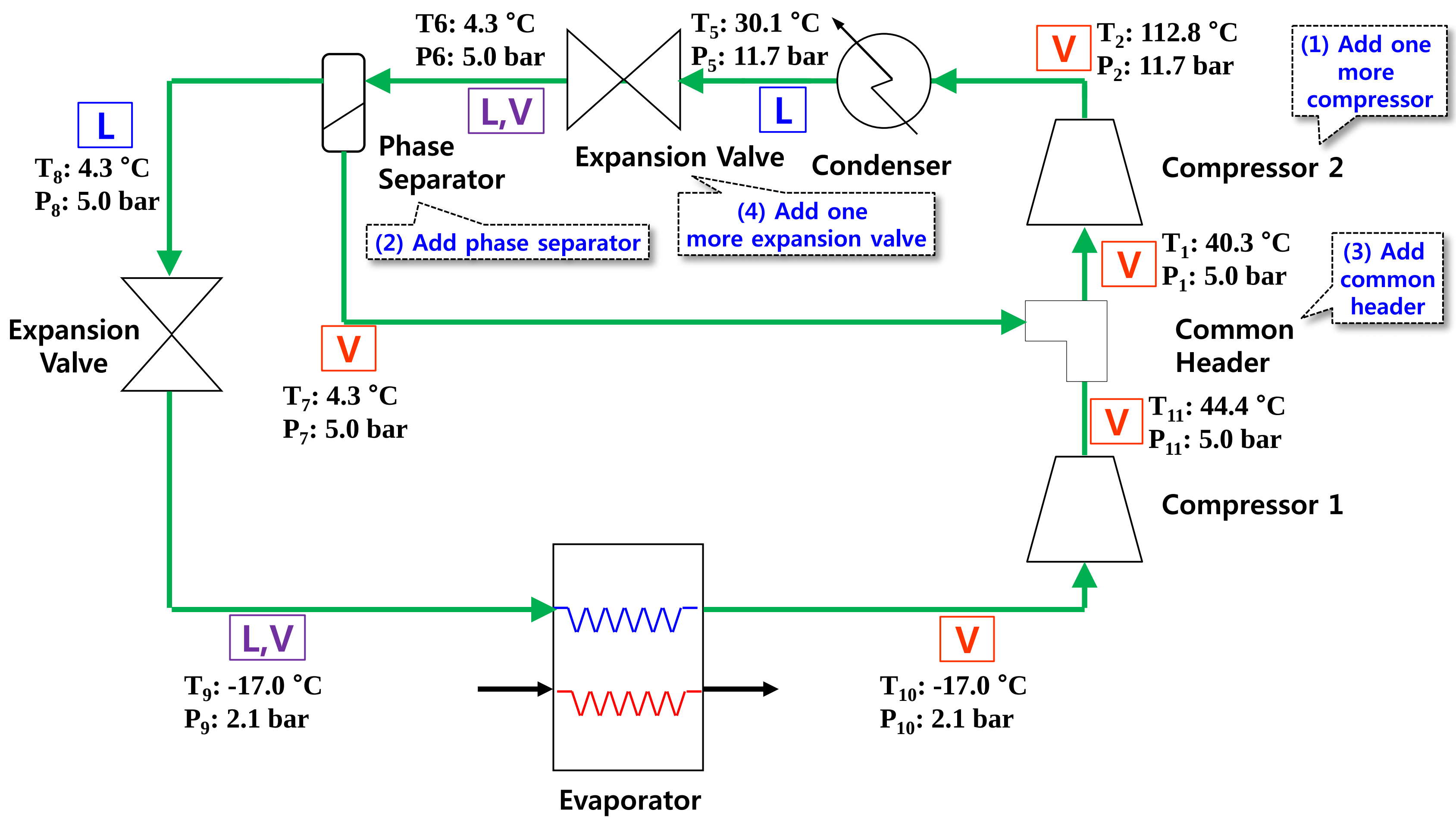
"Adiabatic process": There is no heat transfer between system and it surroundings , because there is no sufficient time to transfer much heat.
"Isentropic process: "Entropy" does not change. "Adiabatic process" and "Reversible"
"Isobaric process": There is no pressure drop

• An example of Simplified Liquefaction Cycle



1) Compressor brings the vapor refrigerant to a high pressure, which raises its temperature as well.
2) The hot vapor refrigerant passes through the condenser, an array of thin tubes that transfer heat from the refrigerant to the cooling medium. As it cools, the vapor refrigerant becomes a liquid under high pressure.
3) The liquid refrigerant goes into the expansion valve, from which it emerges at a lower pressure and temperature. While passing through the expansion valve, high-pressure liquid becomes low-pressure liquid and vapor.
4) In the evaporator, the cool liquid refrigerant completely evaporates by absorbing heat from the warm refrigerant. While passing through the evaporator, the temperature remains constant at the constant pressure during the phase-change process. The low-pressure liquid and vapor becomes low-pressure vapor. The refrigerant leaves the evaporator as saturated vapor and reenters the compressor.
5) In the end flash system, the pressure of LNG is expanded to the atmospheric pressure (1,01 bar) to be stored in the LNG tank.

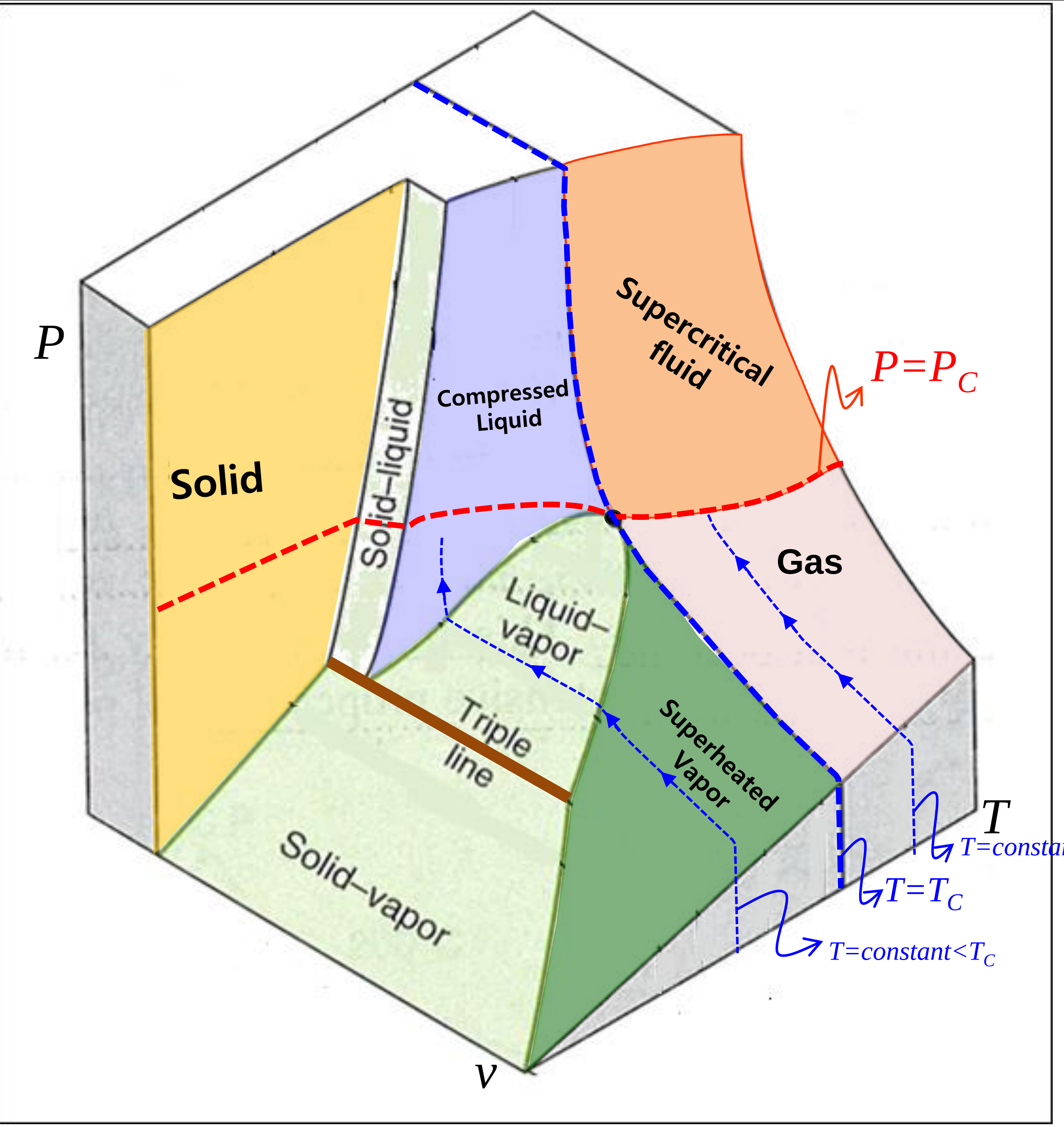
- **Another combination of equipment of a simplified liquefaction cycle**



- **Multistage Compression Refrigeration:**
  2) Phase separator separates a liquid-vapor mixture refrigerant into the vapor and liquid
  3) Common header mixes the saturated vapor from the phase separator and the superheated vapor from the compressor 1, and the cooled mixture enters the compressor 2.

[Thermodynamics] Pressure(P)-Specific Volume(v)-Temperature(T) Surface
- Pure Substance

Tc: critical temperature
Pc: critical pressure



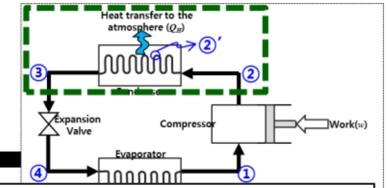
- **Critical point:** The point where the saturated liquid and saturated vapor lines meet.
- **Vapor vs. Gas**
  * **Vapor**: Vapor **can be condensed** either by compression at constant temperature or by cooling at constant pressure. The 'condense' means the change of the state of substance from vapor phase to liquid phase
  * Gas: The vapor phase of a substance is customarily called a gas when it is **above the critical temperature**. **Gas cannot be condensed by compression at constant temperature**, because the motional energies of the molecules are greater than the attractive forces that lead to the liquid state regardless of how much the substance is compressed to bring the molecules closer together[1].
- **Supercritical fluid:**
  - A single phase at and above the critical temperature and pressure
  - Like a gas, it still **expands to fill the confines[1] of its container**. And like liquids, supercritical fluids can behave as **solvents[2], dissolving a wide rage of substances**.
  - Using **supercritical fluid extraction**, the components of mixture, which is composed of the dissoluble substance[3] and non-dissoluble substance, can be separated.
  - For example, supercritical CO_2 is now used to extract sesame[4] oil from the sesame caffeine from coffee, and nicotine from tobacco.

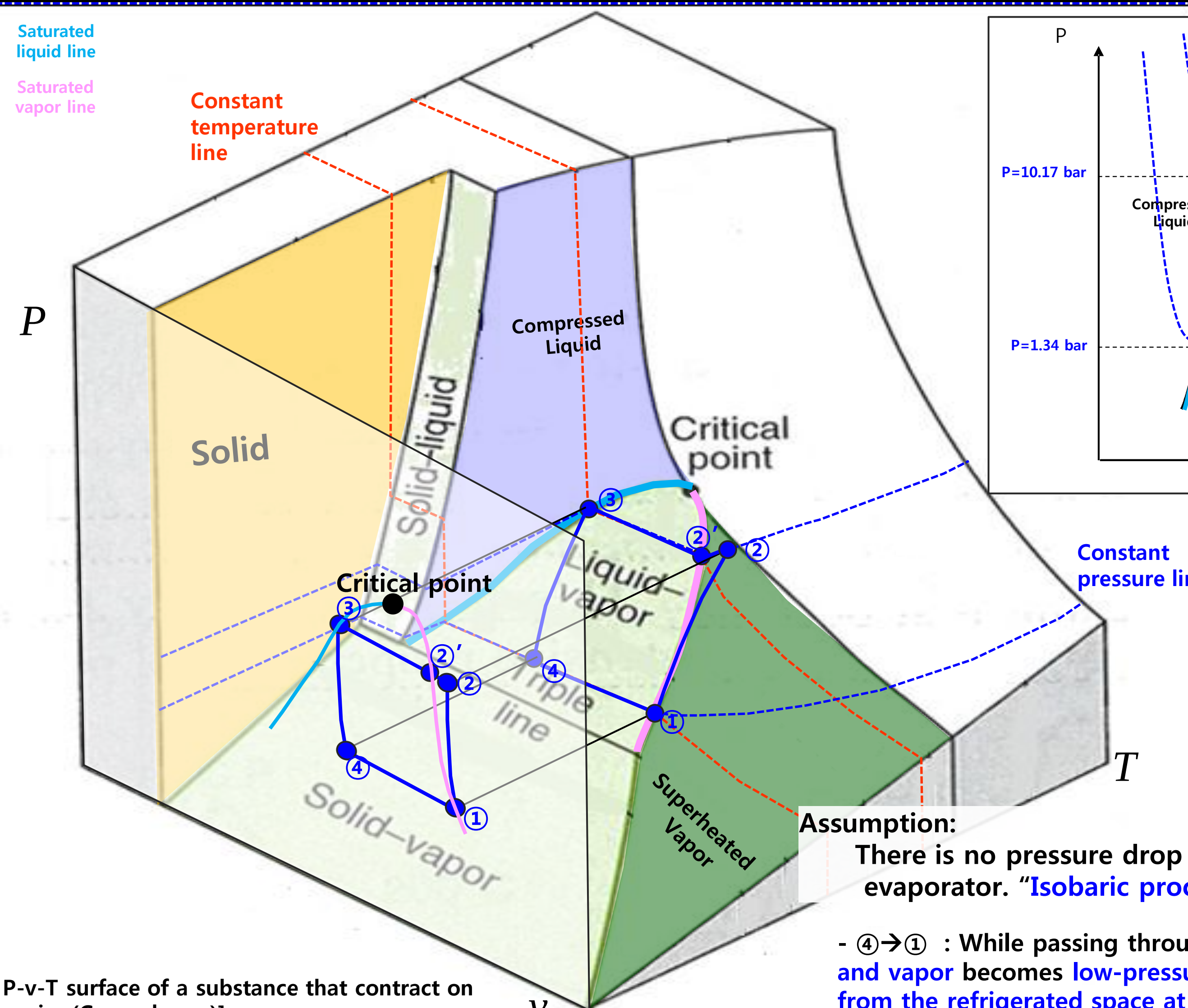
1) confine: place within the closed boundaries
2) solvent: a liquid that can dissolve other substances.
3) dissoluble substance: a substance which can be dissolved
4) sesame: 참깨

- **Triple line:** The line where all three phases(Vapor, Liquid, and Solid) of a pure substances coexist
  The triple line marks the lowest pressure at which a liquid phase of a substance can exist.

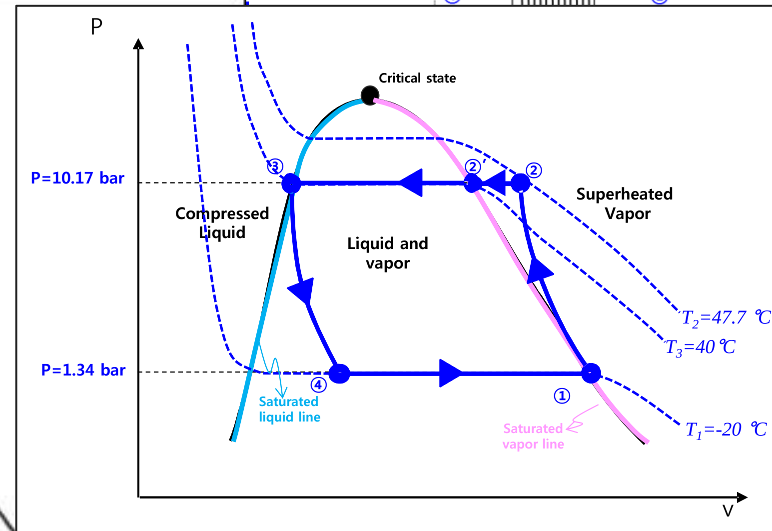
Reference:
Brown, T.L., LeMay, Jr,., H.E. and Bursten, B.E., Chemistry the central science, 10th edition, Prentice Hall, 2006, p.458

3. Thermodynamics in the Liquefaction Cycle
- Process of the Refrigerator – Evaporation
Pressure(P)-Specific Volume(v)-Temperature(T) Surface





[P-v-T surface of a substance that contract on freezing(General case)]



The evaporator has the same concept of the condenser.

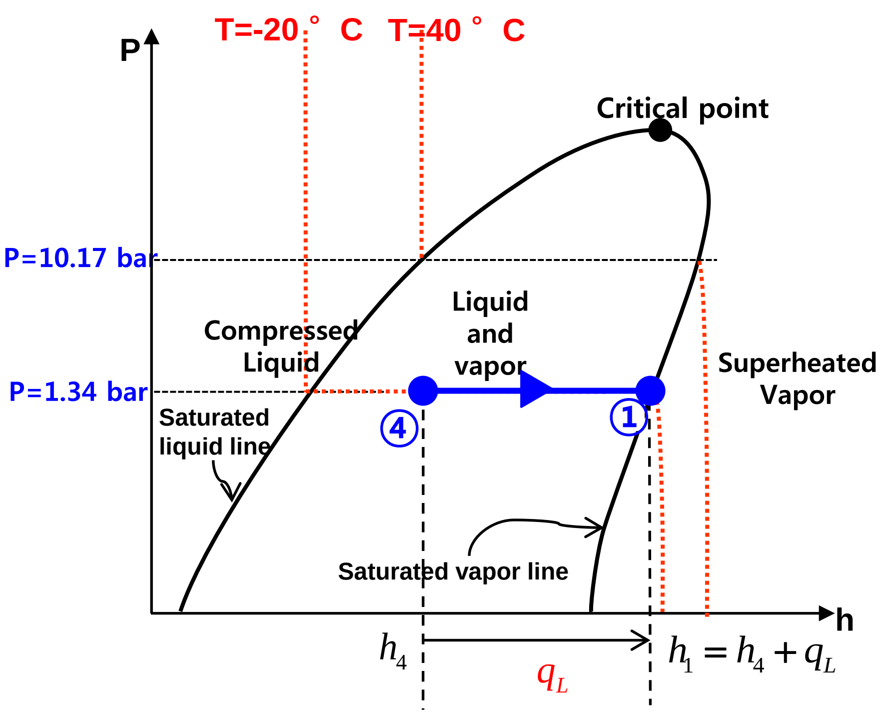
Assumption:
There is no pressure drop of the refrigerant through the evaporator. "Isobaric process"

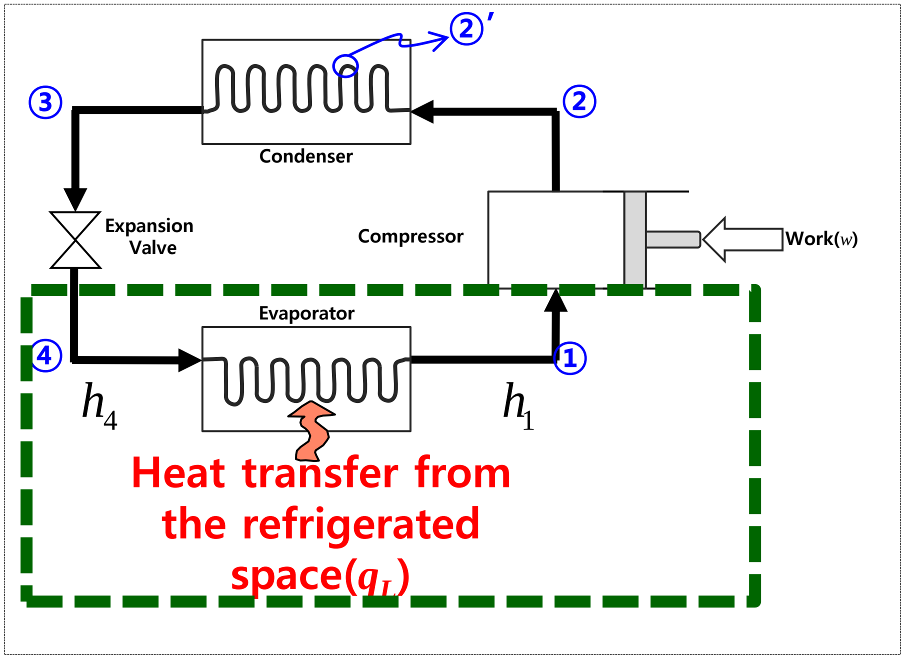
- ④→① : While passing through evaporator, low pressure liquid and vapor becomes low-pressure vapor by absorbing the heat from the refrigerated space at constant temperature in the two phase.

- Process of the Refrigerator – Evaporation

Pressure(P)-Specific Enthalpy(h) Diagram

The evaporator has the same concept of the condenser.





Assumption:

There is no pressure drop of the refrigerant through the evaporator. "Isobaric process"

According to the first law of thermodynamics

$$\begin{bmatrix}\text{Energy of the}\\ \text{refrigerant entering}\\ \text{the evaporator}\end{bmatrix} + \begin{bmatrix}\text{Energy transferred from the}\\ \text{refrigerated space as Heat}\end{bmatrix} = \begin{bmatrix}\text{Energy of the}\\ \text{refrigerant leaving}\\ \text{the evaporator}\end{bmatrix}$$

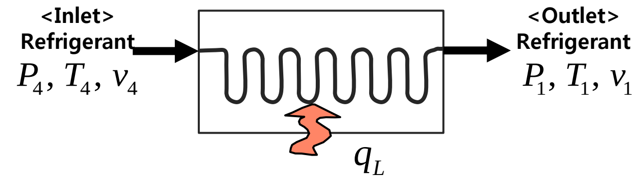
$$h_4(P_4, v_4, T_4) + q_L = h_1(P_1, v_1, T_1)$$

- Mathematical Model of the Refrigerator – Evaporator (1/2)

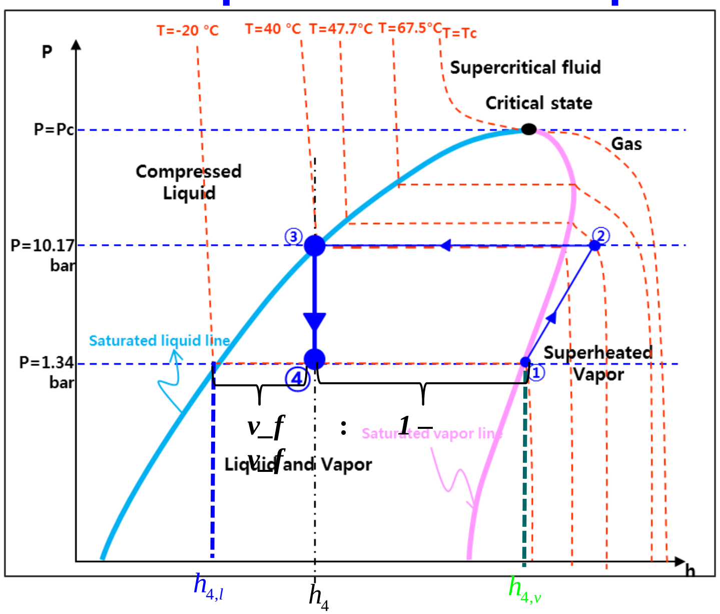
3. Equality constraints

2) The first law of the thermodynamics(Energy conservation)

$$M \cdot h_4\left(P_4, v_4, T_4\right) + \dot{Q}_L = M \cdot h_1\left(P_1, v_1, T_1\right)$$



The ratio of the mass of vapor to the total mass of the mixture of saturated liquid and saturated vapor is called " vapor fraction".



[Given]

$$\dot{Q}_L = 20[kW]$$

[Find]

Operating Conditions [20]:

P_i, T_i, v_i, T_s, v_s, $v_{4,l}$, $v_{4,v}$, v_f, w, M, q_L, q_H (i=1,2,3,4)

$$\text{Minimize}\left(\dot{m} \cdot w\right)$$

Refrigerant: Ammonia

P: pressure [bar]

T: temperature [K]

v: specific volume [m³/kg]

T_s: temperature of the refrigerant in the compressor at isentropic process [K]

v_s: specific volume of the refrigerant in the compressor at isentropic process [m³/kg]

v_f: vapor fraction

q_H: specific heat transfer from the refrigerant to the atmosphere [kJ/kg]

q_L: specific heat transfer from the refrigerated space to the refrigerant [kJ/kg]

M: mass flow rate of the refrigerant [kg/s]

$\dot{Q}_L$: heat transfer from the refrigerated space to the refrigerant [kW]

: Design variables

$h_{4,l}$: enthalpy at the saturated liquid

$h_{4,v}$: enthalpy at the saturated vapor

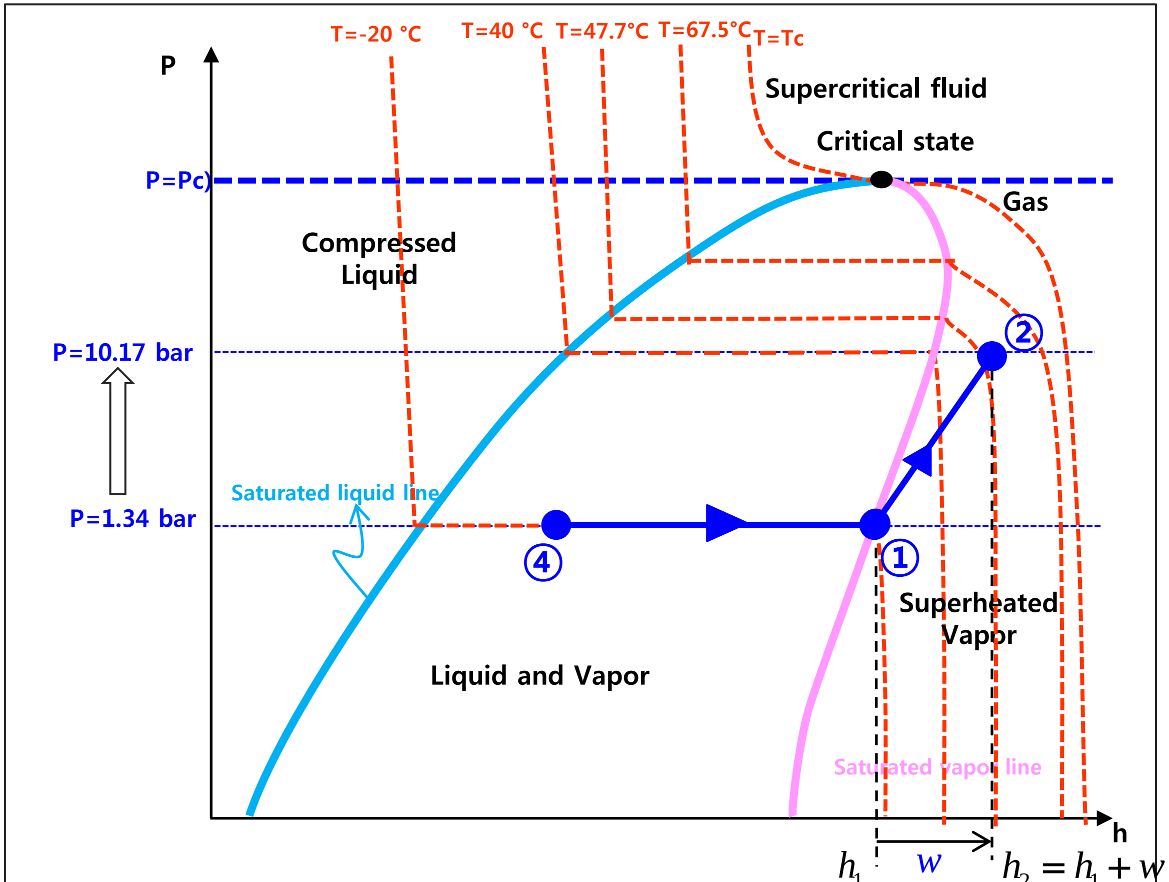
$$M \cdot \left[\left(1 - v\_f\right) \cdot h_{4,l}\left(P_4, v_{4,l}, T_4\right) + v\_f \cdot h_{4,v}\left(P_4, v_{4,v}, T_4\right)\right] + \underline{\dot{Q}_L} = M \cdot h_1\left(P_1, v_1, T_1\right)$$

Rate of heat transfer from the refrigerated space to the refrigerant(Given)

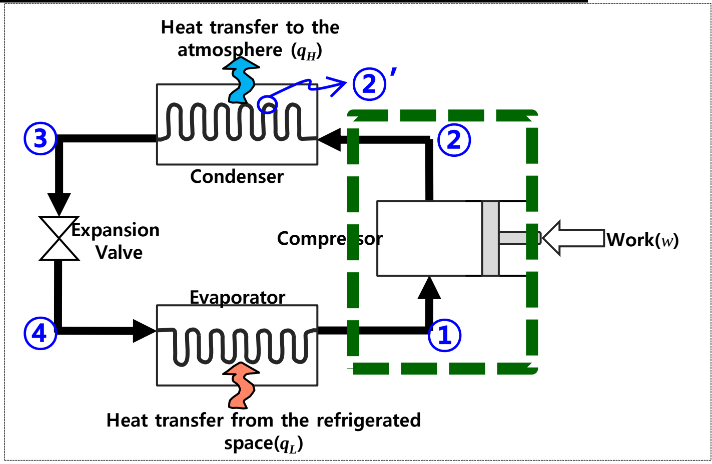
- Process of the Refrigerator – Compression

1. Evaporator → Compressor → Condenser → Expansion valve 순으로 자료 작성

Pressure(P)-Specific Enthalpy(h) Diagram (1/5)



Example) Mollier Diagram of the Ammonia



The compressor is a device in which work is done on the substance flowing through it in order to increase the pressure. In compressor(①→②), low-pressure vapor becomes high-pressure vapor and its temperature is raised as well.

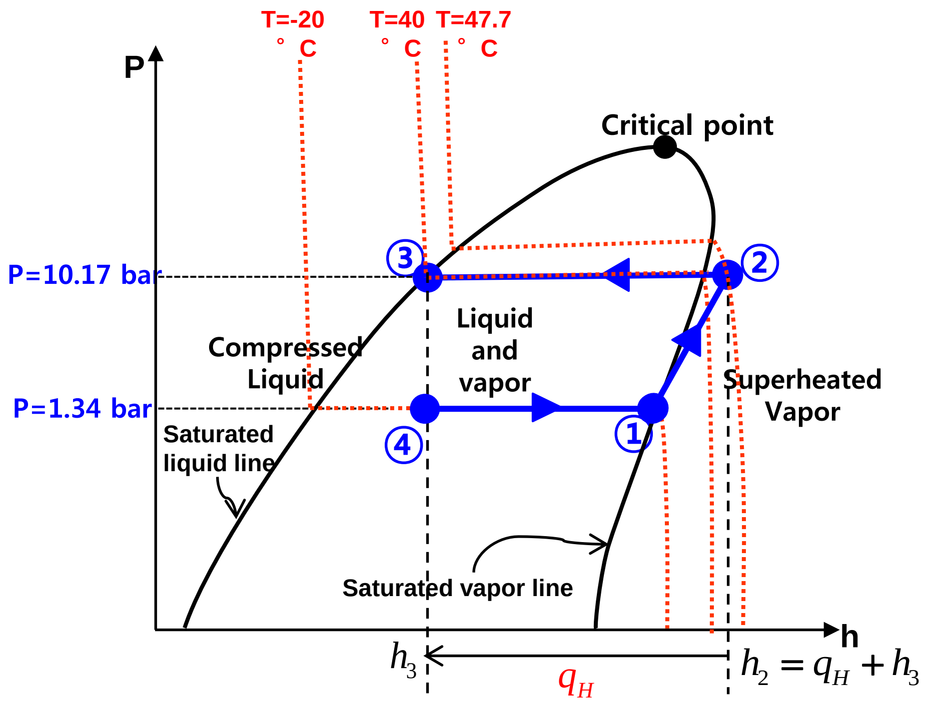
- Natural Phenomena: To compress the refrigerant, the work(w) is provided to the refrigerant. And the energy of the refrigerant is increased by the work(w).
- According to the first law of thermodynamics(The total quantity of energy is constant)

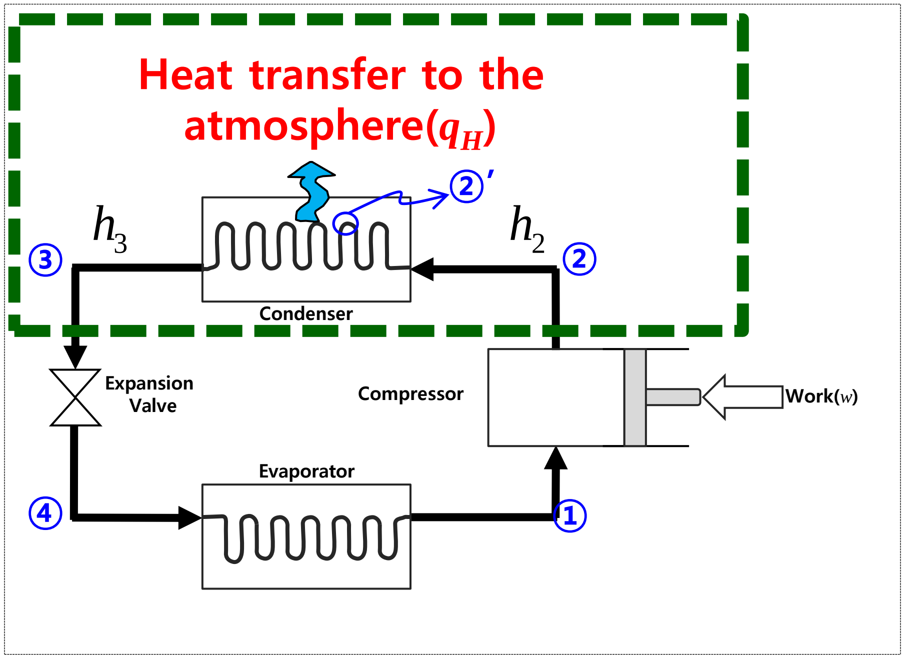
$$\begin{bmatrix}\text{Total energy of the refrigerant}\\ \text{entering the compressor}\end{bmatrix} + \begin{bmatrix}\text{Total entering energy}\\ \text{as Work in the compressor}\end{bmatrix} = \begin{bmatrix}\text{Total energy of the refrigerant}\\ \text{leaving the compressor}\end{bmatrix}$$

$$h_1 + w = h_2$$

→The enthalpy of the refrigerant is increased by the work(w).

Pressure(P)-Specific Enthalpy(h) Diagram





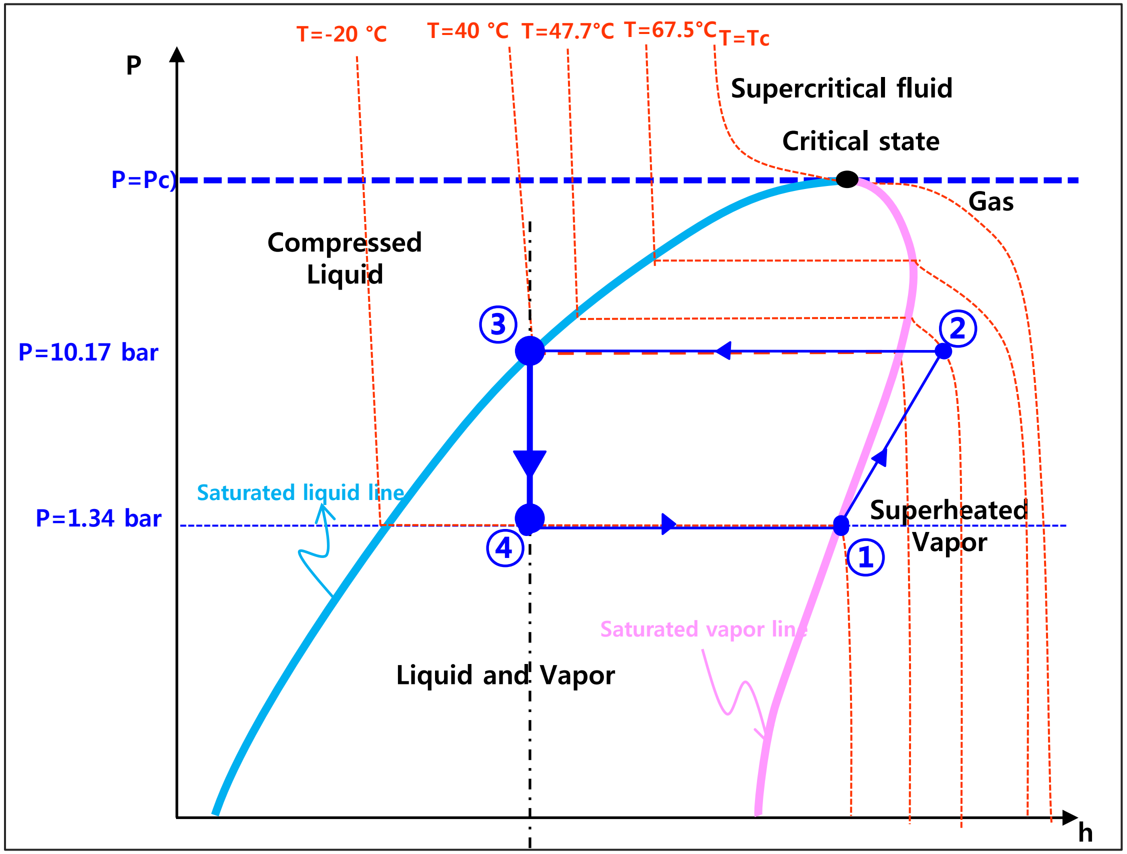
Assumption:
There is no pressure drop of the refrigerant through the condenser. "Isobaric process"

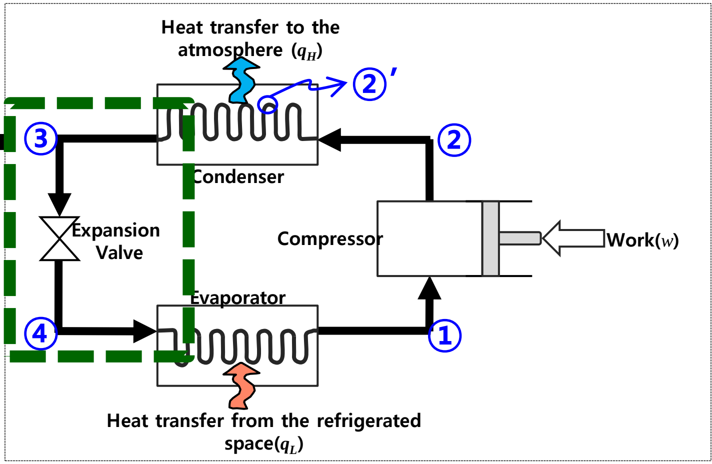
According to the first law of thermodynamics

$$\begin{bmatrix}\text{Energy of the} \\ \text{refrigerant entering} \\ \text{the condenser}\end{bmatrix} = \begin{bmatrix}\text{Energy transferred} \\ \text{to the atmosphere as Heat}\end{bmatrix} + \begin{bmatrix}\text{Energy of the} \\ \text{refrigerant leaving} \\ \text{the condenser}\end{bmatrix}$$

$$h_2(P_2, v_2, T_2) = q_H + h_3(P_3, v_3, T_3)$$

- Process of the Refrigerator – Expansion
Pressure(P)-Specific Enthalpy(h) Diagram (1/3)





1. Natural Phenomena:

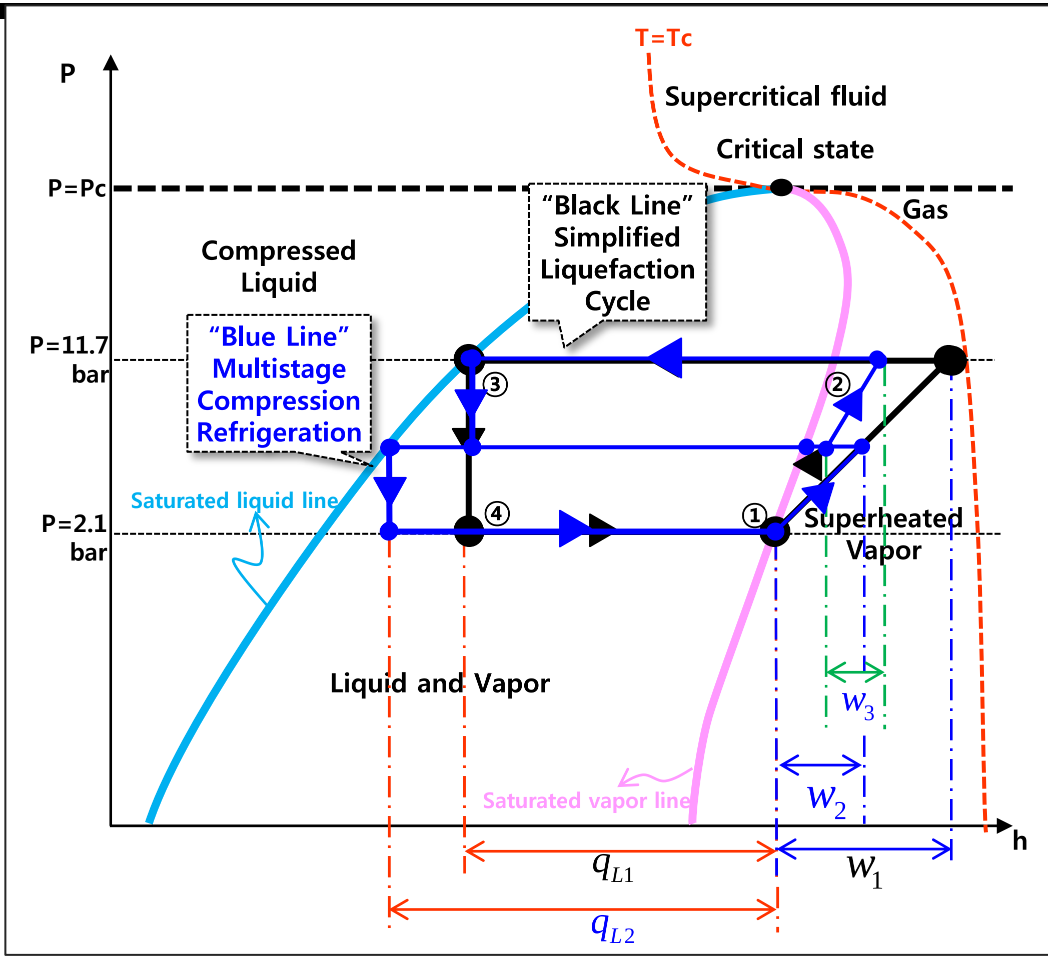
Expansion valves are any kind of flow-restricting devices that cause a significant pressure drop in the fluid. Therefore, there is no work done to decrease the pressure.

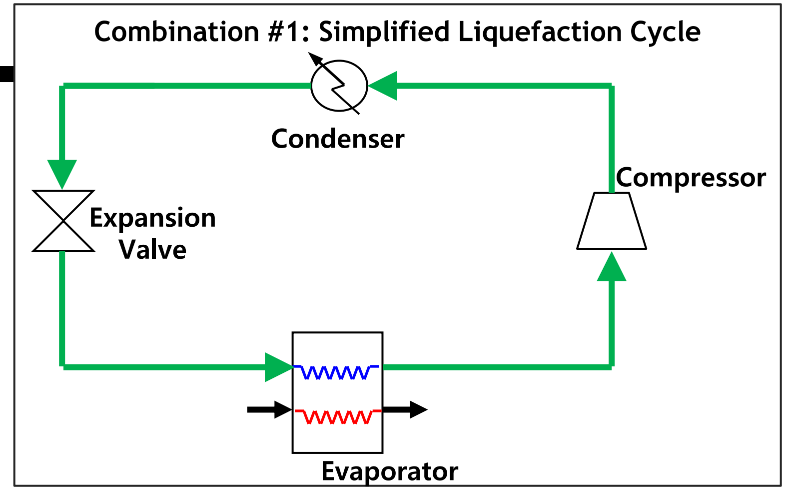
2. Assumption:

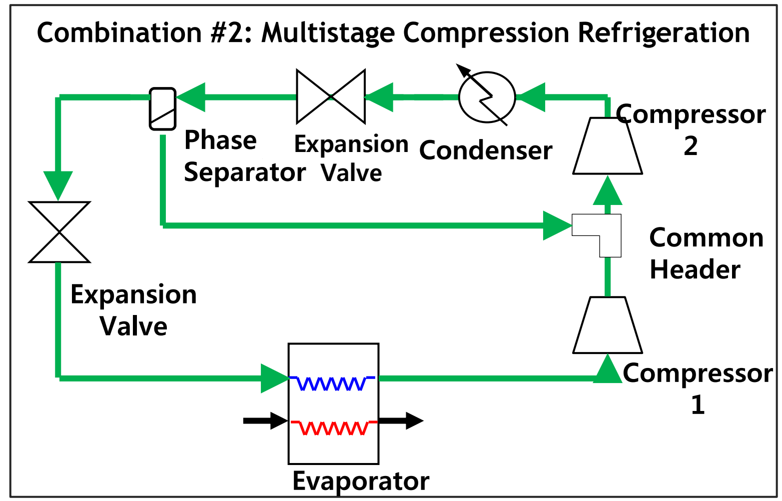
There is not sufficient time to transfer much heat from the atmosphere to the refrigerant in the expansion valve, "Adiabatic process".

Therefore, the energy values at the inlet and outlet "enthalpy" of the expansion valve are the same (③→④).

What is "Optimal Synthesis(Design)" of a Liquefaction Cycle of a LNG FPSO?
- Optimal Synthesis







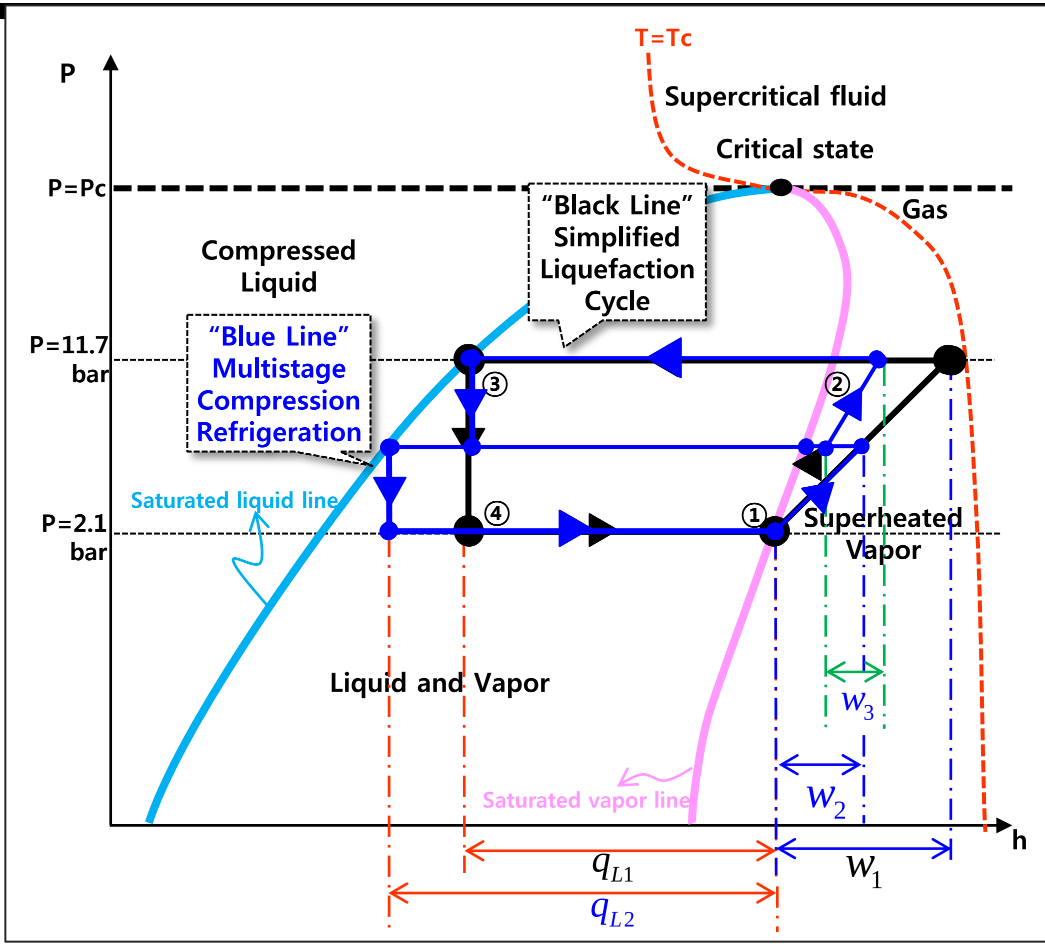
What is the efficiency of liquefaction cycle? (CP: coefficient of performance)

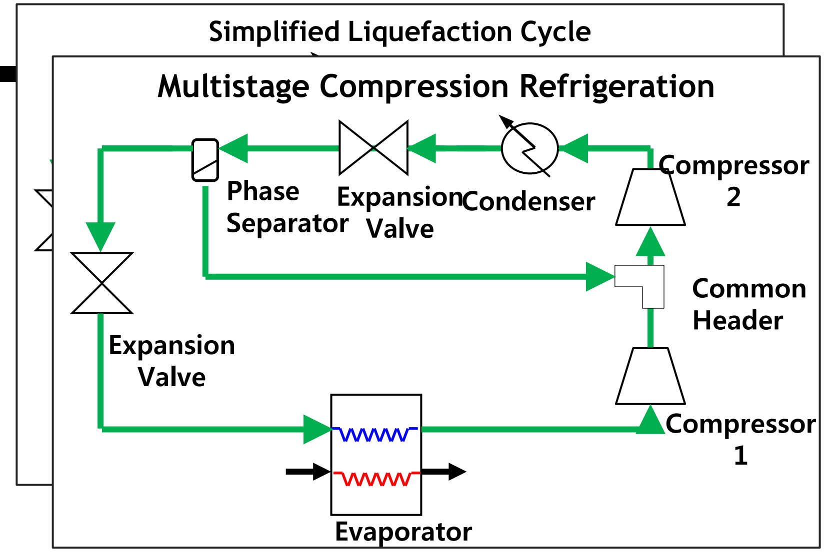
$$CP = \frac{\text{What we want}}{\text{What we pay for}} = \frac{|q_L|}{|w|}$$

1) CP of Simplified Liquefaction Cycle: $CP_1 = \frac{|q_{L1}|}{|w_1|}$

2) CP of Multistage Compression Refrigeration: $CP_2 = \frac{|q_{L2}|}{|w_2 + w_3|}$

- To increase the efficiency of a liquefaction cycle, when q_L is given, we have to determine the operating conditions such as pressure, temperature, specific volume by minimizing the specific work[J/g] provided to the compressor, "objective function".





Various combination of equipments of liquefaction cycle

- **Given:** The quantity of the specific heat transfer from the refrigerated space to the refrigerant(q_L) in the evaporator.

Constraint

1. Optimal synthesis of liquefaction cycles

Design variables

- **Find:** Various combination of equipments that makes up the liquefaction cycle minimizing the power provided to the compressors. → Objective function

2. To calculate the minimized power, we have to determine the optimal operating condition

Design variables

- **Find:** The operating conditions such as the pressure, temperature and specific volume minimizing the power provided to the compressors. → Objective function

Introduction to the Liquefaction Cycle

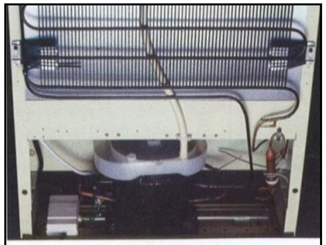
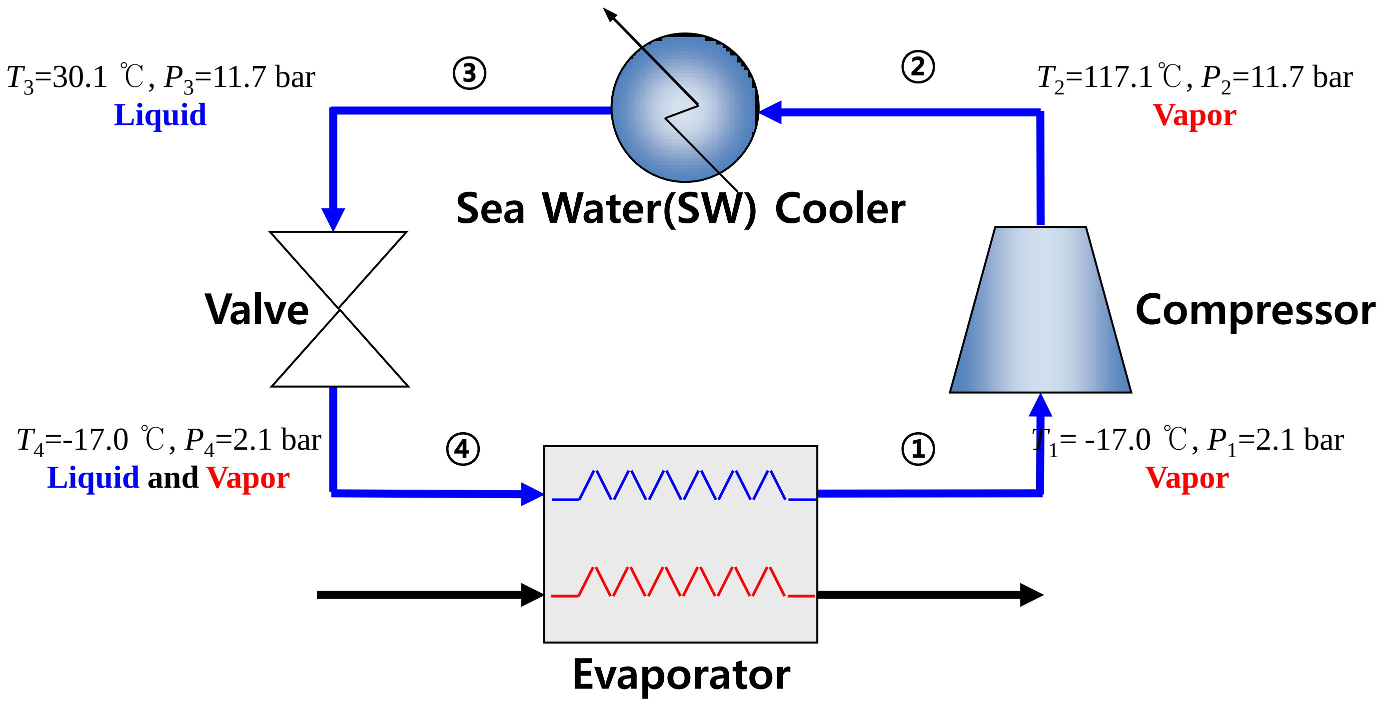


FIGURE 21.19 Condenser coils of a refrigerator.

- **Goal of the LNG Liquefaction Cycle**

  To liquefy NG to LNG for decreasing the volume of the NG

- **An example of Simplified LNG Liquefaction Cycle**

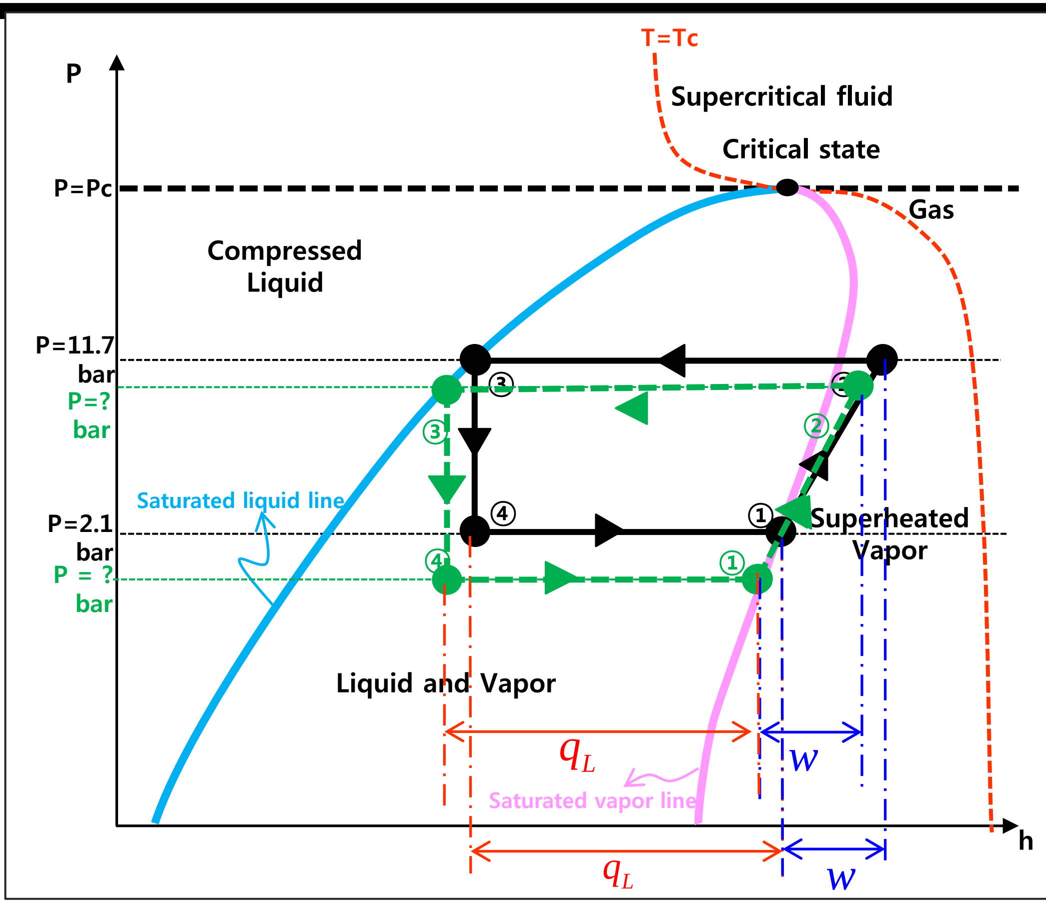


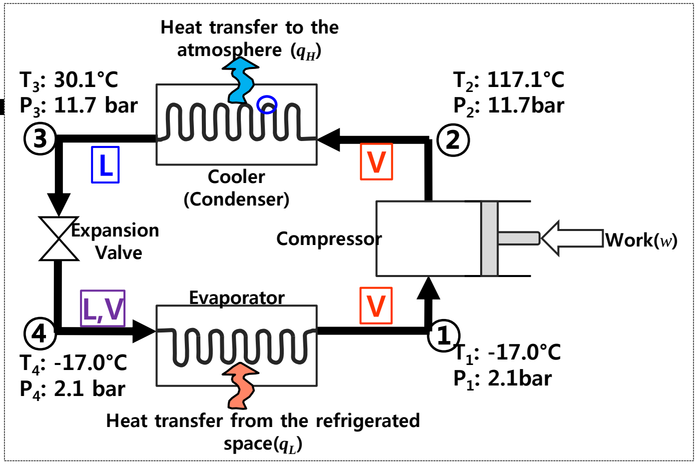
Equipment used in the cycle

1) Compressor: brings the vapor refrigerant to a high pressure, which raises its temperature as well
2) Sea Water Cooler(a kind of condenser): transfer heat from the hot vapor refrigerant to the sea water
3) Valve: decreases the pressure of the liquid refrigerant, which decreases its temperature as well
4) Heat Exchanger(a kind of evaporator): absorbs heat from the natural gas to cool down the NG, while the refrigerant is vaporized

(1) The temperature and pressure of the natural gas and liquefied natural gas are the values of the general case.
(2) In the end flash system, the pressure of LNG expanded to the atmospheric pressure (1,01 bar) to be stored in the LNG tank.

- Optimal Design of Liquefaction Cycles for LNG FPSO





- q_L: specific heat transfer from the refrigerated space to the refrigerant(Given)
- w: work provided to the compressor(Minimizing)

- **Given:** The quantity of the specific heat transfer from the refrigerated space to the refrigerant(q_L) in the evaporator.

Constraint

Design variables

1. Determination of the optimal operating condition

- **Find:** The operating conditions such as the pressure, temperature and specific volume minimizing the work or power provided to the compressor. Objective function

2. Optimal synthesis on liquefaction cycles

Design variables

- **Find:** The number of the combination of equipment that make up the liquefaction cycle minimizing the work provided to the compressor. Objective function

• Physical Constraint based on Thermodynamics #1

Energy conservation

Calculation of the specific enthalpy(h)

Many tables of thermodynamics properties does not give values for internal energy. To allow calculation of enthalpy from the pressure, specific volume and temperature, the following equation is derived by using the definition($h=u+Pv$), equation of state and experiment.

$$h = h^{IG} + h^{R} \quad [J/g]$$

h^{IG}: Ideal gas value of the specific enthalpy
h^{R}: Residual specific enthalpy(correction of the ideal gas state values to the real gas values)

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

Calculation of Specific Enthalpy(h)

$$h = u + P \cdot v$$

- **Calculation of the specific enthalpy(h) for a pure substance**

Many tables of thermodynamics properties does not give values for internal energy. To allow calculation of enthalpy from the pressure, specific volume and temperature, the following equation is derived by using the definition($h=u+Pv$), equation of state and experiment.

P[Pa]
v[m^3]
T[K]

$$h = h^{IG} + h^{R}$$

h^{IG}: **Ideal gas value of the specific enthalpy**
h^{R}: **Residual specific enthalpy**(correction of the ideal gas state values to the real gas values)

$$h^{IG} = h^{IG}(T) = a + b\cdot T + c\cdot T^2 + d\cdot T^3 + e\cdot T^4 + f\cdot T^5$$

where

a, b, c, d, e **and** ***f*****: constants characteristic of the particular substance**
h^{ig}[J/g], T: **temperature**[K]

Example) Ammonia:
a = **-1.8514**, b = **1.9937**, c = **-5.3266x10^{-4}**
d = **2.0615x10^{-6}**, e = **-1.3386x10^{-9}**,
f = **3.0533x10^{-13}**

$$h^{R} = h^{R}(P,v,T) = RT(Z-1) + \frac{T\left(\frac{da(T)}{dT}\right) - a(T)}{b}\frac{1}{(\sigma-\varepsilon)}\ln\left[\frac{Z+\sigma\cdot\beta}{Z+\varepsilon\cdot\beta}\right]$$

R: **Gas constant** (=8.314 J/(mol*K)
= 8.314m^{3}Pa(mol*K))

* 1 bar = 100kPa

where

$$a(T) = \psi\frac{\alpha(T_r,\omega)R^2T_c^2}{P_c} \qquad T_r = \frac{T}{T_c} \qquad Z = \frac{v}{v-b} - \frac{v}{R\cdot T}\cdot\frac{a(T)}{(v+\varepsilon\cdot b)(v+\sigma\cdot b)} \qquad \beta = \frac{bP}{RT} \qquad b = \Omega\frac{RT_c}{P_c}$$

1) The values of parameters $a, \sigma, \varepsilon, \Omega, \Psi$ **are depending on the type of the cubic equation of state. For example, the value of the parameters for the Soave-Redlich-Kwong(SRK) equation of state are given in the following table.**

$\alpha(T_r, \omega)$	σ	ε	Ω	ψ
$\alpha(T_r;\omega) = \left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-(T_r)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

2) The values of ω, **critical pressure**(Pc), **and temperature**(Tc) **are depending on the substance.**

Example) Ammonia:
ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

3) Since the unit of h^{IG} **is** J/g **and** h^{R} **is** J/mol, h^{R} **is devided by molar mass(**M, g/mol**).**

$$h^{R}[J/g] = \frac{h^{R}[J/mol]}{M[g/mol]}$$

Example) Ammonia
$M_{Ammonia}$ = **17.031** (g/mol)

4) Central difference approximation

$$\left(\frac{da}{dT}\right) = \frac{a(T+e) - a(T-e)}{2\cdot e}, \ (e = 10^{-6})$$

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

• Physical Constraint based on Thermodynamics #2

Equation of state

$$P_1 v_1 = RT_1$$

[Equation of state for an ideal gas]

The equation of state for the liquids and vapors is constructed by considering experimental results based on the equation of state for an ideal gas.

Example) Soave, Redlich, Kwong(SRK) equation

$$\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$$

$$v = \frac{RT}{P} + b - \frac{a(T)}{P} \frac{v-b}{(v-\varepsilon b)(v-\sigma b)}$$

s: specific entropy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ *for SRK equation*

R: gas constant (=8.314 Jmol⁻¹K⁻¹)

P_c: critical pressure of the refrigerant

T_c: critical temperature of the refrigerant

$$b = \Omega \frac{RT_C}{P_C}$$

$\Omega = 0.08664$ *for SRK equation*

$\varepsilon = 0$ *for SRK equation*

$\sigma = 1$ *for SRK equation*

Example) Ammonia:

$\omega = 0.253$, $Pc=112.80$ (bar), $Tc=405.7$ (K)

4) SRK EoS -> 5) PR EoS
분자 간의 인력에 의한 압력의 감소량에서 부피 의존성인
V(V+b)대신 V(V+b)+b(V-b)를 사용하면 실제 값에 보다
접근함을 실험적으로 증명.

*EoS: Equation of State

To improve the equation of state for the liquids and vapors, the equation of state for an ideal gas is modified by using the experiment and experience.

(1) Ideal gas EoS* (1802)

(1) $$Pv = RT$$

(2) van der Waals EoS(1873)

(2) $$\left(P + \frac{a}{v^2}\right)(v-b) = RT$$

$$a = \frac{24}{64}\cdot\frac{R^2\cdot T_c^{\ 2}}{P_c},\ b = \frac{1}{8}\cdot\frac{R\cdot T_c}{P_c}$$

(3) Redlich-Kwong EoS(1949)

(3) $$\left(P + \frac{a(T)}{v\cdot(v+b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.42748\cdot(T/T_c)^{-1/2}\cdot R^2\cdot T_c^{\ 2}}{P_c},\ b = \frac{0.08664\cdot R\cdot T_c}{P_c}$$

(4) Soave-Redlich-Kwong EoS(1972)

(4) $$\left(P + \frac{a(T)}{v\cdot(v+b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.42748\cdot\alpha_{SRK}(T/T_c;\omega)\cdot R^2\cdot T_c^{\ 2}}{P_c}$$

$$\alpha_{SRK}(T/T_c;\omega) =$$

$$\left[1 + (0.480 + 1.574\cdot\omega - 0.176\cdot\omega^2)\cdot(1-(T/T_c)^{1/2})\right]^2$$

$$b = \frac{0.08664\cdot R\cdot T_c}{P_c}$$

(5) Peng-Robinson EoS(1976)

(5) $$\left(P + \frac{a(T)}{(v+(1-\sqrt{2})\cdot b)\cdot(v+(1+\sqrt{2})\cdot b)}\right)(v-b) = RT$$

$$a(T) = \frac{0.45724\cdot\alpha_{PR}(T/T_c;\omega)\cdot R^2\cdot T_c^{\ 2}}{P_c}$$

$$\alpha_{PR}(T/T_c;\omega) =$$

$$\left[1 + (0.37464 + 1.54226\cdot\omega - 0.26992\cdot\omega^2)\cdot(1-(T/T_c)^{1/2})\right]^2$$

$$b = \frac{0.07780\cdot R\cdot T_c}{P_c}$$

T: temperature[K]
P: pressure{Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure{Pa]
ω: acentric factor
R: gas constant(=8.314[m3Pa/(mol·K)]

(4) Soave-Redlich-Kwong EoS → (5) Peng-Robinson EoS

① Modify the pressure reduction due to the attractive forces
: The pressure reduction depending on the molar volume(v) is modified by $(v+(1-\sqrt{2})\cdot b)\cdot(v+(1+\sqrt{2})\cdot b)$ instead of $v(v+b)$.

General Form of the Cubic Equations of State for Liquids and Vapors

The van der Waals(vdW), Redlich-Kwong(RK), Soave-Redlich-Kwong(SRK) and Peng-Robinson(PR) equation of state are represented as the following cubic equations form.

$$\left(P+\frac{a(T)}{(v+\varepsilon b)(v+\sigma b)}\right)(v-b)=RT \quad \Rightarrow \quad P=\frac{RT}{v-b}-\frac{a(T)}{(v+\varepsilon b)(v+\sigma b)}$$

$$a(T)=\Psi\frac{\alpha(T_r)R^2T_c^{\,2}}{P_c}, \quad b=\Omega\frac{RT_c}{P_c}$$

EoS	$\alpha(T_r)$	σ	ε	Ω	Ψ	Z_c
vdW(1873)	1	0	0	1/8	24/64	3/8
RK(1949)	$T_r^{-0.5}$	1	0	0.08664	0.42748	1/3
SRK(1972)	$\alpha_{SPK}(T_r, \omega)$	1	0	0.08664	0.42748	1/3
PR(1976)	$\alpha_{SR}(T_r, \omega)$	$1+\sqrt{2}$	$1-\sqrt{2}$	0.07780	0.45724	0.30740
$\alpha_{SPK}(T_r;\omega)=\left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-T_r^{\,0.5}\right)\right]^2$						
$\alpha_{PR}(T_r;\omega)=\left[1+(0.37464+1.54226\omega-0.26992\omega^2)\left(1-T_r^{\,0.5}\right)\right]^2$						

$$T_r=T/T_c$$

T: temperature[K]
P: pressure[Pa]
v: molar volume[m³/mol]
T_C: critical temperature[K]
P_C: critical pressure[Pa]
ω: acentric factor

R: gas constant(=8.314[m3Pa/(mol·K)]

• Physical Constraint based on Thermodynamics #3

Criteria for quality of the energy

Calculation of the specific entropy(s)

To allow calculation of entropy from the pressure, specific volume and temperature, the following equation is derived by using the definition($ds=dq/T$), equation of state and experiment.

$$s = s^{IG} + s^{R} \; [J/(K \cdot g)]$$

s^{ig}: entropy for the ideal gas
s^{R}: Residual entropy(correction of the ideal gas values for the real gas)

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

Calculation of the specific entropy(s) for a pure substance

To allow calculation of entropy from the pressure, specific volume and temperature, the following equation is derived by using the definition($ds=dq/T$), equation of state and experiment.

P[Pa]
v[m^3]
T[K]

$$s = s^{IG} + s^{R}$$

s^{ig}: Ideal gas value of the entropy
s^{R}: Residual entropy(correction of the ideal gas state values to the real gas values)

$$S^{IG} = g + b\cdot\ln(T) + 2\cdot c\cdot T + \frac{3}{2}\cdot d\cdot T^2 + \frac{4}{3}\cdot e\cdot T^3 + \frac{5}{4}\cdot f\cdot T^4$$

where

a, b, c, d, e and f: coefficients of the ideal gas Enthalpy equation
s^{ig}[J/(g·K)], T: temperature[K]
g : Entropy coefficient (i.e. the Entropy of the ideal gas at T=0 K) = 1.00

Example) Ammonia:
a = **-1.8514**, b = **1.9937**, c = **-5.3266x10^{-4}**
d = **2.0615x10^{-6}**, e = **-1.3386x10^{-9}**,
f = **3.0533x10^{-13}**

$$s^R = s^R(P,v,T) = R\ln(Z-\beta) - R\ln\left(\frac{P}{P_0}\right) + \frac{\left(\frac{\partial a}{\partial T}\right)}{b}\frac{1}{(\sigma-\varepsilon)}\ln\left[\frac{Z+\sigma\cdot\beta}{Z+\varepsilon\cdot\beta}\right]$$

P_c: critical pressure of the substance
T_c: critical temperature of the substance
Z: compressible factor

where

$$a(T) = \psi\frac{\alpha(T_r,\omega)R^2T_c^2}{P_c} \qquad T_r = \frac{T}{T_c} \qquad Z = \frac{v}{v-b} - \frac{v}{R\cdot T}\cdot\frac{a(T)}{(v+\varepsilon\cdot b)(v+\sigma\cdot b)} \qquad \beta = \frac{bP}{RT} \qquad b = \Omega\frac{RT_c}{P_c}$$

1) The values of parameters $a, \sigma, \varepsilon, \Omega, \Psi$ are depending on the type of the cubic equation of state. For example, the value of the parameters for the Soave-Redlich-Kwong(SRK) equation of state are given in the following table.

$\alpha(T_r, \omega)$	σ	ε	Ω	ψ
$\alpha(T_r;\omega) = \left[1+(0.480+1.574\omega-0.176\omega^2)\left(1-(T_r)^{0.5}\right)\right]^2$	1	0	0.08664	0.42748

2) The values of ω, critical pressure(Pc), and temperature(Tc) are depending on the substance.

Example) Ammonia:
ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

* 1 bar = 100kPa

3) Since the unit of s^{IG} is J/(g·K) and h^R is J/(mol·K), h^R is devided by molar mass(M, g/mol).

$$s^R[\mathrm{J/(g\cdot K)}] = \frac{s^R[\mathrm{J/(mol\cdot K)}]}{M[g/mol]}$$

Example) Ammonia
$M_{Ammonia}$ = 17.031 (g/mol)

4) Central difference approximation

$$\left(\frac{da}{dT}\right) = \frac{a(T+e)-a(T-e)}{2\cdot e}, \left(e=10^{-6}\right)$$

- **Physical Constraint based on Thermodynamics #4**

Physical assumptions for the liquefaction process

“Isobaric process”

- There is no pressure drop

“Adiabatic process”

- There is no heat transfer between system and it surroundings , because there is no sufficient time to transfer much heat.

“Isentropic process”

- “Entropy“ does not change
- “Adiabatic process” and “Reversible”

• Physical Constraint based on Thermodynamics #1

Energy conservation

Calculation of the specific enthalpy(h)

Many tables of thermodynamics properties does not give values for internal energy. To allow calculation of enthalpy from the pressure, specific volume and temperature, the following equation is derived by using the definition($h=u+Pv$), equation of state and experiment.

$$h = h^{IG} + h^{R} \quad [J/g]$$

h^{IG}: Ideal gas value of the specific enthalpy
h^{R}: Residual specific enthalpy(correction of the ideal gas state values to the real gas values)

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

- **Physical Constraint based on Thermodynamics #2**

Equation of state

$$P_1 v_1 = RT_1$$

[Equation of state for an ideal gas]

the equation of state for the liquids and vapors is constructed considering experimental modification based on the equation of state for an ideal gas

Example) Soave, Redlich, Kwong(SRK) equation

$$\left(P + \frac{a(T)}{v \cdot (v+b)}\right)(v-b) = RT$$

s: specific entropy

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ *for SRK equation*

R: gas constant (=8.314 Jmol⁻¹K⁻¹)

P_c: critical pressure of the refrigerant

T_c: critical temperature of the refrigerant

$$b = \Omega \frac{RT_C}{P_C}$$

$\Omega = 0.08664$ *for SRK equation*

$\varepsilon = 0$ *for SRK equation*

$\sigma = 1$ *for SRK equation*

Example) Ammonia:

ω = 0.253, Pc=112.80 (bar), Tc=405.7 (K)

• Physical Constraint based on Thermodynamics #3

Criteria for quality of the energy

Calculation of the specific entropy(s)

To allow calculation of entropy from the pressure, specific volume and temperature, the following equation is derived by using the definition($ds=dq/T$), equation of state and experiment.

$$S = S^{IG} + S^{R} \; [J/(K \cdot g)]$$

s^{ig}: entropy for the ideal gas
s^{R}: Residual entropy(correction of the ideal gas values for the real gas)

Reference: Smith, J.M., Introduction to Chemical Engineering Thermodynamics, 7th edition, McGraw-Hill, 2005, pp.199-253

2. Determination of the Optimal Operating Conditions for the Refrigerator
- Mathematical Model of the Refrigerator

T: temperature, P: pressure
h: specific enthalpy[J/g]
s: specific entropy [J/(K*g)]
v: specific volume [m^3/g]
w: Power provided to the compressor per mass [J/g]
q_H : specific Heat transfer from the refrigerant to the atmosphere[J/g]
q_L: specific heat transfer from the refrigerated space to the refrigerant[J/g]

1. Compressor

1) Design Variables: $P_1, v_1, T_1, P_2, v_2, T_2, T_S, w$

2) Constraint:

$$h_1(P_1,v_1,T_1)+w=h_2(P_2,v_2,T_2)$$ [The first law of the thermodynamics]

$$\eta=\frac{h_2(P_2,v_2,T_2)-h_1(P_1,v_1,T_1)}{h_S(P_2,v_S,T_S)-h_1(P_1,v_1,T_1)}$$ [Efficiency of the compressor]

$$s_1(P_1,v_1,T_1)=s_2(P_2,v_S,T_S)$$ [The second law of the thermodynamics]

$$v_1=\frac{RT_1}{P}+b-\frac{a(T_1)}{P_1}\frac{v_1-b}{(v_1-\varepsilon b)(v_1-\sigma b)}$$ [Equation of state]

$$v_2=\frac{RT_2}{P}+b-\frac{a(T_2)}{P_2}\frac{v_2-b}{(v_2-\varepsilon b)(v_2-\sigma b)}$$ [Equation of state]

2. Condenser

단위 표기함

1) Design Variables: $P_2, v_2, T_2, P_3, v_3, T_3, q_H$

2) Constraint:

$$h_2(P_2,v_2,T_2)=q_H+h_3(P_3,v_3,T_3)$$ [The first law of the thermodynamics]

$$P_2=P_3$$ [Isobaric process]

$$v_2=\frac{RT_2}{P}+b-\frac{a(T_2)}{P_2}\frac{v_2-b}{(v_2-\varepsilon b)(v_2-\sigma b)}$$ [Equation of state]

$$v_3=\frac{RT_3}{P}+b-\frac{a(T_3)}{P_3}\frac{v_3-b}{(v_3-\varepsilon b)(v_3-\sigma b)}$$ [Equation of state]

$$P_3=10^{A-\frac{B}{T_3+C-273.15}}$$ [Saturated pressure and temperature]

$$T_3>T_{amb}+\Delta T_{\min}$$ [Outlet temperature of the condenser]

3. Expansion Valve

1) Design Variables: $P_3, v_3, T_3, P_4, v_4, T_4, v_{4,l}, v_{4,v}, v\_f$

2) Constraint:

$$h_3(P_3,v_3,T_3)=(1-v\_f)\cdot h_{4,l}(P_4,v_{4,l},T_4)+v\_f\cdot h_{4,v}(P_4,v_{4,v},T_4)$$ [The first law of the thermodynamics]

$$P_4=10^{A-\frac{B}{T_4+C-273.15}}$$ [Saturated pressure and temperature]

$$v_3=\frac{RT_3}{P}+b-\frac{a(T_3)}{P_3}\frac{v_3-b}{(v_3-\varepsilon b)(v_3-\sigma b)}$$ [Equation of state]

$$v_4=(1-v\_f)\cdot v_{4,l}+v\_f\cdot v_{4,v}$$

$$v_{4,l}=\frac{RT_4}{P}+b-\frac{a(T_4)}{P_4}\frac{v_{4,l}-b}{(v_{4,l}-\varepsilon b)(v_{4,l}-\sigma b)}$$ [Equation of state]

$$v_{4,v}=\frac{RT_4}{P}+b-\frac{a(T_4)}{P_4}\frac{v_{4,v}-b}{(v_{4,v}-\varepsilon b)(v_{4,v}-\sigma b)}$$ [Equation of state]

4. Evaporator

1) Design Variables: $P_4, v_4, T_4, P_1, v_1, T_1, v_{4,l}, v_{4,v}, v\_f, \dot{m}, q_L$

2) Constraint:

$$\dot{m}\cdot(1-v\_f)\cdot h_{4,l}(P_4,v_{4,l},T_4)+\dot{m}\cdot v\_f\cdot h_{4,v}(P_4,v_{4,v},T_4)+\dot{m}\cdot q_L=\dot{m}\cdot h_1(P_1,v_1,T_1)$$ [The first law of the thermodynamics]

$$P_4=P_1$$ [Isobaric process]

$$v_1=\frac{RT_1}{P}+b-\frac{a(T_1)}{P_1}\frac{v_1-b}{(v_1-\varepsilon b)(v_1-\sigma b)}$$ [Equation of state]

$$v_4=(1-v\_f)\cdot v_{4,l}+v\_f\cdot v_{4,v}$$

$$v_{4,l}=\frac{RT_4}{P}+b-\frac{a(T_4)}{P_4}\frac{v_{4,l}-b}{(v_{4,l}-\varepsilon b)(v_{4,l}-\sigma b)}$$ [Equation of state]

$$v_{4,v}=\frac{RT_4}{P}+b-\frac{a(T_4)}{P_4}\frac{v_{4,v}-b}{(v_{4,v}-\varepsilon b)(v_{4,v}-\sigma b)}$$ [Equation of state]

$$P_1=10^{A-\frac{B}{T_1+C-273.15}}$$ [Saturated pressure and temperature]

[Given]

$$\dot{m}\cdot q_L=20[kJ/s]$$

4. Determination of the Optimal Operating Conditions for the Refrigerator
- Summary of Mathematical Model of This research for Refrigerator

T: temperature, h: specific enthalpy, s: specific entropy,
P: pressure
v: specific volume
$\dot{W}$: Power provided to the compressor per mass
$\dot{q}_H$: Specific heat transfer from the refrigerant to the atmosphere
$\dot{q}_L$: Specific heat transfer from the refrigerated space to the refrigerant(Given)
U: Heat transfer coefficient of the evaporator
A: Area of the evaporator

1. Design Variables(Operating Conditions, 21): P_i, T_i, v_i, T_s, v_S, $v_{4,l}$, $v_{4,v}$, v_f, w, $\dot{m}$, q_H, q_L (i=1,2,3,4)

2. Equality constraints(19)

1) Compressor(6)

$$\dot{m}\cdot h_1(P_1,v_1,T_1)+\dot{m}\cdot w=\dot{m}h_2(P_2,v_2,T_2)$$ [The first law of the thermodynamics]

$$\eta=\frac{h_S(P_2,v_S,T_S)-h_1(P_1,v_1,T_1)}{h_2(P_2,v_2,T_2)-h_1(P_1,v_1,T_1)}$$ [The second law of the thermodynamics]

$$s_1(P_1,v_1,T_1)=s_2(P_2,v_S,T_S)$$

$$v_1=\frac{RT_1}{P_1}+b-\frac{a(T_1)}{P_1}\frac{v_1-b}{(v_1-\varepsilon b)(v_1-\sigma b)}$$ [Equation of state]

$$v_2=\frac{RT_2}{P_2}+b-\frac{a(T_2)}{P_2}\frac{v_2-b}{(v_2-\varepsilon b)(v_2-\sigma b)} \qquad v_S=\frac{RT_S}{P_2}+b-\frac{a(T_S)}{P_2}\frac{v_S-b}{(v_S-\varepsilon b)(v_S-\sigma b)}$$

2) Condenser(4)

$$h_2(P_2,v_2,T_2)=q_H+h_3(P_3,v_3,T_3)$$ [The first law of the thermodynamics]

$$P_2=P_3$$ [Isobaric process]

$$v_3=\frac{RT_3}{P_3}+b-\frac{a(T_3)}{P_3}\frac{v_3-b}{(v_3-\varepsilon b)(v_3-\sigma b)}$$ [Equation of state]

$$\frac{P_3}{10^5}=10^{A-\frac{B}{T_3+C-273.15}}$$ [Saturated pressure and temperature]

3) Expansion valve(5)

$$h_3(P_3,v_3,T_3)=(1-v\_f)\cdot h_{4,l}(P_4,v_{4,l},T_4)+v\_f\cdot h_{4,v}(P_4,v_{4,v},T_4)$$ [The first law of the thermodynamics]

$$\frac{P_4}{10^5}=10^{A-\frac{B}{T_4+C-273.15}}$$ [Saturated pressure and temperature]

$$v_{4,l}=\frac{RT_4}{P_4}+b-\frac{a(T_4)}{P_4}\frac{v_{4,l}-b}{(v_{4,l}-\varepsilon b)(v_{4,l}-\sigma b)} \qquad v_{4,v}=\frac{RT_4}{P_4}+b-\frac{a(T_4)}{P_4}\frac{v_{4,v}-b}{(v_{4,v}-\varepsilon b)(v_{4,v}-\sigma b)}$$

$$v_4=(1-v\_f)\cdot v_{4,l}+v\_f\cdot v_{4,v}$$

4) Evaporator (4)

$$\dot{m}\cdot(1-v\_f)\cdot h_{4,l}(P_4,v_{4,l},T_4)+\dot{m}\cdot v\_f\cdot h_{4,v}(P_4,v_{4,v},T_4)+\dot{m}\cdot q_L=\dot{m}\cdot h_1(P_1,v_1,T_1)$$ [The first law of the thermodynamics]

$$P_4=P_1$$ [Isobaric process]

$$\frac{P_1}{10^5}=10^{A-\frac{B}{T_1+C-273.15}}$$ [Saturated pressure and temperature]

$$\dot{m}\cdot q_L=20[kJ/s]$$ [Heat transfer in the evaporator]

3. Inequality constraint(1)

$$T_3>T_{amb}+\Delta T_{\min}$$ [Outlet temperature of the condenser]

Since the number of equality constraints is less than the number of design variables, these equations form indeterminate systems.

We need a certain criteria to determine the proper solution. By introducing the criteria(objective function), this problem can be formulated as an optimization problem.

4. Objective function(f)

Minimize the power provided to the compressor.

$$f=\dot{m}\cdot w$$

 # Mole (Mol)

Mass of $\underbrace{6.02\text{x}10^{23}}_{\text{1 mole}}$ atoms of Carbon = $\boxed{12.01 \text{ g}}$

Mass of $\underbrace{6.02\text{x}10^{23}}_{\text{1 mole}}$ atoms of Methane = $\boxed{16.09 \text{ g}}$

- The mass in grams per mole of a substance is called the molar mass.

Why $6.02\text{x}10^{23}$ atoms?

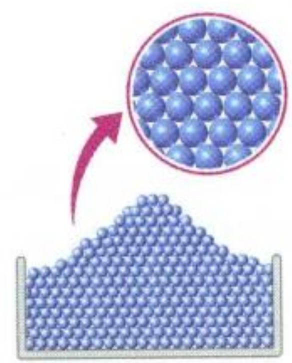
- In everyday life, we use counting units like dozen and tray to deal with large quantities.

<A dozen of pencils>

<A tray of eggs>

- A mole is the amount of matter that contains as many objects such as atoms or molecules as the number of atoms in exactly 12 g of C(carbon), “Avogadro’s number”.

1 mole of atoms or molecules = $\underbrace{6.02\text{x}10^{23}}_{\text{Avogadro's number}}$



<A mole of atoms>

- Amedeo Avogadro (Italian sientist,1776-1856)

[Thermodynamics] Molar Mass, Molar Volume, Density and Specific Volume

- **Molar mass(M): The mass in grams per mole of a substance**

Name	Molar mass [g/mol]
Carbon(C)	12.01
Hydrogen(H)	1.008
Iron(Fe)	55.85
Water(H_2O)	18.02
Methane(CH_4)	16.04

- **Molar volume($\bar{v}$): The volume per mole of a substance**

  $[m^3 / kmol]$

- **Density(ρ)** $= \dfrac{\cancel{\text{Molar}}\ \text{mass}}{\cancel{\text{Molar}}\ \text{volume}}$

  $[kg / m^3]$

Example) Density of the water at 1 atm and 0 °C ?

- **Molar mass of the water = 18.02 g/mol**
- **Molar volume of the water at 1 atm and 0 °C = 1.8×10^{-5} m³/mol[1]**

$$\rho = \frac{18.02}{1.8 \times 10^{-5}} \left[\frac{g}{mol} \times \frac{mol}{m^3} \right]$$
$$\approx 1.001 \times 10^6 \, [g / m^3]$$
$$\approx 1.001 \, [ton / m^3]$$

- **Specific volume(v)** $= \dfrac{1}{\rho} = \dfrac{\text{Volume}}{\text{Mass}}$

  $[m^3 / kg]$

  :The volume per unit mass of a substance

- **Specific~: "~ per unit mass"**
  Molar~: "~ per mole"

Reference: 1)김봉래, 김득호, High Top 화학 II, 두산동아,, 2009, p. 98

- Verification of the Mathematical Model of This research for Refrigerator (3)

Indeterminate systems

1. Design Variables(Operating Conditions, 21)
   : $P_i, T_i, v_i, T_s, v_S, v_{4,l}, v_{4,v}, v\_f, w, \dot{m}, q_L, q_H$ (i=1,2,3,4)
2. Equality constraints [19]
   1) Compressor (6)
   2) Condenser (4)
   3) Expansion valve (5)
   4) Evaporator (4)

- **Assumption of the values of the free variables** [2 = 21 - 19]
  : $P_1 = 2.115[bar]$
  $P_2 = 11.698[bar]$

System of nonlinear equations

- **Design variables** [19]
- **Equality constraints** [19]

→ **Use of Newton-Raphson method**

	Result obtained by this paper
P_1[bar]	2.115
T_1[K]	256.152
v_1[m^3/mol]	0.0098128
P_2[bar]	11.717
T_2[K]	390.278
v_2[m^3/mol]	0.0026551
P_3[bar]	11.717
T_3[K]	303.273
v_3[m^3/mol]	0.0000365
P_4[bar]	2.114
T_4[K]	256.151
v_4[m^3/mol]	0.0017003
v_f	0.1705
v4v[m^3/mol]	0.0098130
v4l[m^3/mol]	0.0000327
T_s[K]	384.793
v_s[m^3/mol]	0.0026119
$\dot{m}$ [g/s]	17.6
w[J/g]	265.698
q_L[J/g]	1136.36
q_H[J/g]	1402.676
Objective function(W)[kW]	4.676 265.698x17.6/1000 =4.6768

Molar mass of ammonia: 17.031 g/mol

2. Determination of the Optimal Operating Conditions for the Refrigerator

- Verification of the Mathematical Model of This research for Refrigerator (1)

How can we verify the mathematical model of this research for refrigerator?

<Input>

Assume the 2 free variables

Mathematical model:
- Design variables(21)
- Equality constraints(19)

<Output>

The values of the other design variables

	Result obtained by this paper
P_1[bar]	**2.115**
T_1[K]	256.152
v_1[m^3/mol]	0.0098128
P_2[bar]	**11.717**
T_2[K]	390.278
v_2[m^3/mol]	0.0026551
P_3[bar]	11.717
T_3[K]	303.273
v_3[m^3/mol]	0.0000365
P_4[bar]	2.114
T_4[K]	256.151
v_4[m^3/mol]	0.0017003
v_f	0.1705
v4v[m^3/mol]	0.0098130
v4l[m^3/mol]	0.0000327
T_s[K]	384.793
v_s[m^3/mol]	0.0026119
$\dot{m}$ [g/s]	17.6
w[J/g]	265.698
q_L[J/g]	1136.36
q_H[J/g]	1402.676
Objective function (W)[kW]	4.676 265.698x17.6/1000 =4.6768

Molar mass of ammonia: 17.031 g/mol

- Mathematical Model of this research
  1. Design variables(Operating conditions, 21)
  2. Equality Constraints(19)

  → indeterminate systems

To verify the mathematical model this research, we assume the values of the two design variables, solve and compare the result with that of the Aspen HYSYS.

- Verification of the Mathematical Model of This research for Refrigerator (2)

How can we verify the mathematical model of this research for refrigerator?

<Input>

Assume the 2 free variables

Mathematical model:
- Design variables(21)
- Equality constraints(19)

<Output>

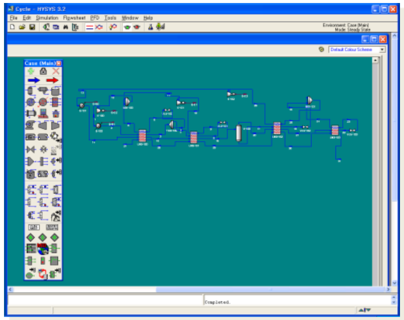
The values of the other design variables

	Result obtained by this paper
P_1[bar]	2.115
T_1[K]	256.152
v_1[m^3/mol]	0.0098128
P_2[bar]	11.717
T_2[K]	390.278
v_2[m^3/mol]	0.0026551
P_3[bar]	11.717
T_3[K]	303.273
v_3[m^3/mol]	0.0000365
P_4[bar]	2.114
T_4[K]	256.151
v_4[m^3/mol]	0.0017003
v_f	0.1705
v4v[m^3/mol]	0.0098130
v4l[m^3/mol]	0.0000327
T_s[K]	384.793
v_s[m^3/mol]	0.0026119
$\dot{m}$ [g/s]	17.6
w[J/g]	265.698
q_L[J/g]	1136.36
q_H[J/g]	1402.676
Objective function(W)[kW]	265.698x17.6/1000 =4.6768

Comparison for verifying the mathematical model

<Input>

Assume the 2 free variables



Aspen HYSYS[1)]

There are equality constraints, but it is difficult to find the equality constraints(Black box).

<Output>

The values of the other design variables

	Aspen HYSYS	Difference
P_1[bar]	2.115	0.000%
T_1[K]	256.133	0.008%
v_1[m^3/mol]	0.0098078	0.051%
P_2[bar]	11.717	0.000%
T_2[K]	390.059	0.056%
v_2[m^3/mol]	0.0026532	0.071%
P_3[bar]	11.717	0.000%
T_3[K]	303.202	0.023%
v_3[m^3/mol]	3.654E-05	0.099%
P_4[bar]	2.115	0.047%
T_4[K]	256.130	0.008%
v_4[m^3/mol]	0.0017274	1.591%
v_f	0.1733642	1.680%
v4v[m^3/mol]	0.0098078	0.053%
v4l[m^3/mol]	3.273E-05	0.090%
T_s[K]	384.793	0.114%
v_s[m^3/mol]	0.0026086	0.127%
$\dot{m}$ [g/s]	17.6543	0.308%
w[J/g]	268.80043	1.168%
q_L[J/g]	1132.87	0.308%
q_H[J/g]	1401.6706	0.072%
Objective function(W)[kW]	268.80043 x17.6/1000 =4.745	1.573%

$$\dot{m}\cdot q_L = 20[kJ/s]$$
$$\dot{m}\cdot w = 4.745[kW] \Rightarrow CP = \frac{q_L}{w} = 4.215$$

1) It is a kind of simulation program to solve the simultaneous equations by considering equality constraint related with thermodynamics and made by Aspentech.

The mathematical model can be verified by comparing between the values of the other design variables obtained by mathematical model of this research and Aspen HYSYS.

2. Determination of the Optimal Operating Conditions for the Baseline Liquefaction Cycle

Optimization Problem

1. Design Variables(Operating Conditions, 21)

: $P_i, T_i, v_i, T_s, v_S, v_{4,l}, v_{4,v}, v\_f, w, M, q_L, q_H$ (i=1,2,3,4)

2. Equality constraints [20]

1) Compressor (6)
2) Condenser (4)
3) Expansion valve (5)
4) Evaporator (5)

3. Inequality constraints [1]

4. Objective function: Minimize the compressors power

$$\text{Minimize} \quad W = M \cdot w$$

w : work input to the compressor per mass [J / kg]

M : mass flow rate of refrigerant [kg / s]

c.f) Heat transfer in the evaporator

$$M \cdot q_L = U \cdot A \cdot (T_C - T_4)$$

U: Heat transfer coefficients[W/m²K]

A: Area of equipment temperature[m²]

T_C: Room temperature[K]

Procedure of finding optimum solution

- **Free variables** [1 = 21 - 20]

  : P_1

- **Inequality constraints** [1]

Calculation of objective function

$$\text{Minimize} \quad W = M \cdot w$$

① 1st Step: Find the free variable P_1, by minimizing the compressor power subject to the inequality constraint using sequential quadratic programming(SQP) method. .

② 2nd Step: determine the 20dependent variables by solving the system of the nonlinear equations.

System of nonlinear equations

- **Design variables** [20]
- **Equality constraints** [20]

→ **Determine the 20 variables using Newton-Raphson method**

- Result of the Optimal Operating Conditions for the Refrigerator

Problem[1)]

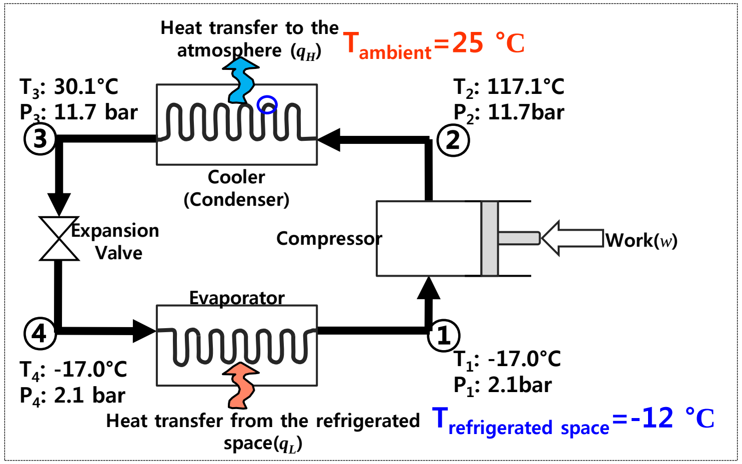
Given: $\dot{m} \cdot q_L = 20[kJ/s], U = 1,000[W/m^2K],\ A = 4.0\ [m^2],$

$T_C = -12°C, \eta = 95\%,\ T_{amb} = 25°C, \Delta T_{\min} = 5°C$

where $\dot{m} \cdot q_L$: Rate of heat transfer from the refrigerated space to the refrigerant
Tc: temperature of the refrigerated space
η: efficiency of the compressorwhere
T_{amb}: ambient temperature
ΔT_{min}: minimum value of the difference between the ambient temperature and outlet temperature in the condenser

Find: Operating condition

Minimize $W = \dot{m} \cdot w$ $\dot{m} \cdot w$: Power provided to the compressor[kW]



°C = K – 273.15

Optimization result:

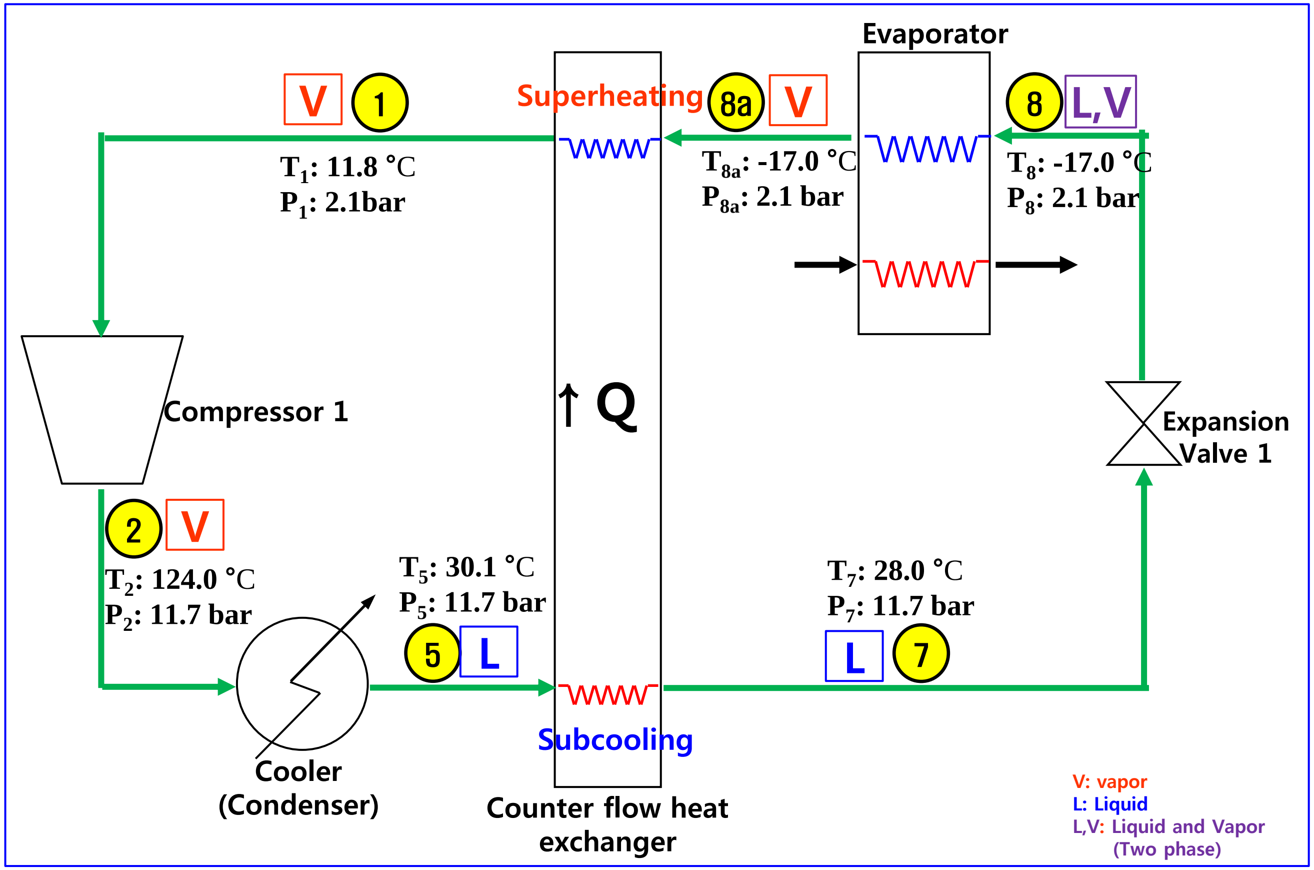
	Result obtained by this paper
P_1[bar]	2.115
T_1[K]	256.152
v_1[m^3/mol]	0.0098128
P_2[bar]	11.717
T_2[K]	390.278
v_2[m^3/mol]	0.0026551
P_3[bar]	11.717
T_3[K]	303.273
v_3[m^3/mol]	0.0000365
P_4[bar]	2.114
T_4[K]	256.151
v_4[m^3/mol]	0.0017003
v_f	0.1705
v4v[m^3/mol]	0.0098130
v4l[m^3/mol]	0.0000327
T_s[K]	384.793
v_s[m^3/mol]	0.0026119
$\dot{m}$ [g/s]	17.6
w[J/g]	265.698
q_L[J/g]	1136.36
q_H[J/g]	1402.676
Objective function(W)[kW]	265.698x17.6/1000=4.6768

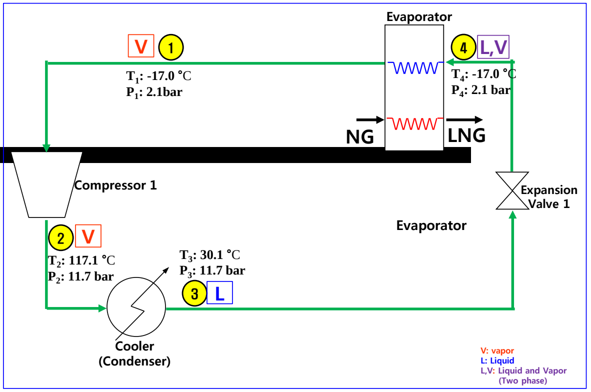
Molar mass of ammonia: 17.031 g/mol

Reference: 1) Jensen, J.B., 2008, Optimal Operation of Refrigeration Cycles, Ph.D. thesis, Norweglan University of Science and Technology.

9.4. Various Combination of Equipment for the Liquefaction Cycle

3. 1 Single Cycle with Regeneration (1)





c.f)
Total work of Single Cycle without "Regeneration":
4.745kW(p.24) $CP = 4.215$

Total work:
4.800kW

$\dot{m} \cdot q_L = 20[kJ/s]$

$\dot{m} \cdot w = 4.800[kW]$

$CP = \dfrac{q_L}{w} = 4.167$

$\dot{m} = 17.48[g/s]$,

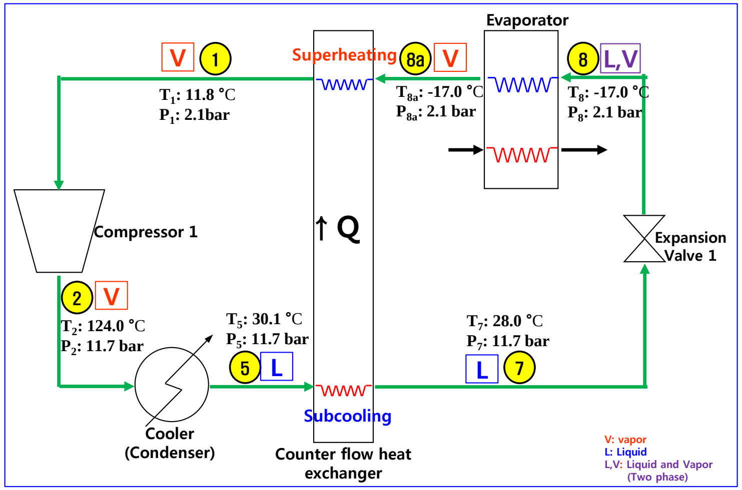
$q_L = 1144.16[J/g]$,

$w = 274.60[J/g]$

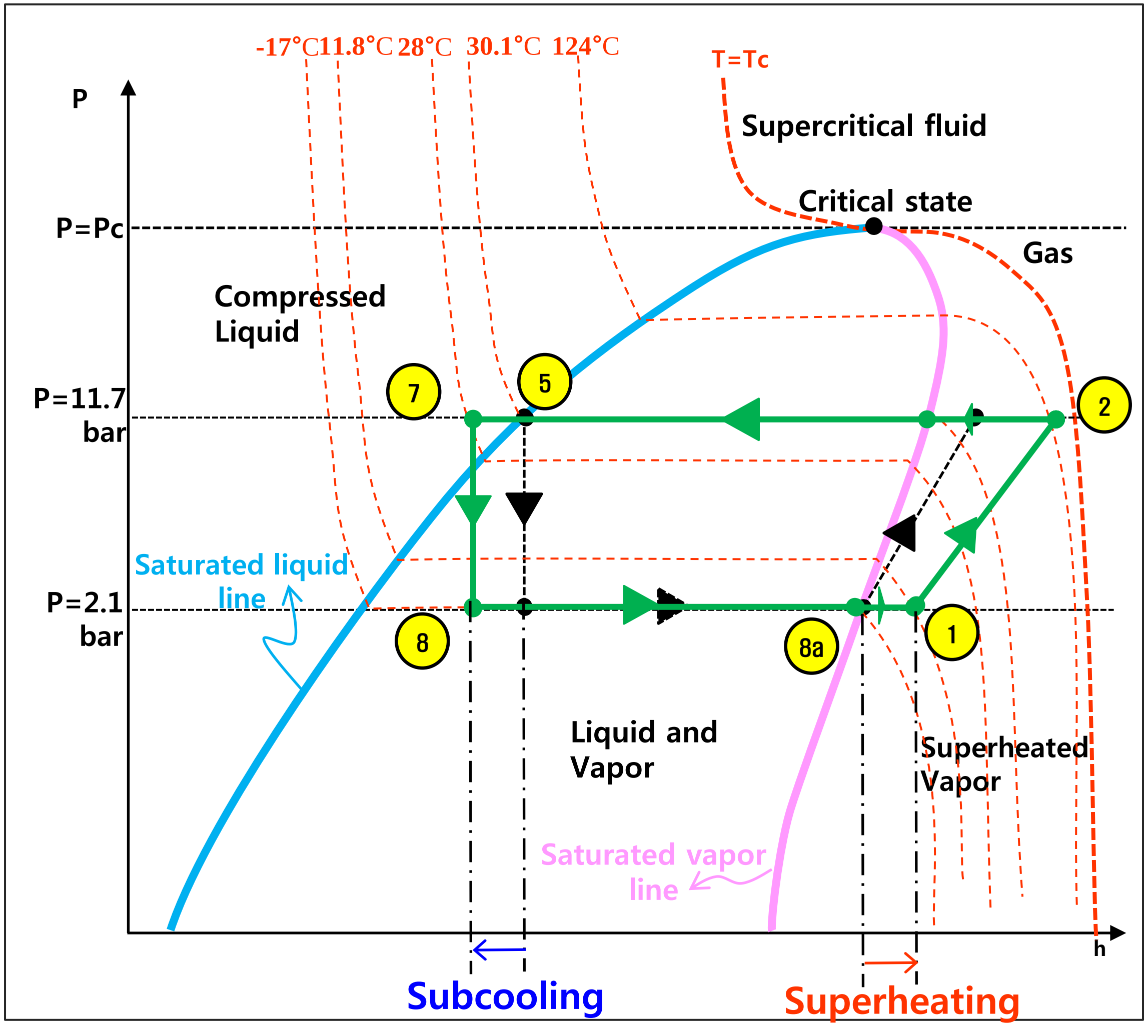
Molar mass of ammonia: 17.031 g/mol

- **Regenerative cooling(Q):** By inserting a counter-flow heat exchanger into the cycle, the high pressure liquid refrigerant after condenser is further cooled before expanding in the expansion valve.

- Single Cycle with Regeneration (2)



V: vapor
L: Liquid
L and V: Liquid and Vapor(Two phase

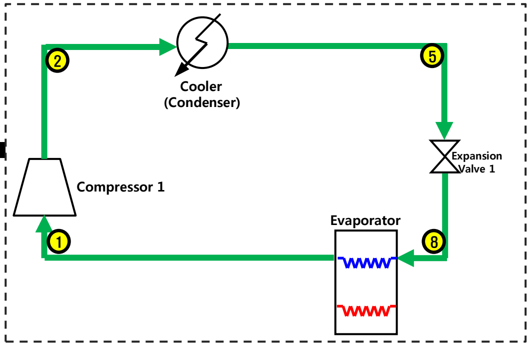


- Advantage:

1) Since the refrigerant is subcooled before expanding, the cooling capacity(Qp) of the refrigerator is increased.

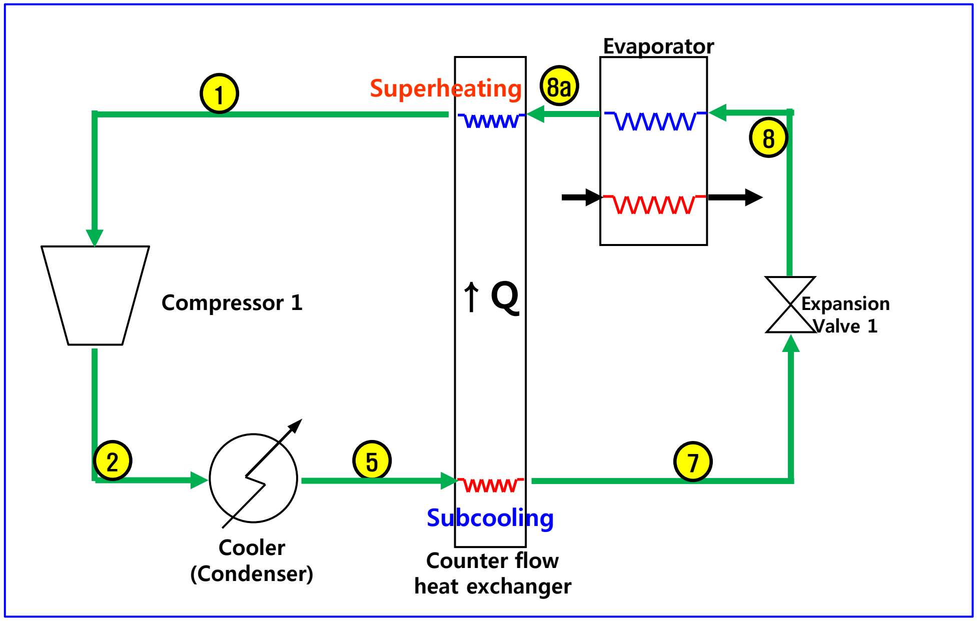
2) Since the compressors are designed as vapor pumps, if any amount of liquid is allowed to enter the compressor, serious mechanical damage to the compressor may result. However, it is difficult to control the state of the refrigerant as the saturated vapor state. In this case, superheating the refrigerant prevents the liquid refrigerant from entering the compressor.

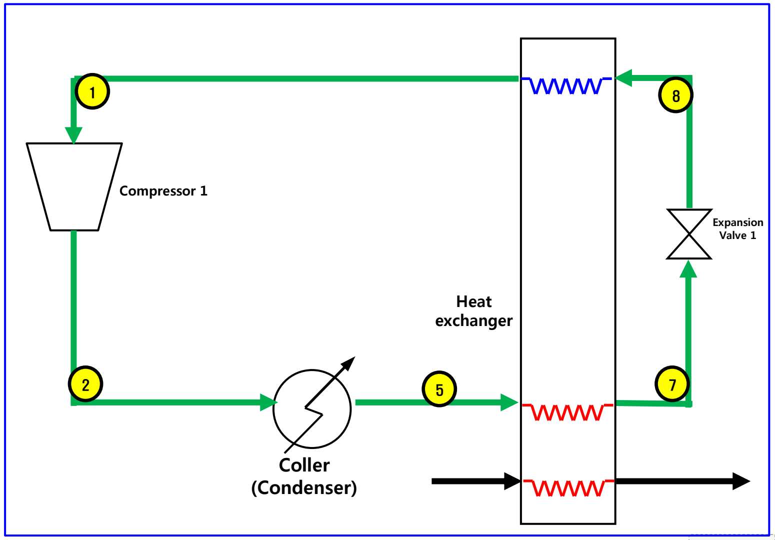
- Single Cycle with Regeneration (3)



- **Refrigerator**

V: vapor
L: Liquid
L and V: Liquid and Vapor(Two phase)





In case of the liquefaction cycle, the counter flow heat exchanger and evaporator can be combined and represented as a single evaporator.

c.f)
Total work of Single Cycle **without** "Multistage Compression with Intercooling": **4.745kW(p.24)**

$CP = 4.215$

Work generated by compressor 1: 2.14W

Work generated by compressor 2: 2.48W

Total work: **4.622kW**

$\dot{m} \cdot q_L = 20[kJ/s]$

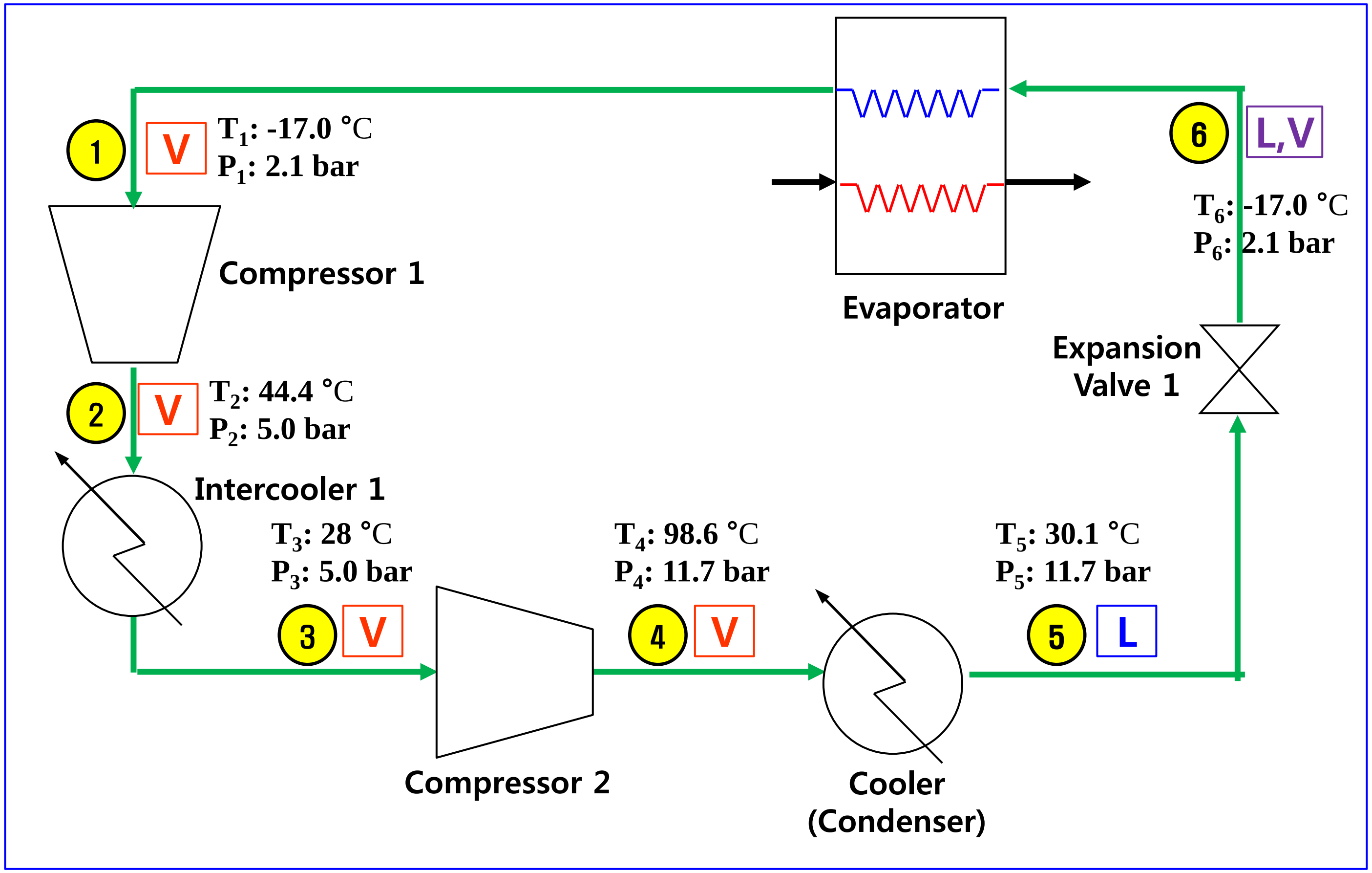
$\dot{m} \cdot w = 4.622[kW]$

$CP = \frac{q_L}{w} = 4.327$

$\dot{m} = 17.6[g/s]$,

$q_L = 1136.36[J/g]$,

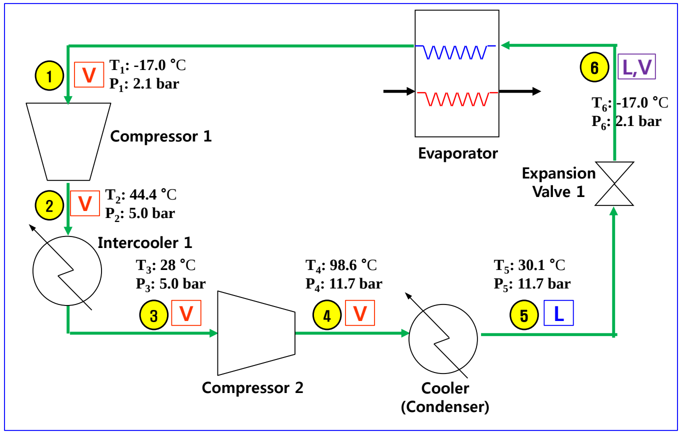
$w = 262.61[J/g]$

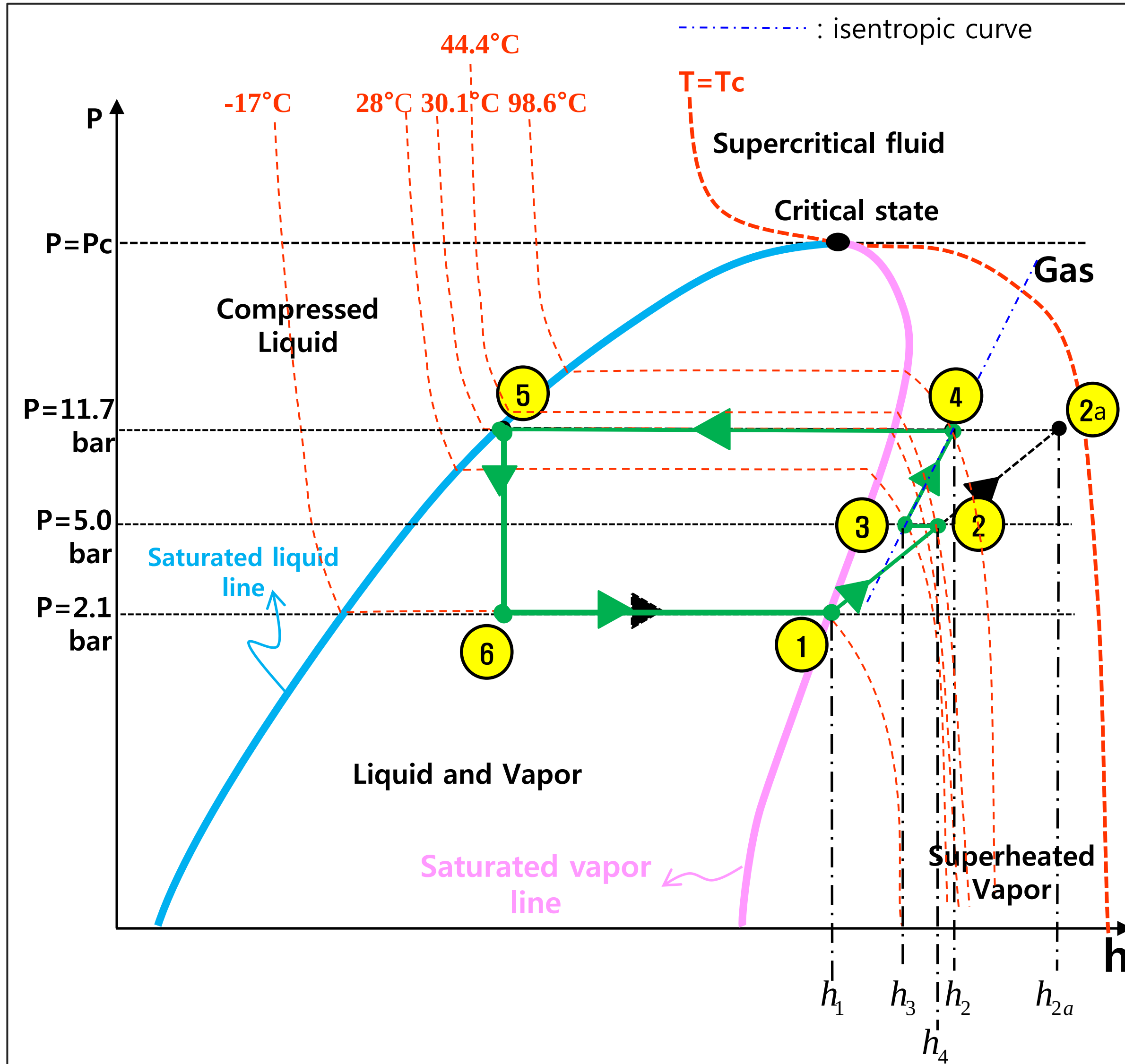


Molar mass of ammonia: 17.031 g/mol

- **Multistage Compression with Intercooling:**
  The refrigerant is compressed in **multistage** and cooled down between each stage by passing through an intercooler

- Single Cycle with Multistage Compression with Intercooling (2)





- Advantage:

The compressor work to be provided can be ~~is~~ reduced.

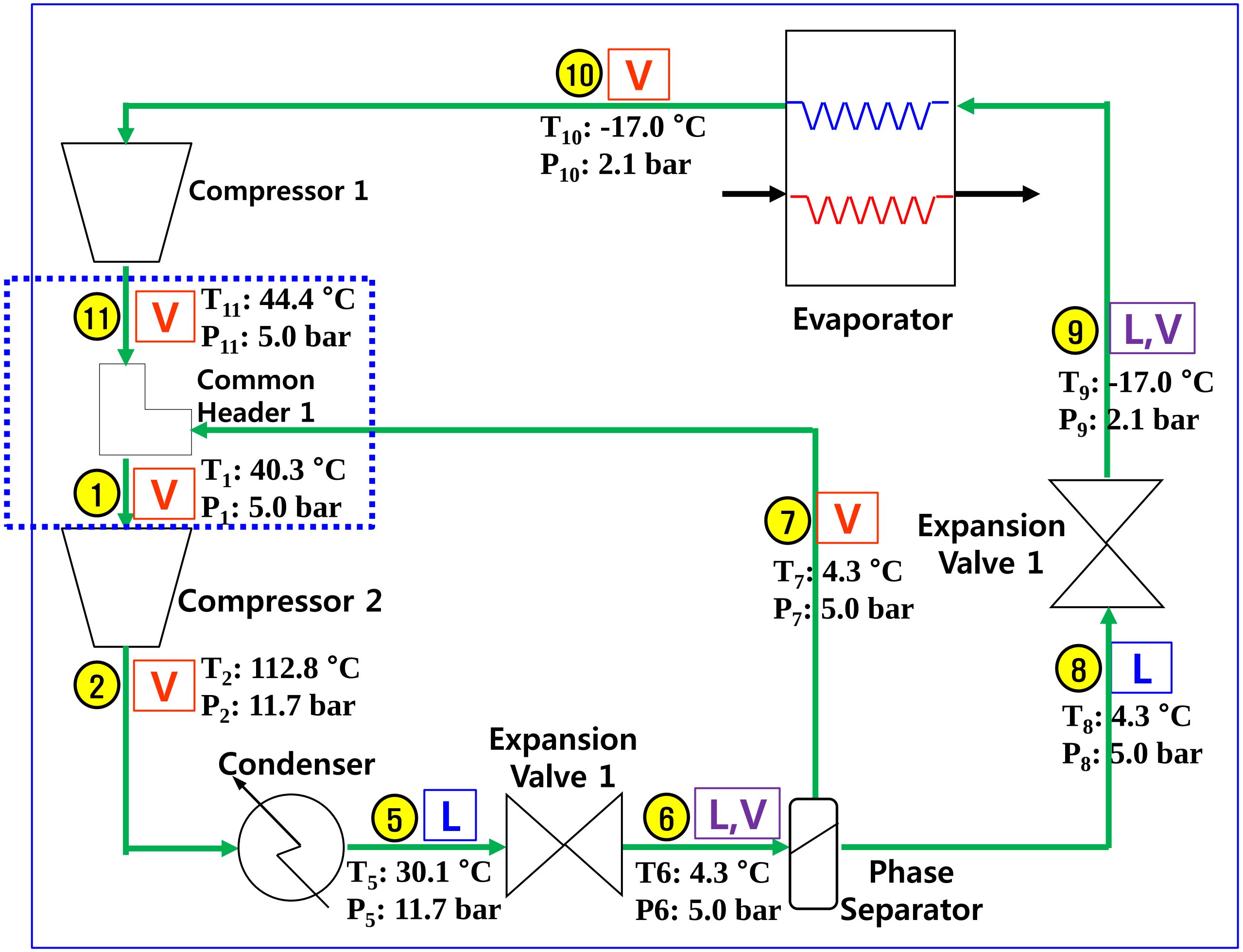
$$W_1 > W_2$$

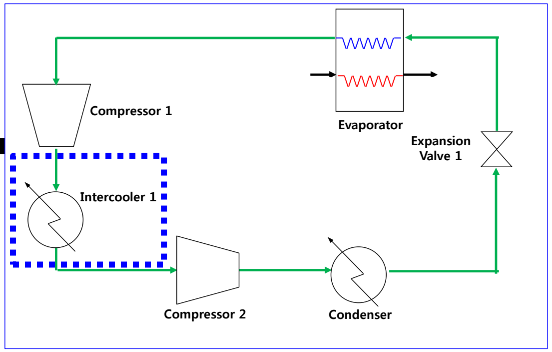
Single stage compression

$$W_1 = h_{2a} - h_1$$
$$= (h_2 - h_1) + (h_{2a} - h_2)$$

Two stage compression with intercooler

$$W_2 = (h_2 - h_1) + (h_4 - h_3)$$





c.f)
Total work of Single Cycle without "Multistage Compression with Intercooling": 4.745kW(p.24)

$$CP = 4.215$$

Work generated by compressor 1: 1.92kW

Work generated by compressor 2: 2.58kW

Total work:
4.50kW

$$\dot{m} \cdot q_L = 20[kJ/s]$$

$$\dot{m} \cdot w = 4.50[kW] \qquad CP = \frac{q_L}{w} = 4.444$$

$$\dot{m} = 15.8[g/s],$$

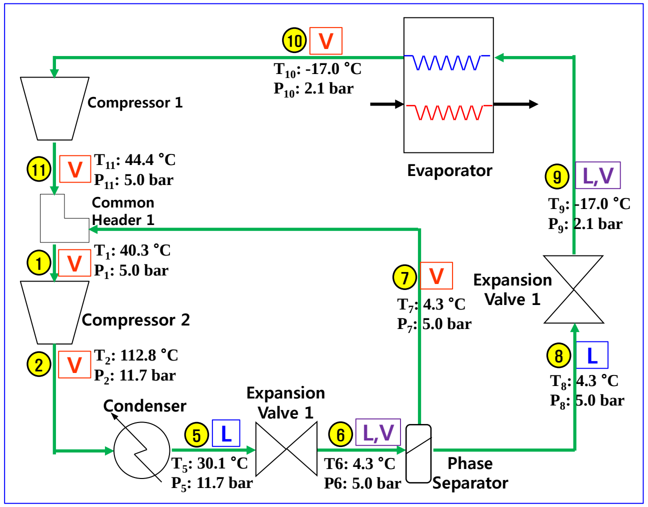
$$q_L = 1265.82[J/g],$$

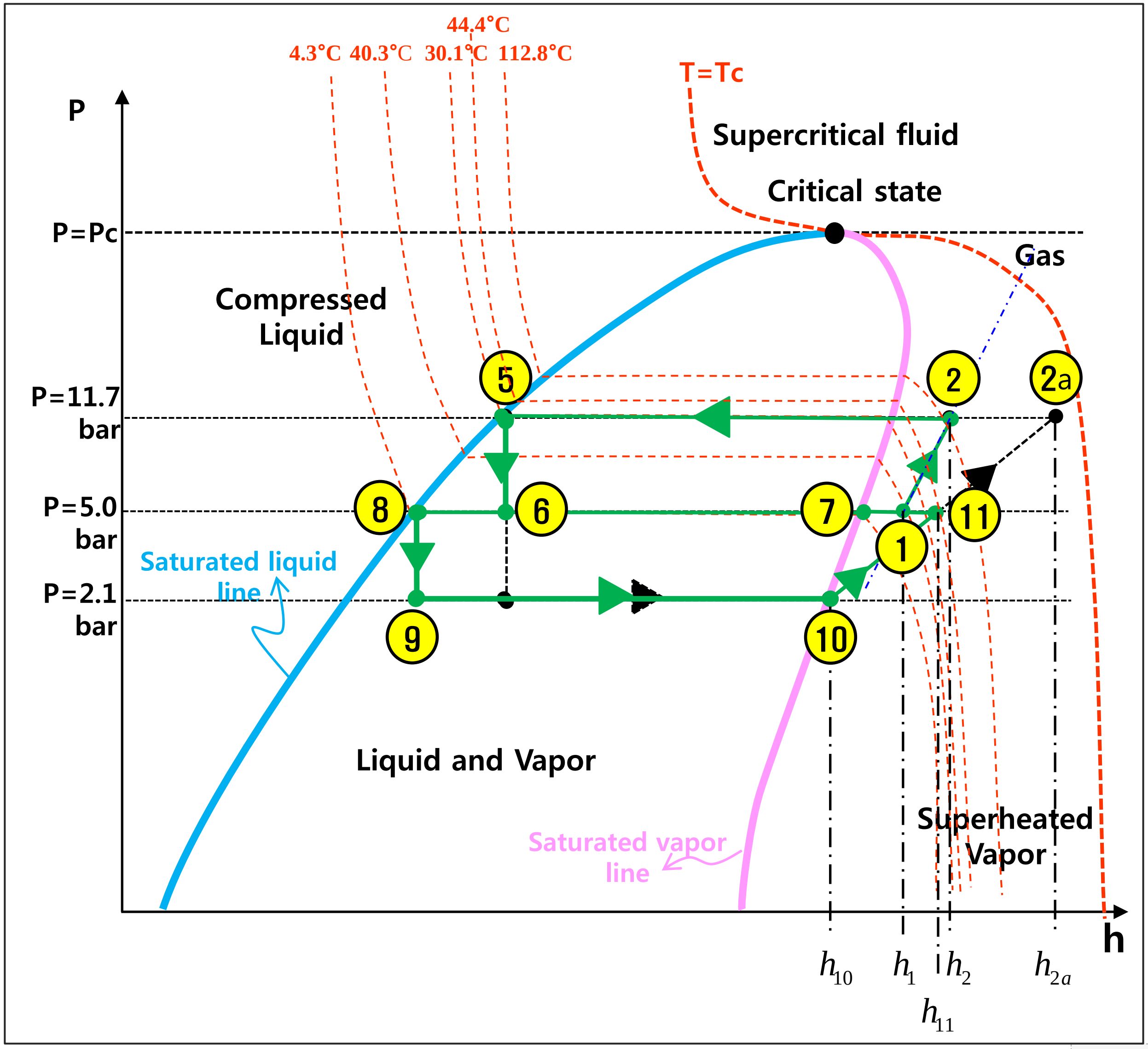
$$w = 284.81[J/g]$$

Molar mass of ammonia: 17.031 g/mol

- **Multistage Compression Refrigeration:**
  1) Phase separator: separates a liquid-vapor mixture refrigerant into the vapor and liquid
  2) The saturated vapor(stream 7) is mixed with the superheated vapor from the compressor 1(stream 11), and the cooled mixture(stream 1) enters the compressor 2.

- Single Cycle with Multistage Compression Refrigeration (2)





- Advantage:

The compressor work to be provided is reduced.

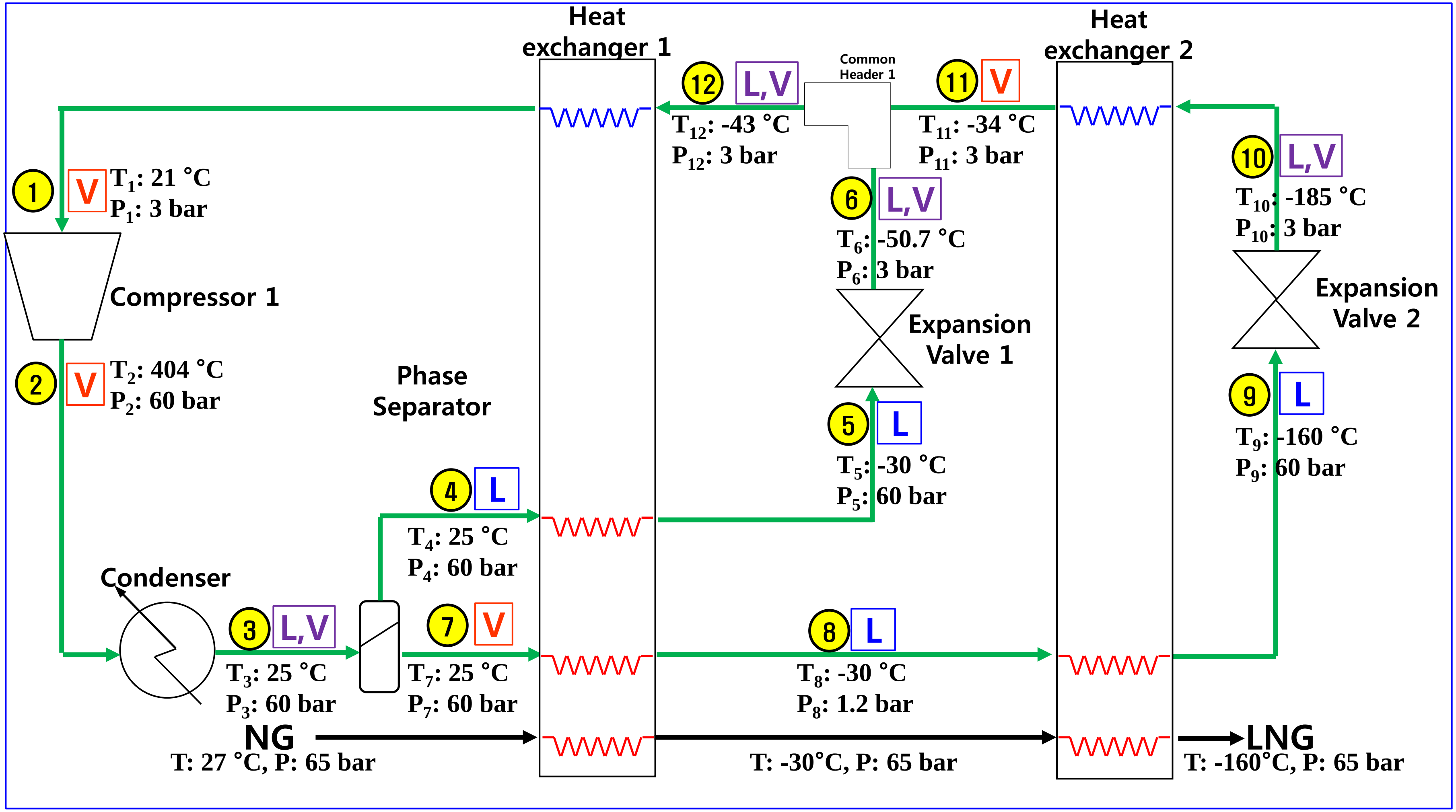
$$W_1 > W_2$$

Single stage compression

$$W_1 = h_{2a} - h_{10}$$
$$= (h_{11} - h_{10}) + (h_{2a} - h_{11})$$

Two stage compression with intercooler

$$W_2 = (h_{11} - h_{10}) + (h_2 - h_1)$$

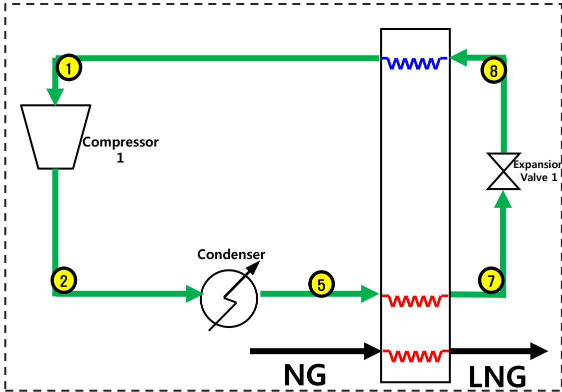


Molar mass of nitrogen (N_2): 28.013 g/mol $\dot{m} = 19378.41[g/s]$,

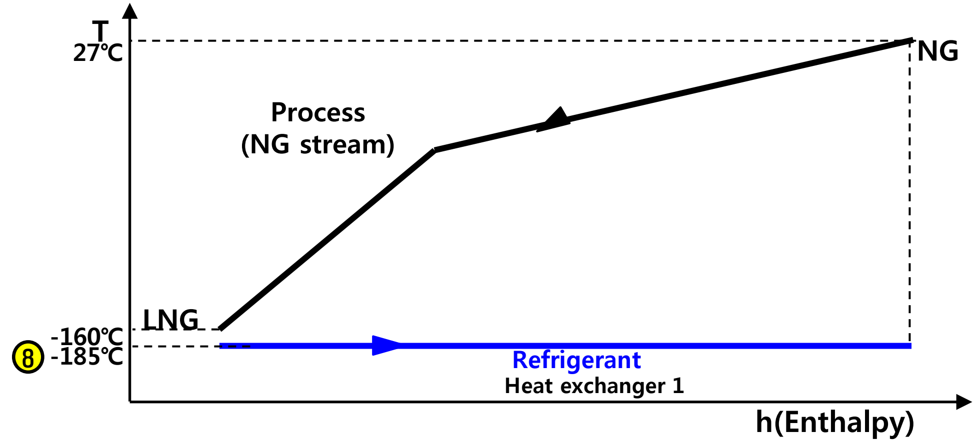
- **Multistage Refrigeration: Repeated partial condensation and separation of the refrigerant**

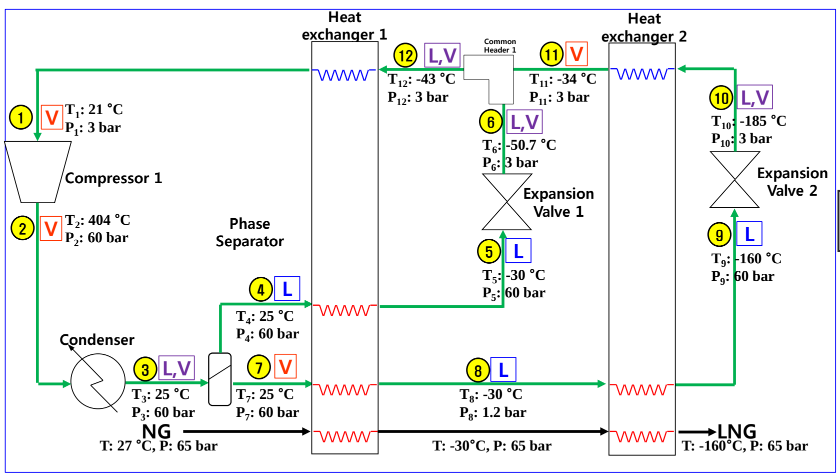
Nogal, F.D., "Optimal Design and Integration of Refrigeration and Power Systems", PhD thesis, Univ. of Manchester, p. 38.

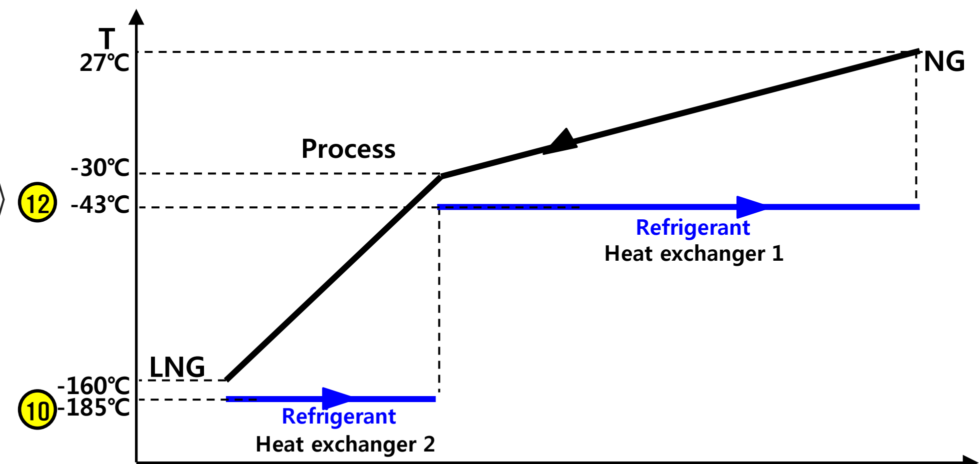
- Single Cycle with Regeneration and Multistage Refrigeration (2)



Singlestage Refrigeration (pure refrigerant)







Multistage Refrigeration (pure refrigerant)

Better performance

- Advantage:

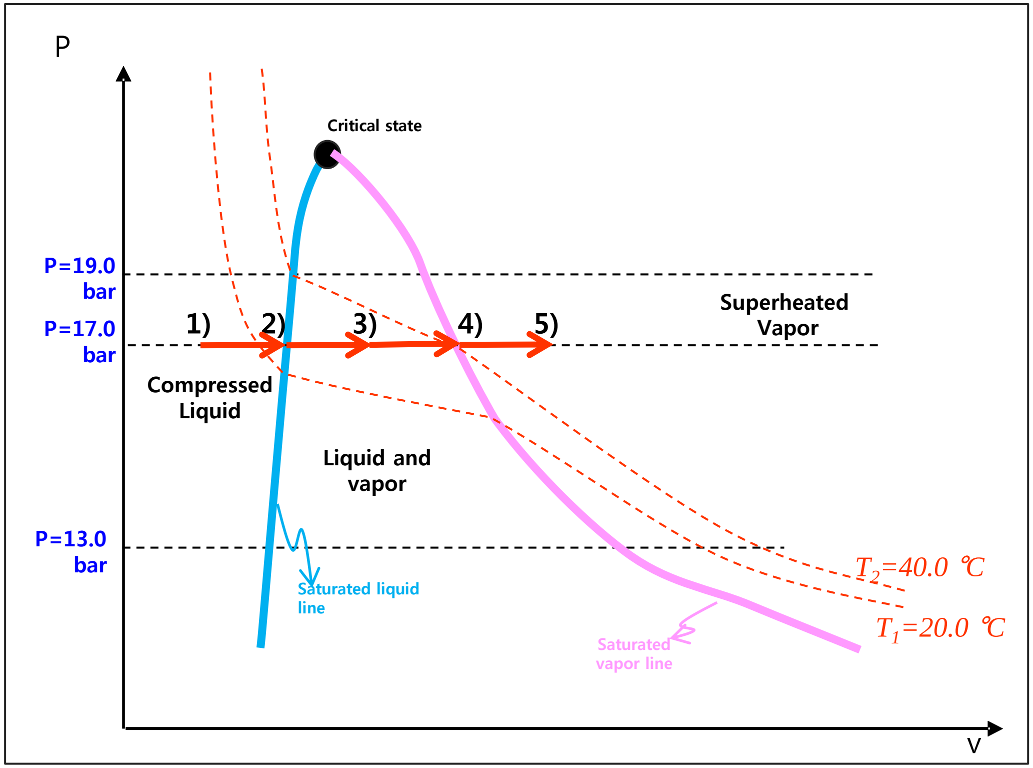
Achieving better match between the temperature profiles of refrigerant and natural gas

→ The compressor work to be provided is reduced.

Nogal, F.D., "Optimal Design and Integration of Refrigeration and Power Systems", PhD thesis, Univ. of Manchester, p. 37

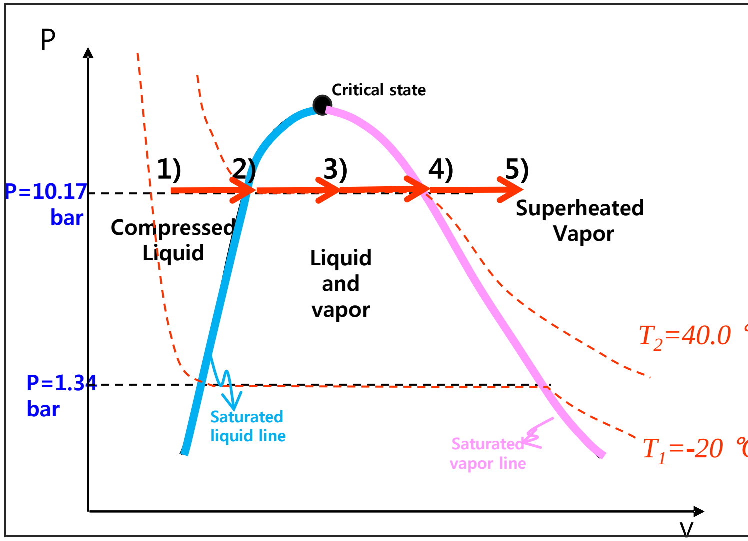
- Pure Substance vs. Mixture

- **Pressure(P)-Specific Volume(v) Diagram of the Mixture**
  - Example: Mixture composed Ethane(C_2H_6, 22.02%), Propane(C_3H_8, 65.30%), and n-Butane(C_3H_{10}, 12.68%)



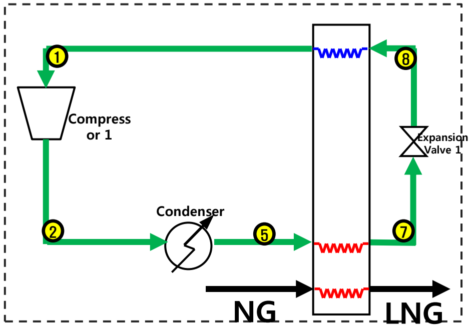
In the two phase region of the mixture, the temperature *is not constant* at the constant pressure.

- **Pressure(P)-Specific Volume Diagram of the Pure Substance**
  - Example: Ammonia

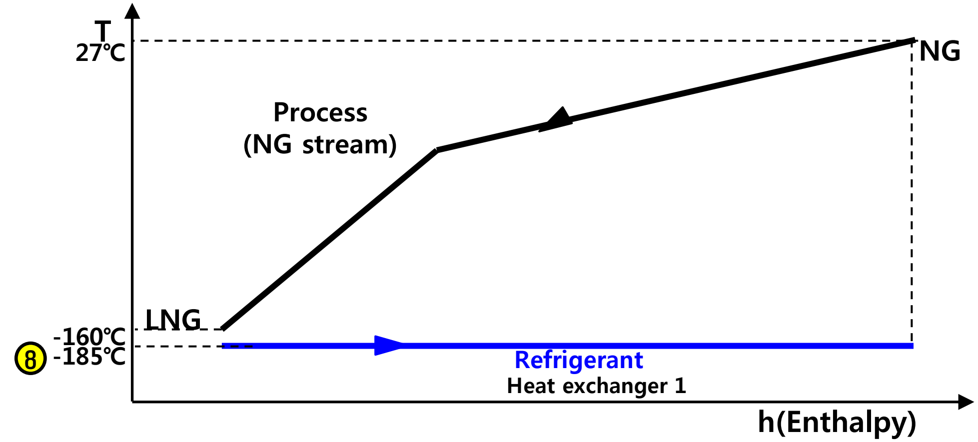


In the two phase region, the temperature is constant at the constant pressure.

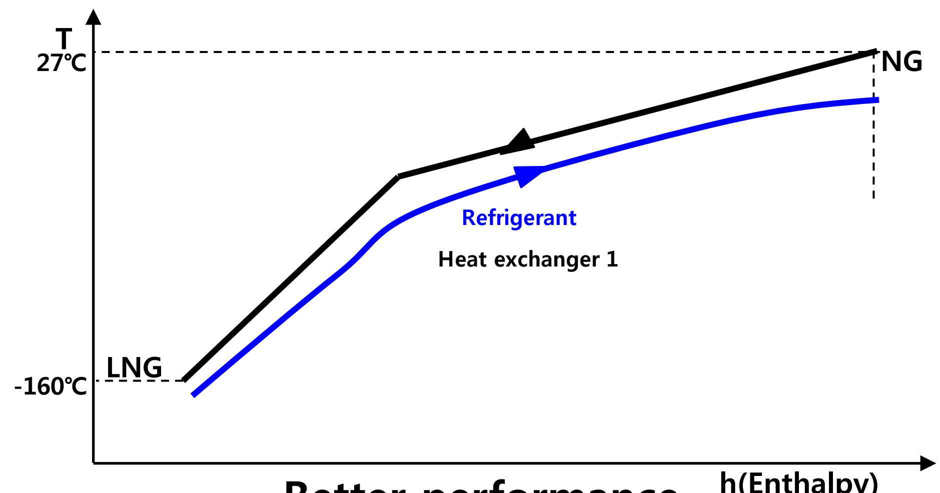
Advantage of mixed refrigerant



Singlestage Refrigeration
(pure refrigerant)



Singlestage Refrigeration
(mixed refrigerant)

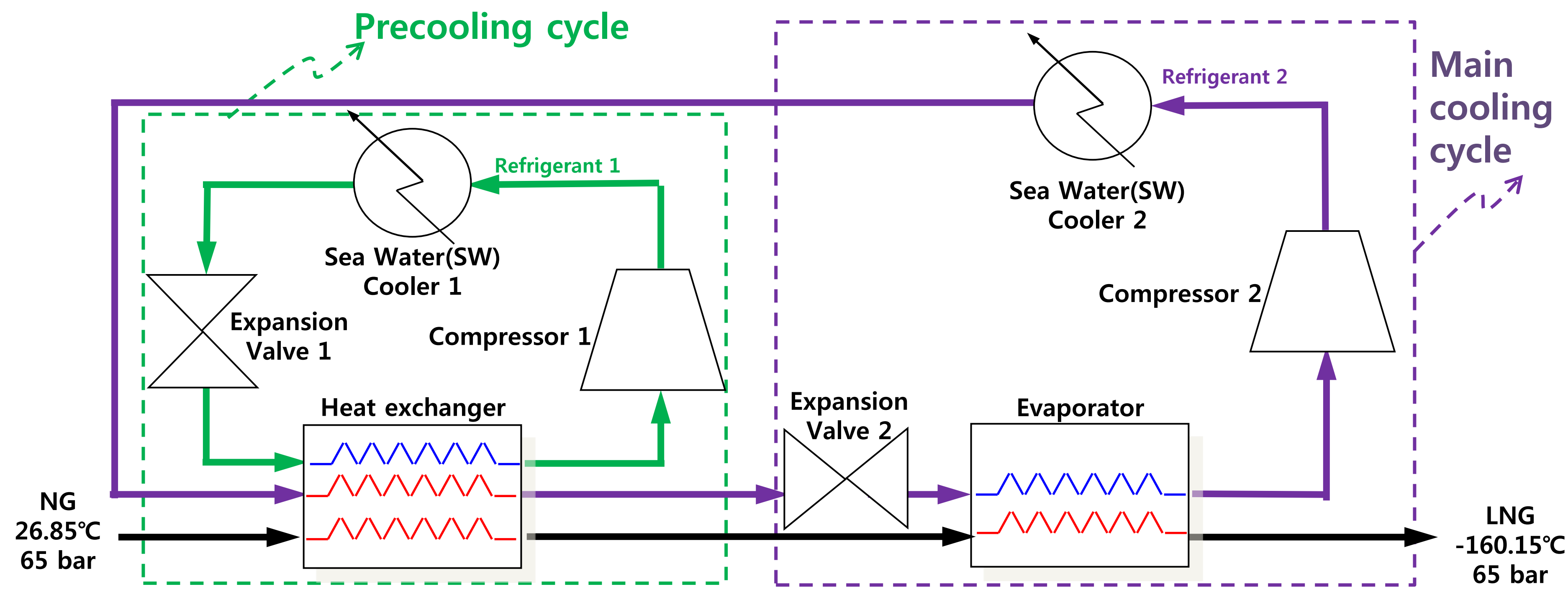


Better performance

- Advantage:
  Achieving better match between the temperature profiles of refrigerant and natural gas
  → The compressor work to be provided is reduced.

Nogal, F.D., "Optimal Design and Integration of Refrigeration and Power Systems", PhD thesis, Univ. of Manchester, p. 38

Dual Cycle for Liquefaction Process



1. Since the currently used C3MR on land is dual cycle, and DMR on offshore(LNG FPSO) is also dual cycle, This research proposes generic model regarding dual cycle.

2. This research used mixed refrigerant.

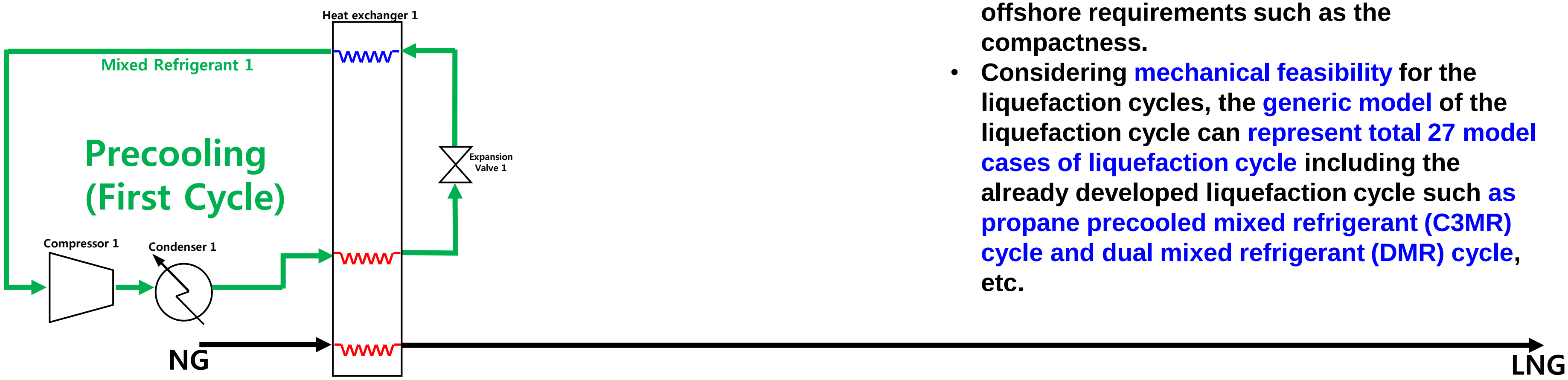
9.5. Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration

Single stage compression refrigeration

- Precooling 3 stage compression refrigeration
- Main cooling 3 stage compression, 3 stage refrigeration

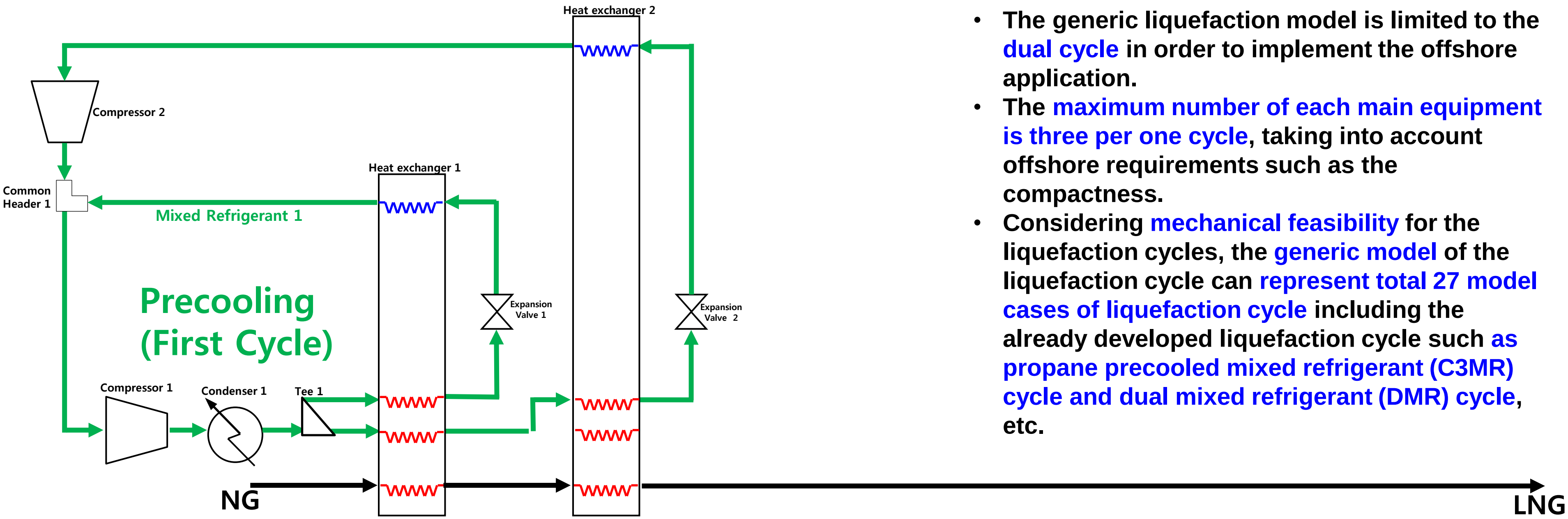


- The generic liquefaction model is limited to the dual cycle in order to implement the offshore application.
- The maximum number of each main equipment is three per one cycle, taking into account offshore requirements such as the compactness.
- Considering mechanical feasibility for the liquefaction cycles, the generic model of the liquefaction cycle can represent total 27 model cases of liquefaction cycle including the already developed liquefaction cycle such as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle, etc.

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration

2 stage compression refrigeration

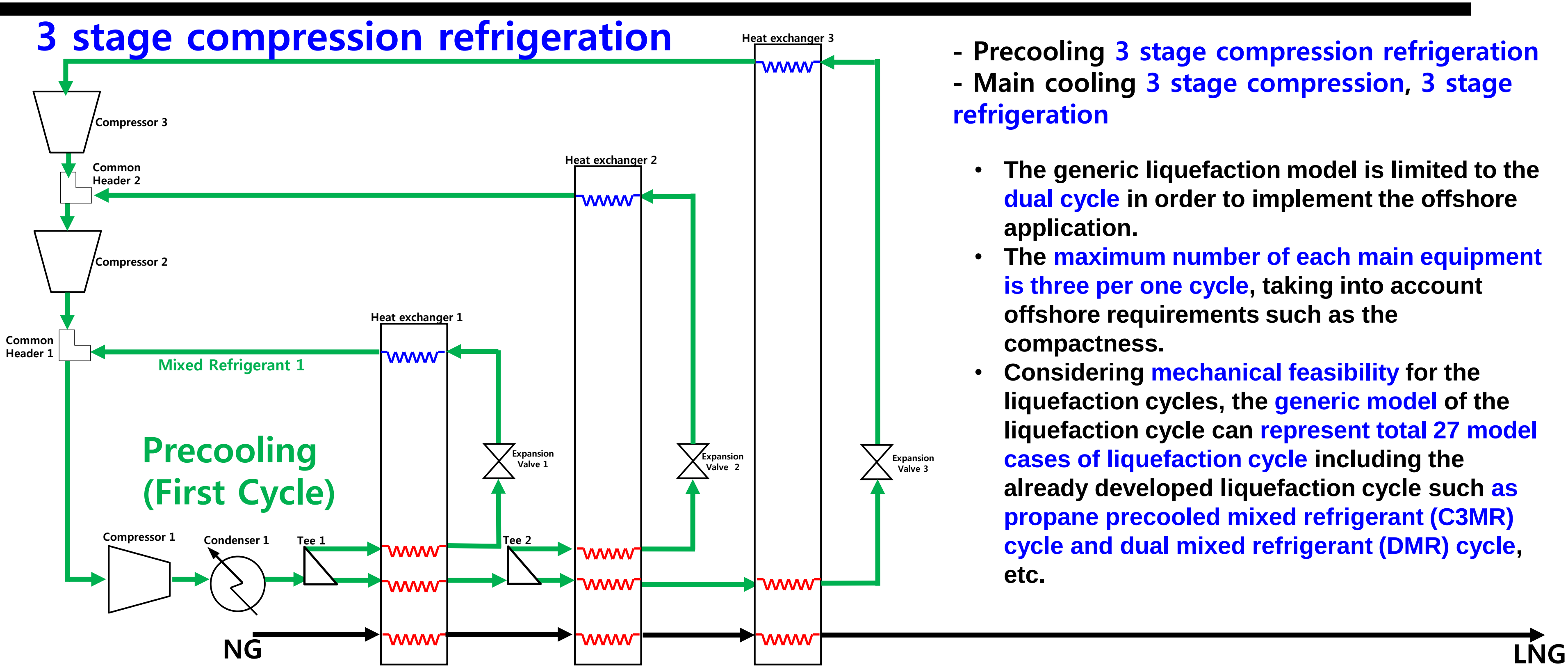


- Precooling 3 stage compression refrigeration
- Main cooling 3 stage compression, 3 stage refrigeration

- The generic liquefaction model is limited to the dual cycle in order to implement the offshore application.
- The maximum number of each main equipment is three per one cycle, taking into account offshore requirements such as the compactness.
- Considering mechanical feasibility for the liquefaction cycles, the generic model of the liquefaction cycle can represent total 27 model cases of liquefaction cycle including the already developed liquefaction cycle such as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle, etc.

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration



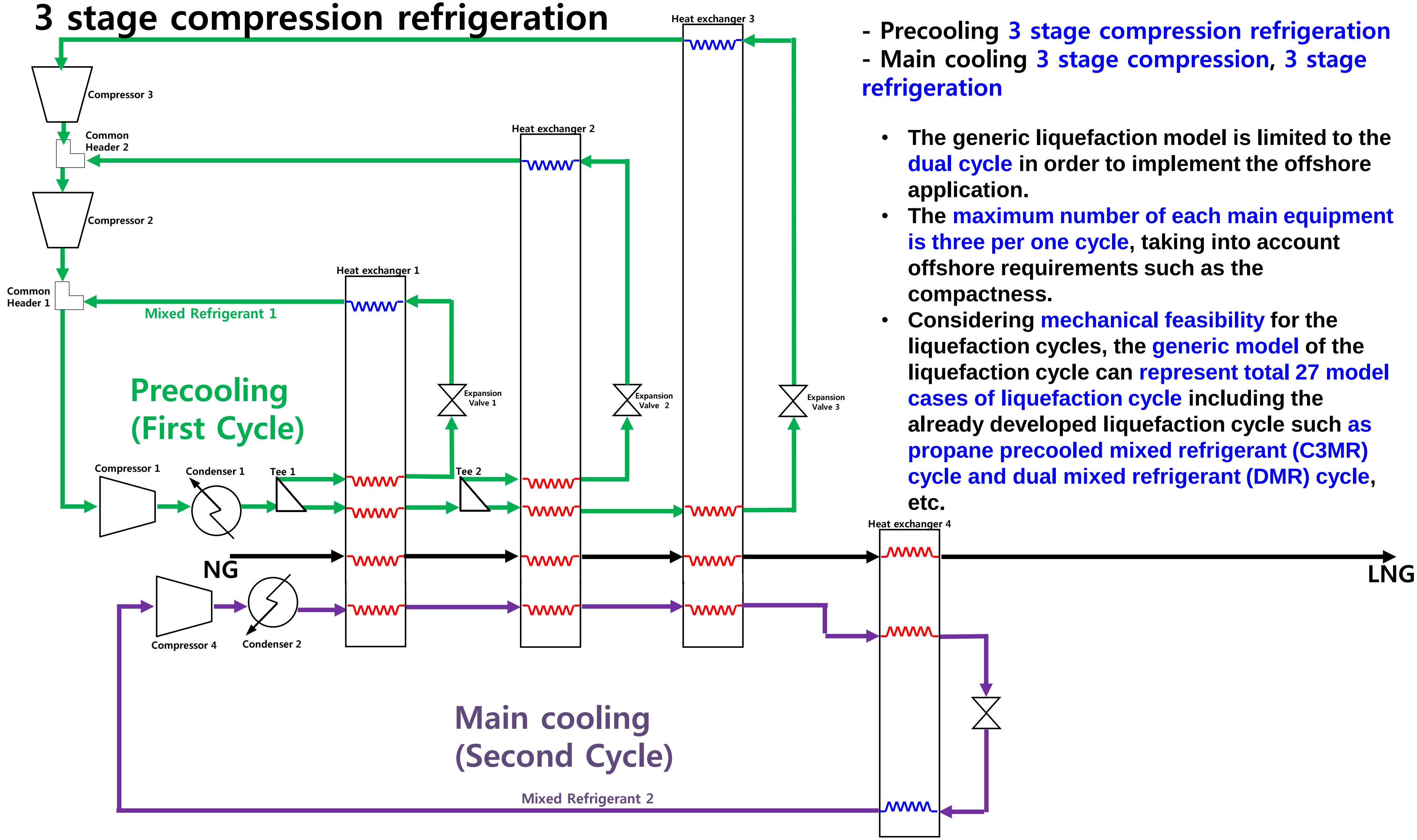
- Precooling 3 stage compression refrigeration
- Main cooling 3 stage compression, 3 stage refrigeration

- The generic liquefaction model is limited to the dual cycle in order to implement the offshore application.
- The maximum number of each main equipment is three per one cycle, taking into account offshore requirements such as the compactness.
- Considering mechanical feasibility for the liquefaction cycles, the generic model of the liquefaction cycle can represent total 27 model cases of liquefaction cycle including the already developed liquefaction cycle such as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle, etc.

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration

3 stage compression refrigeration



- Precooling 3 stage compression refrigeration
- Main cooling 3 stage compression, 3 stage refrigeration

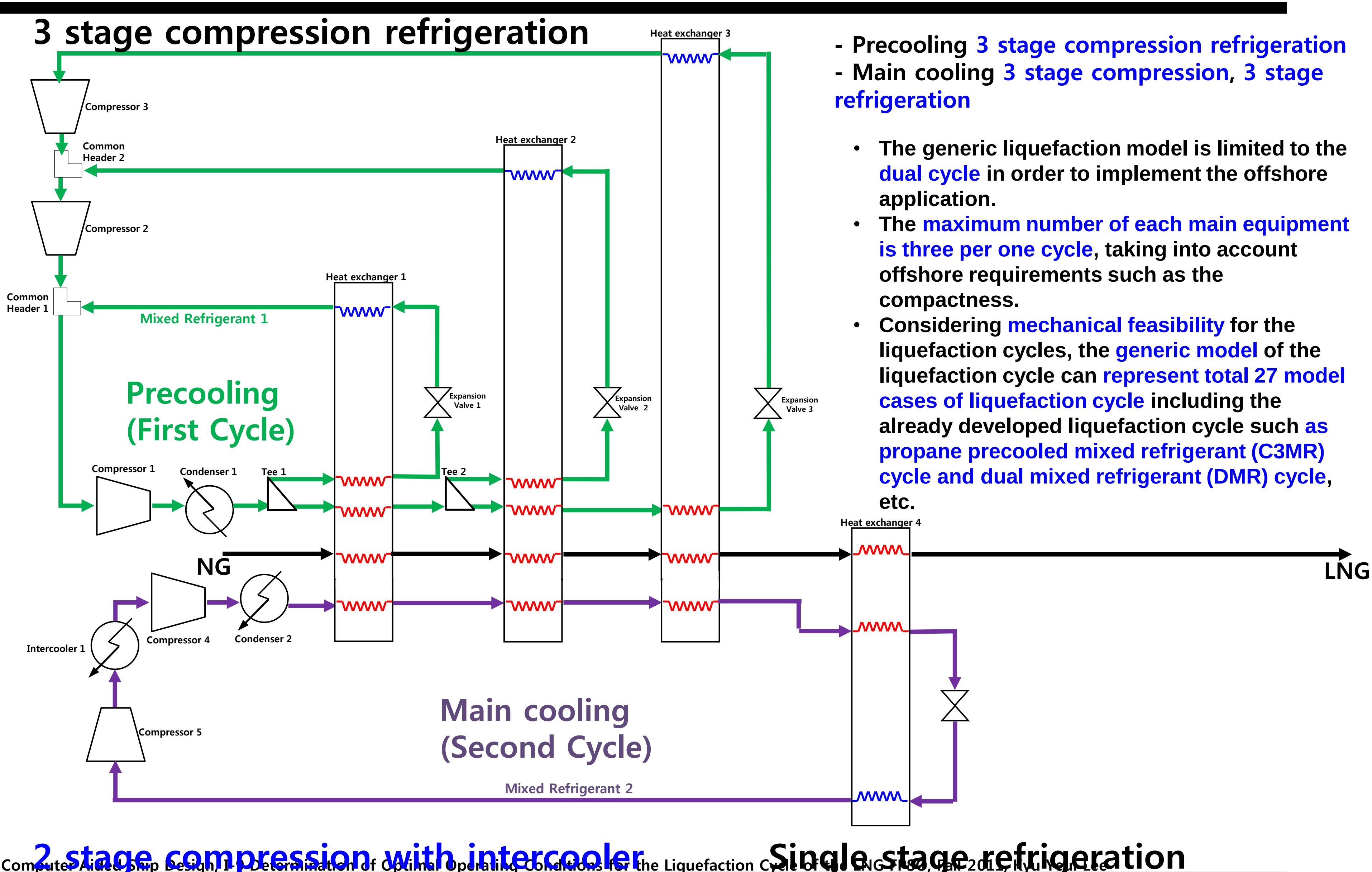
- The generic liquefaction model is limited to the dual cycle in order to implement the offshore application.
- The maximum number of each main equipment is three per one cycle, taking into account offshore requirements such as the compactness.
- Considering mechanical feasibility for the liquefaction cycles, the generic model of the liquefaction cycle can represent total 27 model cases of liquefaction cycle including the already developed liquefaction cycle such as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle, etc.

Single stage compression

Single stage refrigeration

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration

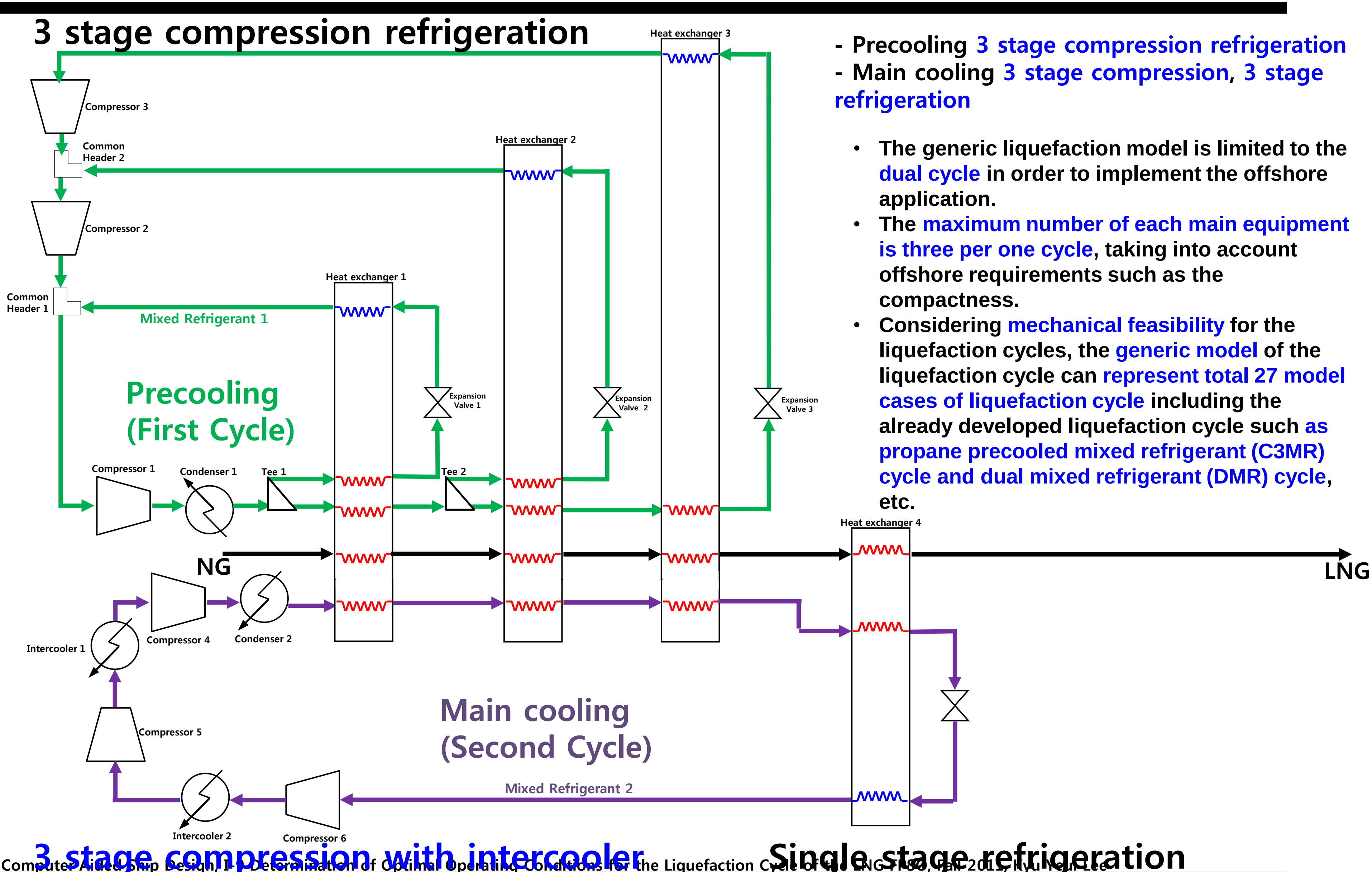


- Precooling 3 stage compression refrigeration
- Main cooling 3 stage compression, 3 stage refrigeration

- The generic liquefaction model is limited to the dual cycle in order to implement the offshore application.
- The maximum number of each main equipment is three per one cycle, taking into account offshore requirements such as the compactness.
- Considering mechanical feasibility for the liquefaction cycles, the generic model of the liquefaction cycle can represent total 27 model cases of liquefaction cycle including the already developed liquefaction cycle such as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle, etc.

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration



- Precooling **3 stage compression refrigeration**
- Main cooling **3 stage compression**, **3 stage refrigeration**

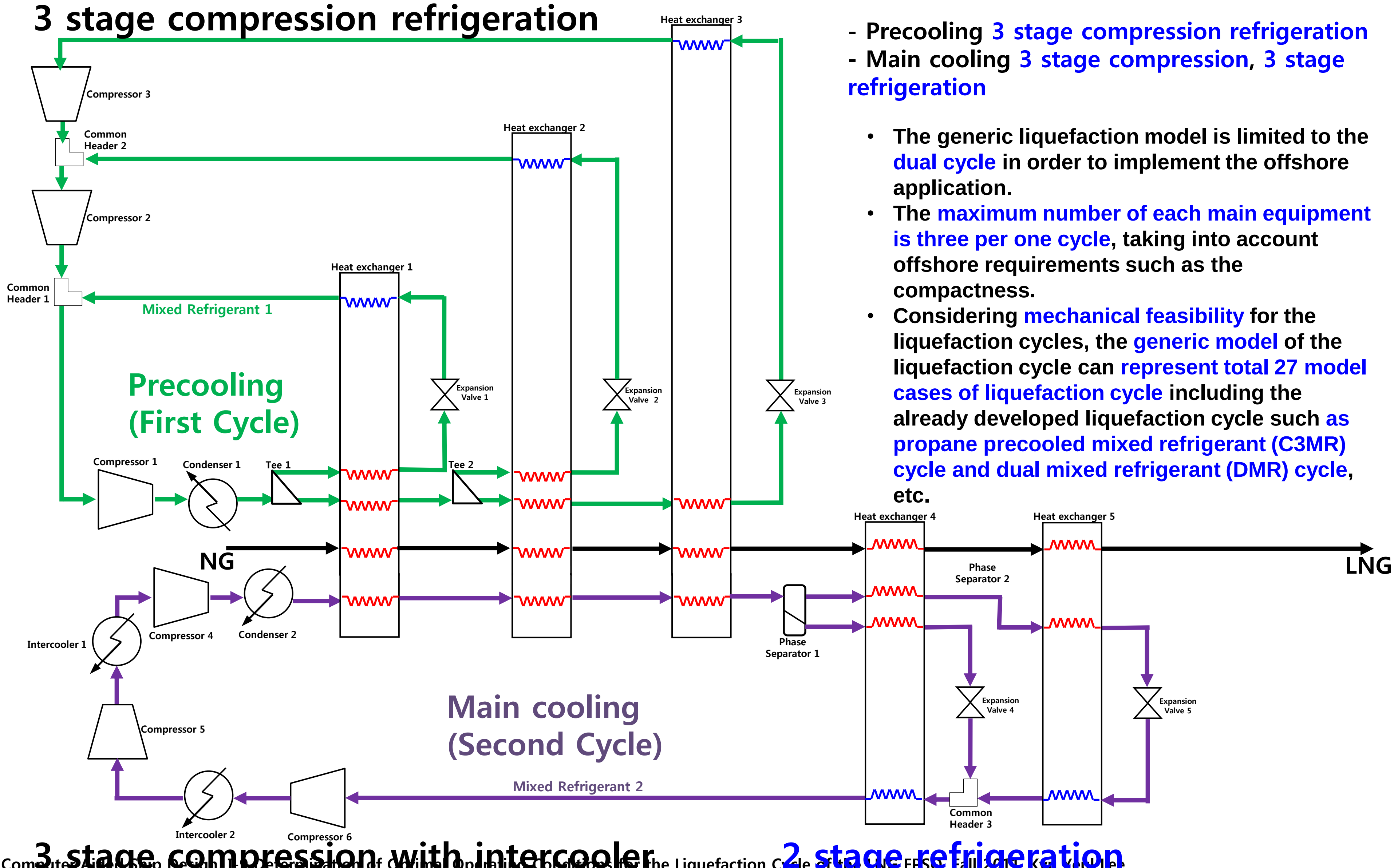
- The generic liquefaction model is limited to the **dual cycle** in order to implement the offshore application.
- The **maximum number of each main equipment is three per one cycle**, taking into account offshore requirements such as the compactness.
- Considering **mechanical feasibility** for the liquefaction cycles, the **generic model** of the liquefaction cycle can **represent total 27 model cases of liquefaction cycle** including the already developed liquefaction cycle such **as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle**, etc.

Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration

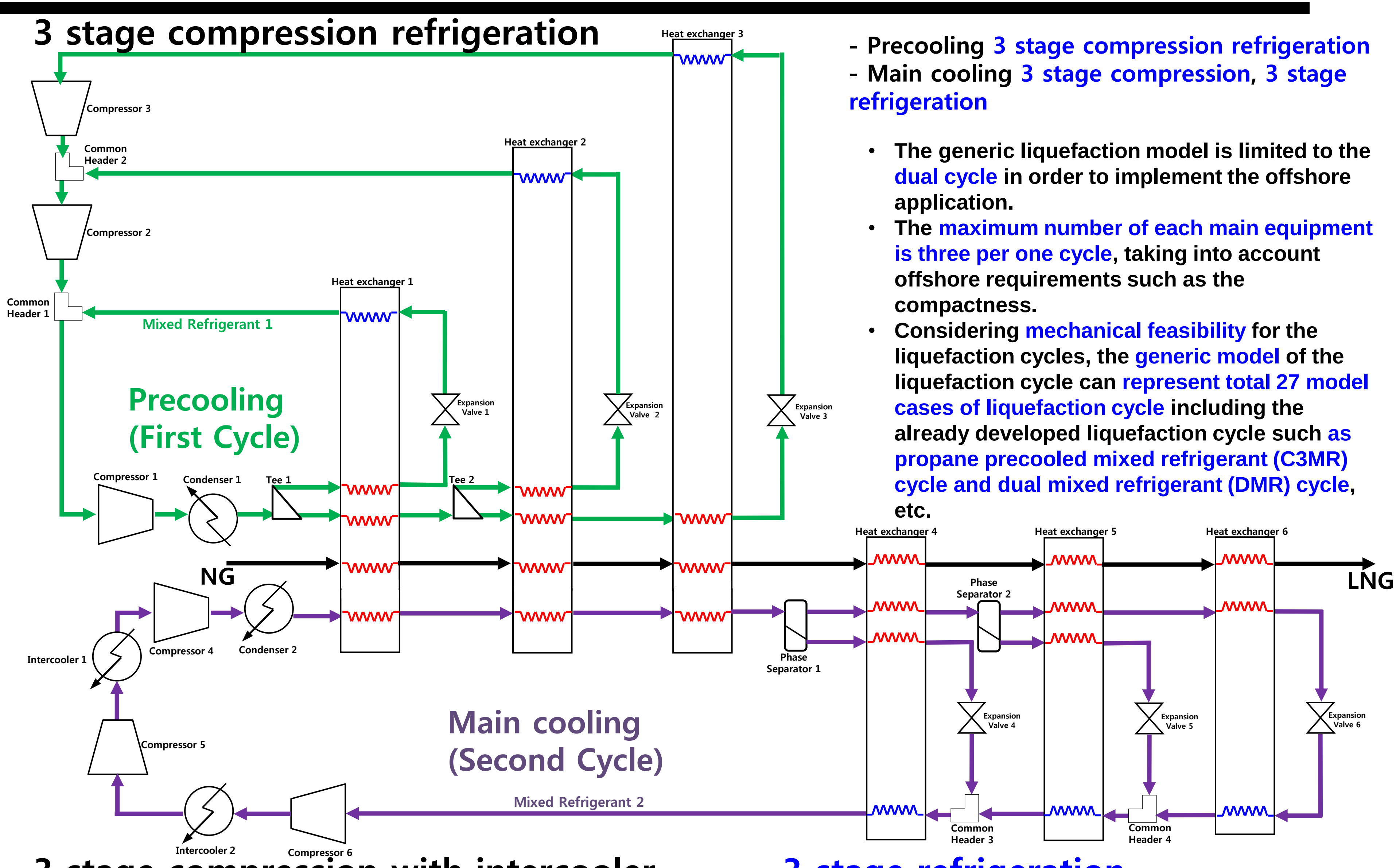
- Precooling **3 stage compression refrigeration**
- Main cooling **3 stage compression**, **3 stage refrigeration**

- The generic liquefaction model is limited to the **dual cycle** in order to implement the offshore application.
- The **maximum number of each main equipment is three per one cycle**, taking into account offshore requirements such as the compactness.
- Considering **mechanical feasibility** for the liquefaction cycles, the **generic model** of the liquefaction cycle can **represent total 27 model cases of liquefaction cycle** including the already developed liquefaction cycle such **as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle**, etc.



Proposed Generic Model of the Liquefaction Cycle of a LNG FPSO

:Dual Cycle with Regeneration + Multistage Compression with Intercooling + Multistage Compression Refrigeration + Multistage Refrigeration



- Precooling **3 stage compression refrigeration**
- Main cooling **3 stage compression**, **3 stage refrigeration**

- The generic liquefaction model is limited to the **dual cycle** in order to implement the offshore application.
- The **maximum number of each main equipment is three per one cycle**, taking into account offshore requirements such as the compactness.
- Considering **mechanical feasibility** for the liquefaction cycles, the **generic model** of the liquefaction cycle can **represent total 27 model cases of liquefaction cycle** including the already developed liquefaction cycle such **as propane precooled mixed refrigerant (C3MR) cycle and dual mixed refrigerant (DMR) cycle**, etc.

Mathematical Model of Generic Liquefaction Cycle (1)

1. Design variables(Operating Conditions) [187]

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,1_m,\ldots,26_m,1_{NG},\ldots,5_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

***T*: Temperature / *P*: Pressure / *v*: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant**

***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

2. Equality constraints [165]

2.1 Equality constraints of precooling part [83]

1) Compressor 1: [6]

$$h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)+w_1=h_{2p}\left(P_{2p},T_{2p},v_{2p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,2p}\left(P_{2p},T_{S,2p},v_{S,2p},z_{j,pre}\right)-h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)}{h_{2p}\left(P_{2p},T_{2p},v_{2p},z_{j,pre}\right)-h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)}$$

$$s_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)=s_{2p}\left(P_{2p},T_{S,2p},v_{S,2p},z_{j,pre}\right)$$

$$v_{1p}=v_{1p}\left(P_{1p},T_{1p},z_{j,pre}\right)$$

$$v_{S,2p}=v_{S,2p}\left(P_{2p},T_{S,2p},z_{j,pre}\right)$$

$$v_{2p}=v_{2p}\left(P_{2p},T_{2p},z_{j,pre}\right)$$

2) Condenser 1: [3]

The temperature of the outlet of the sea water cooler is usually given. T=310K

$$P_{2p}=P_{3p}$$

$$v_{3p}=v_{3p}(T_{3p},P_{3p},z_{j,pre})$$

3) Tee 1: [6]

$$h_{3p}\left(P_{3p},T_{3p},v_{3p},z_{j,pre}\right)$$
$$=c_1\cdot h_{4p}\left(P_{4p},T_{4p},v_{4p},z_{j,pre}\right)+\left(1-c_1\right)\cdot h_{8p}\left(P_{8p},T_{8p},v_{8p},z_{j,pre}\right)$$

$$P_{3p}=P_{4p},\ \ P_{3p}=P_{8p}$$

$$T_{4p}=T_{8p}$$

$$v_{4p}=v_{4p}\left(T_{4p},P_{4p},z_{j,pre}\right),\ v_{8p}=v_{8p}\left(T_{8p},P_{8p},z_{j,pre}\right)$$

4) Heat exchanger 1: [14]

$$c_1\cdot\dot{m}_{pre}\cdot h_{4p}\left(P_{4p},T_{4p},v_{4p},z_{j,pre}\right)+c_1\cdot\dot{m}_{pre}\cdot h_{6p}\left(P_{6p},T_{6p},v_{6p},z_{j,pre}\right)$$
$$+\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{8p}\left(P_{8p},T_{8p},v_{8p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{7m}\left(P_{7m},T_{7m},v_{7m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{NG}\left(P_{NG},T_{NG},v_{NG},z_{l,NG}\right)$$
$$=c_1\cdot\dot{m}_{pre}\cdot h_{5p}\left(P_{5p},T_{5p},v_{5p},z_{j,pre}\right)+c_1\cdot\dot{m}_{pre}\cdot h_{7p}\left(P_{7p},T_{7p},v_{7p},z_{j,pre}\right)$$
$$+\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{9p}\left(P_{9p},T_{9p},v_{9p},z_{j,pre}\right)+\dot{m}_{main}\cdot h_{8m}\left(P_{8m},T_{8m},v_{8m},z_{k,main}\right)$$
$$+\dot{m}_{NG}\cdot h_{1NG}\left(P_{1NG},T_{1NG},v_{1NG},z_{l,NG}\right)$$

$$P_{4p}=P_{5p},\ P_{6p}=P_{7p},P_{8p}=P_{9p},P_{7m}=P_{8m},P_{NG}=P_{1NG}$$

$$T_{5p}=T_{9p},T_{5p}=T_{8m},T_{5p}=T_{1NG}$$

$$v_{5p}=v_{5p}\left(T_{5p},P_{5p},z_{j,pre}\right),\ v_{7p}=v_{7p}\left(T_{7p},P_{7p},z_{j,pre}\right),$$

$$v_{9p}=v_{9p}\left(T_{9p},P_{9p},z_{j,pre}\right),\ v_{8m}=v_{8m}\left(T_{8m},P_{8m},z_{k,main}\right),$$

$$v_{1NG}=v_{1NG}\left(T_{1NG},P_{1NG},z_{l,NG}\right)$$

5) Expansion valve 1: [2]

$$h_{5p}\left(P_{5p},T_{5p},v_{5p},z_{j,pre}\right)=h_{6p}\left(P_{6p},T_{6p},v_{6p},z_{j,pre}\right)$$

$$v_{6p}=v_{6p}\left(T_{6p},P_{6p},z_{j,pre}\right)$$

6) Tee 2: [6]

$$\left(1-c_1\right)\cdot h_{9p}\left(P_{9p},T_{9p},v_{9p},z_{j,pre}\right)$$
$$=c_2\cdot\left(1-c_1\right)\cdot h_{10p}\left(P_{10p},T_{10p},v_{10p},z_{j,pre}\right)$$
$$+\left(1-c_2\right)\cdot\left(1-c_1\right)\cdot h_{14p}\left(P_{14p},T_{14p},v_{14p},z_{j,pre}\right)$$

$$P_{9p}=P_{10p},\ \ P_{9p}=P_{14p}$$

$$T_{10p}=T_{14p}$$

$$v_{10p}=v_{10p}\left(T_{10p},P_{10p},z_{j,pre}\right),$$

$$v_{14p}=v_{14p}\left(T_{14p},P_{14p},z_{j,pre}\right)$$

Mathematical Model of Generic Liquefaction Cycle (2)

1. Design variables(Operating Conditions) [187]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:
$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,1_m,\ldots,26_m,1_{NG},\ldots,5_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [165]

2.1 Equality constraints of precooling part [83]

7) Heat exchanger 2: [14]

$$c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{10p}\left(P_{10p},T_{10p},v_{10p},z_{j,pre}\right)+c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{12p}\left(P_{12p},T_{12p},v_{12p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{14p}\left(P_{14p},T_{14p},v_{14p},z_{j,pre}\right)+$$
$$+\dot{m}_{main}\cdot h_{8m}\left(P_{8m},T_{8m},v_{8m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{1NG}\left(P_{1NG},T_{1NG},v_{1NG},z_{l,NG}\right)$$
$$=c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{11p}\left(P_{11p},T_{11p},v_{11p},z_{j,pre}\right)+c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{13p}\left(P_{13p},T_{13p},v_{13p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{15p}\left(P_{15p},T_{15p},v_{15p},z_{j,pre}\right)+$$
$$+\dot{m}_{main}\cdot h_{9m}\left(P_{9m},T_{9m},v_{9m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{2NG}\left(P_{2NG},T_{2NG},v_{2NG},z_{l,NG}\right)$$

$$P_{10p}=P_{11p},\ P_{12p}=P_{13p},P_{14p}=P_{15p},P_{8m}=P_{9m},P_{1NG}=P_{2NG}$$

$$T_{11p}=T_{15p},T_{11p}=T_{9m},T_{11p}=T_{2NG}$$

$$v_{11p}=v_{11p}\left(T_{11p},P_{11p},z_{j,pre}\right),\ v_{13p}=v_{13p}\left(T_{13p},P_{13p},z_{j,pre}\right),$$
$$v_{15p}=v_{15p}\left(T_{15p},P_{15p},z_{j,pre}\right),\ v_{9m}=v_{9m}\left(T_{9m},P_{9m},z_{k,main}\right),$$
$$v_{2NG}=v_{2NG}\left(T_{2NG},P_{2NG},z_{l,NG}\right)$$

8) Expansion valve 2: [2]

$$h_{11p}\left(P_{11p},T_{11p},v_{11p},z_{j,pre}\right)=h_{12p}\left(P_{12p},T_{12p},v_{12p},z_{j,pre}\right)$$
$$v_{12p}=v_{12p}\left(T_{12p},P_{12p},z_{j,pre}\right)$$

9) Heat exchanger 3: [11]

$$(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{15p}\left(P_{15p},T_{15p},v_{15p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{17p}\left(P_{17p},T_{17p},v_{17p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{9m}\left(P_{9m},T_{9m},v_{9m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{2NG}\left(P_{2NG},T_{2NG},v_{2NG},z_{l,NG}\right)$$
$$=(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{16p}\left(P_{16p},T_{16p},v_{16p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{10m}\left(P_{10m},T_{10m},v_{10m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{3NG}\left(P_{3NG},T_{3NG},v_{3NG},z_{l,NG}\right)$$

$$P_{15p}=P_{16p},P_{17p}=P_{18p},P_{9m}=P_{10m},P_{2NG}=P_{3NG}$$

$$T_{16p}=T_{10m},T_{16p}=T_{3NG}$$

$$v_{16p}=v_{16p}\left(T_{16p},P_{16p},z_{j,pre}\right),\ v_{18p}=v_{18p}\left(T_{18p},P_{18p},z_{j,pre}\right),$$
$$v_{10m}=v_{10m}\left(T_{10m},P_{10m},z_{k,main}\right),\ v_{3NG}=v_{3NG}\left(T_{3NG},P_{3NG},z_{l,NG}\right)$$

10) Expansion valve 3: [2]

$$h_{16p}\left(P_{16p},T_{16p},v_{16p},z_{j,pre}\right)=h_{17p}\left(P_{17p},T_{17p},v_{17p},z_{j,pre}\right)$$
$$v_{17p}=v_{17p}\left(T_{17p},P_{17p},z_{j,pre}\right)$$

Mathematical Model of Generic Liquefaction Cycle (3)

1. Design variables(Operating Conditions) [187]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\dots,21_p,1_m,\dots,26_m,1_{NG},\dots,5_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [165]

2.1 Equality constraints of precooling part [83]

11) Compressor 3: [5]

$$\left(1-c_2\right)\cdot\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)+w_3$$
$$=\left(1-c_2\right)\cdot\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,19p}\left(P_{19p},T_{S,19p},v_{S,19p},z_{j,pre}\right)-h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)}{h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right)-h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)}$$

$$s_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)=s_{19p}\left(P_{19p},T_{S,19p},v_{S,19p},z_{j,pre}\right)$$

$$v_{19p}=v_{19p}\left(P_{19p},T_{19p},z_{j,pre}\right)$$

$$v_{S,19p}=v_{S,19p}\left(P_{19p},T_{S,19p},z_{j,pre}\right)$$

12) Common Header 2: [4]

$$c_2\cdot\left(1-c_1\right)\cdot h_{13p}\left(P_{13p},T_{13p},v_{13p},z_{j,pre}\right)+\left(1-c_2\right)\cdot\left(1-c_1\right)\cdot h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right)$$
$$=\left(1-c_1\right)\cdot h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)$$

$$P_{13p}=P_{19p},\ P_{13p}=P_{20p}$$

$$v_{20p}=v_{20p}\left(P_{20p},T_{20p},z_{j,pre}\right)$$

13) Compressor 2: [5]

$$\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)+w_2$$
$$=\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,21p}\left(P_{21p},T_{S,21p},v_{S,21p},z_{j,pre}\right)-h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)}{h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)-h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)}$$

$$s_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)=s_{21p}\left(P_{21p},T_{S,21p},v_{S,21p},z_{j,pre}\right)$$

$$v_{21p}=v_{21p}\left(P_{21p},T_{21p},z_{j,pre}\right)$$

$$v_{S,21p}=v_{S,21p}\left(P_{21p},T_{S,21p},z_{j,pre}\right)$$

14) Common Header 1: [3]

$$\left(1-c_1\right)\cdot h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)+c_1\cdot h_{7p}\left(P_{7p},T_{7p},v_{7p},z_{j,pre}\right)$$
$$=h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)$$

$$P_{7p}=P_{21p},\ P_{7p}=P_{1p}$$

Mathematical Model of Generic Liquefaction Cycle (4)

1. Design variables(Operating Conditions) [187]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:
$$P_i,\ T_i,\ v_i\left(i=1_p,\dots,21_p,1_m,\dots,26_m,1_{NG},\dots,5_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [165]

2.2 Equality constraints of main cooling part [80]

1) Compressor 6: [6]

$$h_{1m}\left(P_{1m},T_{1m},v_{1m},z_{k,main}\right)+w_6=h_{2m}\left(P_{2m},T_{2m},v_{2m},z_{k,main}\right)$$
$$\eta=\frac{h_{S,2m}\left(P_{2m},T_{S,2m},v_{S,2m},z_{k,main}\right)-h_{1m}\left(P_{1m},T_{1m},v_{1m},z_{k,main}\right)}{h_{2m}\left(P_{2m},T_{2m},v_{2m},z_{k,main}\right)-h_{1m}\left(P_{1m},T_{1m},v_{1m},z_{k,main}\right)}$$
$$s_{1m}\left(P_{1m},T_{1m},v_{1m},z_{k,main}\right)=s_{2m}\left(P_{2m},T_{S,2m},v_{S,2m},z_{k,main}\right)$$
$$v_{1m}=v_{1m}\left(P_{1m},T_{1m},z_{k,main}\right)$$
$$v_{S,2m}=v_{S,2m}\left(P_{2m},T_{S,2m},z_{k,main}\right)$$
$$v_{2m}=v_{2m}\left(P_{2m},T_{2m},z_{k,main}\right)$$

2) Intercooler 2: [3]

The temperature of the outlet of the sea water cooler is usually given.
T=305K

$$P_{2m}=P_{3m}$$
$$v_{3m}=v_{3m}(T_{3m},P_{3m},z_{k,main})$$

3) Compressor 5: [5]

$$h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)+w_5=h_{4m}\left(P_{4m},T_{4m},v_{4m},z_{k,main}\right)$$
$$\eta=\frac{h_{S,4m}\left(P_{4m},T_{S,4m},v_{S,4m},z_{k,main}\right)-h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)}{h_{4m}\left(P_{4m},T_{4m},v_{4m},z_{k,main}\right)-h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)}$$
$$s_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)=s_{4m}\left(P_{4m},T_{S,4m},v_{S,4m},z_{k,main}\right)$$
$$v_{S,4m}=v_{S,4m}\left(P_{4m},T_{S,4m},z_{k,main}\right)$$
$$v_{4m}=v_{4m}\left(P_{4m},T_{4m},z_{k,main}\right)$$

4) Intercooler 1: [3]

The temperature of the outlet of the sea water cooler is usually given.
T=305K

$$P_{4m}=P_{5m}$$
$$v_{5m}=v_{5m}(T_{5m},P_{5m},z_{k,main})$$

5) Compressor 4: [5]

$$h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)+w_4=h_{6m}\left(P_{6m},T_{6m},v_{6m},z_{k,main}\right)$$
$$\eta=\frac{h_{S,6m}\left(P_{6m},T_{S,6m},v_{S,6m},z_{k,main}\right)-h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)}{h_{6m}\left(P_{6m},T_{6m},v_{6m},z_{k,main}\right)-h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)}$$
$$s_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)=s_{6m}\left(P_{6m},T_{S,6m},v_{S,6m},z_{k,main}\right)$$
$$v_{S,6m}=v_{S,6m}\left(P_{6m},T_{S,6m},z_{k,main}\right)$$
$$v_{6m}=v_{6m}\left(P_{6m},T_{6m},z_{k,main}\right)$$

6) Condenser 2: [3]

The temperature of the outlet of the sea water cooler is usually given.
T=305K

$$P_{6m}=P_{7m}$$
$$v_{7m}=v_{7m}(T_{7m},P_{7m},z_{k,main})$$

Mathematical Model of Generic Liquefaction Cycle (5)

1. Design variables(Operating Conditions) [187]

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,1_m,\ldots,26_m,1_{NG},\ldots,5_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

***T*: Temperature / *P*: Pressure / *v*: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant**

***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

2. Equality constraints [165]

2.2 Equality constraints of main cooling part [80]

7) Phase Separator 1: [7]

$$h_{10m}\left(P_{10m},T_{10m},v_{10m},z_{k,main}\right)$$
$$=v\_f_{10}\cdot h_{14m}\left(P_{14m},T_{14m},v_{14m},v\_f_{10}\cdot z_{k,main}\right)$$
$$+\left(1-v\_f_{10}\right)\cdot h_{11m}\left(P_{11m},T_{11m},v_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)$$

$$P_{10m}=P_{11m},\ P_{10m}=P_{14m}$$

$$T_{10m}=T_{11m},\ T_{11m}=T_{14m}$$

$$v_{11m}=v_{11m}\left(P_{11m},T_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right),\ v_{14m}=v_{14m}\left(P_{14m},T_{14m},v\_f_{10}\cdot z_{k,main}\right)$$

8) Heat exchanger 4: [10]

$$\left(1-v\_f_{10}\right)\cdot\dot{m}_{main}\cdot h_{11m}\left(P_{11m},T_{11m},v_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)$$
$$+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{14m}\left(P_{14m},T_{14m},v_{14m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{main}\cdot h_{26m}\left(P_{26m},T_{26m},v_{26m},z_{k,main}\right)$$
$$+\dot{m}_{NG}\cdot h_{3NG}\left(P_{3NG},T_{3NG},v_{3NG},z_{l,NG}\right)$$
$$=\left(1-v\_f_{10}\right)\cdot\dot{m}_{main}\cdot h_{12m}\left(P_{12m},T_{12m},v_{12m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)$$
$$+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{15m}\left(P_{15m},T_{15m},v_{15m},v\_f_{10}\cdot z_{k,main}\right)+h_{1m}\left(P_{1m},T_{1m},v_{1m},z_{k,main}\right)$$
$$+\dot{m}_{NG}\cdot h_{4NG}\left(P_{4NG},T_{4NG},v_{4NG},z_{l,NG}\right)$$

$$P_{11m}=P_{12m},P_{14m}=P_{15m},P_{26m}=P_{1m},P_{3NG}=P_{4NG}$$

$$T_{12m}=T_{15m},T_{12m}=T_{4NG}$$

$$v_{12m}=v_{12m}\left(T_{12m},P_{12m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right),\ v_{15m}=v_{15m}\left(T_{15m},P_{15m},v\_f_{10}\cdot z_{k,main}\right),$$

$$v_{4NG}=v_{4NG}\left(T_{4NG},P_{4NG},z_{l,NG}\right)$$

9) Phase Separator 2: [7]

$$v\_f_{10}\cdot h_{15m}\left(P_{15m},T_{15m},v_{15m},v\_f_{10}\cdot z_{k,main}\right)$$
$$=v\_f_{15}\cdot v\_f_{10}\cdot h_{19m}\left(P_{19m},T_{19m},v_{19m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$+\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot h_{16m}\left(P_{16m},T_{16m},v_{16m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)$$

$$P_{15m}=P_{16m},\ P_{15m}=P_{19m}$$

$$T_{15m}=T_{16m},\ T_{16m}=T_{19m}$$

$$v_{16m}=v_{16m}\left(P_{16m},T_{16m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right),\ v_{19m}=v_{19m}\left(P_{19m},T_{19m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$

10) Heat exchanger 5: [11]

$$\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{16m}\left(P_{16m},T_{16m},v_{16m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$+v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{19m}\left(P_{19m},T_{19m},v_{19m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{24m}\left(P_{24m},T_{24m},v_{24m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{NG}\cdot h_{4NG}\left(P_{4NG},T_{4NG},v_{4NG},z_{l,NG}\right)$$
$$=\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{17m}\left(P_{17m},T_{17m},v_{17m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$+v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{20m}\left(P_{20m},T_{20m},v_{20m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{25m}\left(P_{25m},T_{25m},v_{25m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{NG}\cdot h_{5NG}\left(P_{5NG},T_{5NG},v_{5NG},z_{l,NG}\right)$$

$$P_{16m}=P_{17m},P_{19m}=P_{20m},P_{24m}=P_{25m},P_{4NG}=P_{5NG}$$

$$T_{17m}=T_{20m},T_{17m}=T_{5NG}$$

$$v_{17m}=v_{17m}\left(P_{17m},T_{17m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right),\ v_{20m}=v_{20m}\left(P_{20m},T_{20m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right),$$

$$v_{25m}=v_{25m}\left(P_{25m},T_{25m},v\_f_{10}\cdot z_{k,main}\right),\ v_{5NG}=v_{5NG}\left(T_{5NG},P_{5NG},z_{l,NG}\right)$$

Mathematical Model of Generic Liquefaction Cycle (6)

1. Design variables(Operating Conditions) [187]

T: Temperature / *P*: Pressure / *v*: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
*Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant

:
$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,1_m,\ldots,26_m,1_{NG},\ldots,5_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [165]

2.2 Equality constraints of main cooling part [80]

11) Heat exchanger 6: [6]

$$\begin{aligned}&v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{20m}\left(P_{20m},T_{20m},v_{20m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)\\&+v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{22m}\left(P_{22m},T_{22m},v_{22m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)\\&+\dot{m}_{NG}\cdot h_{5NG}\left(P_{5NG},T_{5NG},v_{5NG},z_{l,NG}\right)\\&=v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{21m}\left(P_{21m},T_{21m},v_{21m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)\\&\quad+v\_f_{15}\cdot v\_f_{10}\cdot\dot{m}_{main}\cdot h_{23m}\left(P_{23m},T_{23m},v_{23m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)\\&\quad+\dot{m}_{NG}\cdot h_{LNG}\left(P_{LNG},T_{LNG},v_{LNG},z_{l,NG}\right)\end{aligned}$$

$$P_{20m}=P_{21m},P_{22m}=P_{23m}$$

$$T_{21m}=T_{LNG}$$

$$v_{21m}=v_{21m}\left(P_{21m},T_{21m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right),$$

$$v_{23m}=v_{23m}\left(P_{23m},T_{23m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$

12) Expansion valve 4: [2]

$$\begin{aligned}&h_{12m}\left(P_{12m},T_{12m},v_{12m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)\\&\quad=h_{13m}\left(P_{13m},T_{13m},v_{13m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)\end{aligned}$$

$$v_{13m}=v_{13m}\left(P_{13m},T_{13m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)$$

13) Expansion valve 5: [2]

$$h_{17m}\left(P_{17m},T_{17m},v_{17m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)=h_{18m}\left(P_{18m},T_{18m},v_{18m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)$$

$$v_{18m}=v_{18m}\left(P_{18m},T_{18m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)$$

14) Expansion valve 6: [2]

$$h_{21m}\left(P_{21m},T_{21m},v_{21m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)=h_{22m}\left(P_{22m},T_{22m},v_{22m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$
$$v_{22m}=v_{22m}\left(P_{22m},T_{22m}v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)$$

15) Common Header 3: [4]

$$\begin{aligned}&\left(1-v\_f_{10}\right)\cdot h_{13m}\left(P_{13m},T_{13m},v_{13m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)+v\_f_{10}\cdot h_{25m}\left(P_{25m},T_{25m},v_{25m},v\_f_{10}\cdot z_{k,main}\right)\\&\quad=h_{26m}\left(P_{26m},T_{26m},v_{26m},z_{k,main}\right)\end{aligned}$$

$$P_{13m}=P_{25m},\ P_{13m}=P_{26m}\ \ v_{26m}=v_{26m}\left(P_{26m},T_{26m},z_{k,main}\right)$$

16) Common Header 4: [4]

$$\begin{aligned}&\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot h_{18m}\left(P_{18m},T_{18m},v_{18m},\left(1-v\_f_{15}\right)\cdot v\_f_{10}\cdot z_{k,main}\right)\\&+v\_f_{15}\cdot v\_f_{10}\cdot h_{23m}\left(P_{23m},T_{23m},v_{23m},v\_f_{15}\cdot v\_f_{10}\cdot z_{k,main}\right)\\&\quad=v\_f_{10}\cdot h_{24m}\left(P_{24m},T_{24m},v_{24m},v\_f_{10}\cdot z_{k,main}\right)\end{aligned}$$

$$P_{18m}=P_{23m},\ P_{18m}=P_{24m}\quad v_{24m}=v_{24m}\left(T_{24m},P_{24m},v\_f_{10}\cdot z_{k,main}\right),$$

$$\sum_{j=1}^{3}z_{j,pre}=1,\quad\sum_{k=1}^{4}z_{k,main}=1$$

Summary of the Mathematical Model of Generic Liquefaction Cycle

1. Design variables(Operating Conditions) [187]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,1_m,\ldots,26_m,1_{NG},\ldots,5_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,2m},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,2m},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,\ w_6,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},v\_f_{15},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [165]

2.1 Equality constraints of precooling part [83]

2.2 Equality constraints of main cooling part [80]

→ indeterminate systems

3. Objective Function: Minimize the compressors power

$$\text{Minize}\ \ \dot{m}_{pre}\cdot w_1+\dot{m}_{pre}\cdot w_2+\dot{m}_{pre}\cdot w_3+\dot{m}_{main}\cdot w_4+\dot{m}_{main}\cdot w_5+\dot{m}_{main}\cdot w_6$$

→ Optimization Problem!

4. Free variables [22 = 187 – 165]

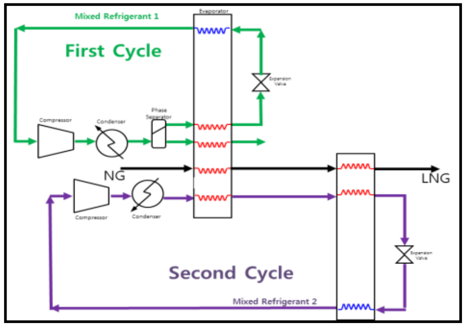
:

$$P_{1p},\ P_{2p},\ P_{12p},\ P_{17p},\ T_{5p},\ T_{11p},\ T_{16p},\ c_1,\ c_2,\ z_{1,pre},\ z_{2,pre},\ \dot{m}_{pre},$$

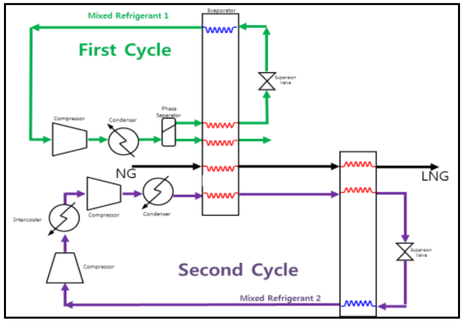
$$P_{1m},\ P_{2m},\ P_{4m},\ P_{6m},\ T_{12m},\ T_{17m},\ z_{1,main},\ z_{2,main},\ z_{3,main},\ \dot{m}_{main}$$

Feasible Liquefaction Cycle from the Generic Model (Case 1 ~ Case 9)

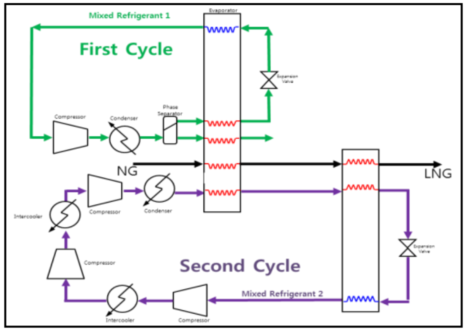
FEASIBLE LIQUEFACTION MODEL (CASE 1)



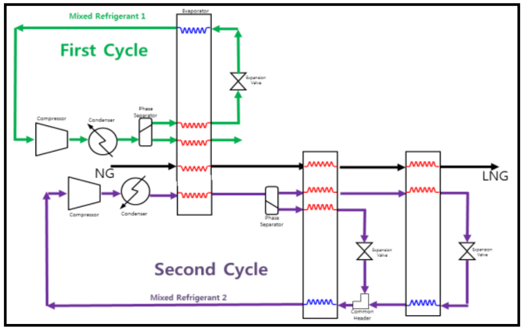
FEASIBLE LIQUEFACTION MODEL (CASE 2)



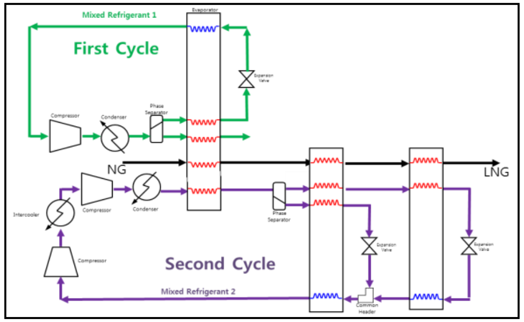
FEASIBLE LIQUEFACTION MODEL (CASE 3)



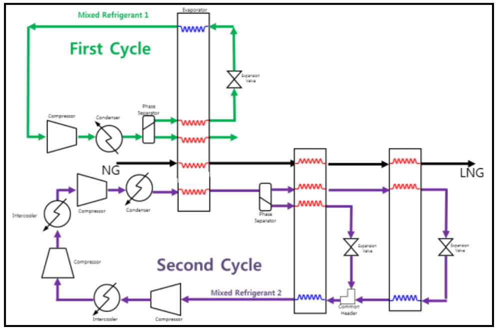
FEASIBLE LIQUEFACTION MODEL (CASE 4)



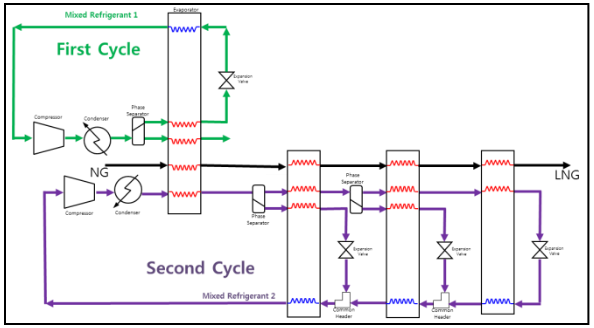
FEASIBLE LIQUEFACTION MODEL (CASE 5)



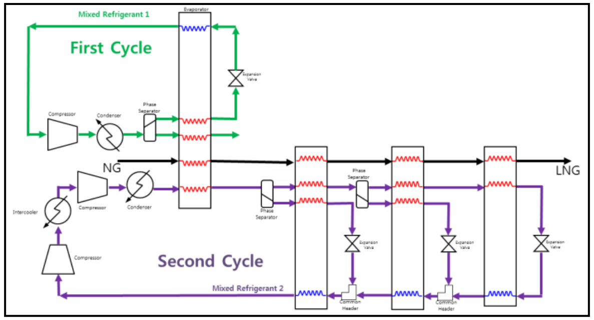
FEASIBLE LIQUEFACTION MODEL (CASE 6)



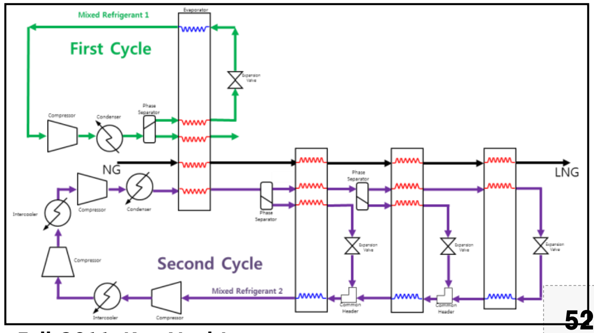
FEASIBLE LIQUEFACTION MODEL (CASE 7)



FEASIBLE LIQUEFACTION MODEL (CASE 8)

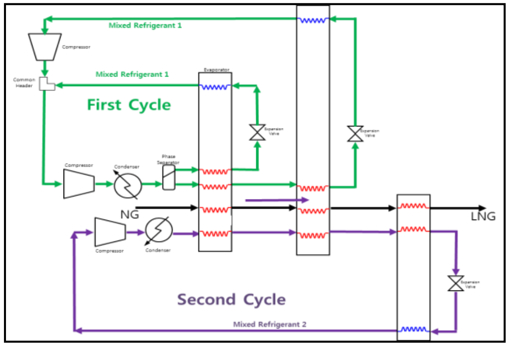


FEASIBLE LIQUEFACTION MODEL (CASE 9)

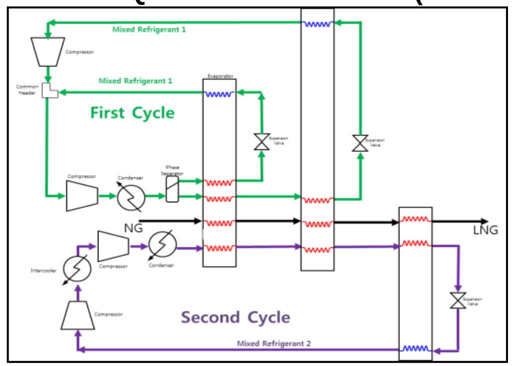


Feasible Liquefaction Cycle from the Generic Model (Case 10 ~ Case 18)

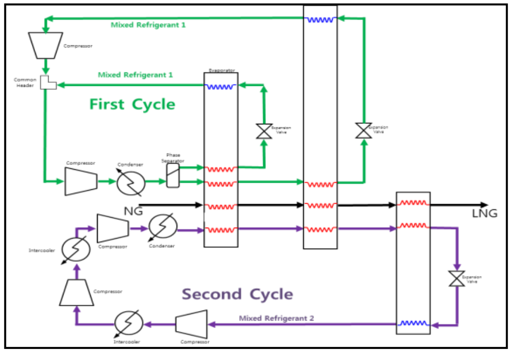
FEASIBLE LIQUEFACTION MODEL (CASE 10)



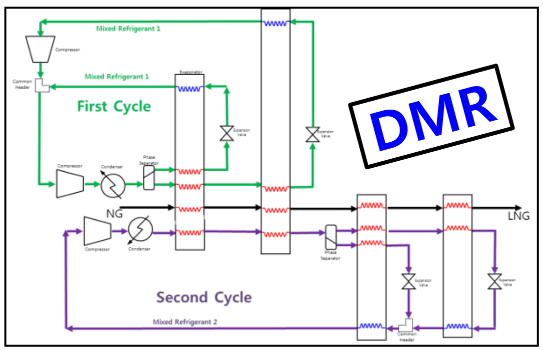
FEASIBLE LIQUEFACTION MODEL (CASE 11)



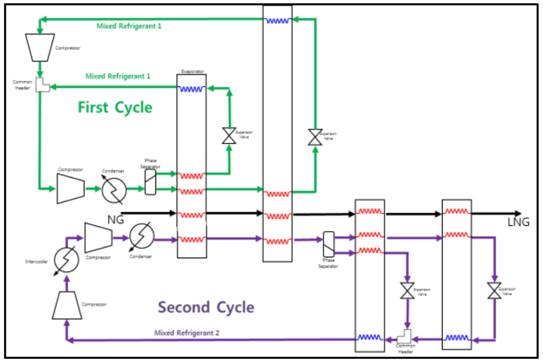
FEASIBLE LIQUEFACTION MODEL (CASE 12)



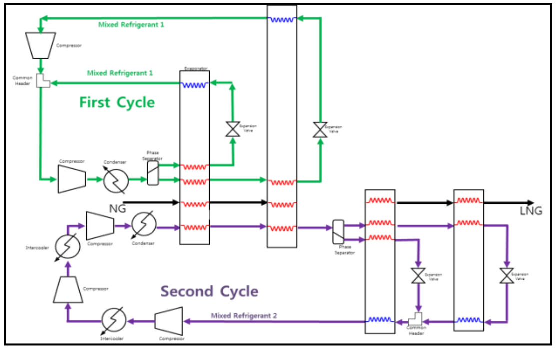
FEASIBLE LIQUEFACTION MODEL (CASE 13)



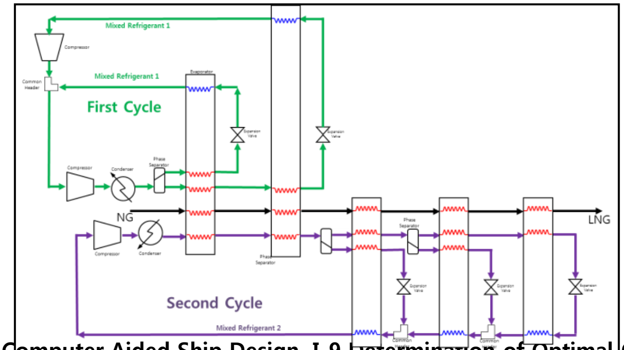
FEASIBLE LIQUEFACTION MODEL (CASE 14)



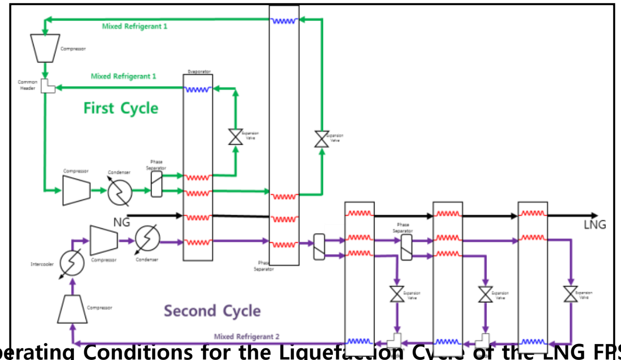
FEASIBLE LIQUEFACTION MODEL (CASE 15)



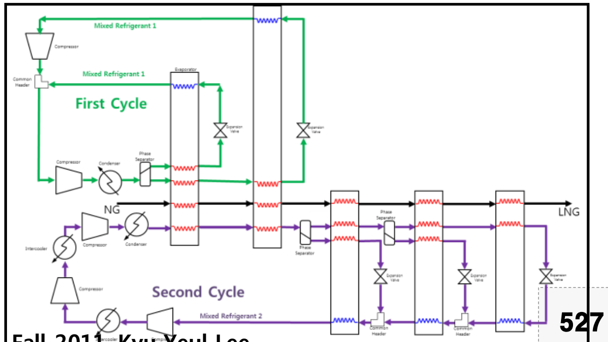
FEASIBLE LIQUEFACTION MODEL (CASE 16)



FEASIBLE LIQUEFACTION MODEL (CASE 17)

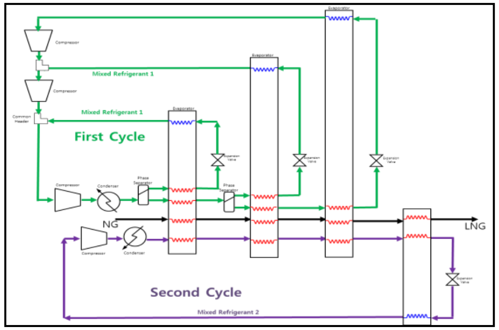


FEASIBLE LIQUEFACTION MODEL (CASE 18)

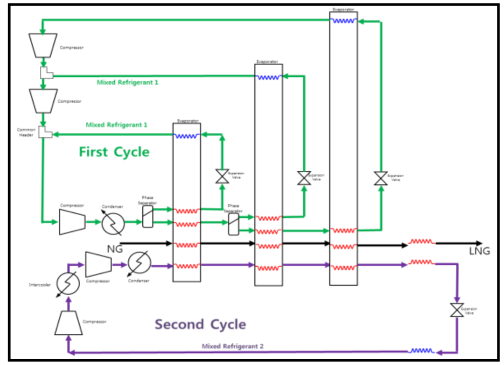


Feasible Liquefaction Cycle from the Generic Model (Case 19 ~ Case 27)

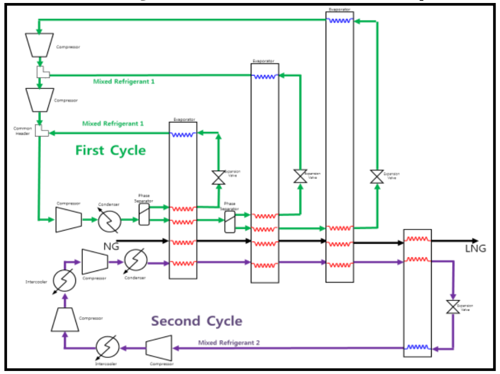
FEASIBLE LIQUEFACTION MODEL (CASE 19)



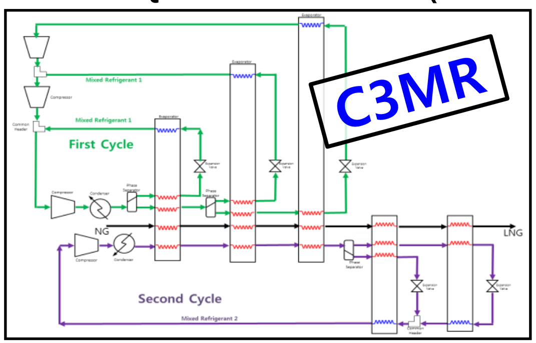
FEASIBLE LIQUEFACTION MODEL (CASE 20)



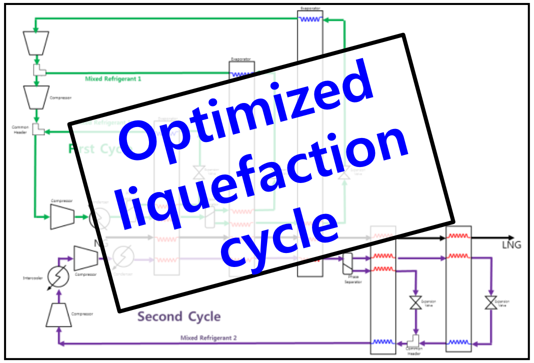
FEASIBLE LIQUEFACTION MODEL (CASE21)



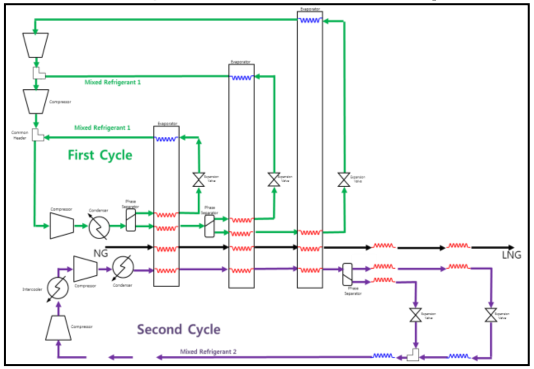
FEASIBLE LIQUEFACTION MODEL (CASE 22)



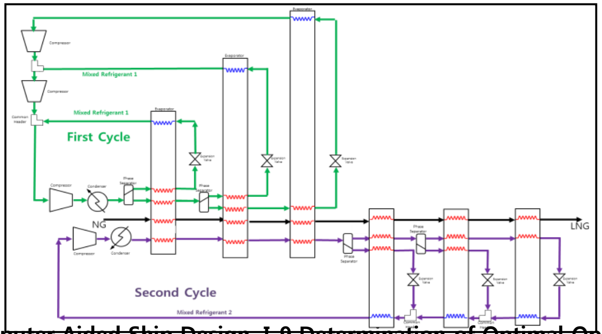
FEASIBLE LIQUEFACTION MODEL (CASE 23)



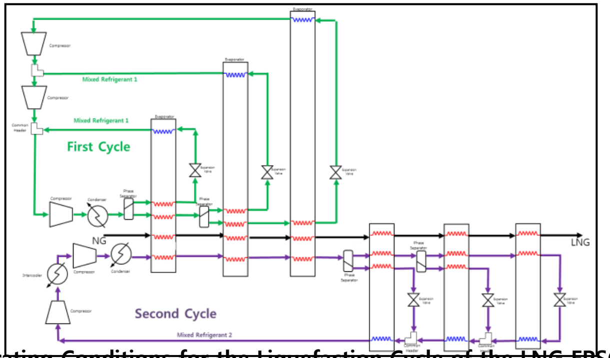
FEASIBLE LIQUEFACTION MODEL (CASE 24)



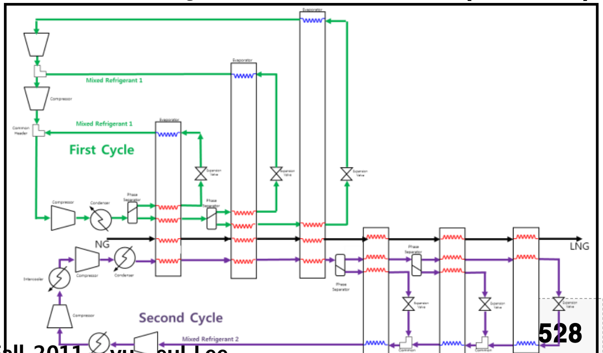
FEASIBLE LIQUEFACTION MODEL (CASE 25)



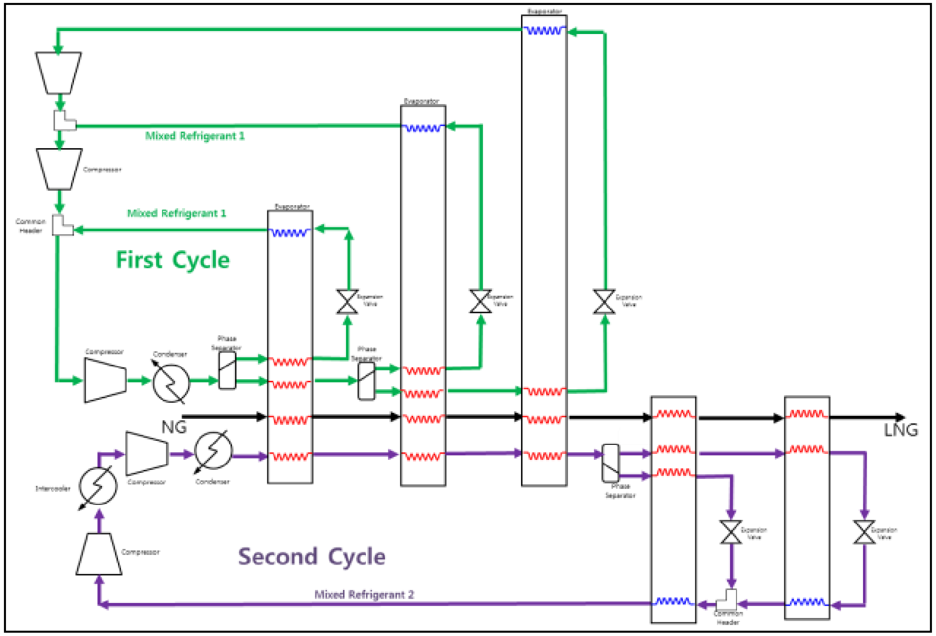
FEASIBLE LIQUEFACTION MODEL (CASE 26)



FEASIBLE LIQUEFACTION MODEL (CASE 27)



FEASIBLE LIQUEFACTION MODEL (CASE 23)



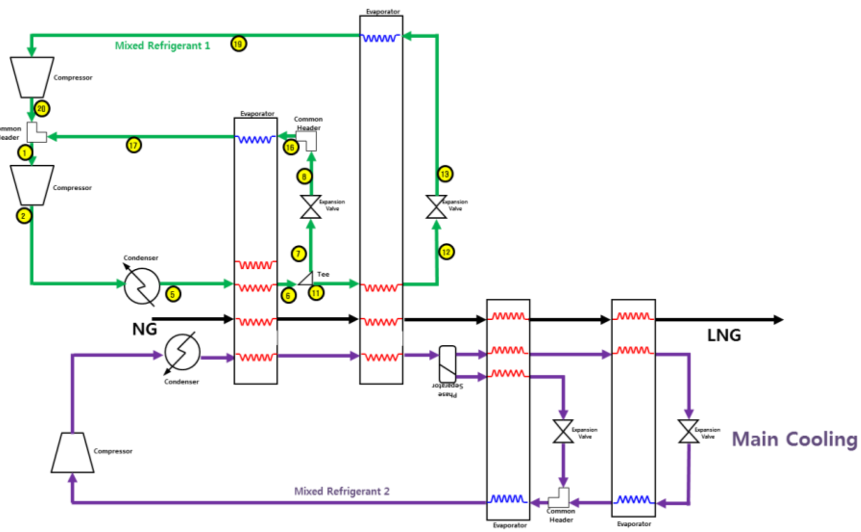
9.6. Calculation Result of the Dual Mixed Refrigerant(DMR) Cycle and Proposed Liquefaction Cycle[1)] of LNG FPSO

1) Proposed Liquefaction Cycle (CASE 23)

- Precooling 3 stage compression refrigeration
- Main cooling 2 stage compassion, 2 stage refrigeration

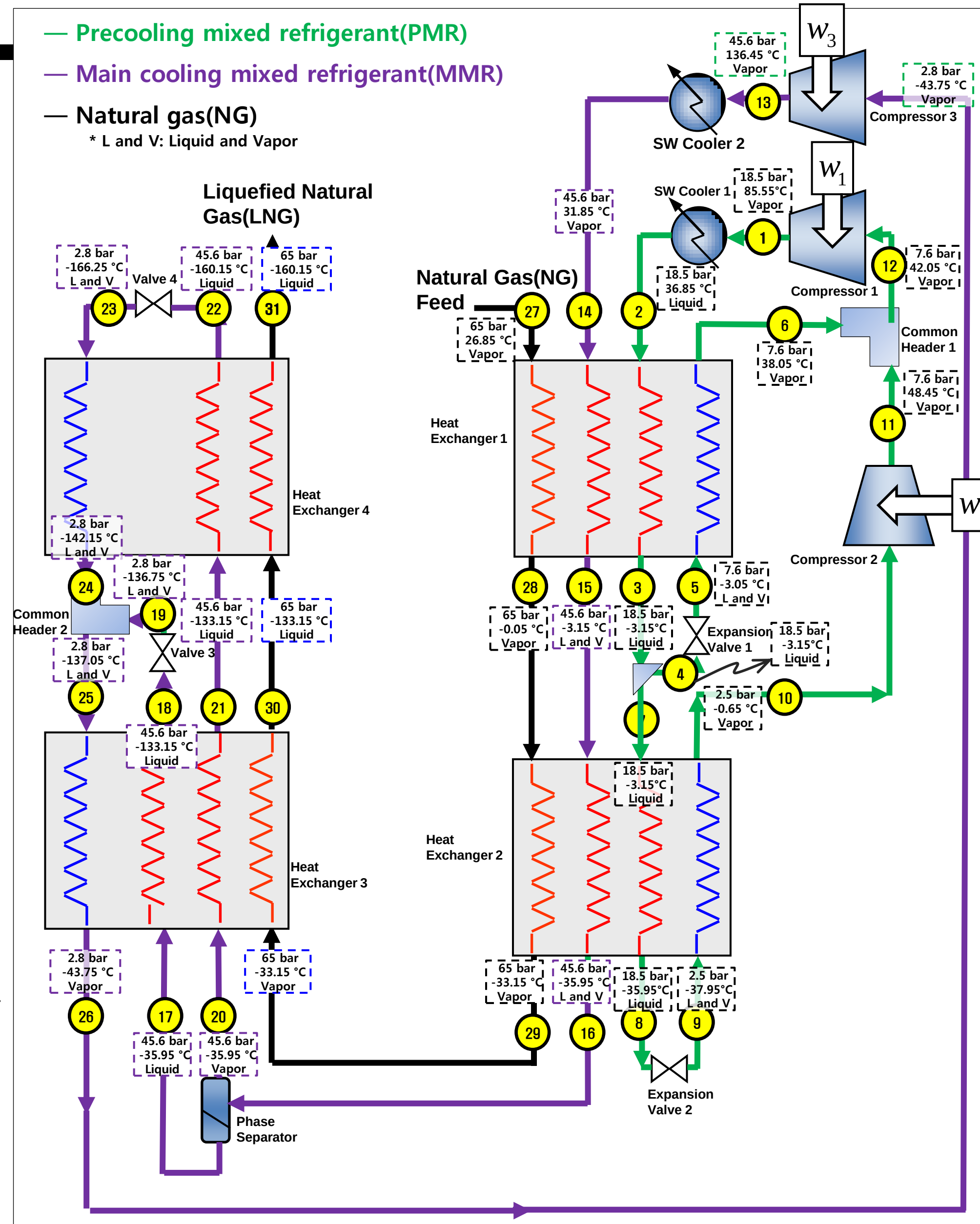
Configuration of the Dual Mixed Refrigerant(DMR) Cycle

FEASIBLE LIQUEFACTION MODEL (CASE 13)



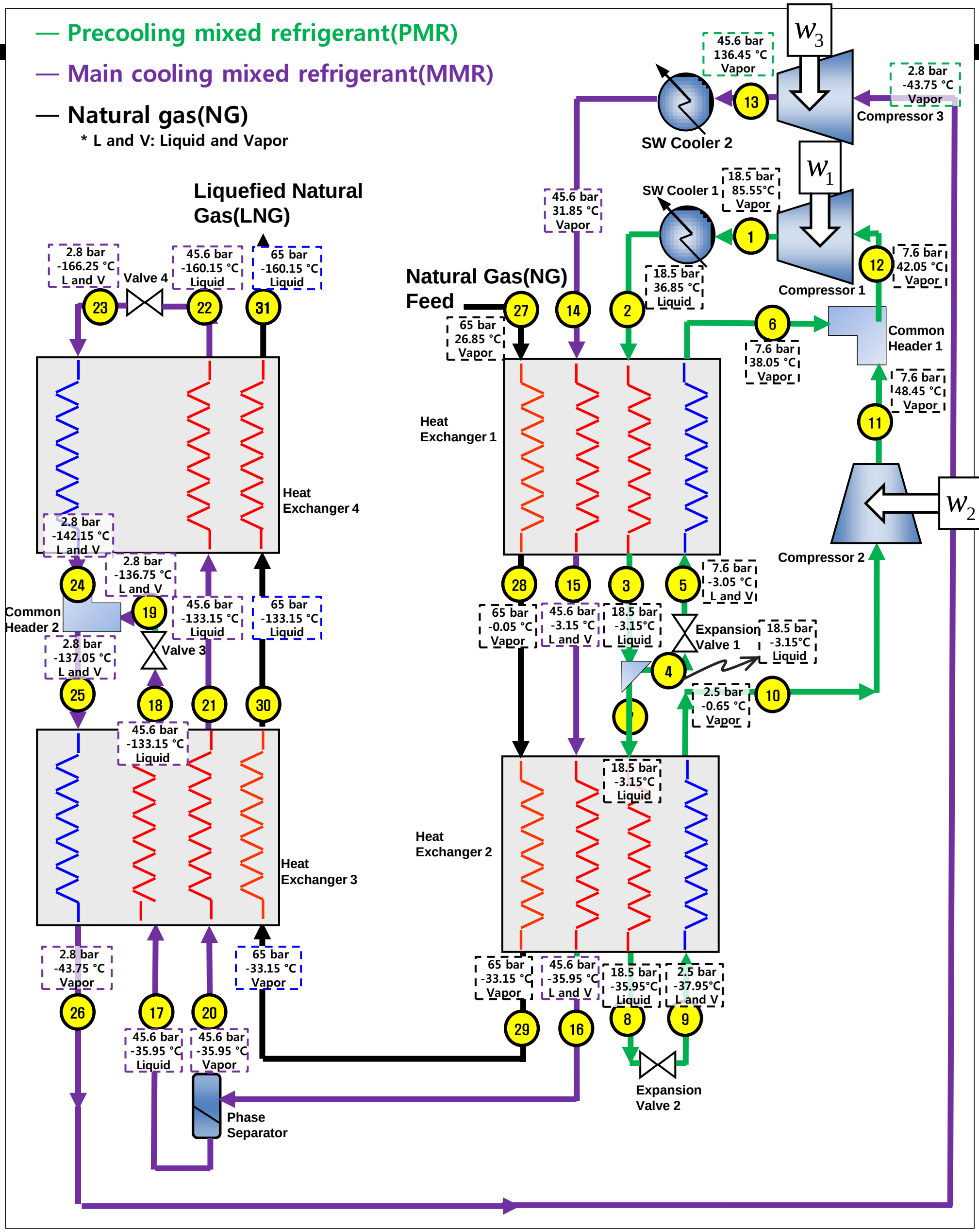
- Precooling 2 stage compression refrigeration
- Main cooling 1 stage compression, 2 stage refrigeration

- Tee: separates an inlet stream of refrigerant into the two outlet streams
- Common Header : combines the two refrigerant streams separated by tee or phase separator
- Phase Separator : separates a liquid-vapor mixture refrigerant into the vapor and liquid



[Figure] Configuration of the Dual Mixed Refrigerant Cycle

Configuration of the Dual Mixed Refrigerant(DMR) Cycle (1)



[Figure] Configuration of the Dual Mixed Refrigerant Cycle

- **Purpose:** Liquefying the natural gas by using two kind of mixed refrigerants

- **Refrigerant:**
  - Mixed refrigerant composed of Ethane(C_2H_6), Propane(C_3H_8), n-Butane(C_4H_{10}) for precooling
  - Mixed refrigerant composed of Nitrogen(N_2), Methane(C_1H_4), Ethane(C_2H_6), Propane(C_3H_8) for main cooling

- **Problem Statement:**

  [Given]:

  NG(27) T=26.85°C, P=65 bar,
  LNG(31) T=-160.15°C, P=65 bar

  $$\dot{m}_{NG} = 49.21\ kg/h$$
  $$\left(= 4.0\times10^{-4}\ MTPA\right)$$

  [Find]:

  The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

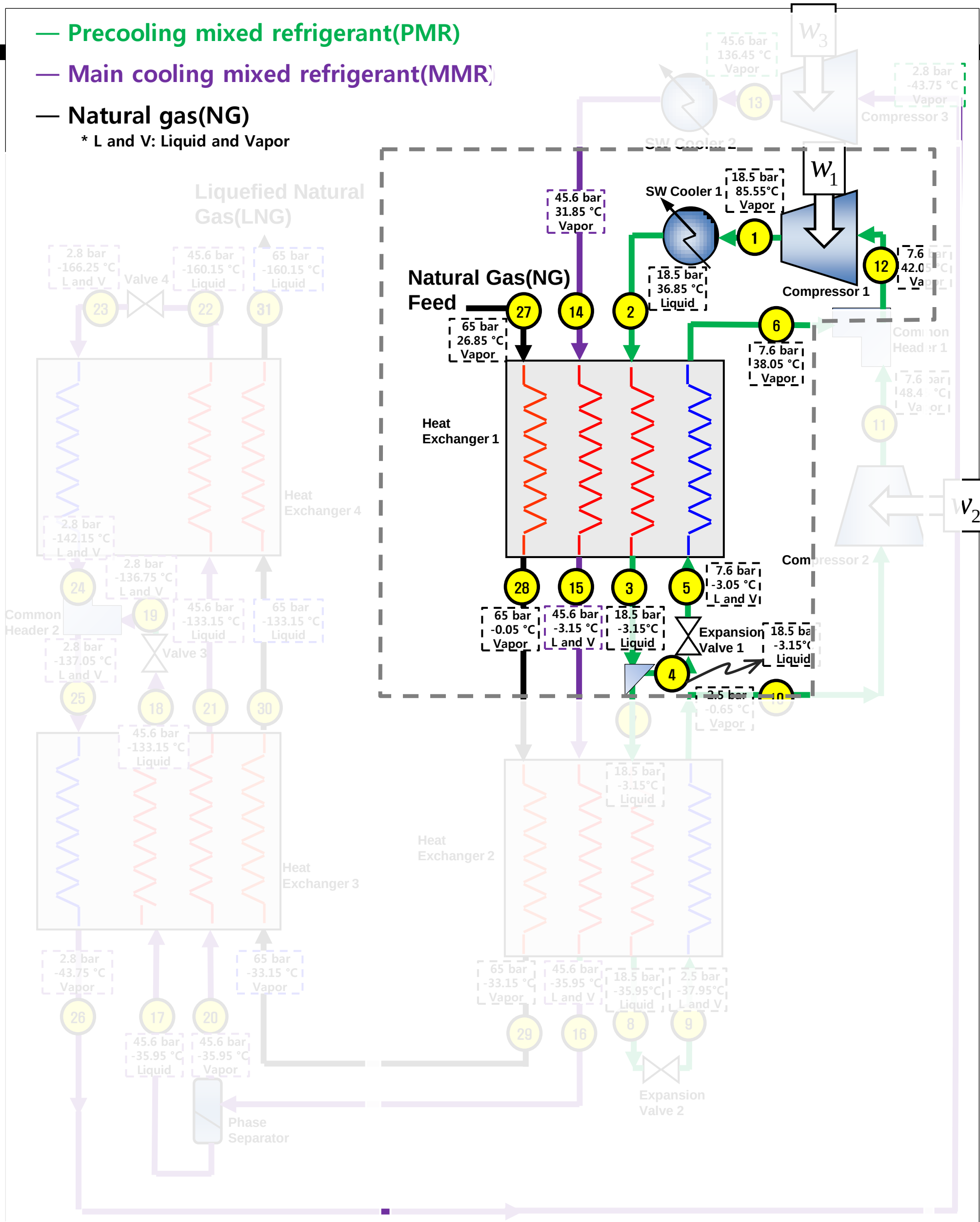
MTPA: Million Tonnes Per Annum

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

Configuration of the Dual Mixed Refrigerant(DMR) Cycle (2)

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

:stream number(St.)



[Figure] Configuration of the Dual Mixed Refrigerant Cycle

Mixed refrigerant composed of Ethane(C_2H_6), Propane(C_3H_8), n-Butane(C_4H_{10}) for precooling

[St. 12 → St. 1]
The compressor 1, usually driven by a steam turbine, brings the precooling mixed refrigerant(PMR) to a high pressure, which raises its temperature as well.

[St. 1 → St. 2]
Then, the hot PMR is cooled by sea water and is fully condensed in the sea water(SW) cooler 1.

[St. 2 → St. 3, St. 14→ St. 15, St. 27→28]
The condensed PMR is subcooled in the heat exchanger 1. The heat exchanger 1 also provides the first precool of the main mixed refrigerant(MMR) and natural gas(NG) circuits.

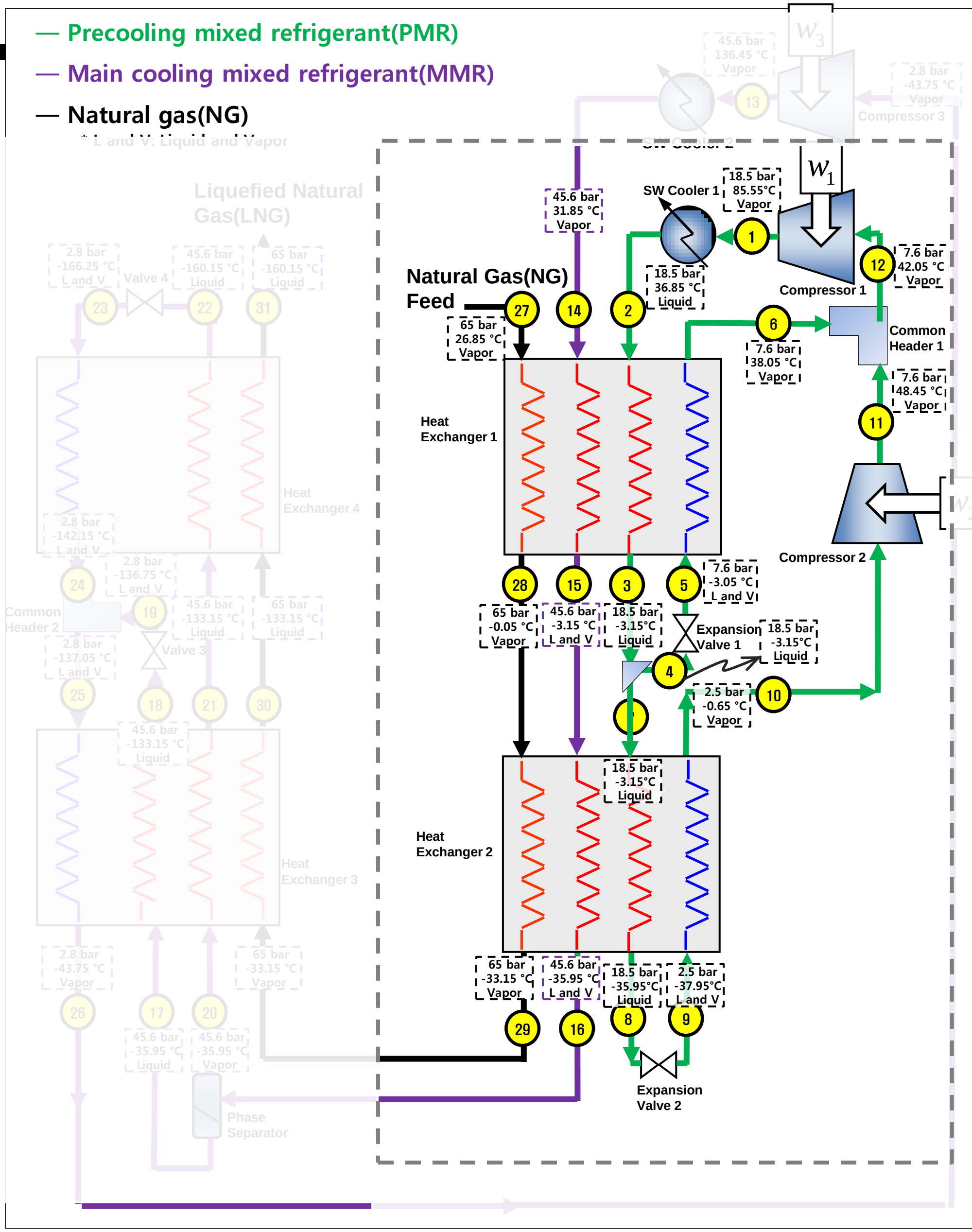
[St. 3 → St. 4 → St. 5]
At the outlet of the heat exchanger, part of the subcooled PMR is let down in pressure through the expansion valve 1.

[St. 5 → St. 6]
The resulting PMR flow of the expansion valve 1 returns to the heat exchanger 1 to be vaporized and heated, thus serving as cooling medium for the heat exchanger 1.

Configuration of the Dual Mixed Refrigerant(DMR) Cycle (3)

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

:stream number(St.)



[Figure] Configuration of the Dual Mixed Refrigerant Cycle

Mixed refrigerant composed of Ethane(C_2H_6), Propane(C_3H_8), n-Butane(C_4H_{10}) for precooling

[St. 3 → St. 7]
The remaining subcooled PMR from the heat exchanger 1 is routed to the heat exchanger 2 for further precooling of the NG and MMR.

[St. 7 → St. 8 → St. 9]
Following the same concept as in the heat exchanger 1, the PMR is subcooled in the heat exchanger 2 and is let down in pressure through the expansion valve 2 before returning in the heat exchanger 2.

[St. 9 → St. 10]
The resulting mixed flow of the expansion valve 2 returns to the heat exchanger 2 to be vaporized and heated, thus serving as cooling medium for the heat exchanger 2.

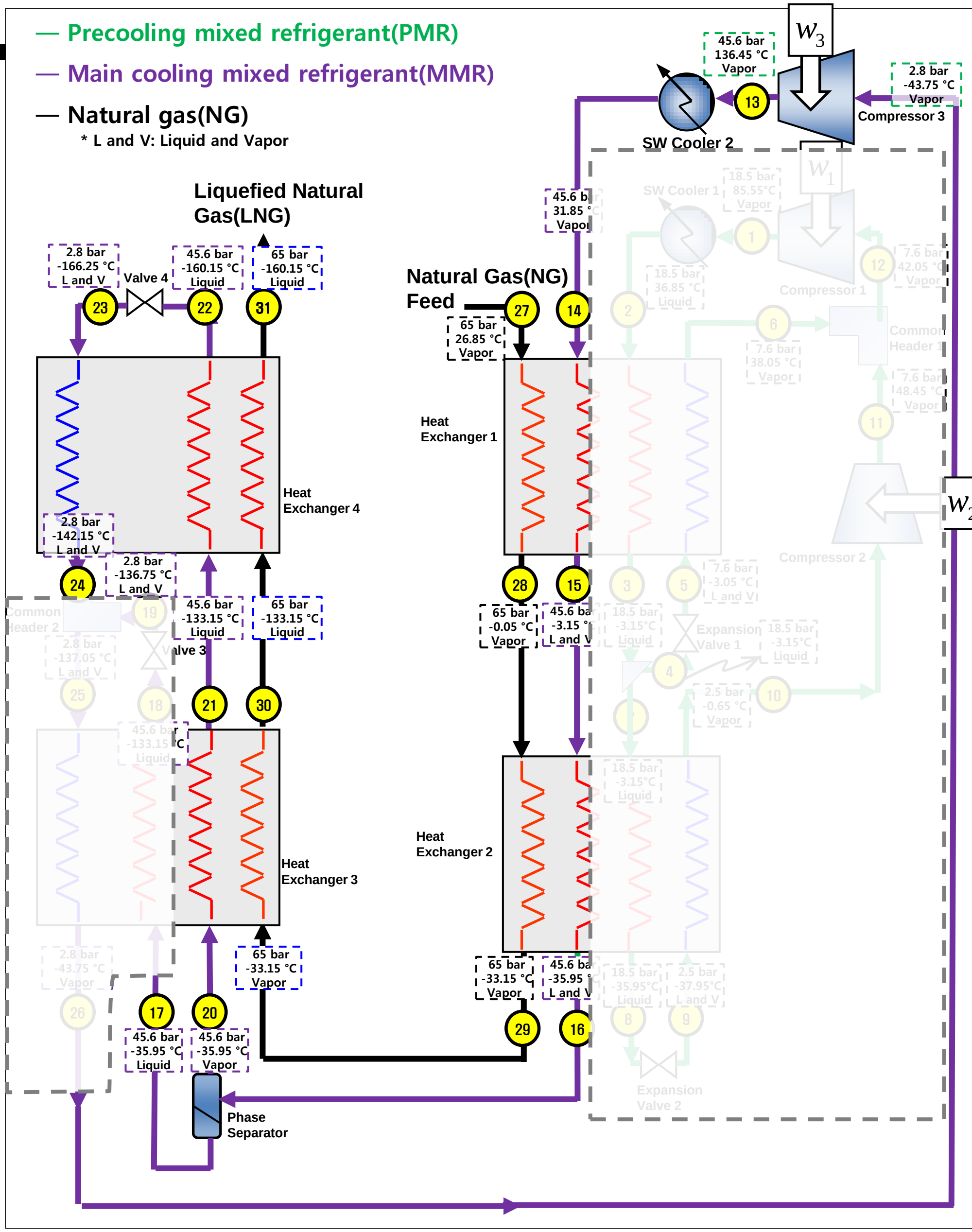
[St. 10 → St. 11]
The compressor 2 brings the PMR to a middle pressure, which raises its temperature as well.

[St. 6 and St. 11 → St. 12]
In the common header, the separated PMR streams are combined and the combined PMR returns to the compressor 1.

Configuration of the Dual Mixed Refrigerant(DMR) Cycle (4)

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

:stream number(St.)



Mixed refrigerant composed of Nitrogen(N_2), Methane(C_1H_4), Ethane(C_2H_6), Propane(C_3H_8) for main cooling

[St. 26 → St. 13]
The compressor 3 brings the main mixed refrigerant(MMR) to a high pressure, which raises its temperature as well.

[St. 13 → St. 14]
Then, the hot MMR is cooled by sea water through the sea water cooler 2.

[St. 14 → St. 15 → St. 16]
The MMR is further cooled and partially condensed successively in the heat exchanger 1 and 2.

[St. 16 → St. 17 and St. 20]
The two phase MMR flow is separated in the vapor light MMR and liquid heavy MMR in the phase separator.

[St. 20 → St. 21 → St. 22, St. 29 → St. 30]
The light MMR is fully condensed in the heat exchanger 3 and further subcooled in the heat exchanger 4. The NG is also condensed in the heat exchanger 3

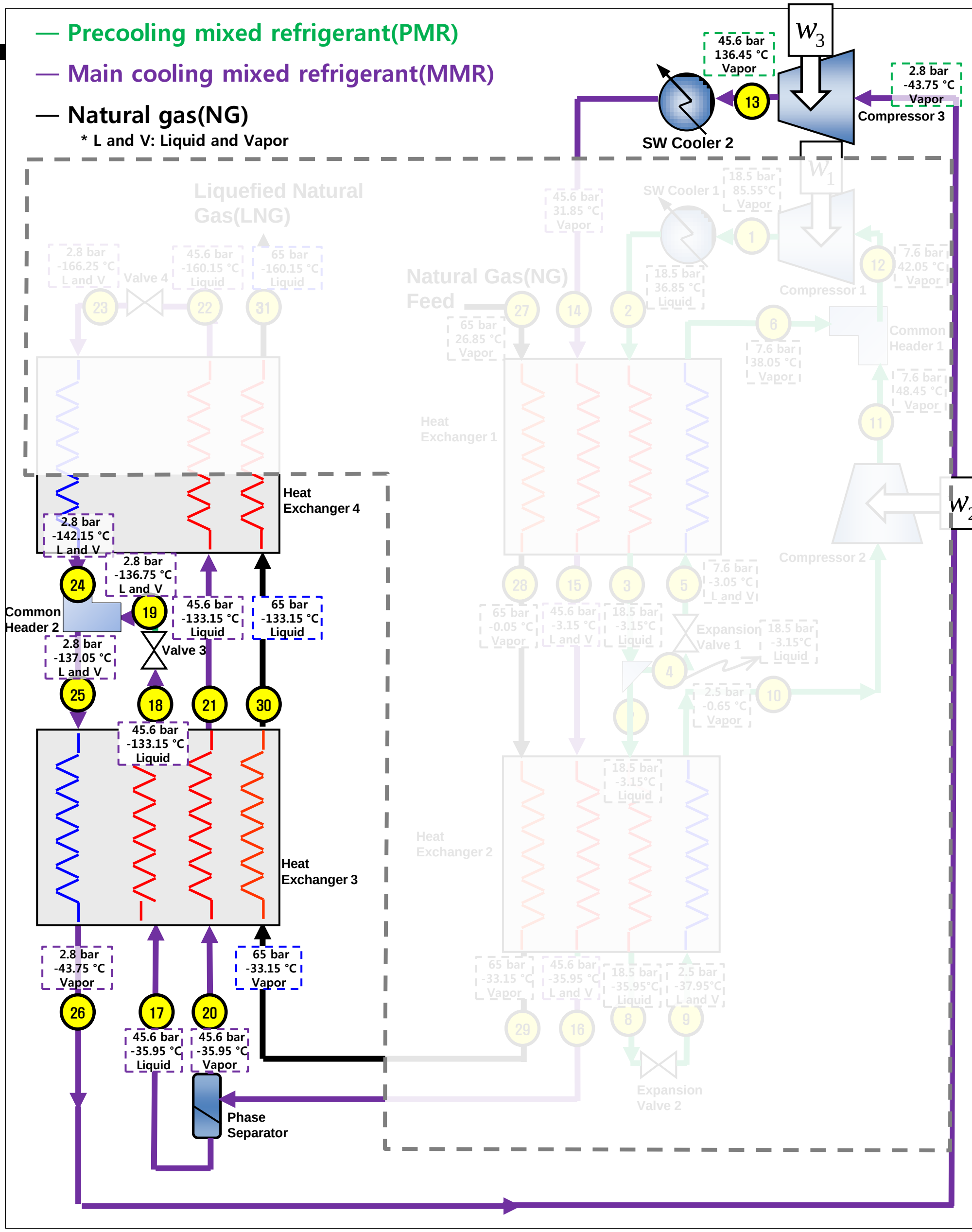
[St. 22 → St. 23 → St. 24, St. 30 → St. 31]
The subcooled light MMR is then expanded through the expansion valve 4 and returns in the heat exchanger 4 ensuring subcooling of the LNG.

[Figure] Configuration of the Dual Mixed Refrigerant Cycle

Configuration of the Dual Mixed Refrigerant(DMR) Cycle (5)

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

:stream number(St.)



[Figure] Configuration of the Dual Mixed Refrigerant Cycle

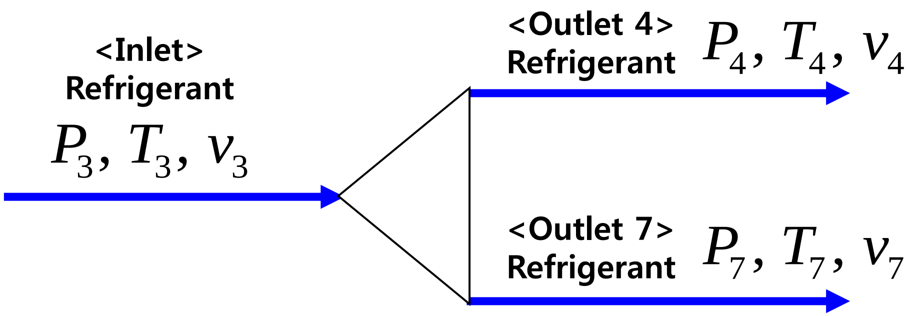
Mixed refrigerant composed of Nitrogen(N_2), Methane(C_1H_4), Ethane(C_2H_6), Propane(C_3H_8) for main cooling

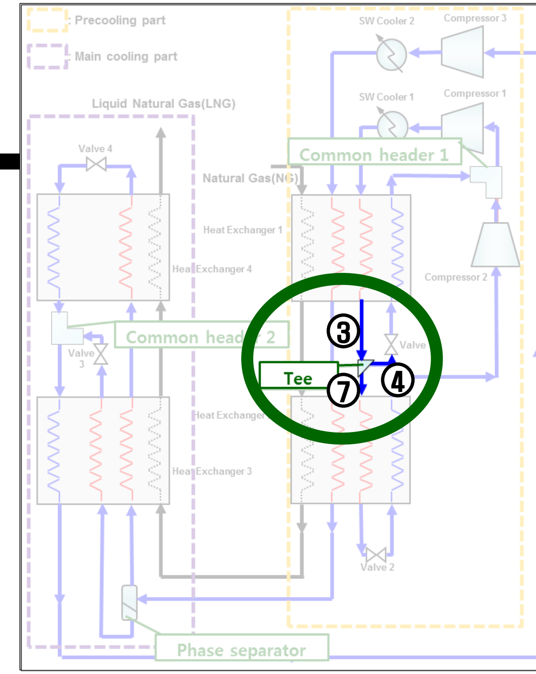
[St. 17 → St. 18 → St. 19]
The heavy MMR is subcooled in the heat exchanger 3 and is let down in pressure through the expansion valve 3.

[St. 19 and St. 24 → St. 25 → St. 26]
Then, the heavy MMR returns in the heat exchanger 3 for serving as cooling medium for the heat exchanger 3 mixed with the light MMR from the heat exchanger 4 in the common header 2 and returns to the compressor 3.

Mathematical Model for Tee

- **Tee** : separates an inlet stream of refrigerant into the two outlet streams





[Figure 3] Configuration of the Dual Mixed Refrigerant Cycle

[Given]:

NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{NG}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:

The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

1. Design variables(Operating Conditions)

: $P_3, T_3, v_3, P_4, T_4, v_4, P_7, T_7, v_7, c$

2. Assumption:

- There is no pressure drop of the refrigerant through the tee. "Isobaric process"
- There is no heat transfer between the refrigerant and surroundings "Adiabatic process"

3. Equality constraints

1) The first law of the thermodynamics(Energy conservation)

$$h_3(P_3,v_3,T_3) = \boxed{c}\left(h_4(P_4,v_4,T_4)\right) + (1-\boxed{c})\left(h_7(P_7,v_7,T_7)\right)$$

2) Isobaric process

$$P_3 = P_4, \quad P_3 = P_7$$

3) Conditions for temperature of the outlet

$$T_4 = T_7$$

4) Equations of state(Soave, Redlich, Kwong(SRK) equation)

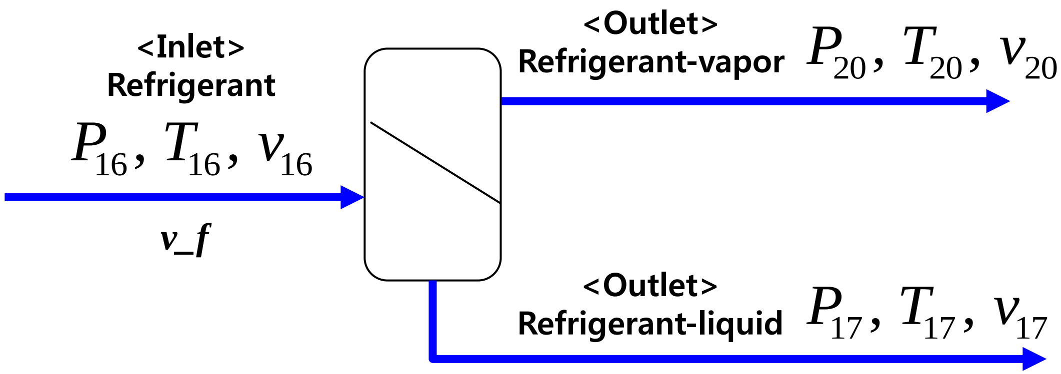
$$v_3 = \frac{RT_3}{P} + b - \frac{a(T_3)}{P_3}\frac{v_3 - b}{(v_3-\varepsilon b)(v_3-\sigma b)} \qquad v_4 = \frac{RT_4}{P} + b - \frac{a(T_4)}{P_4}\frac{v_4 - b}{(v_4-\varepsilon b)(v_4-\sigma b)}$$

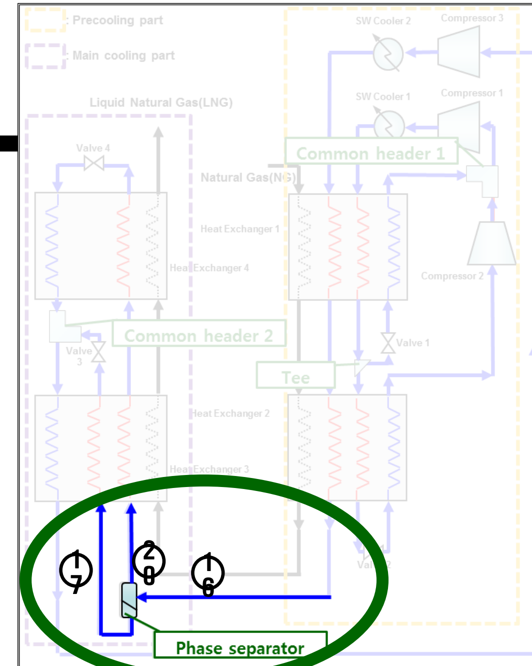
$$v_7 = \frac{RT_7}{P} + b - \frac{a(T_7)}{P_7}\frac{v_7 - b}{(v_7-\varepsilon b)(v_7-\sigma b)}$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy
c: flow rate ratio between inlet and outlet 4

Mathematical Model for Phase Separator

- **Phase Separator**: separates a liquid-vapor mixture refrigerant into the vapor and liquid





[Figure 3] Configuration of the Dual Mixed Refrigerant Cycle

[Given]:
NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{NG}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:
The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

1. **Design variables(Operating Conditions)**
   : $P_{16}, T_{16}, v_{16}, P_{17}, T_{17}, v_{17}, P_{20}, T_{20}, v_{20}$, v_f

2. **Assumption:**
- There is no pressure drop of the refrigerant through the phase separator. "Isobaric process"
- There is no heat transfer between the refrigerant and surroundings "Adiabatic process".

3. **Equality constraints**

1) The first law of the thermodynamics(Energy conservation)

$$h_{16}(P_{16}, v_{16}, T_{16}) = \boxed{v\_f} \cdot h_{20}(P_{20}, v_{20}, T_{20}) + (1 - \boxed{v\_f}) \cdot h_{17}(P_{17}, v_{17}, T_{17})$$

2) Isobaric process

$$P_{16} = P_{17}, \quad P_{16} = P_{20}$$

3) Conditions for temperature of the outlet

$$T_{17} = T_{20}$$

4) Equations of state(Soave, Redlich, Kwong(SRK) equation)

$$v_{16} = \frac{RT_{16}}{P} + b - \frac{a(T_{16})}{P_{16}} \frac{v_{16} - b}{(v_{16} - \varepsilon b)(v_{16} - \sigma b)}$$

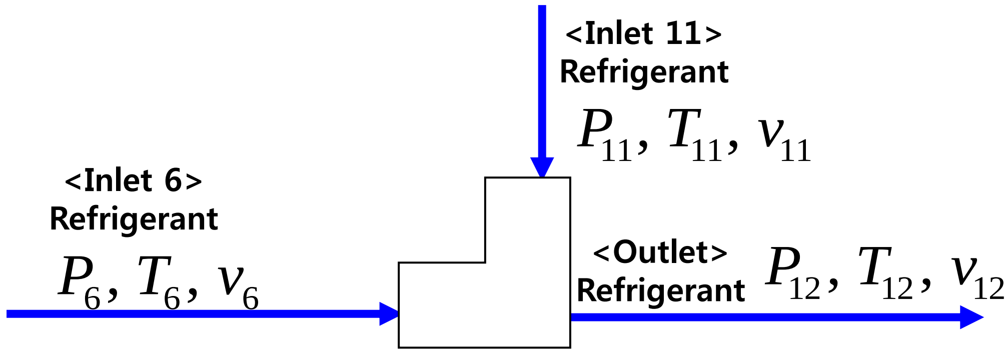
$$v_{17} = \frac{RT_{17}}{P} + b - \frac{a(T_{17})}{P_{17}} \frac{v_{17} - b}{(v_{17} - \varepsilon b)(v_{17} - \sigma b)}$$

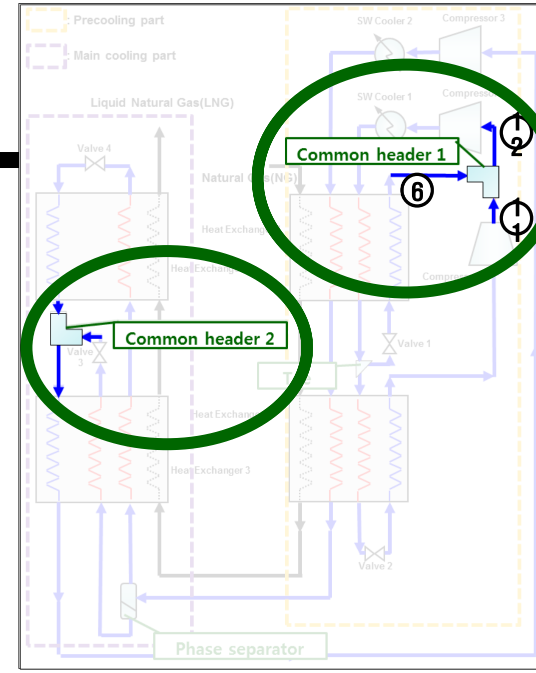
$$v_{20} = \frac{RT_{20}}{P} + b - \frac{a(T_{20})}{P_{20}} \frac{v_{20} - b}{(v_{20} - \varepsilon b)(v_{20} - \sigma b)}$$

T: temperature
P: pressure
v: specific volume
h: specific enthalpy
v_f: vapor fraction at 16 stream

Mathematical Model for Common Header (1/2)

- **Common Header**: combines the two refrigerant streams separated by tee





[Figure 3] Configuration of the Dual Mixed Refrigerant Cycle

[Given]:
NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{NG}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:
The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

1. Design variables(Operating Conditions)

: $P_6, T_6, v_6, P_{11}, T_{11}, v_{11}, P_{12}, T_{12}, v_{12}, c$

2. Assumption:

- To prevent a backflow in the common headers, the pressures of the inlet streams are the same.
- There is no pressure drop of the refrigerant through the common header. "Isobaric process"
- There is no heat transfer between the refrigerant and surroundings, "Adiabatic process".

T: temperature
P: pressure
v: specific volume
h: specific enthalpy
c: flow rate ratio between inlet and outlet 4

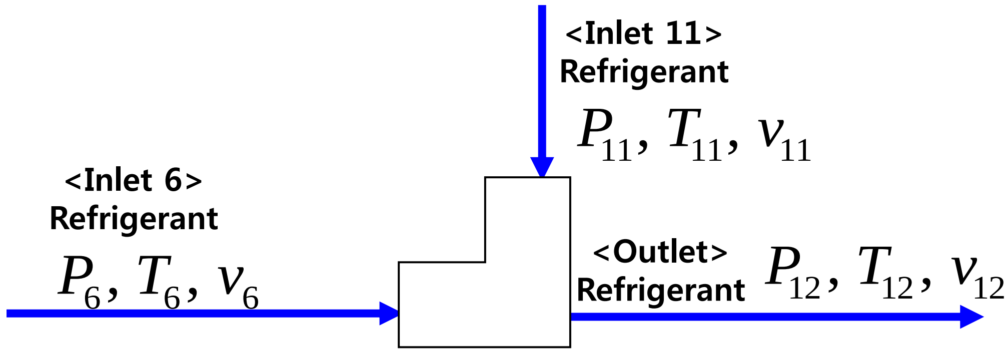
3. Equality constraints

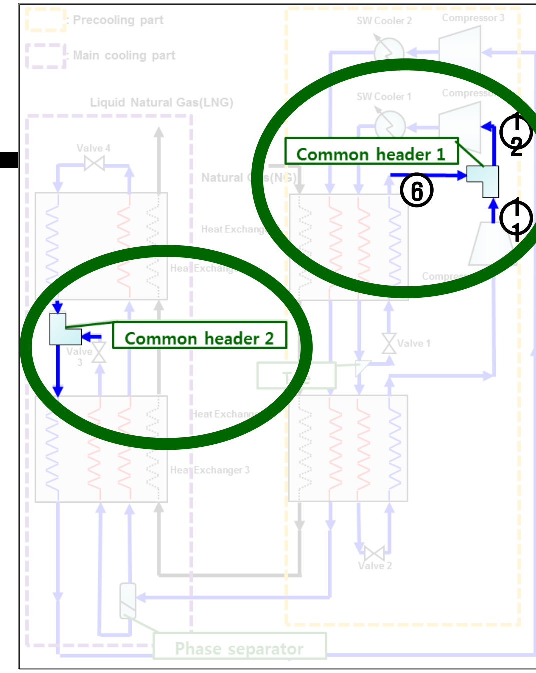
1) The first law of the thermodynamics(Energy conservation)

$$\boxed{c \cdot h_6(P_6, v_6, T_6) + (1-c) \cdot h_{11}(P_{11}, v_{11}, T_{11}) = h_{12}(P_{12}, v_{12}, T_{12})}$$

Mathematical Model for Common Header (2/2)

- **Common Header**: combines the two refrigerant streams separated by tee





[Figure 3] Configuration of the Dual Mixed Refrigerant Cycle

[Given]:

NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{NG}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:

The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

3. Equality constraints

2) Conditions for pressure of the inlet to prevent backflow

$$\boxed{P_6 = P_{11}}$$

3) Isobaric process

$$\boxed{P_6 = P_{12}}$$

4) Equations of state(Soave, Redlich, Kwong(SRK) equation)

$v_6 = \frac{RT_6}{P} + b - \frac{a(T_6)}{P_6}\frac{v_6 - b}{(v_6-\varepsilon b)(v_6-\sigma b)}$	$v_{12} = \frac{RT_{12}}{P} + b - \frac{a(T_{12})}{P_{12}}\frac{v_{12} - b}{(v_{12}-\varepsilon b)(v_{12}-\sigma b)}$
$v_{11} = \frac{RT_{11}}{P} + b - \frac{a(T_{11})}{P_{11}}\frac{v_{11} - b}{(v_{11}-\varepsilon b)(v_{11}-\sigma b)}$	

***T*:** temperature
***P*:** pressure
***v*:** specific volume

$$a(T) = \psi \frac{\alpha(T_r) R^2 T_C^2}{P_C}$$

$\psi = 0.42748$ *for SRK equation*

***R*:** gas constant (=8.314 Jmol⁻¹K⁻¹)

***P*$_c$:** critical pressure of the refrigerant

***T*$_c$:** critical temperature of the refrigerant

$$b = \Omega \frac{RT_C}{P_C}$$

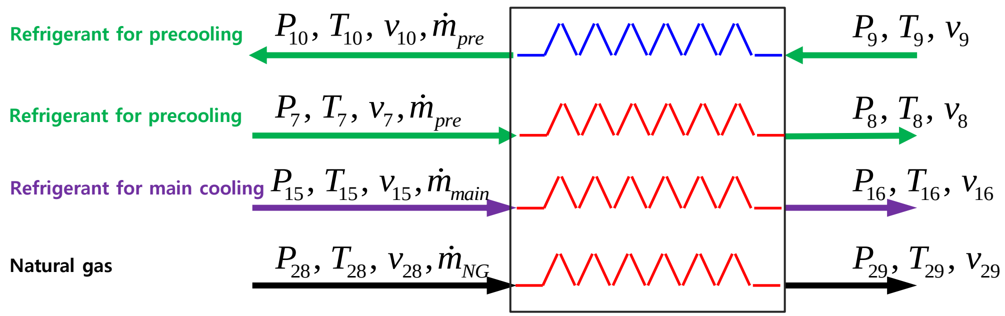
$\Omega = 0.08664$ *for SRK equation*

$\varepsilon = 0$ *for SRK equation*

$\sigma = 1$ *for SRK equation*

Mathematical Model for Heat Exchanger (1/2)

- **Heat Exchanger** : devices where several moving fluid streams exchange heat without mixing



1. Design variables(Operating Conditions)

: P_i, T_i, v_i, $\dot{m}_{pre}$, $\dot{m}_{main}$, $\dot{m}_{NG}$ $(i = 7, 8, 9, 10, 15, 16, 28, 29)$

2. Assumption:

- There is no pressure drop of the refrigerant through the heat exchanger. "Isobaric process"

3. Equality constraints

1) The first law of the thermodynamics(Energy conservation)

$$\left[\dot{m}_{pre}\cdot h_7+\dot{m}_{pre}\cdot h_9+\dot{m}_{main}\cdot h_{15}+\dot{m}_{NG}\cdot h_{28}\right] = \left[\dot{m}_{pre}\cdot h_8+\dot{m}_{pre}\cdot h_{10}+\dot{m}_{main}\cdot h_{16}+\dot{m}_{NG}\cdot h_{29}\right]$$

2) Isobaric process

$$\boxed{P_7 = P_8},\ \boxed{P_9 = P_{10}},\ \boxed{P_{15} = P_{16}},\ \boxed{P_{28} = P_{29}}$$

3) Conservation Condition of the Output Temperature

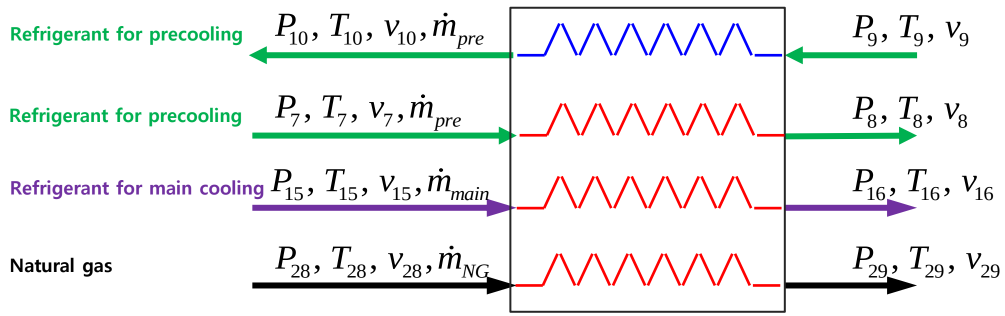
$$\boxed{T_8 = T_{16}},\ \boxed{T_8 = T_{29}}$$

[Given]:
NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{NG}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:
The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

Mathematical Model for Heat Exchanger (2/2)

- **Heat Exchanger** : devices where several moving fluid streams exchange heat without mixing



1. **Design variables(Operating Conditions)**

: P_i, T_i, v_i, $\dot{m}_{\text{pre}}$, $\dot{m}_{\text{main}}$, $\dot{m}_{\text{NG}}$ $(i = 7, 8, 9, 10, 15, 16, 28, 29)$

2. **Assumption:**
- There is no pressure drop of the refrigerant through the heat exchanger. "Isobaric process"

3. **Equality constraints**

4) Equation of state for each stream

4. **Inequality constraints**
- Minimum temperature difference in the heat exchanger

$T_7 - T_{10} \geq 3.0$
$T_8 - T_9 \geq 3.0$

[Given]:

NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar
$\dot{m}_{\text{NG}}$ = 49.21 kg/h (=0.0004 MMTA)

[Find]:

The operating conditions such as the pressure, temperature and specific volume, mass flow rate and composition of the refrigerants minimizing the work provided to the compressor.

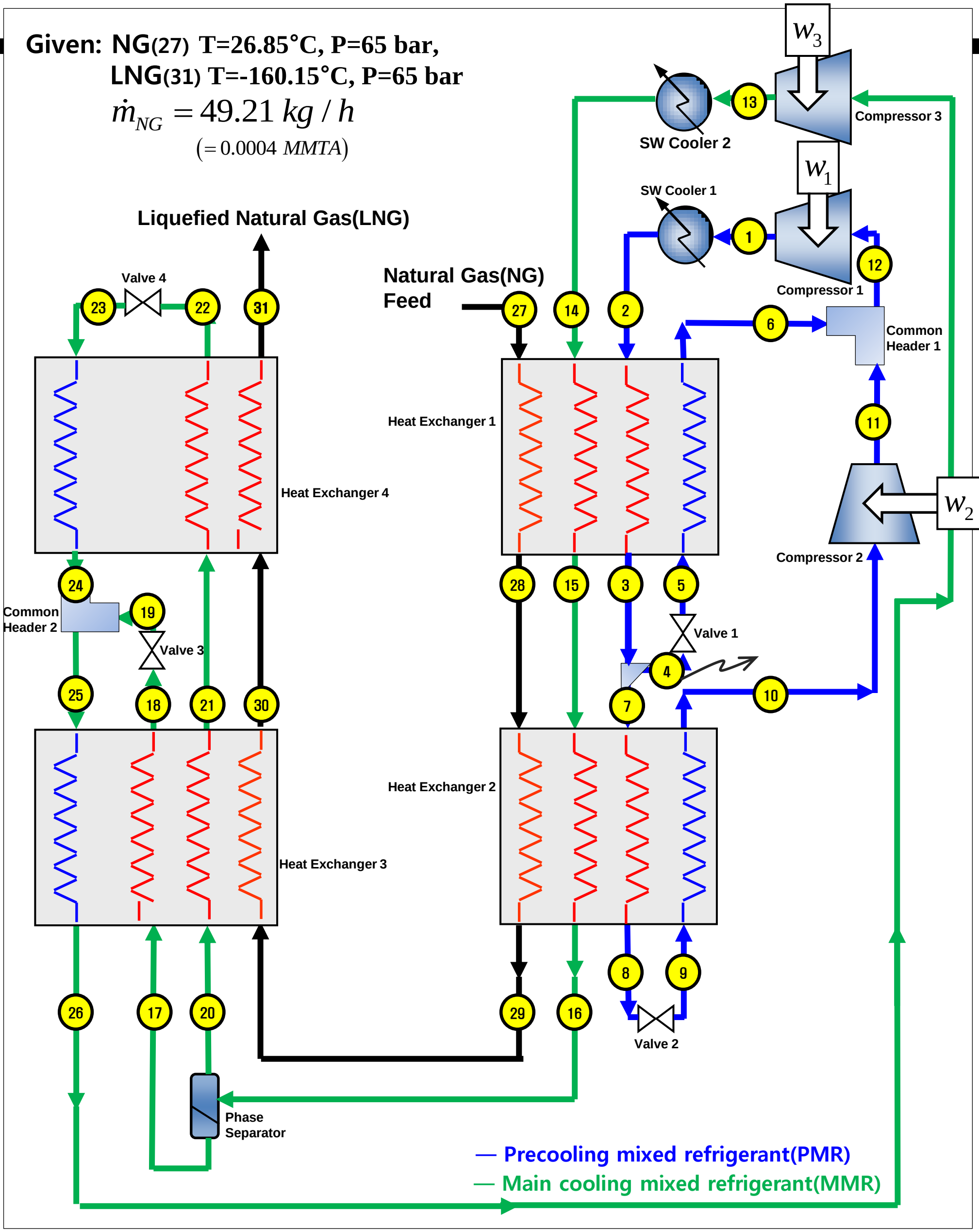
Mathematical Model of the DMR cycle

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

Given: NG(27) T=26.85°C, P=65 bar,
LNG(31) T=-160.15°C, P=65 bar

$\dot{m}_{NG} = 49.21\ kg/h$

$(=0.0004\ MMTA)$



Configuration of the Dual Mixed Refrigerant Cycle[1)]

1. Design variables(operating conditions) [107]

$: P_i,\ T_i,\ v_i\ (i=1, \ldots, 26, 28, 29, 30),$

$T_{S,1}, T_{S,11}, T_{S,13}, v_{S,1}, v_{S,11}, v_{S,13}, w_1,\ w_2,\ w_3,\ c,$

$\dot{m}_{pre},\ \dot{m}_{main}, v\_f, z_{j,pre}\ (j=1,2,3), z_{k,main}\ (k=1,2,3,4)$

2. Equality constraints [91]

1) Composition of the refrigerant [2]

2) Precooling part [49]

- Compressor 1[6], Sea water cooler 1[3], Heat exchanger 1 [11], Tee [6], Expansion Valve 1 [2], Heat exchanger 2 [11], Expansion Valve 2 [2], Compressor2 [5], Common header1 [3]

3) Main cooling part [40]

- Compressor 3 [6], Sea water cooler 2 [3], Phase separator [6], Heat exchanger 3 [10], Expansion Valve [2], Heat exchanger 4 [6], Valve [2], Common header2 [4]

→ indeterminate systems

3. Inequality constraints [11]

1) Minimum temperature approach in heat exchanger [8]

2) Inlet condition of the compressor [3]

4. Objective function: Minimize the compressors power

$$\text{Minize } \dot{m}_{pre} \cdot w_1 + \dot{m}_{pre} \cdot w_2 + \dot{m}_{main} \cdot w_3$$

→ Optimization Problem!

T: Temperature / **P**: Pressure / **v**: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / M_{pre}: mass flow rate at the precooling part

*Subscript 'NG': natural gas, Subscript 'main': main cooling part

Mathematical Model of the Precooling part of the DMR cycle

Composition of the refrigerant(z)
Precooling: Ethane, Propane, n-Butane
Main cooling: Nitrogen, Methane, Ethane, Propane
Natural Gas: Methane(87.5%), Ethane(5.5%), Nitrogen(4.0%), Propane(2.1%), n-Butane(0.5%), i-Butane(0.3%), i-Pentane(0.1%)

1. Design variables(Operating Conditions)[107]

: $P_i,\ T_i,\ v_i\left(i=1,\ \ldots,26,28,29,30\right), T_{S,1}, T_{S,11}, T_{S,13}, v_{S,1}, v_{S,11}, v_{S,13}, w_1,\ w_2,\ w_3,\ c,\ \dot{m}_{pre},\ \dot{m}_{main}, v\_f, z_{j,pre}\left(j=1,2,3\right), z_{k,main}\left(k=1,2,3,4\right)$

2. Equality constraints[91]

$$\sum_{j=1}^{3} z_{j,pre}=1,\quad \sum_{k=1}^{4} z_{k,main}=1$$

***T*: Temperature / *P*: Pressure / *v*: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant**
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

2.1 Equality constraints of Precooling part[49]

1) Compressor 1: [6]

$$h_{12}\left(P_{12},T_{12},v_{12},z_{j,pre}\right)+w_1=h_1\left(P_1,T_1,v_1,z_{j,pre}\right)$$

$$\eta=\frac{h_S\left(P_1,T_{S,1},v_{S,1},z_{j,pre}\right)-h_{12}\left(P_{12},T_{12},v_{12},z_{j,pre}\right)}{h_1\left(P_1,T_1,v_1,z_{j,pre}\right)-h_{12}\left(P_{12},T_{12},v_{12},z_{j,pre}\right)}$$

$$s_{12}\left(P_{12},T_{12},v_{12},z_{j,pre}\right)=s_1\left(P_1,T_{S,1},v_{S,1},z_{j,pre}\right)$$

$$v_{12}=v_{12}(T_{12},P_{12},z_{j,pre})$$

$$v_{S,1}=v_{S,1}(T_{S,1},P_1,z_{j,pre})$$

$$v_1=v_1(T_1,P_1,z_{j,pre})$$

4) Tee: [6]

$$h_3\left(P_3,T_3,v_3,z_{j,pre}\right)=c\cdot h_4\left(P_4,T_4,v_4,z_{j,pre}\right)+(1-c)\cdot h_7\left(P_7,T_7,v_7,z_{j,pre}\right)$$

$$P_3=P_4,\ \ P_3=P_7$$

$$T_4=T_7$$

$$v_4=v_4\left(T_4,P_4,z_{j,pre}\right),\ v_7=v_7\left(T_7,P_7,z_{j,pre}\right)$$

5) Expansion Valve 1: [2]

$$h_4\left(P_4,T_4,v_4\right)=h_5\left(P_5,T_5,v_5\right)$$

$$v_5=v_5\left(T_5,P_5\right)$$

8) Compressor 2: [5]

$$(1-c)\cdot h_{10}\left(P_{10},T_{10},v_{10},z_{j,pre}\right)+w_2=(1-c)\cdot h\left(P_{11},T_{11},v_{11},z_{j,pre}\right)$$

$$\eta=\frac{h_S\left(P_{11},T_{S,11},v_{S,11},z_{j,pre}\right)-h_{10}\left(P_{10},T_{10},v_{10},z_{j,pre}\right)}{h_{11}\left(P_{11},T_{11},v_{11},z_{j,pre}\right)-h_{10}\left(P_{10},T_{10},v_{10},z_{j,pre}\right)}$$

$$s_{10}\left(P_{10},T_{10},v_{10},z_{j,pre}\right)=s_{11}\left(P_{11},T_{S,11},v_{S,11},z_{j,pre}\right)$$

$$v_{S,11}=v_{S,11}\left(P_{11},T_{S,11},z_{j,pre}\right)$$

$$v_{11}=v_{11}(T_{11},P_{11},z_{j,pre})$$

2) Sea water cooler 1: [3]

The temperature of the outlet of the sea water cooler is usually given. T=310K

$$P_1=P_2$$

$$v_2=v_2(T_2,P_2,z_{j,pre})$$

6) Heat exchanger 2: [11]

$(1-c)\cdot$ [redacted]

$$=(1-c)\cdot\dot{m}_{pre}\cdot h_8\left(P_8,T_8,v_8,z_{j,pre}\right)+(1-c)\cdot\dot{m}_{pre}\cdot h_{10}\left(P_{10},T_{10},v_{10},z_{j,pre}\right)+\dot{m}_{main}\cdot h_{16}\left(P_{16},T_{16},v_{16},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{29}\left(P_{29},T_{29},v_{29},z_{l,NG}\right)$$

$$P_7=P_8,\ P_9=P_{10},P_{15}=P_{16},P_{28}=P_{29}\qquad T_8=T_{16},\ \ T_8=T_{29}$$

$$v_8=v_8\left(T_8,P_8,z_{j,pre}\right),\ v_{10}=v_{10}\left(T_{10},P_{10},z_{j,pre}\right),\ v_{16}=v_{16}\left(T_{16},P_{16},z_{k,main}\right),\ v_{29}=v_{29}\left(T_{29},P_{29},z_{l,NG}\right)$$

3) Heat exchanger 1: [11]

[redacted])

= [redacted]

$$P_2=P_3,\ \ P_5=P_6,\ \ P_{14}=P_{15},\ \ P_{NG}=P_{28}\qquad T_3=T_{15},\ \ T_3=T_{28}$$

$$v_3=v_3\left(T_3,P_3,z_{j,pre}\right),\ v_6=v_6\left(T_6,P_6,z_{j,pre}\right),\ v_{15}=v_{15}\left(T_{15},P_{15},z_{k,main}\right),\ v_{28}=v_{28}\left(T_{28},P_{28},z_{l,NG}\right)$$

7) Expansion Valve 2: [2]

$$h_8\left(P_8,T_8,v_8,z_{j,pre}\right)=h_9\left(P_9,T_9,v_9,z_{j,pre}\right)$$

$$v_9=v_9\left(T_9,P_9,z_{j,pre}\right)$$

9) Common header 1: [3]

$$c\cdot h_6\left(P_6,T_6,v_6,z_{j,pre}\right)+(1-c)\cdot h_{11}\left(P_{11},T_{11},v_{11},z_{j,pre}\right)$$

$$=h_{12}\left(P_{12},T_{12},v_{12},z_{j,pre}\right)\qquad P_6=P_{11},\ P_6=P_{12}$$

Mathematical Model of the Main Cooling Part of the DMR cycle

Composition of the refrigerant
Precooling: Ethane, Propane, n-Butane
Main cooling: Nitrogen, Methane, Ethane, Propane
Natural Gas: Methane(87.5%), Ethane(5.5%), Nitrogen(4.0%), Propane(2.1%), n-Butane(0.5%), i-Butane(0.3%), i-Pentane(0.1%)

1. Design variables(Operating Conditions)[107]

$$: P_i,\ T_i,\ v_i\left(i=1,\ ...,26,28,29,30\right), T_{S,1}, T_{S,11}, T_{S,13}, v_{S,1}, v_{S,11}, v_{S,13}, w_1,\ w_2,\ w_3,\ c,\ \dot{m}_{pre},\ \dot{m}_{main}, v\_f, z_{j,pre}\left(j=1,2,3\right), z_{k,main}\left(k=1,2,3,4\right)$$

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / $\dot{m}_{pre}$: mass flow rate at the precooling part
***Subscript 'NG': natural gas, Subscript 'main': main cooling part**
$*\ v\_f$: **vapor fraction at stream 16**

2. Equality constraints[91]

2.2 Equality constraints of Main cooling part[40]

10) Compressor 3: [6]

$$h_{26}\left(P_{26},T_{26},v_{26},z_{k,main}\right)+w_3=h\left(P_{13},T_{13},v_{13},z_{k,main}\right)$$

$$\eta=\frac{h_{S,13}\left(P_{13},T_{S,13},v_{S,13},z_{k,main}\right)-h_{26}\left(P_{26},T_{26},v_{26},z_{k,main}\right)}{h_{13}\left(P_{13},T_{13},v_{13},z_{k,main}\right)-h_{26}\left(P_{26},T_{26},v_{26},z_{k,main}\right)}$$

$$s_{26}\left(P_{26},T_{26},v_{26},z_{k,main}\right)=s_{13}\left(P_{13},T_{S,13},v_{S,13},z_{k,main}\right)$$

$$v_{26}=v_{26}(T_{26},P_{26},z_{k,main})$$

$$v_{S,13}=v_{13}(T_{S,13},P_{13},z_{k,main})$$

$$v_{13}=v_{13}(T_{13},P_{13},z_{k,main})$$

11) Sea water cooler 2: [3]

The temperature of the outlet of the sea water cooler is usually given.
T=305K

$$P_{13}=P_{14}\qquad v_{14}=v_{14}(T_{14},P_{14},z_{k,main})$$

12) Phase separator: [7]

$$h_{16}\left(P_{16},T_{16},v_{16},z_{k,main}\right)=v\_f\cdot h_{20}\left(P_{20},T_{20},v_{20},v\_f\cdot z_{k,main}\right)+\left(1-v\_f\right)\cdot h_{17}\left(P_{17},T_{17},v_{17},\left(1-v\_f\right)z_{k,main}\right)$$

$$P_{16}=P_{17},\ P_{16}=P_{20}\quad T_{17}=T_{20}\quad T_{16}=T_{17}$$

$$v_{17}=v_{17}\left(T_{17},P_{17},\left(1-v\_f\right)z_{k,main}\right),\ v_{20}=v_{20}\left(T_{20},P_{20},v\_f\cdot z_{k,main}\right)$$

13) Heat exchanger 3: [10]

$$v\_f\cdot\dot{m}_{main}\cdot h_{20}\left(P_{20},T_{20},v_{20},v\_f\cdot z_{k,main}\right)+\left(1-v\_f\right)\cdot\dot{m}_{main}\cdot h_{17}\left(P_{17},T_{17},v_{17},\left(1-v\_f\right)z_{k,main}\right)$$
$$+\dot{m}_{main}\cdot h_{25}\left(P_{25},T_{25},v_{25},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{29}\left(P_{29},T_{29},v_{29},z_{l,NG}\right)$$
$$=v\_f\cdot\dot{m}_{main}\cdot h_{21}\left(P_{21},T_{21},v_{21},v\_f\cdot z_{k,main}\right)+\left(1-v\_f\right)\cdot\dot{m}_{main}\cdot h_{18}\left(P_{18},T_{18},v_{18}\left(1-v\_f\right)z_{k,main}\right)$$
$$+\dot{m}_{main}\cdot h_{26}\left(P_{26},T_{26},v_{26},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{30}\left(P_{30},T_{30},v_{30},z_{l,NG}\right)$$

$$P_{20}=P_{21},\ P_{17}=P_{18},P_{25}=P_{26},P_{29}=P_{30}\qquad T_{21}=T_{30},\quad T_{21}=T_{18}$$

$$v_{21}=v_{21}\left(T_{21},P_{21},v\_f\cdot z_{k,main}\right),\ v_{18}=v_{18}\left(T_{18},P_{18},\left(1-v\_f\right)z_{k,main}\right),\ v_{30}=v_{30}\left(T_{30},P_{30},z_{l,NG}\right)$$

14) Expansion Valve 3: [2]

$$h_{18}\left(P_{18},T_{18},v_{18},(1-v\_f)z_{k,main}\right)=h_{19}\left(P_{19},T_{19},v_{19},(1-v\_f)z_{k,main}\right)$$

$$v_{19}=v_{19}\left(T_{19},P_{19}\right)$$

16) Expansion Valve 4: [2]

$$h_{22}\left(P_{22},T_{22},v_{22},v\_f\cdot z_{k,main}\right)=h_{23}\left(P_{23},T_{23},v_{23},v\_f\cdot z_{k,main}\right)$$

$$v_{23}=v_{23}\left(T_{23},P_{23},v\_f\cdot z_{k,main}\right)$$

15) Heat exchanger 4: [6]

$v\_f\cdot$ [redacted]

$=v\_f\cdot$ [redacted]

$$P_{21}=P_{22},\ P_{23}=P_{24},\ T_{22}=T_{31}\qquad v_{22}=v_{22}\left(T_{22},P_{22},v\_f\cdot z_{k,main}\right),\ v_{24}=v_{24}\left(T_{24},P_{24},v\_f\cdot z_{k,main}\right)$$

17) Common header 2: [4]

$$\left(1-v\_f\right)\cdot h_{19}\left(P_{19},T_{19},v_{19},\left(1-v\_f\right)z_{k,main}\right)$$
$$+v\_f\cdot h_{24}\left(P_{24},T_{24},v_{24},v\_f\cdot z_{k,main}\right)=h_{25}\left(P_{25},T_{25},v_{25},z_{k,main}\right)$$

$$P_{19}=P_{24},\ P_{19}=P_{25}\quad v_{25}=v_{25}(T_{25},P_{25},z_{k,main})$$

Procedure of the Determination of the Optimal Operating Conditions

Optimization Problem

1. Design Variables(Operating Conditions, 21)
 : $P_i, T_i, v_i, T_s, v_S, v_{4,l}, v_{4,v}$, v_f, $w, \dot{m}, q_L, q_H$ (i=1,2,3,4)

2. Equality constraints [20]
   1) Compressor (6)
   2) Condenser (4)
   3) Expansion valve (5)
   4) Evaporator (5)

3. Inequality constraints [1]

4. Objective function: Minimize the compressors power

$$\text{Minimize} \quad W = \dot{m} \cdot w$$

w : work input to the compressor per mass[J / kg]

$\dot{m}$: mass flow rate of refrigerant[kg / s]

c.f) Heat transfer in the evaporator

$$\dot{m} \cdot q_L = U \cdot A \cdot (T_C - T_4)$$

U : Heat transfer coefficients[W/m^{2}K]

A: Area of equipment temperature[m^2]

T_C: Room temperature[K]

Detailed sequence of finding optimum solution

1. Free variables [1 = 21 - 20]
 : P_1

2. Inequality constraints [1]

3. Objective function:
 Minimize $W = \dot{m} \cdot w$

1) Find the free variable.

Find the free variable P_1, by minimizing the compressor power subject to the inequality constraint using sequential quadratic programming(SQP) method.

2) Determine the dependent variables.

Determine the 20 dependent variables by solving the system of the nonlinear equations.

System of Nonlinear Equations
- Design variables [20]
- Equality constraints [20]

→ Determine the 20 variables by using Newton-Raphson method

3) Calculate objective function

Minimize $W = \dot{m} \cdot w$

- Optimization Method

Modified optimization problem

- **Free variables** [16 = 107 - 91]

  : $P_1, P_5, P_9, P_{13}, P_{19}, T_3, T_8, T_{18}, c, \dot{m}_{pre}, \dot{m}_{main},$

  $z_{C2,pre}, z_{C3,pre}, z_{N2,main}, z_{C1,main}, z_{C2,main}$
- **Inequality constraints** [11]
- **Objective function**

  $$\text{Minize } \dot{m}_{pre}\cdot w_1 + \dot{m}_{pre}\cdot w_2 + \dot{m}_{main}\cdot w_3$$

  → Use of sequential quadratic programming(SQP) method

Values of free variables ↓ ↑ Values of other design variables

System of nonlinear equations

- **Design variables** [91]
- **Equality constraints** [91]

  → Use of Newton-Rapson method

1. Design variables(operating conditions) [107]

$:P_i,\ T_i,\ v_i\ (i=1,\ ...,26,28,29,30),$

$T_{S,1}, T_{S,11}, T_{S,13}, v_{S,1}, v_{S,11}, v_{S,13}, w_1,\ w_2,\ w_3,\ c,$

$\dot{m}_{pre},\ \dot{m}_{main}, v\_f, z_{j,pre}\ (j=1,2,3), z_{k,main}\ (k=1,2,3,4)$

2. Equality constraints [91]

1) Composition of the refrigerant [2]
2) Precooling part [49]
3) Main cooling part [40]

3. Inequality constraints [11]

1) Minimum temperature approach in heat exchanger [8]
2) Inlet condition of the compressor [3]

4. Objective function: Minimize the compressors power

$$\text{Minize } \dot{m}_{pre}\cdot w_1 + \dot{m}_{pre}\cdot w_2 + \dot{m}_{main}\cdot w_3$$

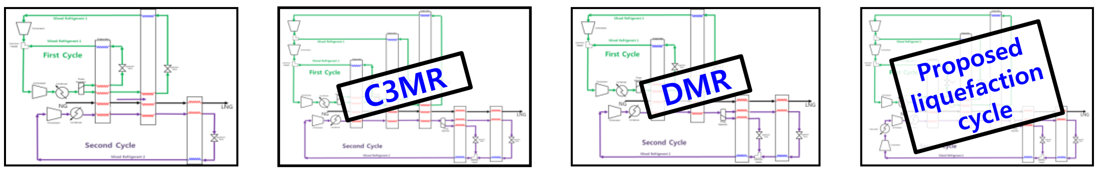
T: Temperature / ***P***: Pressure / ***v***: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / M_{pre}: mass flow rate at the precooling part
*Subscript 'NG': natural gas, Subscript 'main': main cooling part

Configuration of the developed Optimization Program of the Liquefaction Cycle for the LNG FPSO

Similar to Commercial Program "Aspen Hysys"

Optimization Program of the Liquefaction Cycle for the LNG FPSO

Liquefaction Cycles



…

Optimization Kernel

- SQP1)
- SLP2)
- Penalty Function Method
- …

Equipment Modules

- Compressor
- Evaporator
- Expansion Valve
- Heat Exchanger
- Condenser
- Tee
- Phase Separator
- Common Header
- …

Mathematical Models

- Energy Conservation (Specific Enthalpy)
- Criteria for Quality of Energy (Specific Entropy)
- Equation of State (SRK)
- Equation of State (PR)
- …

Physical Assumptions

- Adiabatic Process
- Isobaric Process
- Reversible Process
- Isentropic Process
- …

1) SQP: Sequential Quadratic Programming
2) SLP: Sequential Linear Programming

- Comparison between the Optimal Operating Conditions for the DMR Cycle Based on the Mathematical Model and the Past Relevant research

P[bar], T[K], v[m³/mol], w[J/mol], m[mol/s], W[kW]

Result obtained by this paper:

P1[bar]	**19.64**	P11	8.19	P21	48.92	Ts1	346.97
T1[K]	352.31	T11	313.47	T21	140.36	Ts11	306.24
v1[m3/mol]	0.001208	v11	0.002846	v21	0.000043	Ts13	395.94
P2	19.64	P12	8.19	P22	48.92	vs1	0.001467
T2	310.00	T12	307.89	T22	113.00	vs11	0.003234
v2	0.000092	v12	0.002774	v22	0.000039	vs13	0.000625
P3	19.64	P13	**48.92**	P23	**2.79**	w1[J/mol]	2505.86
T3	275.01	T13	422.24	T23	105.80	w2	1187.98
v3	0.000081	v13	0.000669	v23	0.000360	w3	8746.96
P4	19.64	P14	48.92	P24	2.79	c	**0.584643**
T4	275.01	T14	305.00	T24	137.74	$\dot{m}_{pre}$[mol/s]	**0.932866**
v4	0.000081	v14	0.000379	v24	0.003016	$\dot{m}_{main}$	**0.957021**
P5	**8.19**	P15	48.92	P25	2.79	zpre_Ethane	**0.253895**
T5	272.01	T15	275.01	T25	137.41	zpre_Propane	**0.63883**
v5	0.000142	v15	0.000242	v25	0.001016	zpre_n-Butane	0.107275
P6	8.19	P16	48.92	P26	2.79	zmain_Nitrogen	**0.069317**
T6	303.97	T16	239.64	T26	237.65	zmain_Methane	**0.405874**
v6	0.002722	v16	0.000131	v26	0.006835	zmain_Ethane	**0.2964**
P7	19.64	P17	48.92	P28	65.00	zmain_Propaane	0.228409
T7	275.01	T17	239.64	T28	**275.01**	Objective Function(work) [kW]	11.817
v7	0.000081	v17	0.000068	v28	0.000290		
P8	19.64	P18	48.92	P29	65.00		
T8	239.64	T18	140.36	T29	**239.64**		
v8	0.000074	v18	0.000050	v29	0.000209		
P9	**2.86**	P19	2.79	P30	65.00		
T9	236.58	T19	136.08	T30	**140.36**		
v9	0.000232	v19	0.000344	v30	0.000042		
P10	2.86	P20	48.92				
T10	265.92	T20	239.64				
v10	0.007301	v20	0.000311				

→ The is the result of the optimization for the DMR cycle.

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

Comparison of the calculation result of this study and a relevant research

P[bar], T[K], v[m³/mol], w[J/mol], m[mol/s], W[kW]

Result obtained by this study:

P1[bar]	19.6	P11	8.2	P21	48.9	Ts1	347.4
T1[K]	352.7	T11	313.9	T21	140.4	Ts11	306.7
v1[m3/mol]	0.001209	v11	0.002846	v21	0.000043	Ts13	394.8
P2	19.6	P12	8.2	P22	48.9	vs1	0.001173
T2	310.0	T12	308.3	T22	113.0	vs11	0.002753
v2	0.000092	v12	0.002774	v22	0.000039	vs13	0.000612
P3	19.6	P13	48.9	P23	2.8	w1[J/mol]	2513.4
T3	275.0	T13	421.0	T23	105.7	W2	1190.3
v3	0.000081	v13	0.000669	v23	0.000361	W3	8707.3
P4	19.6	P14	48.9	P24	2.8	C	0.5846
T4	275.0	T14	305.0	T24	137.9	$\dot{m}_{pre}$[mol/s]	0.9329
v4	0.000081	v14	0.000379	v24	0.003018	$\dot{m}_{main}$[mol/s]	0.9570
P5	8.2	P15	48.9	P25	2.8	zpre_Ethane	0.2539
T5	272.0	T15	275.0	T25	137.3	zpre_Propane	0.6388
v5	0.000142	v15	0.000242	v25	0.001017	zpre_n-Butane	0.1073
P6	8.2	P16	48.9	P26	2.8	zmain_Nitrogen	0.0693
T6	304.2	T16	239.6	T26	236.6	zmain_Methane	0.4059
v6	0.002722	v16	0.000131	v26	0.006841	zmain_Ethane	0.2964
P7	19.6	P17	48.9	P28	65.0	zmain_Propaane	0.2284
T7	275.0	T17	239.6	T28	275.0	Objective Function(work) [kW]	11.788
v7	0.000081	v17	0.000068	v28	0.000290		
P8	19.6	P18	48.9	P29	65.0		
T8	239.6	T18	140.4	T29	239.6		
v8	0.000075	v18	0.000050	v29	0.000209		
P9	2.9	P19	2.8	P30	65.0		
T9	236.6	T19	136.0	T30	140.4		
v9	0.000232	v19	0.000344	v30	0.000042		
P10	2.9	P20	48.9				
T10	266.4	T20	239.6				
v10	0.007297	v20	0.000311				

free variables

Result obtained by Venkatarathnam[1]:

[%]: Difference

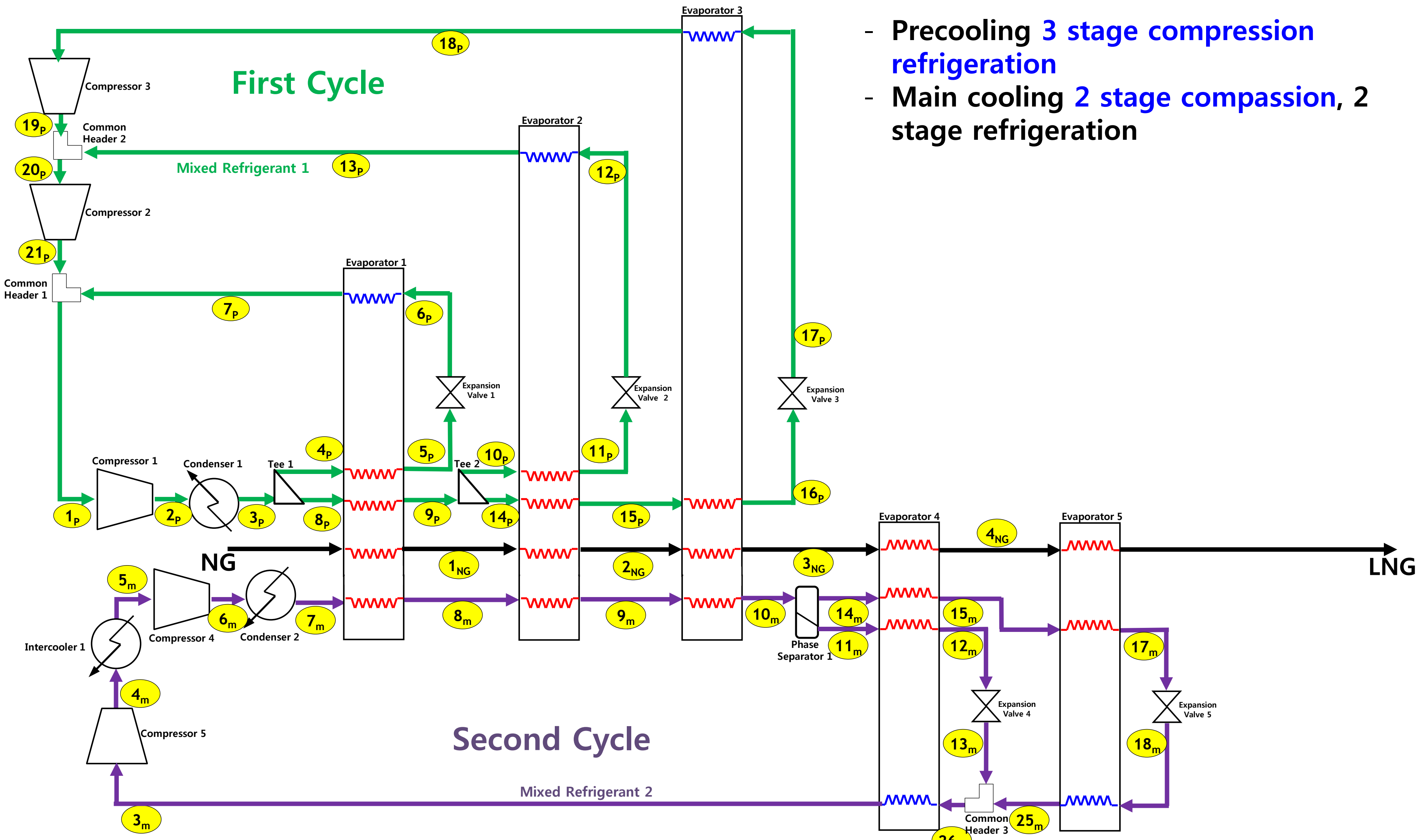
		[%			[%			[%			[%
P1[bar]	19.2	2.29%	P11	7.6	7.76%	P21	48.6	0.66%	Ts1	354.4	1.98%
T1[K]	360.2	2.09%	T11	313.6	0.10%	T21	144.7	3.00%	Ts11	306.6	0.01%
v1[m3/mol]	0.001291	6.42%	v11	0.003089	7.88%	v21	0.000045	3.40%	Ts13	392.3	0.62%
P2	19.2	2.29%	P12	7.6	7.76%	P22	48.6	0.66%	vs1	0.001253	6.40%
T2	310.0	0.00%	T12	313.8	1.76%	T22	113.0	0.00%	vs11	0.002995	8.08%
v2	0.000122	24.92%	v12	0.003092	10.29%	v22	0.000040	1.03%	vs13	0.000613	0.24%
P3	19.2	2.29%	P13	48.6	0.66%	P23	3.0	7.00%	w1[J/mol]	2767.7	9.19%
T3	273.1	0.70%	T13	418.1	0.68%	T23	106.9	1.08%	w2	1103.9	7.83%
v3	0.000087	6.32%	v13	0.000669	0.11%	v23	0.000304	18.52%	w3	8441.3	3.15%
P4	19.2	2.29%	P14	48.6	0.66%	P24	3.0	7.00%	c	0.5980	2.24%
T4	273.1	0.70%	T14	305	0.00%	T24	141.78668 56	2.76%	$\dot{m}_{pre}$[mol/s]	0.9130	2.18%
v4	0.000087	6.32%	v14	0.000389	2.49%	v24	0.002911	3.67%	$\dot{m}_{main}$[mol/s]	1.0000	4.30%
P5	7.6	7.76%	P15	48.6	0.66%	P25	3.0	7.00%	zpre_Ethane	0.2482	2.32%
T5	269.7	0.84%	T15	273.1	0.70%	T25	140.3	2.14%	zpre_Propane	0.6416	0.43%
v5	0.000159	10.74%	v15	0.000248	2.54%	v25	0.001090	6.73%	zpre_n-Butane	0.1103	2.72%
P6	7.6	7.76%	P16	48.6	0.66%	P26	3.0	7.00%	zmain_Nitrogen	0.0700	0.98%
T6	313.9	3.09%	T16	240.0	0.15%	T26	237.0	0.15%	zmain_Methane	0.4180	2.90%
v6	0.003093	12.02%	v16	0.000141	7.04%	v26	0.006363	7.50%	zmain_Ethane	0.2990	0.87%
P7	19.2	2.29%	P17	48.6	0.66%	P28	65.0	0.00%	zmain_Propaane	0.2130	7.23%
T7	273.0	0.74%	T17	240.0	0.15%	T28	273.1	0.70%	Objective Function(work) [kW]	11.976	1.57%
v7	0.000087	6.32%	v17	0.000071	4.25%	v28	0.000286	1.34%			
P8	19.2	2.29%	P18	48.6	0.66%	P29	65.0	0.00%			
T8	240.0	0.15%	T18	144.7	3.00%	T29	240.0	0.15%			
v8	0.000079	5.82%	v18	0.000053	4.59%	v29	0.000210	0.45%			
P9	2.8	2.14%	P19	3.0	7.00%	P30	65.0	0.00%			
T9	236.5	0.03%	T19	139.1	2.29%	T30	144.7	3.00%			
v9	0.000258	10.17%	v19	0.000368	6.51%	v30	0.000044	2.99%			
P10	2.8	2.14%	P20	48.6	0.66%						
T10	268.7	0.86%	T20	240.0	0.15%						
v10	0.007537	3.18%	v20	0.000312	0.47%						

free variables

→ The result of the optimal operating condition of the DMR cycle obtained by this study saves 1.57 % of the total required power consumption compared with the relevant research.

Reference: 1) Venkatarathnam, G., 2008, Cryogenic Mixed Refrigerant Processes, Springer, New York

CASE 23: Proposed Liquefaction Cycle



- Precooling **3 stage compression refrigeration**
- Main cooling **2 stage compassion**, 2 stage refrigeration

Mathematical Model of the Proposed Liquefaction Cycle (Case 23) (1)

1. Design variables(Operating Conditions) [153]

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\dots,21_p,3_m,\dots,15_m,17_m,18_m,25_m,26_m,1_{NG},\dots,4_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant

***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

2. Equality constraints [133]

2.1 Equality constraints of precooling part [83]

1) Compressor 1: [6]

$$h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)+w_1=h_{2p}\left(P_{2p},T_{2p},v_{2p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,2p}\left(P_{2p},T_{S,2p},v_{S,2p},z_{j,pre}\right)-h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)}{h_{2p}\left(P_{2p},T_{2p},v_{2p},z_{j,pre}\right)-h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)}$$

$$s_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)=s_{2p}\left(P_{2p},T_{S,2p},v_{S,2p},z_{j,pre}\right)$$

$$v_{1p}=v_{1p}\left(P_{1p},T_{1p},z_{j,pre}\right)$$

$$v_{S,2p}=v_{S,2p}\left(P_{2p},T_{S,2p},z_{j,pre}\right)$$

$$v_{2p}=v_{2p}\left(P_{2p},T_{2p},z_{j,pre}\right)$$

2) Condenser 1: [3]

The temperature of the outlet of the sea water cooler is usually given. T=310K

$$P_{2p}=P_{3p}$$

$$v_{3p}=v_{3p}(T_{3p},P_{3p},z_{j,pre})$$

3) Tee 1: [6]

$$h_{3p}\left(P_{3p},T_{3p},v_{3p},z_{j,pre}\right)$$
$$=c_1\cdot h_{4p}\left(P_{4p},T_{4p},v_{4p},z_{j,pre}\right)+\left(1-c_1\right)\cdot h_{8p}\left(P_{8p},T_{8p},v_{8p},z_{j,pre}\right)$$

$$P_{3p}=P_{4p},\ \ P_{3p}=P_{8p}$$

$$T_{4p}=T_{8p}$$

$$v_{4p}=v_{4p}\left(T_{4p},P_{4p},z_{j,pre}\right),\ v_{8p}=v_{8p}\left(T_{8p},P_{8p},z_{j,pre}\right)$$

4) Evaporator 1: [14]

$$c_1\cdot\dot{m}_{pre}\cdot h_{4p}\left(P_{4p},T_{4p},v_{4p},z_{j,pre}\right)+c_1\cdot\dot{m}_{pre}\cdot h_{6p}\left(P_{6p},T_{6p},v_{6p},z_{j,pre}\right)$$
$$+\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{8p}\left(P_{8p},T_{8p},v_{8p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{7m}\left(P_{7m},T_{7m},v_{7m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{NG}\left(P_{NG},T_{NG},v_{NG},z_{l,NG}\right)$$
$$=c_1\cdot\dot{m}_{pre}\cdot h_{5p}\left(P_{5p},T_{5p},v_{5p},z_{j,pre}\right)+c_1\cdot\dot{m}_{pre}\cdot h_{7p}\left(P_{7p},T_{7p},v_{7p},z_{j,pre}\right)$$
$$+\left(1-c_1\right)\cdot\dot{m}_{pre}\cdot h_{9p}\left(P_{9p},T_{9p},v_{9p},z_{j,pre}\right)+\dot{m}_{main}\cdot h_{8m}\left(P_{8m},T_{8m},v_{8m},z_{k,main}\right)$$
$$+\dot{m}_{NG}\cdot h_{1NG}\left(P_{1NG},T_{1NG},v_{1NG},z_{l,NG}\right)$$

$$P_{4p}=P_{5p},\ P_{6p}=P_{7p},P_{8p}=P_{9p},P_{7m}=P_{8m},P_{NG}=P_{1NG}$$

$$T_{5p}=T_{9p},T_{5p}=T_{8m},T_{5p}=T_{1NG}$$

$$v_{5p}=v_{5p}\left(T_{5p},P_{5p},z_{j,pre}\right),\ v_{7p}=v_{7p}\left(T_{7p},P_{7p},z_{j,pre}\right),$$

$$v_{9p}=v_{9p}\left(T_{9p},P_{9p},z_{j,pre}\right),\ v_{8m}=v_{8m}\left(T_{8m},P_{8m},z_{k,main}\right),$$

$$v_{1NG}=v_{1NG}\left(T_{1NG},P_{1NG},z_{l,NG}\right)$$

5) Expansion valve 1: [2]

$$h_{5p}\left(P_{5p},T_{5p},v_{5p},z_{j,pre}\right)=h_{6p}\left(P_{6p},T_{6p},v_{6p},z_{j,pre}\right)$$

$$v_{6p}=v_{6p}\left(T_{6p},P_{6p},z_{j,pre}\right)$$

6) Tee 2: [6]

$$\left(1-c_1\right)\cdot h_{9p}\left(P_{9p},T_{9p},v_{9p},z_{j,pre}\right)$$
$$=c_2\cdot\left(1-c_1\right)\cdot h_{10p}\left(P_{10p},T_{10p},v_{10p},z_{j,pre}\right)$$
$$+\left(1-c_2\right)\cdot\left(1-c_1\right)\cdot h_{14p}\left(P_{14p},T_{14p},v_{14p},z_{j,pre}\right)$$

$$P_{9p}=P_{10p},\ \ P_{9p}=P_{14p}$$

$$T_{10p}=T_{14p}$$

$$v_{10p}=v_{10p}\left(T_{10p},P_{10p},z_{j,pre}\right),$$

$$v_{14p}=v_{14p}\left(T_{14p},P_{14p},z_{j,pre}\right)$$

Mathematical Model of the Proposed Liquefaction Cycle (Case 23) (2)

1. Design variables(Operating Conditions) [153]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:
$$P_i,\ T_i,\ v_i\left(i=1_p,...,21_p,3_m,...,15_m,17_m,18_m,25_m,26_m,1_{NG},...,4_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [133]

2.1 Equality constraints of precooling part [83]

7) Evaporator 2: [14]

$$c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{10p}\left(P_{10p},T_{10p},v_{10p},z_{j,pre}\right)+c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{12p}\left(P_{12p},T_{12p},v_{12p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{14p}\left(P_{14p},T_{14p},v_{14p},z_{j,pre}\right)+$$
$$+\dot{m}_{main}\cdot h_{8m}\left(P_{8m},T_{8m},v_{8m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{1NG}\left(P_{1NG},T_{1NG},v_{1NG},z_{l,NG}\right)$$
$$=c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{11p}\left(P_{11p},T_{11p},v_{11p},z_{j,pre}\right)+c_2\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{13p}\left(P_{13p},T_{13p},v_{13p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{15p}\left(P_{15p},T_{15p},v_{15p},z_{j,pre}\right)+$$
$$+\dot{m}_{main}\cdot h_{9m}\left(P_{9m},T_{9m},v_{9m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{2NG}\left(P_{2NG},T_{2NG},v_{2NG},z_{l,NG}\right)$$

$$P_{10p}=P_{11p},\ P_{12p}=P_{13p},P_{14p}=P_{15p},P_{8m}=P_{9m},P_{1NG}=P_{2NG}$$

$$T_{11p}=T_{15p},T_{11p}=T_{9m},T_{11p}=T_{2NG}$$

$$v_{11p}=v_{11p}\left(T_{11p},P_{11p},z_{j,pre}\right),\ v_{13p}=v_{13p}\left(T_{13p},P_{13p},z_{j,pre}\right),$$
$$v_{15p}=v_{15p}\left(T_{15p},P_{15p},z_{j,pre}\right),\ v_{9m}=v_{9m}\left(T_{9m},P_{9m},z_{k,main}\right),$$
$$v_{2NG}=v_{2NG}\left(T_{2NG},P_{2NG},z_{l,NG}\right)$$

8) Expansion valve 2: [2]

$$h_{11p}\left(P_{11p},T_{11p},v_{11p},z_{j,pre}\right)=h_{12p}\left(P_{12p},T_{12p},v_{12p},z_{j,pre}\right)$$
$$v_{12p}=v_{12p}\left(T_{12p},P_{12p},z_{j,pre}\right)$$

9) Evaporator 3: [11]

$$(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{15p}\left(P_{15p},T_{15p},v_{15p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{17p}\left(P_{17p},T_{17p},v_{17p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{9m}\left(P_{9m},T_{9m},v_{9m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{2NG}\left(P_{2NG},T_{2NG},v_{2NG},z_{l,NG}\right)$$
$$=(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{16p}\left(P_{16p},T_{16p},v_{16p},z_{j,pre}\right)$$
$$+(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)$$
$$+\dot{m}_{main}\cdot h_{10m}\left(P_{10m},T_{10m},v_{10m},z_{k,main}\right)+\dot{m}_{NG}\cdot h_{3NG}\left(P_{3NG},T_{3NG},v_{3NG},z_{l,NG}\right)$$

$$P_{15p}=P_{16p},P_{17p}=P_{18p},P_{9m}=P_{10m},P_{2NG}=P_{3NG}$$

$$T_{16p}=T_{10m},T_{16p}=T_{3NG}$$

$$v_{16p}=v_{16p}\left(T_{16p},P_{16p},z_{j,pre}\right),\ v_{18p}=v_{18p}\left(T_{18p},P_{18p},z_{j,pre}\right),$$
$$v_{10m}=v_{10m}\left(T_{10m},P_{10m},z_{k,main}\right),\ v_{3NG}=v_{3NG}\left(T_{3NG},P_{3NG},z_{l,NG}\right)$$

10) Expansion valve 3: [2]

$$h_{16p}\left(P_{16p},T_{16p},v_{16p},z_{j,pre}\right)=h_{17p}\left(P_{17p},T_{17p},v_{17p},z_{j,pre}\right)$$
$$v_{17p}=v_{17p}\left(T_{17p},P_{17p},z_{j,pre}\right)$$

Mathematical Model of the Proposed Liquefaction Cycle (Case 23) (3)

1. Design variables(Operating Conditions) [153]

:

$$P_i,\ T_i,\ v_i\left(i=1_p,...,21_p,3_m,...,15_m,17_m,18_m,25_m,26_m,1_{NG},...,4_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

***T*: Temperature / *P*: Pressure / *v*: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant**

***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

2. Equality constraints [133]

2.1 Equality constraints of precooling part [83]

11) Compressor 3: [5]

$$(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)+w_3 =(1-c_2)\cdot(1-c_1)\cdot\dot{m}_{pre}\cdot h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,19p}\left(P_{19p},T_{S,19p},v_{S,19p},z_{j,pre}\right)-h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)}{h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right)-h_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)}$$

$$s_{18p}\left(P_{18p},T_{18p},v_{18p},z_{j,pre}\right)=s_{19p}\left(P_{19p},T_{S,19p},v_{S,19p},z_{j,pre}\right)$$

$$v_{19p}=v_{19p}\left(P_{19p},T_{19p},z_{j,pre}\right)$$

$$v_{S,19p}=v_{S,19p}\left(P_{19p},T_{S,19p},z_{j,pre}\right)$$

12) Common Header 2: [4]

$$c_2\cdot(1-c_1)\cdot h_{13p}\left(P_{13p},T_{13p},v_{13p},z_{j,pre}\right)+(1-c_2)\cdot(1-c_1)\cdot h_{19p}\left(P_{19p},T_{19p},v_{19p},z_{j,pre}\right) =(1-c_1)\cdot h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)$$

$$P_{13p}=P_{19p},\ P_{13p}=P_{20p}$$

$$v_{20p}=v_{20p}\left(P_{20p},T_{20p},z_{j,pre}\right)$$

13) Compressor 2: [5]

$$(1-c_1)\cdot\dot{m}_{pre}\cdot h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)+w_2 =(1-c_1)\cdot\dot{m}_{pre}\cdot h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)$$

$$\eta=\frac{h_{S,21p}\left(P_{21p},T_{S,21p},v_{S,21p},z_{j,pre}\right)-h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)}{h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)-h_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)}$$

$$s_{20p}\left(P_{20p},T_{20p},v_{20p},z_{j,pre}\right)=s_{21p}\left(P_{21p},T_{S,21p},v_{S,21p},z_{j,pre}\right)$$

$$v_{21p}=v_{21p}\left(P_{21p},T_{21p},z_{j,pre}\right)$$

$$v_{S,21p}=v_{S,21p}\left(P_{21p},T_{S,21p},z_{j,pre}\right)$$

14) Common Header 1: [3]

$$(1-c_1)\cdot h_{21p}\left(P_{21p},T_{21p},v_{21p},z_{j,pre}\right)+c_1\cdot h_{7p}\left(P_{7p},T_{7p},v_{7p},z_{j,pre}\right) =h_{1p}\left(P_{1p},T_{1p},v_{1p},z_{j,pre}\right)$$

$$P_{7p}=P_{21p},\ P_{7p}=P_{1p}$$

Mathematical Model of the Proposed Liquefaction Cycle (Case 23) (4)

1. Design variables(Operating Conditions) [153]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,\text{pre}}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:

$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,3_m,\ldots,15_m,17_m,18_m,25_m,26_m,1_{NG},\ldots,4_{NG}\right),$$

$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$

$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [133]

2.2 Equality constraints of main cooling part [48]

1) Compressor 4: [5]

$$h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)+w_4=h_{6m}\left(P_{6m},T_{6m},v_{6m},z_{k,main}\right)$$

$$\eta=\frac{h_{S,6m}\left(P_{6m},T_{S,6m},v_{S,6m},z_{k,main}\right)-h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)}{h_{6m}\left(P_{6m},T_{6m},v_{6m},z_{k,main}\right)-h_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)}$$

$$s_{5m}\left(P_{5m},T_{5m},v_{5m},z_{k,main}\right)=s_{6m}\left(P_{6m},T_{S,6m},v_{S,6m},z_{k,main}\right)$$

$$v_{S,6m}=v_{S,6m}\left(P_{6m},T_{S,6m},z_{k,main}\right)$$

$$v_{6m}=v_{6m}\left(P_{6m},T_{6m},z_{k,main}\right)$$

2) Condenser 2: [3]

The temperature of the outlet of the sea water cooler is usually given.
T=305K

$$P_{6m}=P_{7m}$$

$$v_{7m}=v_{7m}(T_{7m},P_{7m},z_{k,main})$$

3) Phase Separator 1: [7]

$$\begin{aligned}&h_{10m}\left(P_{10m},T_{10m},v_{10m},z_{k,main}\right)\\&=v\_f_{10}\cdot h_{14m}\left(P_{14m},T_{14m},v_{14m},v\_f_{10}\cdot z_{k,main}\right)\\&\qquad+\left(1-v\_f_{10}\right)\cdot h_{11m}\left(P_{11m},T_{11m},v_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)\end{aligned}$$

$$P_{10m}=P_{11m},\ P_{10m}=P_{14m}$$

$$T_{10m}=T_{11m},\ T_{11m}=T_{14m}$$

$$v_{11m}=v_{11m}\left(P_{11m},T_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right),\ v_{14m}=v_{14m}\left(P_{14m},T_{14m},v\_f_{10}\cdot z_{k,main}\right)$$

4) Evaporator 4: [10]

$$\begin{aligned}&\left(1-v\_f_{10}\right)\cdot\dot{m}_{main}\cdot h_{11m}\left(P_{11m},T_{11m},v_{11m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)\\&+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{14m}\left(P_{14m},T_{14m},v_{14m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{main}\cdot h_{26m}\left(P_{26m},T_{26m},v_{26m},z_{k,main}\right)\\&+\dot{m}_{NG}\cdot h_{3NG}\left(P_{3NG},T_{3NG},v_{3NG},z_{l,NG}\right)\\&=\left(1-v\_f_{10}\right)\cdot\dot{m}_{main}\cdot h_{12m}\left(P_{12m},T_{12m},v_{12m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right)\\&\quad+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{15m}\left(P_{15m},T_{15m},v_{15m},v\_f_{10}\cdot z_{k,main}\right)+h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)\\&\quad+\dot{m}_{NG}\cdot h_{4NG}\left(P_{4NG},T_{4NG},v_{4NG},z_{l,NG}\right)\end{aligned}$$

$$P_{11m}=P_{12m},P_{14m}=P_{15m},P_{26m}=P_{3m},P_{3NG}=P_{4NG}$$

$$T_{12m}=T_{15m},T_{12m}=T_{4NG}$$

$$v_{12m}=v_{12m}\left(T_{12m},P_{12m},\left(1-v\_f_{10}\right)\cdot z_{k,main}\right),\ v_{15m}=v_{15m}\left(T_{15m},P_{15m},v\_f_{10}\cdot z_{k,main}\right),$$

$$v_{4NG}=v_{4NG}\left(T_{4NG},P_{4NG},z_{l,NG}\right)$$

Mathematical Model of the Proposed Liquefaction Cycle (Case 23) (5)

1. Design variables(Operating Conditions) [153]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:
$$P_i,\ T_i,\ v_i\left(i=1_p,...,21_p,3_m,...,15_m,17_m,18_m,25_m,26_m,1_{NG},...,4_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [133]

2.2 Equality constraints of main cooling part [48]

5) Expansion valve 4: [2]

$$h_{12m}\left(P_{12m},T_{12m},v_{12m},(1-v\_f_{10})\cdot z_{k,main}\right)=h_{13m}\left(P_{13m},T_{13m},v_{13m},(1-v\_f_{10})\cdot z_{k,main}\right)$$
$$v_{13m}=v_{13m}\left(P_{13m},T_{13m},(1-v\_f_{10})\cdot z_{k,main}\right)$$

6) Evaporator 5: [6]

$$\begin{aligned}&v\_f_{10}\cdot\dot{m}_{main}\cdot h_{15m}\left(P_{15m},T_{15m},v_{15m},v\_f_{10}\cdot z_{k,main}\right)\\&+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{18m}\left(P_{18m},T_{18m},v_{18m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{NG}\cdot h_{4NG}\left(P_{4NG},T_{4NG},v_{4NG},z_{l,NG}\right)\\&=v\_f_{10}\cdot\dot{m}_{main}\cdot h_{17m}\left(P_{17m},T_{17m},v_{17m},v\_f_{10}\cdot z_{k,main}\right)\\&\quad+v\_f_{10}\cdot\dot{m}_{main}\cdot h_{25m}\left(P_{25m},T_{25m},v_{25m},v\_f_{10}\cdot z_{k,main}\right)+\dot{m}_{NG}\cdot h_{LNG}\left(P_{LNG},T_{LNG},v_{LNG},z_{l,NG}\right)\end{aligned}$$
$$P_{15m}=P_{17m},P_{18m}=P_{25m}$$
$$T_{17m}=T_{LNG}$$
$$v_{17m}=v_{17m}\left(P_{17m},T_{17m},v\_f_{10}\cdot z_{k,main}\right),\ v_{25m}=v_{25m}\left(P_{25m},T_{25m},v\_f_{10}\cdot z_{k,main}\right)$$

7) Common Header 3: [4]

$$\begin{aligned}&(1-v\_f_{10})\cdot h_{13m}\left(P_{13m},T_{13m},v_{13m},(1-v\_f_{10})\cdot z_{k,main}\right)+v\_f_{10}\cdot h_{25m}\left(P_{25m},T_{25m},v_{25m},v\_f_{10}\cdot z_{k,main}\right)\\&\qquad=h_{26m}\left(P_{26m},T_{26m},v_{26m},z_{k,main}\right)\end{aligned}$$
$$P_{13m}=P_{25m},\ P_{13m}=P_{26m}$$
$$v_{26m}=v_{26m}\left(P_{26m},T_{26m},z_{k,main}\right)$$

8) Expansion valve 5: [2]

$$h_{17m}\left(P_{17m},T_{17m},v_{17m},v\_f_{10}\cdot z_{k,main}\right)=h_{18m}\left(P_{18m},T_{18m},v_{18m},v\_f_{10}\cdot z_{k,main}\right)$$
$$v_{18m}=v_{18m}\left(P_{18m},T_{18m},v\_f_{10}\cdot z_{k,main}\right)$$

9) Compressor 5: [6]

$$h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)+w_5=h_{4m}\left(P_{4m},T_{4m},v_{4m},z_{k,main}\right)$$
$$\eta=\frac{h_{S,4m}\left(P_{4m},T_{S,4m},v_{S,4m},z_{k,main}\right)-h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)}{h_{4m}\left(P_{4m},T_{4m},v_{4m},z_{k,main}\right)-h_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)}$$
$$s_{3m}\left(P_{3m},T_{3m},v_{3m},z_{k,main}\right)=s_{4m}\left(P_{4m},T_{S,4m},v_{S,4m},z_{k,main}\right)$$
$$v_{S,4m}=v_{S,4m}\left(P_{4m},T_{S,4m},z_{k,main}\right)$$
$$v_{4m}=v_{4m}\left(P_{4m},T_{4m},z_{k,main}\right)$$
$$v_{3m}=v_{3m}\left(P_{3m},T_{3m},z_{k,main}\right)$$

10) Intercooler 1: [3]

The temperature of the outlet of the sea water cooler is usually given. T=305K

$$P_{4m}=P_{5m}$$
$$v_{5m}=v_{5m}(T_{5m},P_{5m},z_{k,main})$$

$$\sum_{j=1}^{3}z_{j,pre}=1,\quad\sum_{k=1}^{4}z_{k,main}=1$$

Summary of the Mathematical Model of the Proposed Liquefaction Cycle (Case 23)

1. Design variables(Operating Conditions) [153]

T: Temperature / P: Pressure / v: Specific volume / $z_{j,pre}$: mole fraction of the component j at the precooling part/ w: work input to the compressor per mass/ c: flow rate ratio between inlet and outlet 4 / m_{pre}: mass flow rate at the precooling refrigerant
***Subscript 'NG': natural gas, Subscript 'main': main cooling refrigerant**

:
$$P_i,\ T_i,\ v_i\left(i=1_p,\ldots,21_p,3_m,\ldots,15_m,17_m,18_m,25_m,26_m,1_{NG},\ldots,4_{NG}\right),$$
$$T_{S,2p},T_{S,19p},T_{S,21p},T_{S,4m},T_{S,6m},v_{S,2p},v_{S,19p},v_{S,21p},v_{S,4m},v_{S,6m},$$
$$w_1,\ w_2,\ w_3,w_4,\ w_5,c_1,c_2,\dot{m}_{pre},\ \dot{m}_{main},v\_\ f_{10},z_{j,pre}\left(j=1,2,3\right),z_{k,main}\left(k=1,2,3,4\right)$$

2. Equality constraints [133]

2.1 Equality constraints of precooling part [83]

2.2 Equality constraints of main cooling part [48]

→ indeterminate systems

3. Objective Function: Minimize the compressors power

$$\text{Minize}\ \ \dot{m}_{pre}\cdot w_1+\dot{m}_{pre}\cdot w_2+\dot{m}_{pre}\cdot w_3+\dot{m}_{main}\cdot w_4+\dot{m}_{main}\cdot w_5$$


→ Optimization Problem!

4. Free variables [20= 153 – 133]

:
$$P_{1p},\ P_{2p},\ P_{12p},\ P_{17p},\ T_{5p},\ T_{11p},\ T_{16p},\ c_1,\ c_2,\ z_{1,pre},\ z_{2,pre},\ \dot{m}_{pre},$$
$$P_{3m},\ P_{4m},\ P_{6m},\ T_{12m},\ z_{1,main},\ z_{2,main},\ z_{3,main},\ \dot{m}_{main}$$

Comparison between the optimal result of the proposed Cycle and DMR Cycle

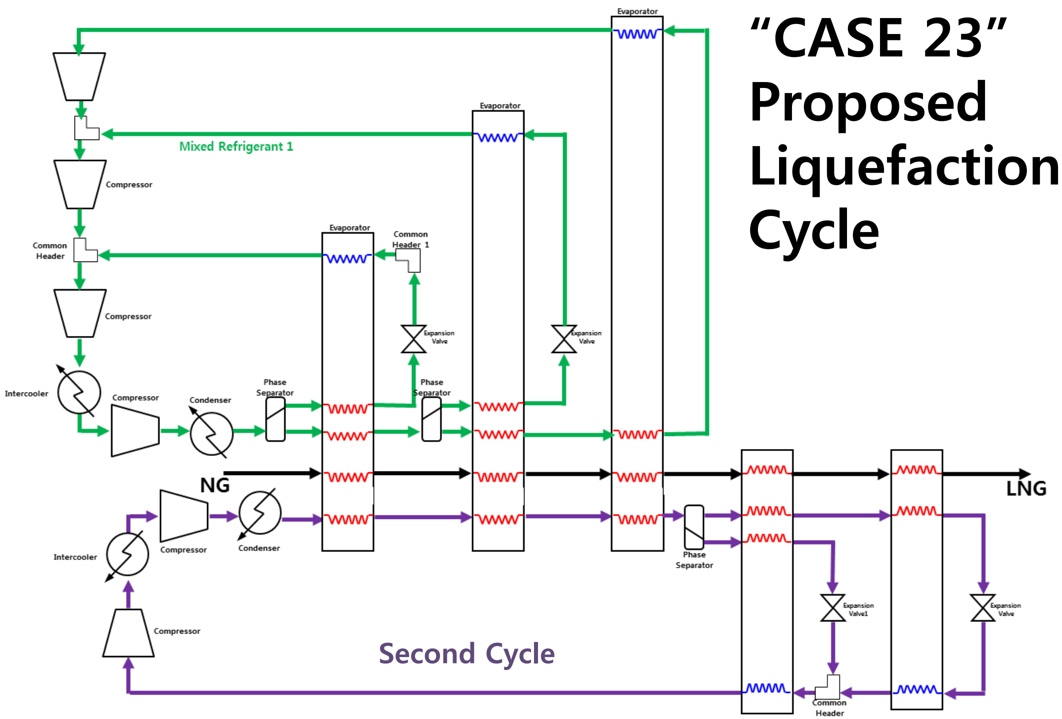
- After the optimal operating conditions for the two cycles are achieved, the power required by the compressors for the two cycles are compared.

- Common condition

1) Given: NG T=26.85°C, P=65 bar,
   LNG T=-160.15°C, P=65 bar
   $\dot{m}_{NG} = 49.21\ kg/h\ (= 0.0004\ MMTA)$

2) Refrigerant:
   - Mixed refrigerant 1 is composed of Ethane(C_2H_6), Propane(C_3H_8), n-Butane(C_4H_{10}) for precooling
   - Mixed refrigerant 2 is composed of Nitrogen(N_2), Methane(C_1H_4), Ethane(C_2H_6), Propane(C_3H_8) for main cooling

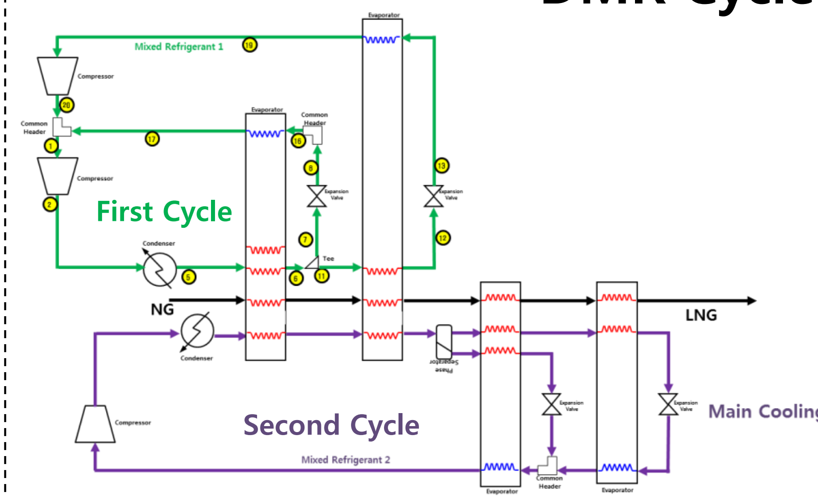
"CASE 23" Proposed Liquefaction Cycle



[Result of the proposed iquefaction Cycle]

- Precooling 3 stage compression refrigeration
- Main cooling 2 stage compassion, 2 stage refrigeration

"CASE 13" DMR Cycle



[Configuration of the Dual Mixed Refrigerant Cycle[1)]]

- Precooling 2 stage compression refrigeration
- Main cooling 1 stage compression, 2 stage refrigeration

Total required power consumption of the compressors and pumps : 11.359 kW

Total required power consumption of the compressors[2)] : 11.976 kW

→ The result of the synthesis of the liquefaction cycle saves 5.2 % of the total required power consumption.

Reference: 1) Venkatarathnam, G., Cryogenic Mixed Refrigerant Processes, Springer, New York, 2008.
2) K. Y. Lee, J. H. Cha, J. C. Lee, M. I. Roh, and J. H. Hwang, Determination of the Optimal Operating Condition of Dual Mixed Refrigerant Cycle at the Pre-FEED stage of LNG FPSO Topside Liquefaction Process, ISOPE Conference, June 2011.

2011년 2학기 전산선박설계 강의자료
(Computer Aided Ship Design Lecture Note)

Part II. Curve and Surface Modeling

서울대학교 조선해양공학과

선박설계자동화연구실

이규열

CONTENTS

Naval Architecture & Ocean Engineering

Computer Aided Ship design

-Part II. Curve and Surface Modeling-

November, 2011

Prof. Kyu-Yeul Lee

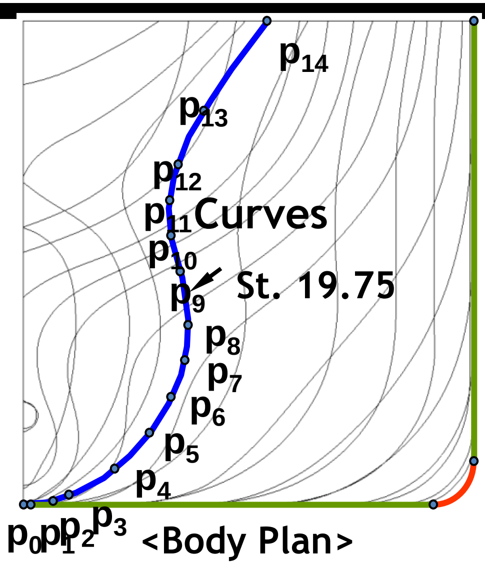
**Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering**

Seoul National Univ.

@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Chapter 0. Summary

Summary:
Represent B-Spline Curve that is passing through given points



<Body Plan>

Given:

- Point p_i on the curve
- Knot u_j of the point on the curve
- Tangent vectors t_0, t_1 of both ends

Find:

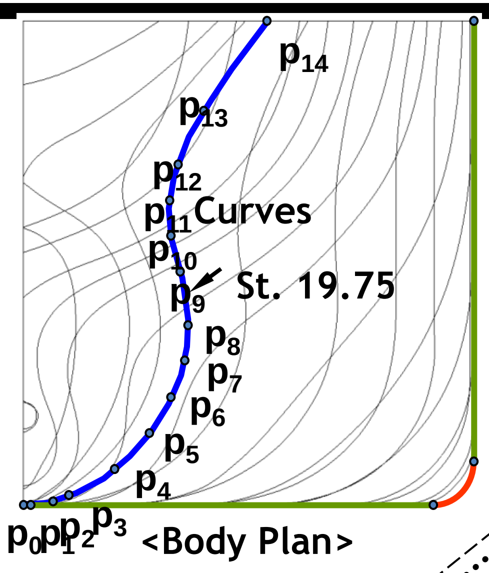
- Cubic B-spline curve $\mathbf{r}(u)$ that is Passing through the point p_i on the curve and satisfying continuity condition C^2
- That is , find control points di of B-spline: curve

$$\mathbf{r}(u) = \begin{bmatrix} x(u) \\ y(u) \\ z(u) \end{bmatrix} = \sum_{i=0}^{D-1} \underbrace{\mathbf{d}_i}_{\text{Coefficient}} \underbrace{N_i^n(u)}_{\text{Basis function}}$$

Linear Combination!

To represent the Curve $\mathbf{r}(u)$, we have to find the coefficients, i.e., the control points $\mathbf{d}_i$

Determine B-Spline control points($\mathbf{d}_i$)



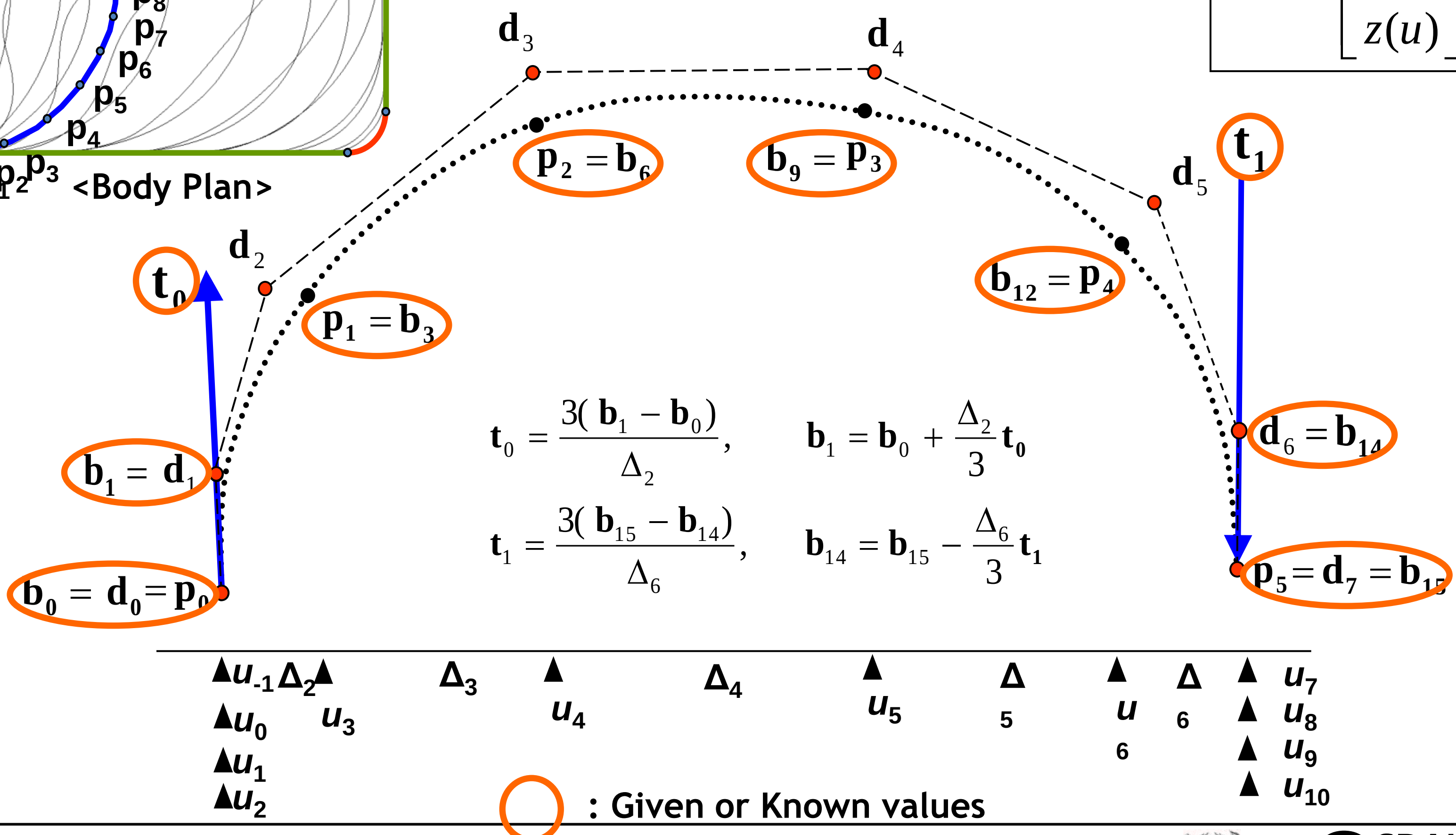
Given:

- Point p_i on the curve
- Knot u_j of the point on the curve
- Tangent vectors t_0, t_1 of both ends

Find:

- Cubic B-spline curve r(u) that is Passing through the point p_i on the curve and satisfying continuity condition C^2
- That is , find control points di of B-spline: curve

$$\mathbf{r}(u) = \begin{bmatrix} x(u) \\ y(u) \\ z(u) \end{bmatrix} = \sum_{i=0}^{D-1} \mathbf{d}_i N_i^n(u)$$



$$\mathbf{t}_0 = \frac{3(\mathbf{b}_1 - \mathbf{b}_0)}{\Delta_2}, \qquad \mathbf{b}_1 = \mathbf{b}_0 + \frac{\Delta_2}{3}\mathbf{t}_0$$

$$\mathbf{t}_1 = \frac{3(\mathbf{b}_{15} - \mathbf{b}_{14})}{\Delta_6}, \qquad \mathbf{b}_{14} = \mathbf{b}_{15} - \frac{\Delta_6}{3}\mathbf{t}_1$$

Determine the control point d_i of B-spline curve by using tri-diagonal matrix solution

Given:

Point p_i on the curve
Knot u_i of point on the curve
Tangent vectors t_0, t_1 of Both ends

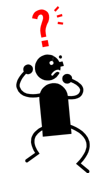
Find:

Cubic B-spline curve r(*u*) whitch Passing through point p_i on the curve and satisfying continuity condition C^2
(Control point of B-spline: d_i)

$$\underbrace{\begin{bmatrix} \mathbf{p}_0 \\ \mathbf{t}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \mathbf{p}_4 \\ \mathbf{t}_1 \\ \mathbf{p}_5 \end{bmatrix}}_{=\mathbf{D}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-3}{\Delta_2} & \frac{3}{\Delta_2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_1 & \beta_1 & \gamma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2 & \beta_2 & \gamma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_3 & \beta_3 & \gamma_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_4 & \beta_4 & \gamma_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-3}{\Delta_6} & \frac{3}{\Delta_6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{d}_3 \\ \mathbf{d}_4 \\ \mathbf{d}_5 \\ \mathbf{d}_6 \\ \mathbf{d}_7 \end{bmatrix}}_{=\mathbf{X}}$$

Known　　　Known　　　Unknown

$$\mathbf{r}(u) = \begin{bmatrix} x(u) \\ y(u) \\ z(u) \end{bmatrix} = \sum_{i=0}^{D-1} \mathbf{d}_i N_i^n(u)$$



What is $N_i^n(u)$

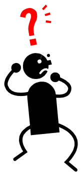
$$\mathbf{D} = \mathbf{A}\mathbf{X}$$

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{D}$$

Since Matrix A is Tri-diagonal matrix, Matrix A^{-1} is easy to obtain.

Basis function of B-spline curve
: Cox-de Boor Recurrence Formula
(B-spline function)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

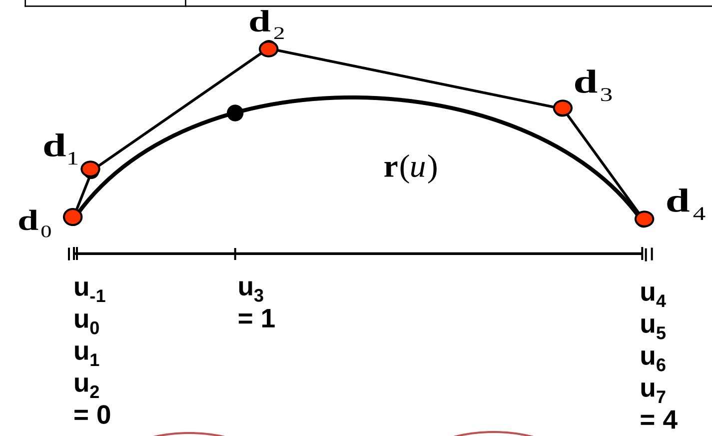


What is $N_i^n(u)$

☑ **B-spline curve**

$$\mathbf{r}(u) = \begin{bmatrix} x(u) \\ y(u) \\ z(u) \end{bmatrix} = \sum_{i=0}^{D-1} \underbrace{\mathbf{d}_i}_{\text{Coefficient}} \underbrace{N_i^n(u)}_{\text{Basis function}}$$

Linear Combination!



$$= \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u)$$

Example of Cubic B-spline curve

- Cox-de Boor Recurrence Formula (B-spline function)

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \; u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

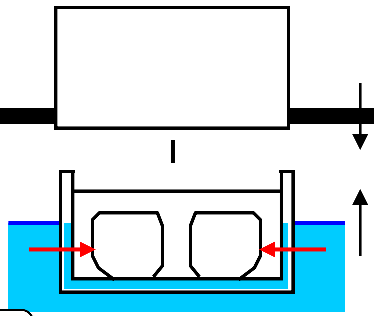
Chapter 1. Introduction

1.1 Application of curves and surfaces to ship design

1.2 Learning Objectives

1.1 Application of curves and surfaces to ship design

Basic Requirements of a ship



Density of steel

= 7.85 ton/m³

about 0.5 ton/m³

Wood
10 ton

||

10 ton

1.025 ton/m³

■ The basic requirements of a ship

1) Ship should float and be stable — Ship stability

→ Weight of the ship is equal to the buoyancy* in static equilibrium

2) Ship should transport carg[o] — Ship compartment layout

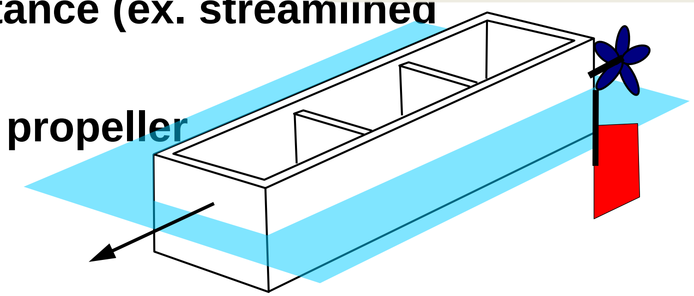
→ The inner space should be large enough for the cargoes

3) Ship should move fast to the destination and be possible to control — Hull form design, Ship hydrodynamic, Propeller design, Ship maneuverability and control

→ shape : should be made to keep low resistance (ex. streamlined shape)

→ propulsion engine : diesel engine, helical propeller

→ Steering engine : Steering gear, Rudder

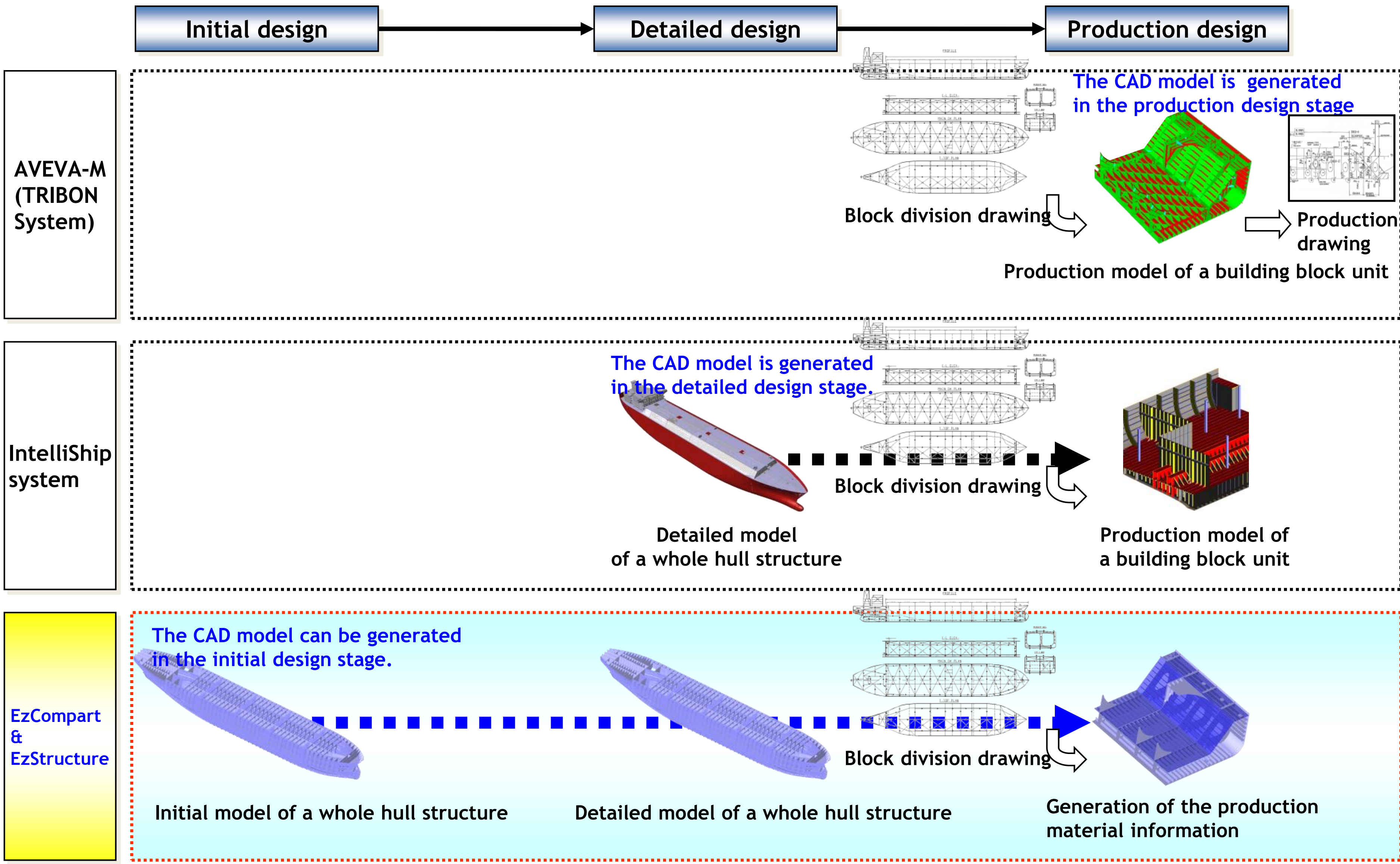


4) Ship should be strong enough in all her life — Ship structural mechanics /Structural design & analysis

→ steel plate (about 10~30mm) and stiffeners welded structure

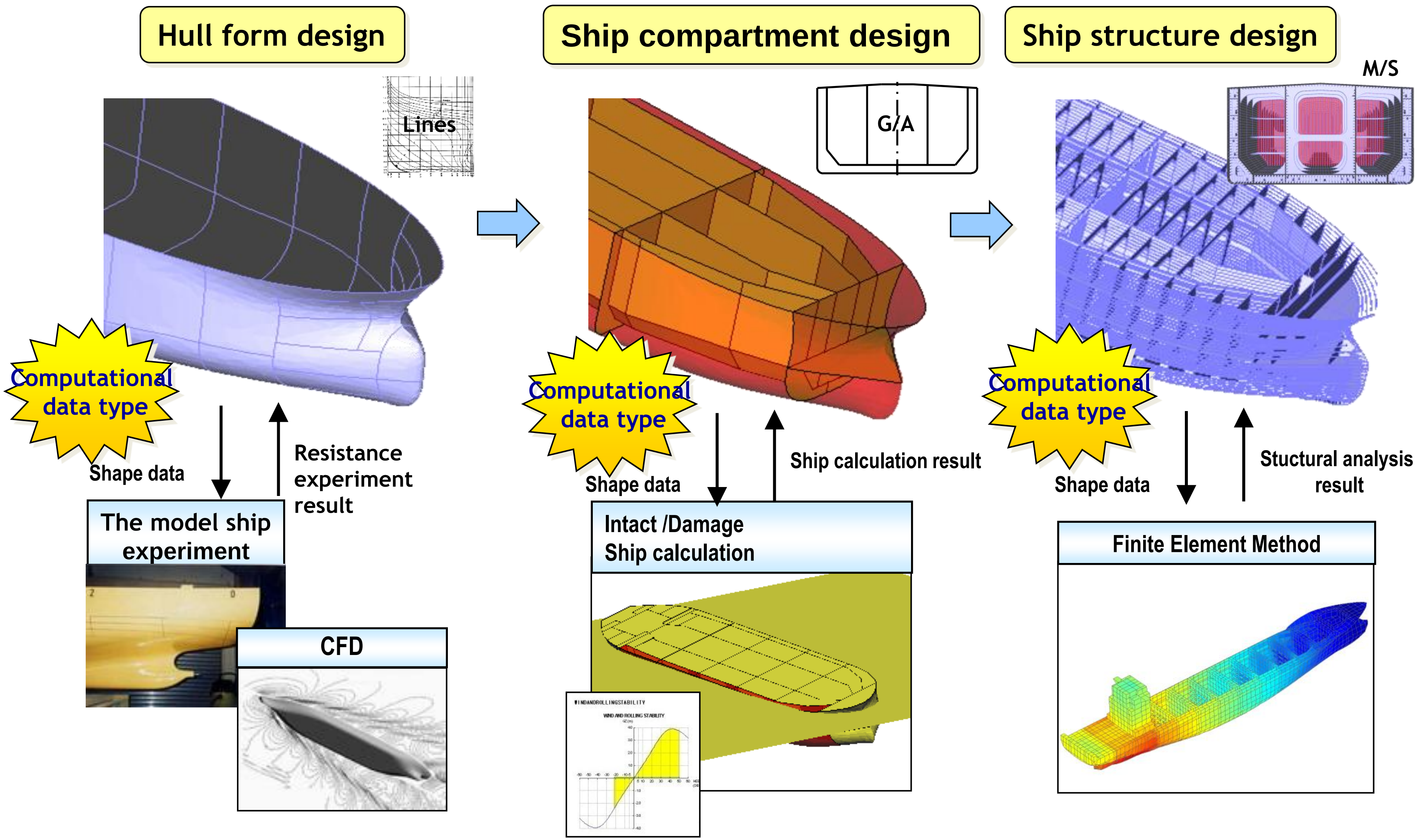
* Archimedes's Principle : the buoyancy of the floating body is equal to the weigh of displaced fluid of the immersed portion of the volume of the ship

Ship Design Stage and used Ship CAD System

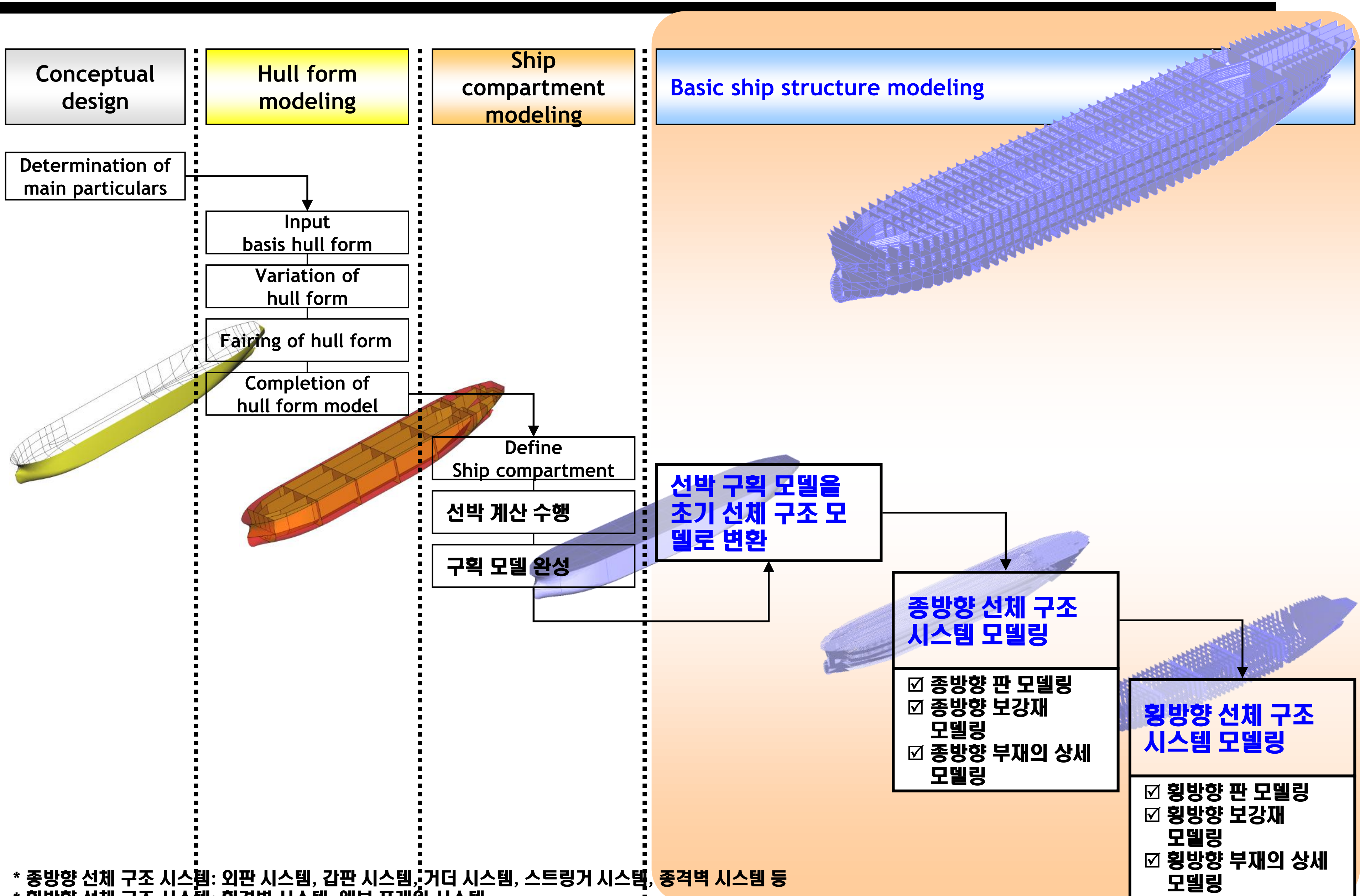


* TRIBON system: CAD system (the exclusive use of shipbuilding), Product of TRIBON Solution Inc. in Sweden
* IntelliShip: M-CAD systme(the exclusive use of shipbuilding), Intergraph Inc., Samsung Heavy Industries, Odense shipyard in denmark, Hitachi shipyard in Japan Cooperative development

Stage of Basic design of a ship

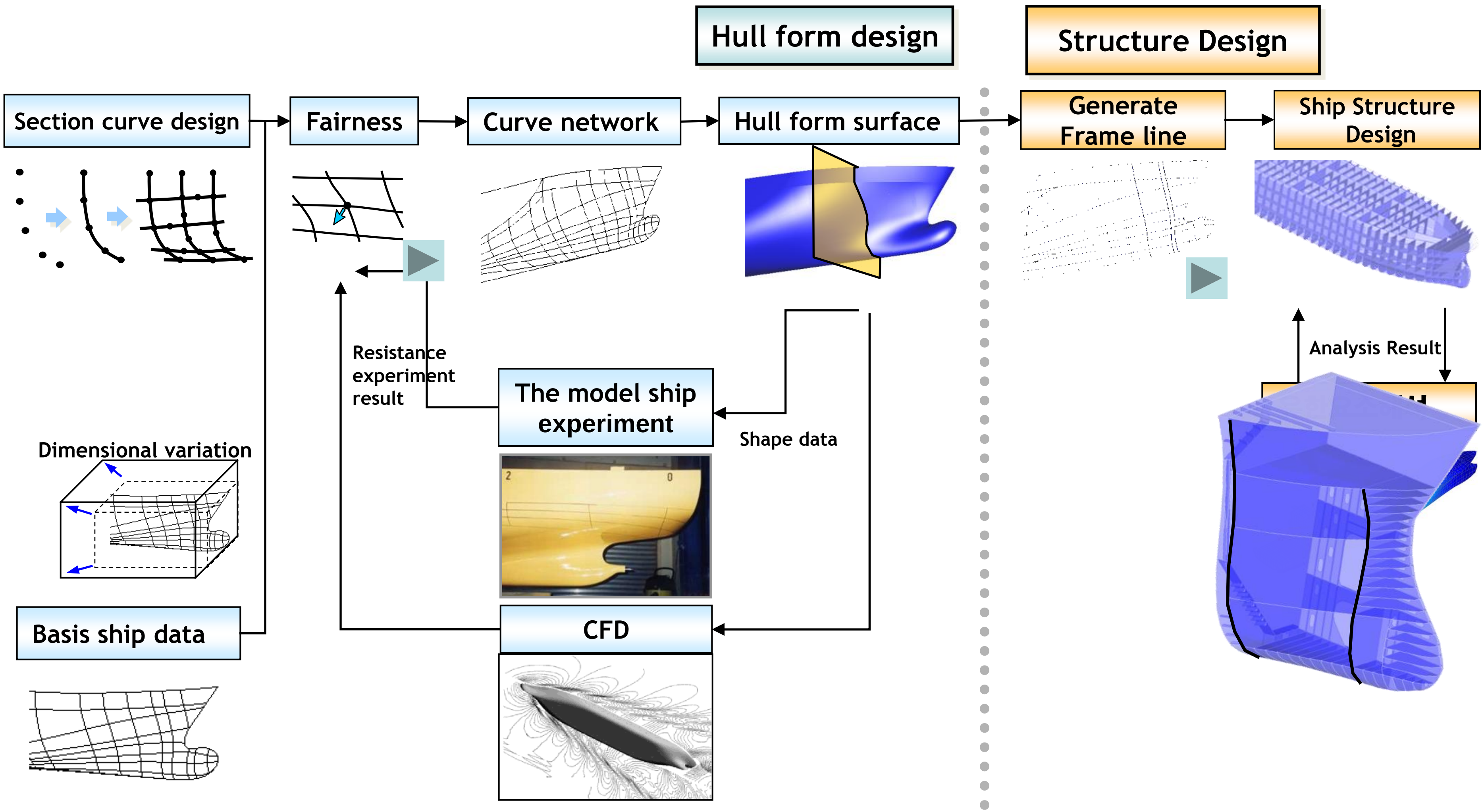


Modeling Stage of Ship Basic Design



* 종방향 선체 구조 시스템: 외판 시스템, 갑판 시스템, 거더 시스템, 스트링거 시스템, 종격벽 시스템 등
* 횡방향 선체 구조 시스템: 횡격벽 시스템, 웨브 프레임 시스템

Ship Shape('Hull Form') Design

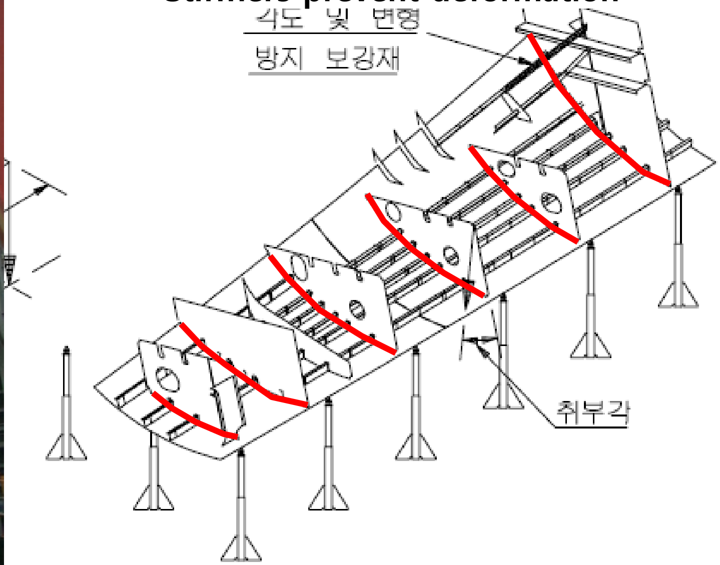


Needs of the hull surface modeling

■ Important production information such as joint length, painting area, weight and CG of the building blocks can be estimated in the initial design stage

- Estimation of the cost, duration of the ship building
- Zig information for fixed round block

Stiffners prevent deformation



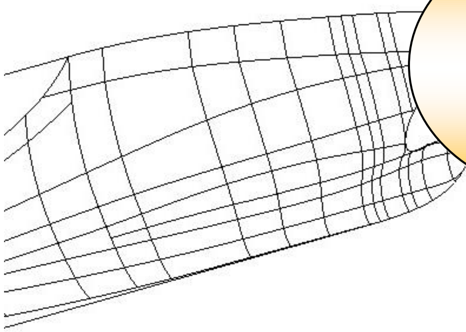
Quality Requirement of hull form surface

Initial hull form design

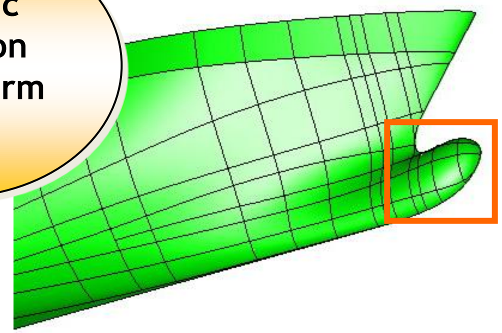
Detailed design / Production design

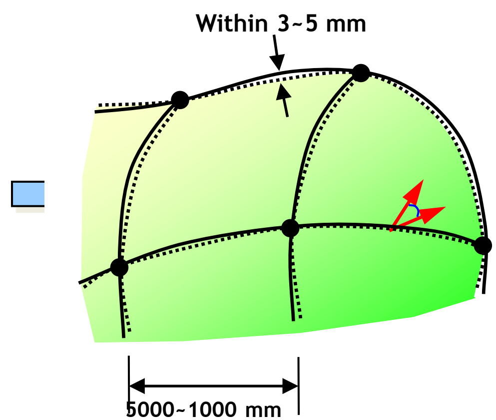
Find : smooth hullform surfaces

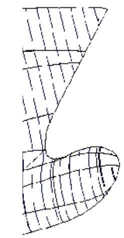
Given : curve network



Automatic generation of hull form surface







- Irregular topology
- In the form of non-uniform B-spline curves

Requirements

- In the form of Bicubic B-spline surface patches
- Max. distance error between given curve network and generated surface < tolerance*
- Smoothness: exact or close to G1**

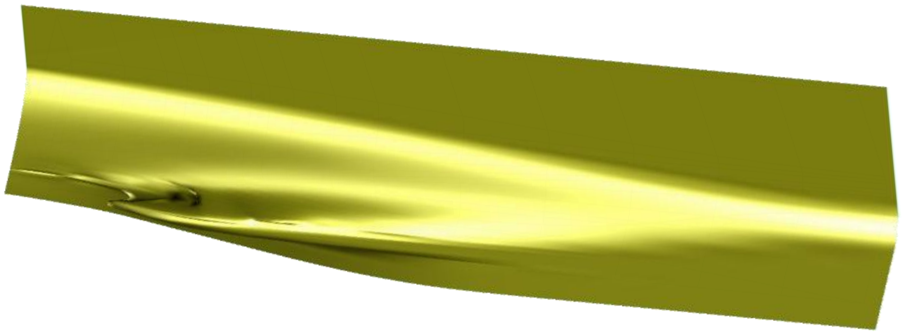
- Intersection between surfaces and plane
- Validation of the fairness

* Acceptable tolerance in shipbuilding industry is about 3~5 mm

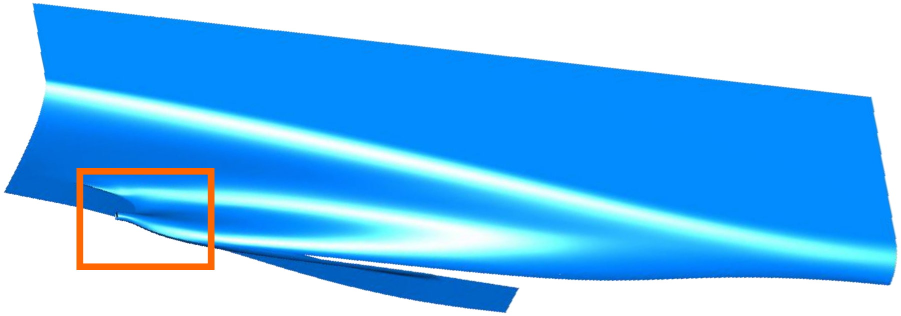
** G^1 means geometric continuity or tangent plane continuity, IntelliShip requires exact G^1 hullform surfaces

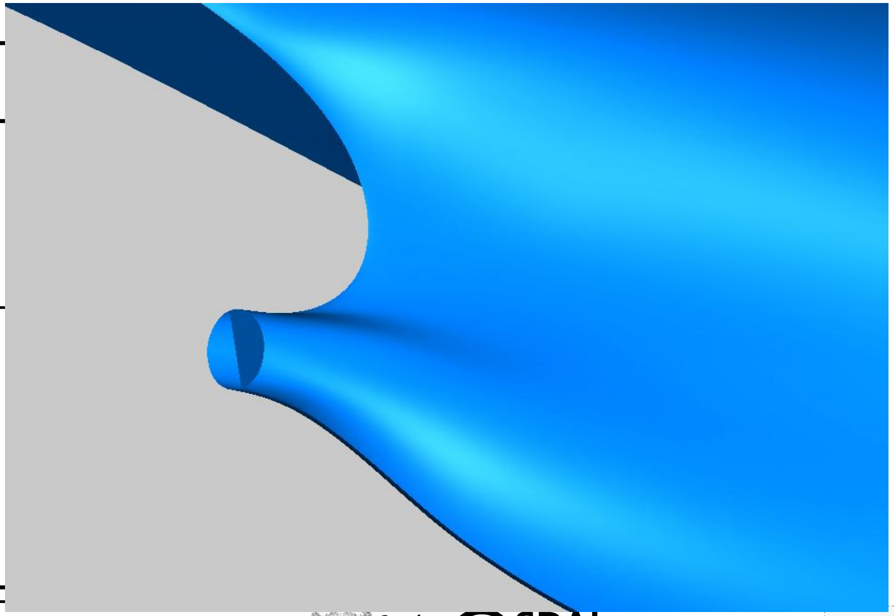
Hull Surface modeling by single patch approach and piecewise patch approach

▪ **Single patch approach**



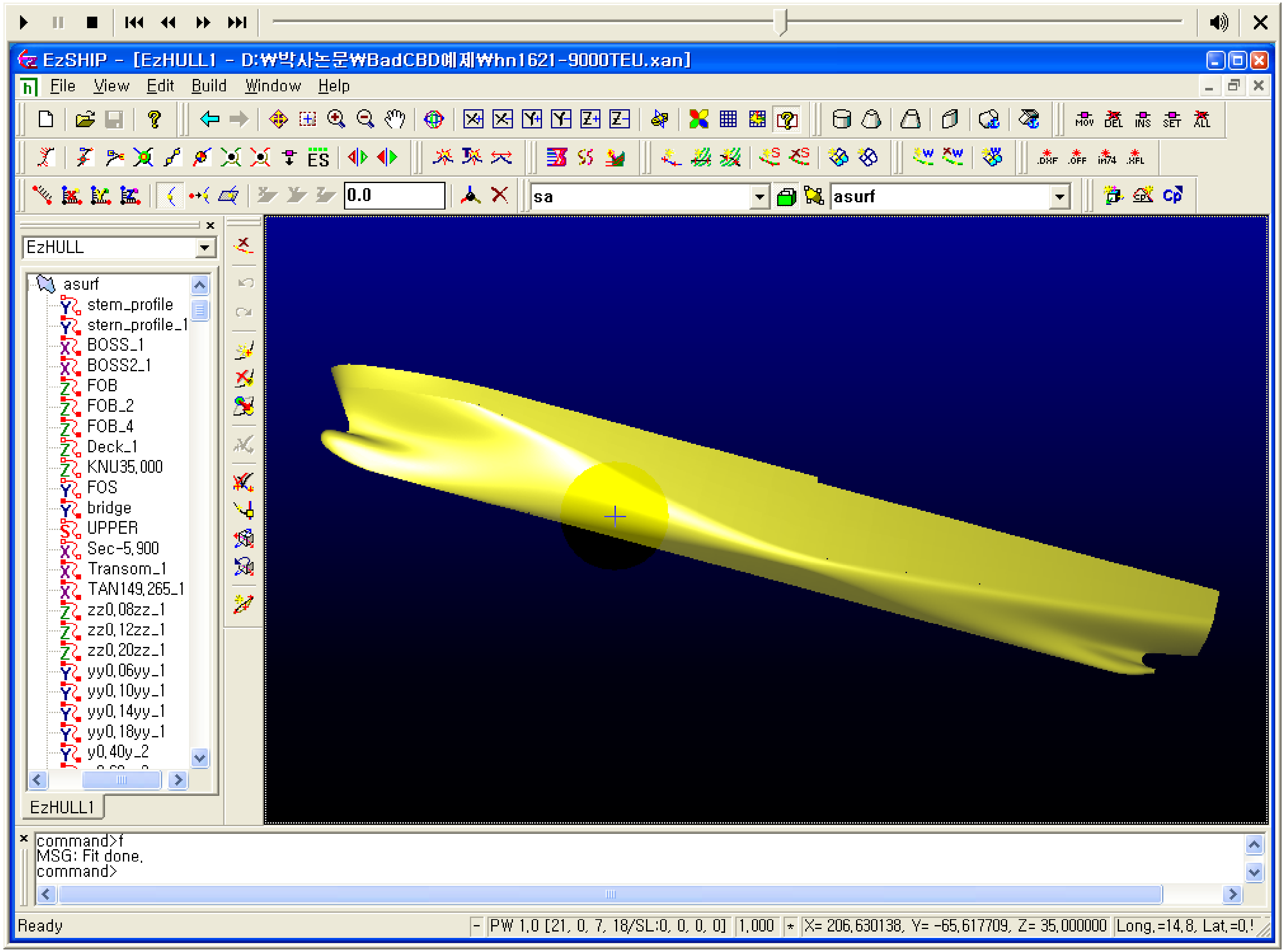
▪ **Piecewise patch approach**





Method	Single patch approach
Advantage	• Easy to represent the hull surface • Mathematically, 2nd derivatives are continuous at all points on the surfaces(C^2)
Dis-advantage	• A single patch approach cannot exactly represent a complex shape in the bow and stern parts and also knuckle curve.

Demonstration

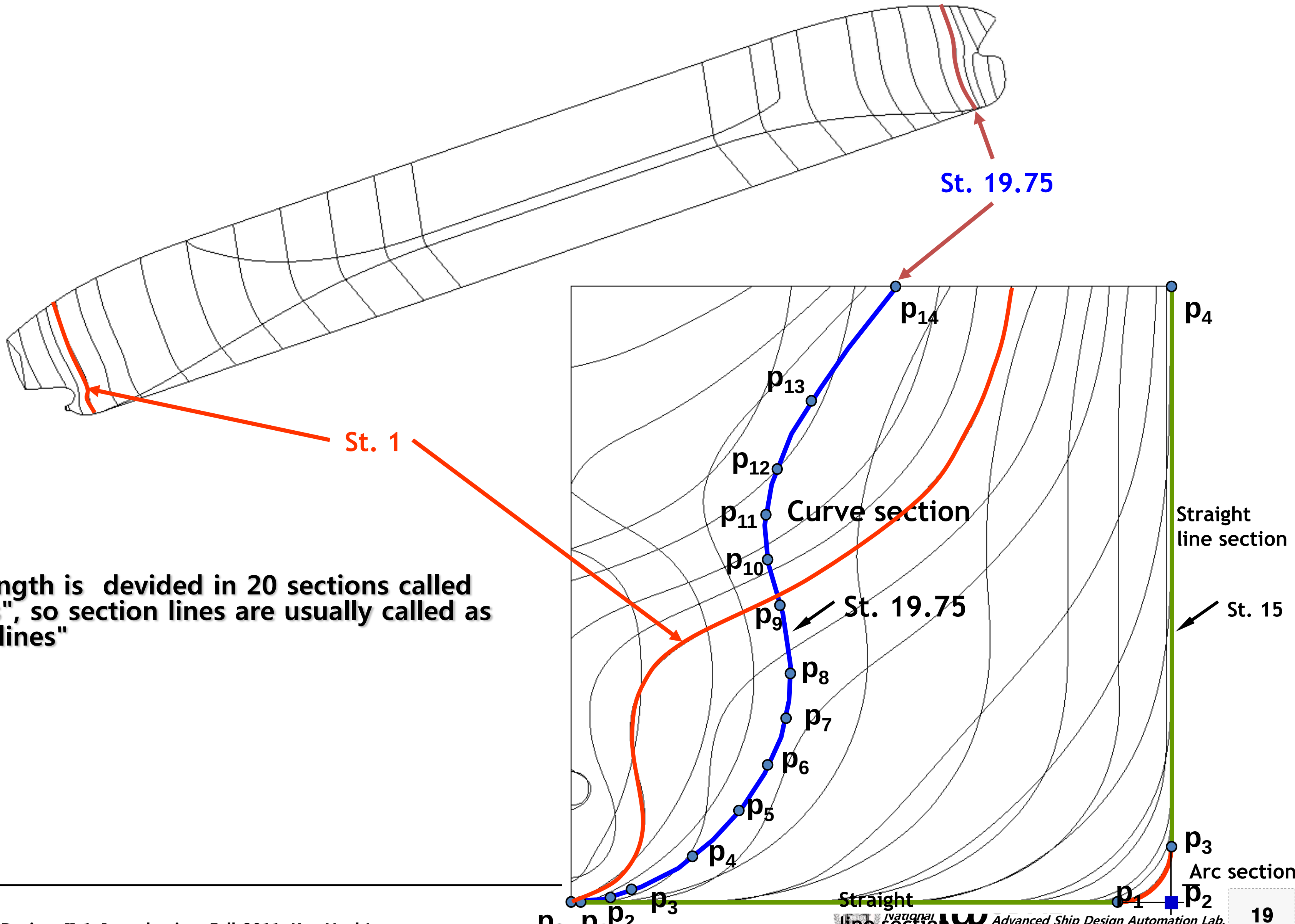


1.2 Learning Objectives

1) Modeling of curve passing through given points

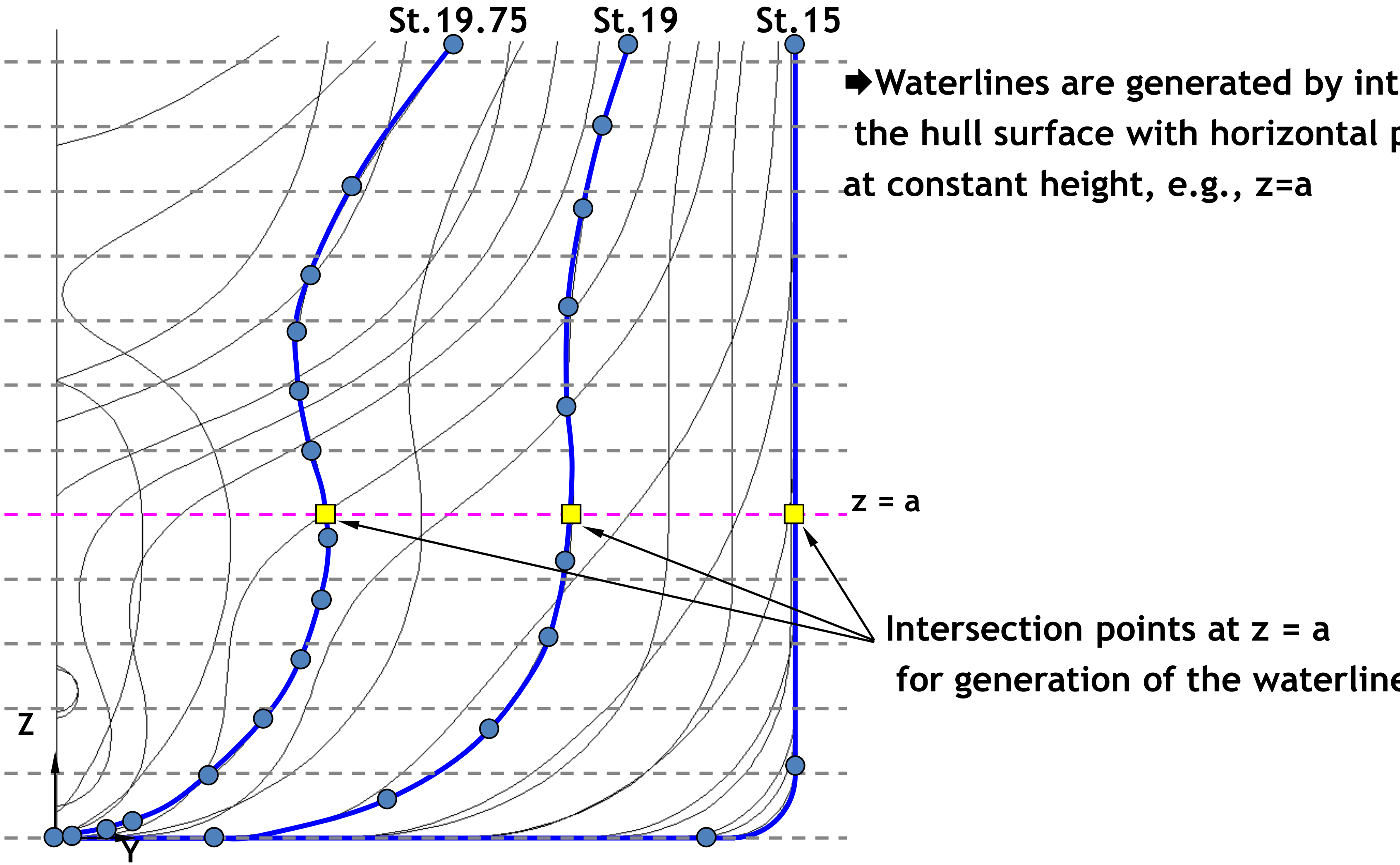
2) Modeling of surface passing through given points

Major curves for Hull Form representation – Section lines



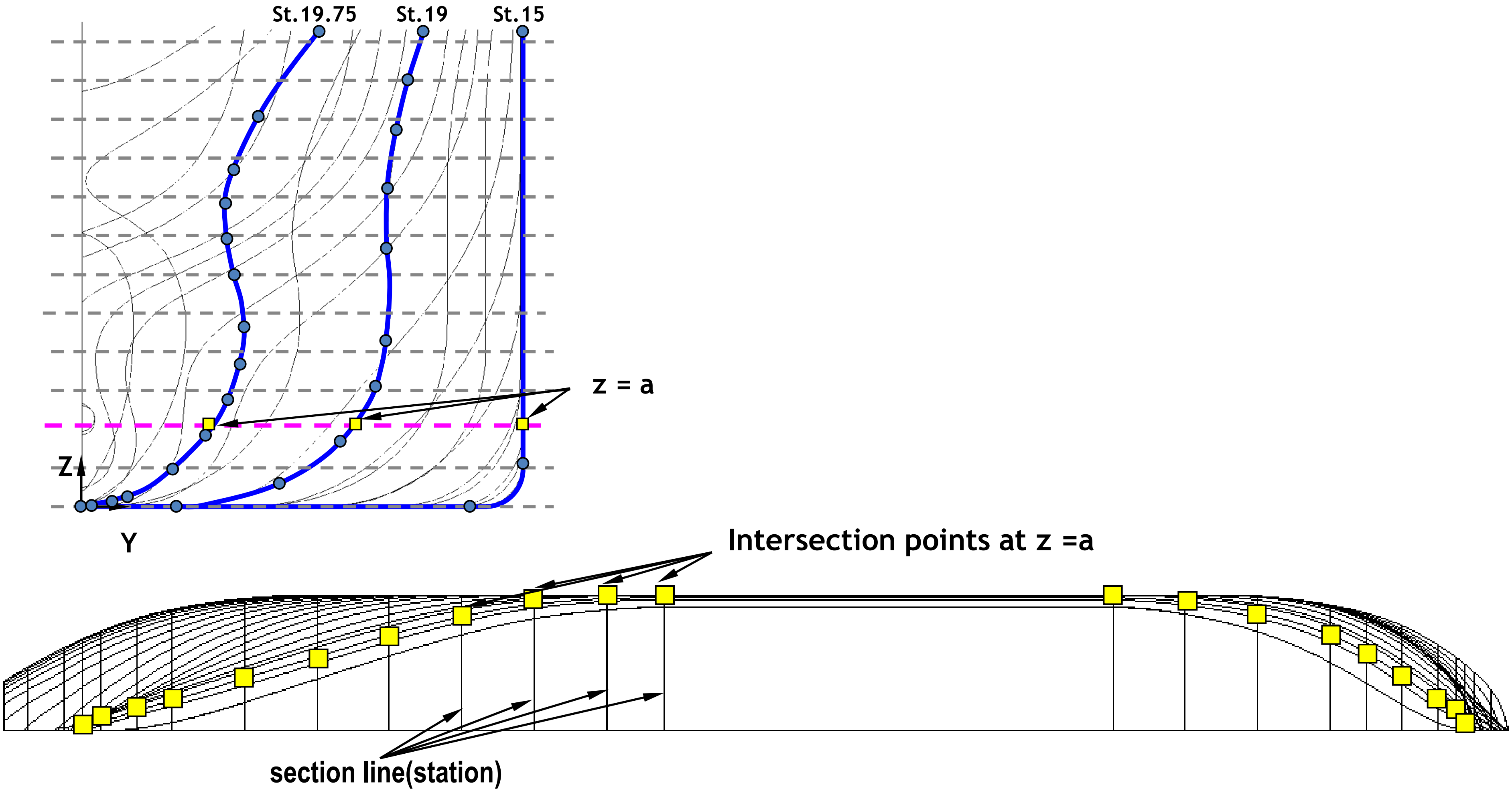
☑ Ship's length is devided in 20 sections called "stations", so section lines are usually called as "station lines"

Major curves for Hull Form representation – Generation of Waterlines(1)



➡Waterlines are generated by intersecting the hull surface with horizontal plane at constant height, e.g., z=a

Major curves for Hull Form representation – Generation of Waterlines(2)



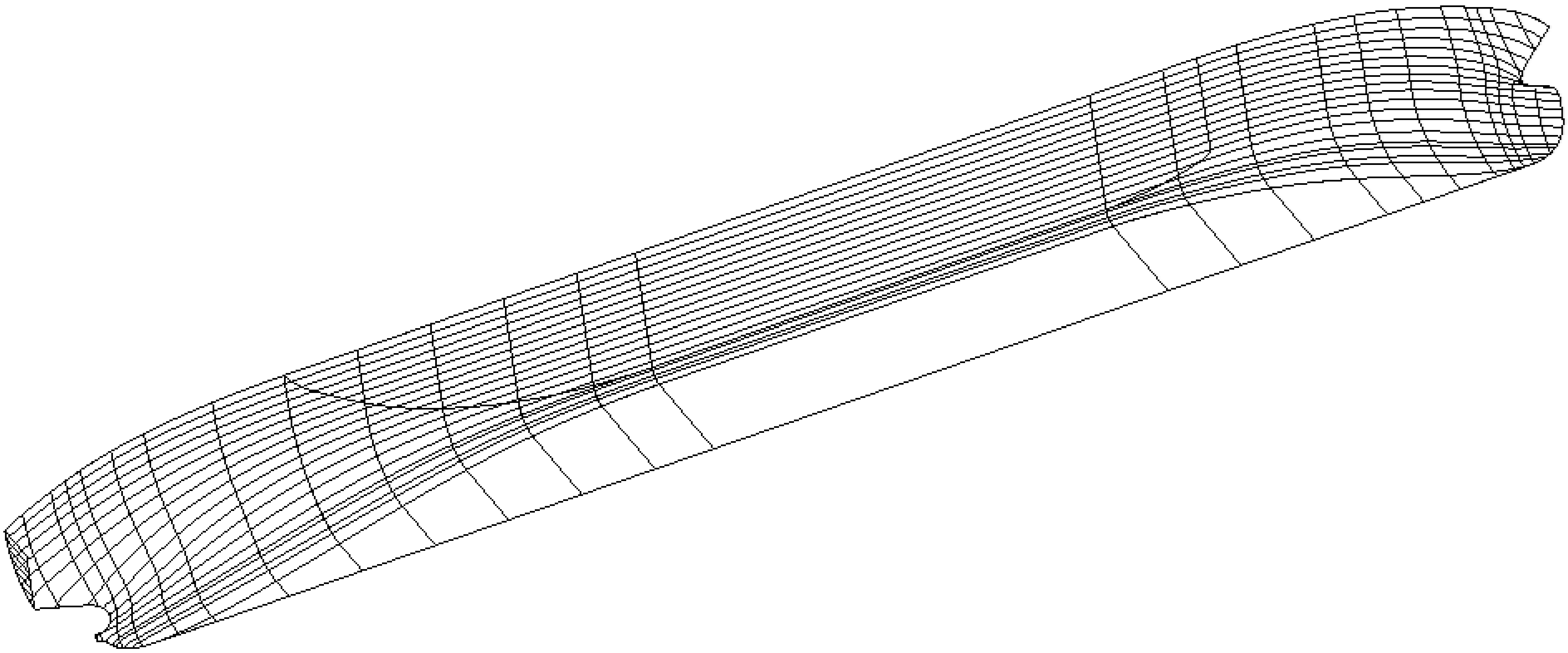
Interpolate all Intersection points at z = a by using a NURB curve
➡ Generation of waterline at z=a

➡

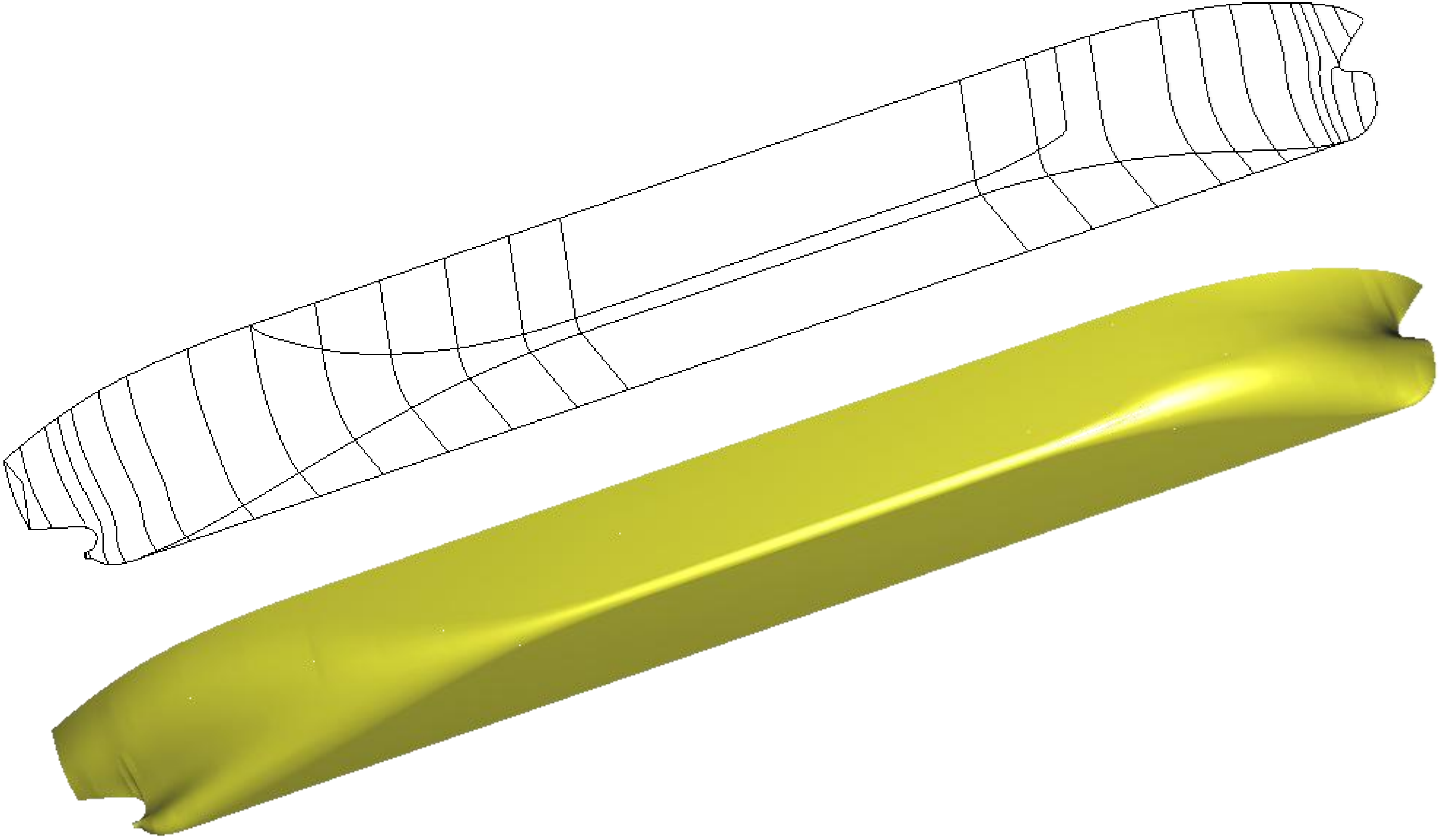
Generate Waterline

Repeat the above steps at different height

Major curves for Hull Form representation – Generation of Waterlines(3)



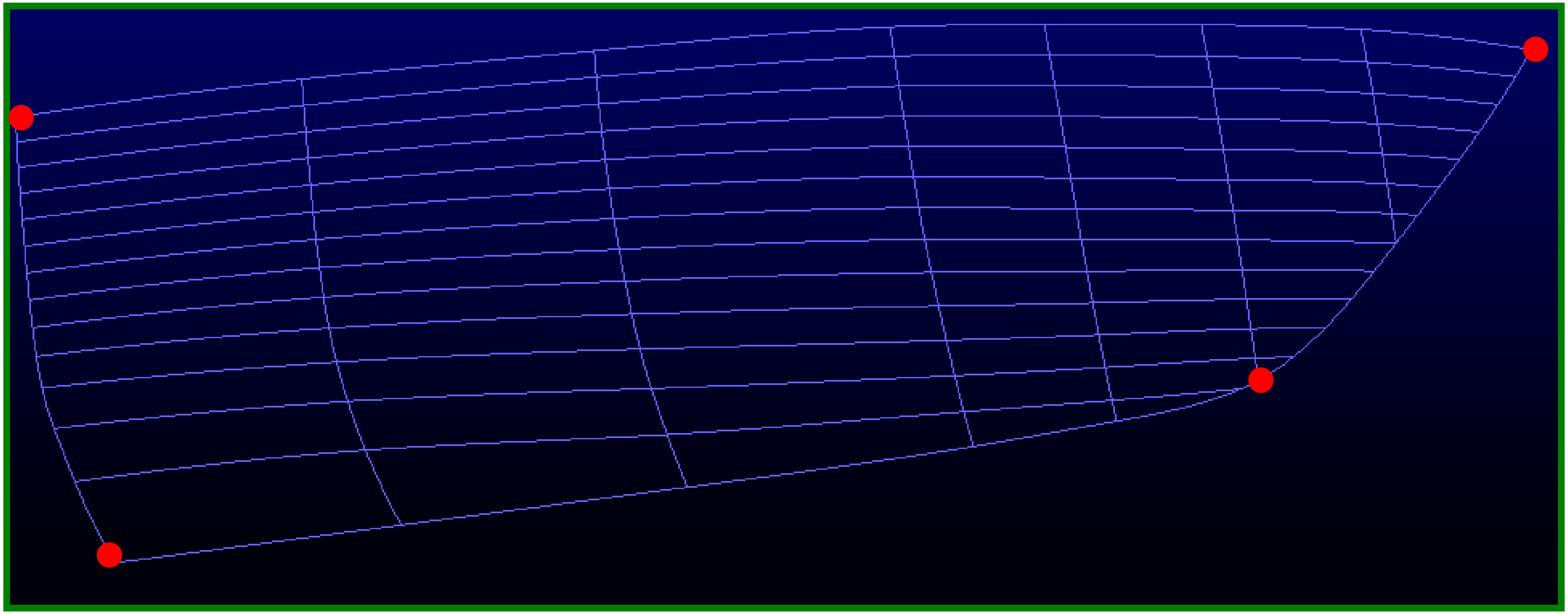
Program implementation of generation of ship hull surface by using single B-spline surface patch



- Chap4. Term Project

Modeling of a yacht surface (1)

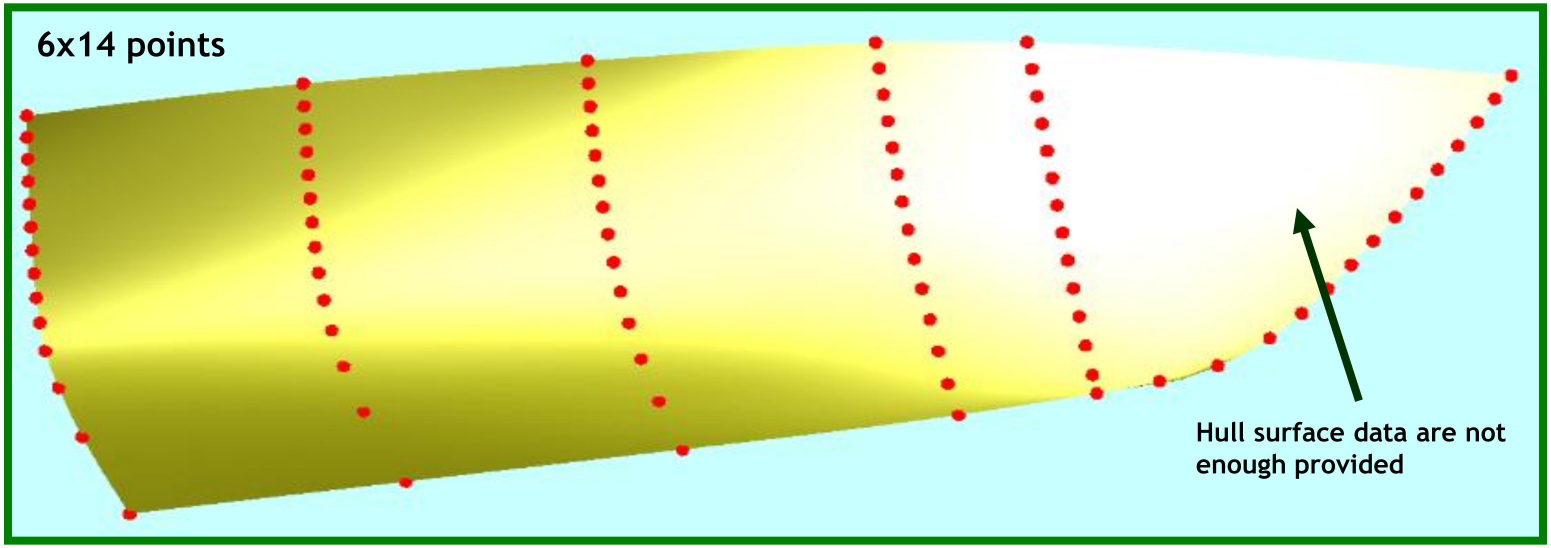
•Example of a yacht surface generated by the Student during the lecture of “Planning Procedure of Naval Architecture and Ocean Engineering, second semester, 2005, Department of Naval Architecture and Ocean Engineering, SNU



Determine the vertexes of tetragonal patch

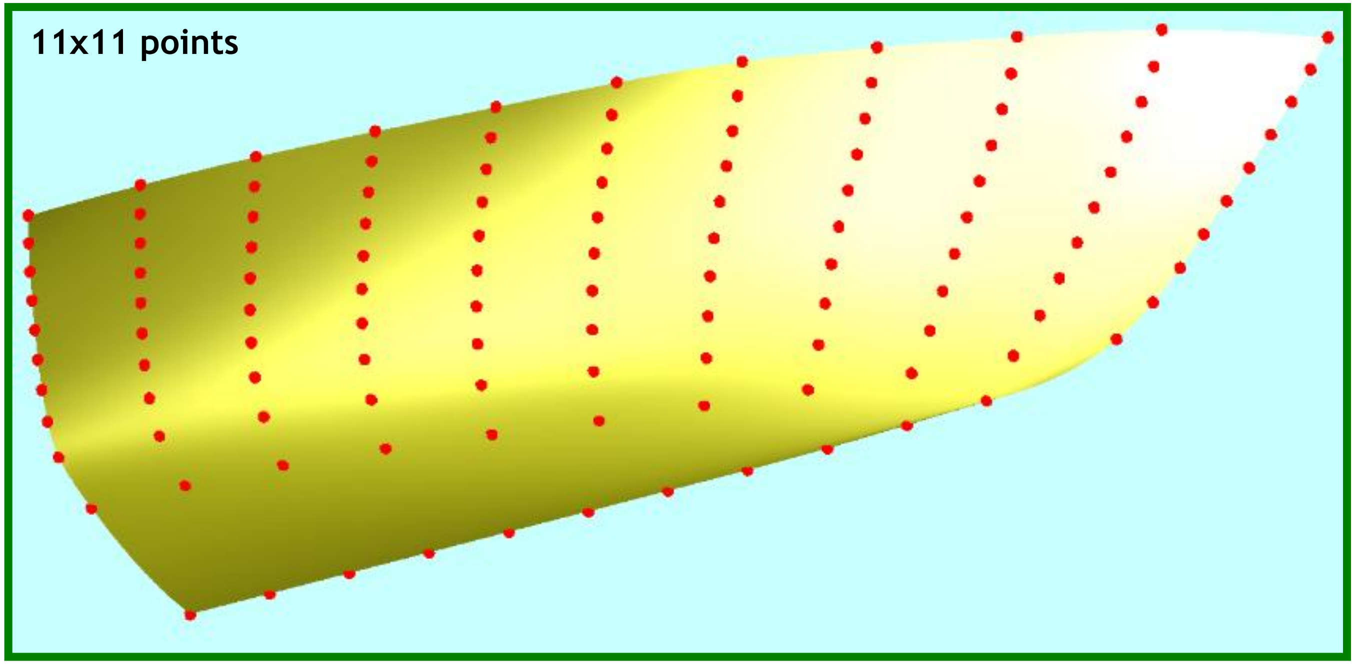
Modeling of a yacht surface (1)

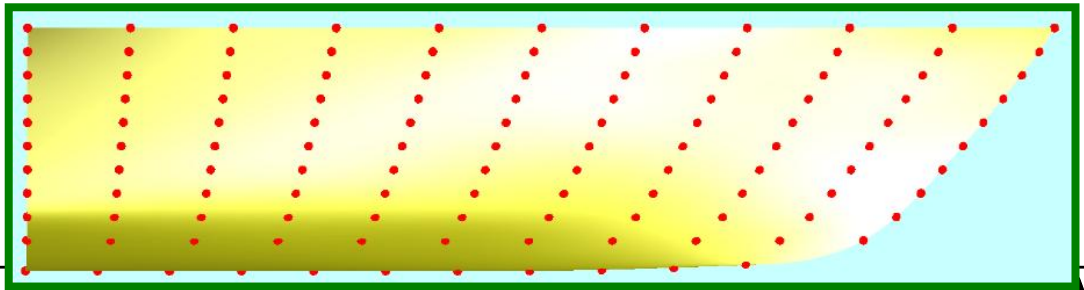
❑ **Modeling result of a yacht surface passing through the given data points that are located irregularly in the longitudinal direction**



Modeling of a yacht surface (2)

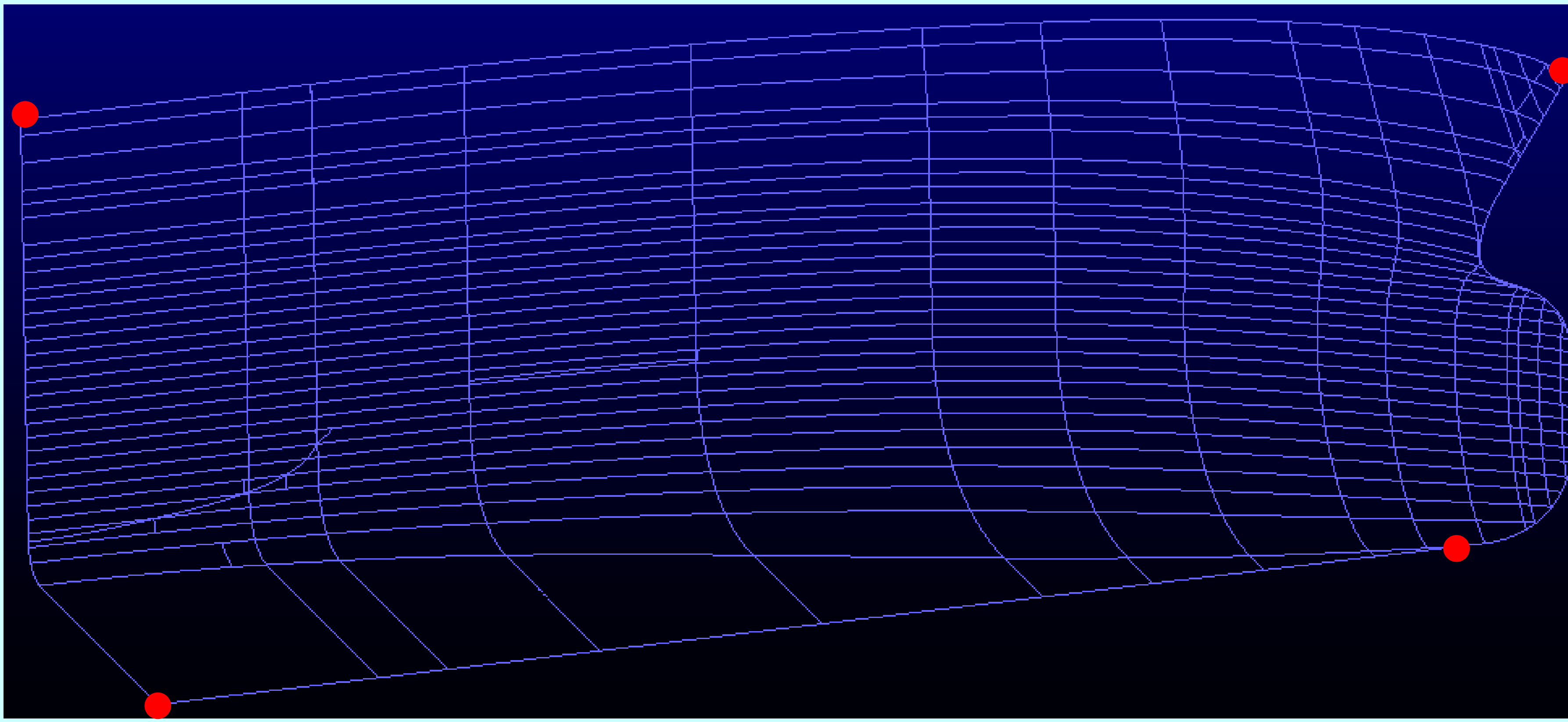
Modeling result of a yacht surface passing through the given data points that are located nearly at same distance in the longitudinal direction



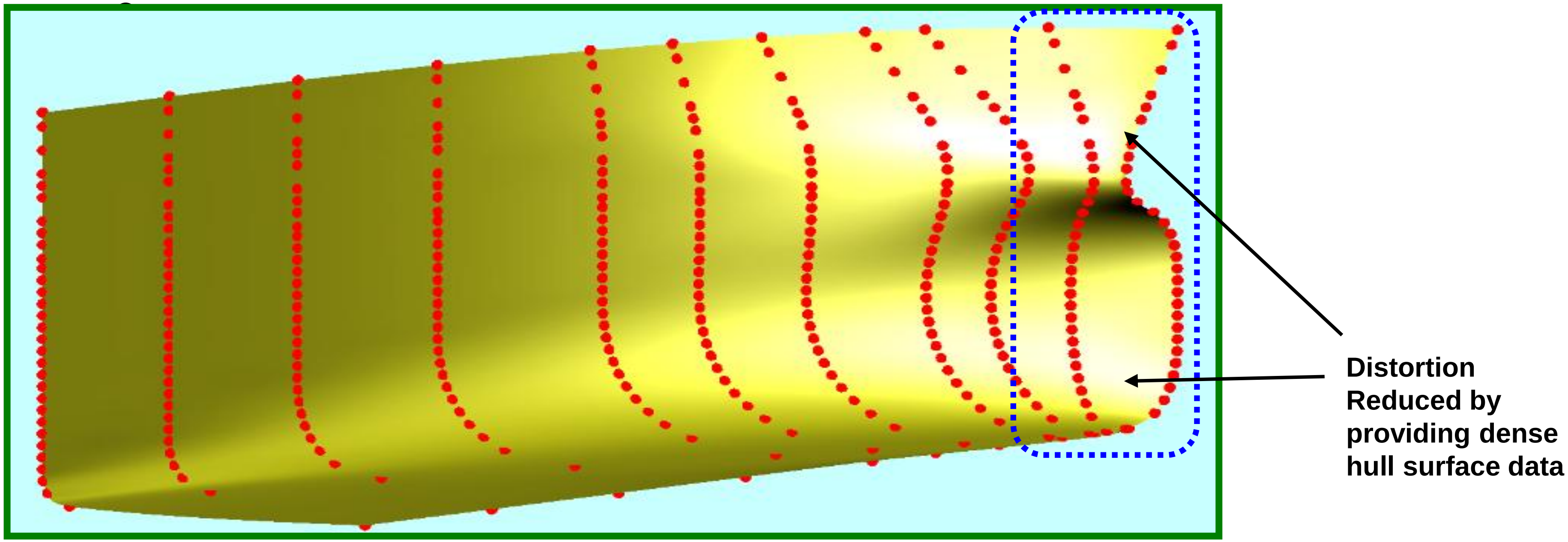


Modeling of the hull surface with a bulbous bow by using only one B-spline surface patch(1)

BOW



Modeling of the hull surface with a bulbous bow by using only one B-spline surface patch(2)



Chapter 2. Bezier Curves

2.1 Parametric Function/Curves

1) Explicit function / Implicit function / Parametric function

2) Characteristics of parametric function

3) Expression of general function by using parametric function

1) Explicit / Implicit / Parametric function

☑ **Explicit function**

- **If the function is expressed by y=f(x), it is called 'Explicit function'**
- **'y' can be obtained easily if x is give.**

ex) $y = \sqrt{r^2 - x^2}$

☑ **Implicit function**

- **For multi variable function, e.g., two variables x ,y , the implicit function is expressed by f(x,y)=0**
- **It is easy to check that the given point is inside or outside, left or right of the curve**
- **Implicit function is not always possible to transform to the explicit form.**

ex) $x^2 + y^2 - r^2 = 0$

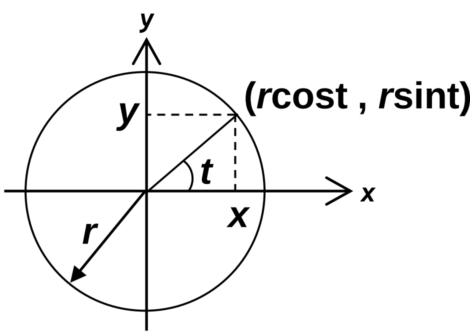
ex) $(0)^2 + (0)^2 - r^2 < 0$

$(r)^2 + (r)^2 - r^2 > 0$

ex) $y = \pm\sqrt{r^2 - x^2}$

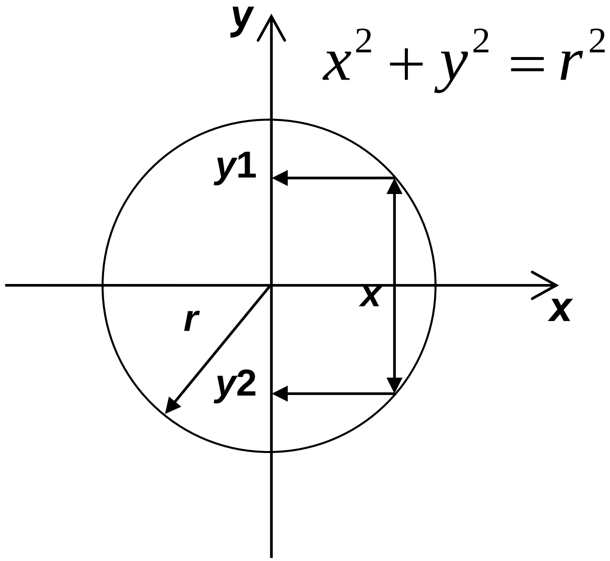
☑ **Parametric function**

- **For multi variable function, e.g., two variables x ,y, the function can be expressed by x=f(t), y=g(t) using parameter 't'. We call it 'parametric function'**
- **Every explicit function is possible to transform to a parametric form.**



ex) $x(t) = r\cos t,\ y(t) = r\sin t$

2) Characteristics of parametric function(1)



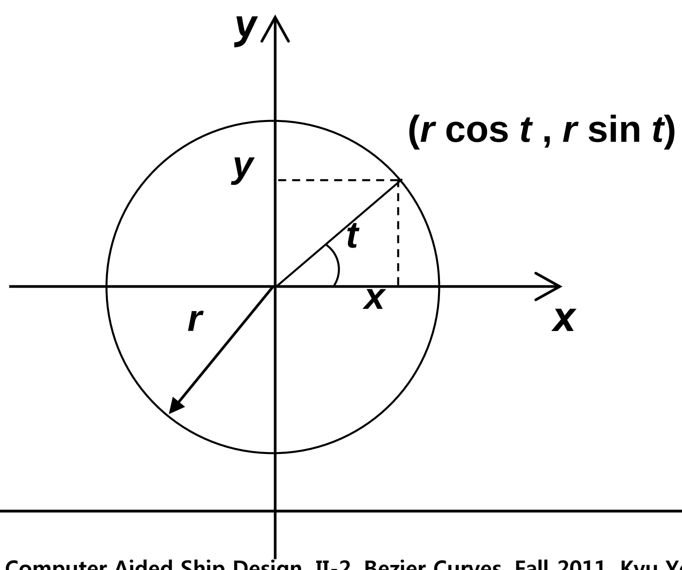
☑ **General Function**

- y value of more than two can be obtained for an x value (multi-value function)

$$x^2 + y^2 = r^2 \qquad y = \pm\sqrt{r^2 - x^2}$$

- It icould be difficult to express derivatives

$$\frac{dy}{dx}_{,x=r} = \infty$$



☑ **Parametric function**

- a parameter value has only one result

$$x(t) = r\cos t, y(t) = r\sin t$$

- It is easy to express derivatives
→Calculate dy/dx dividing each elements: dy/dt, dx/dt

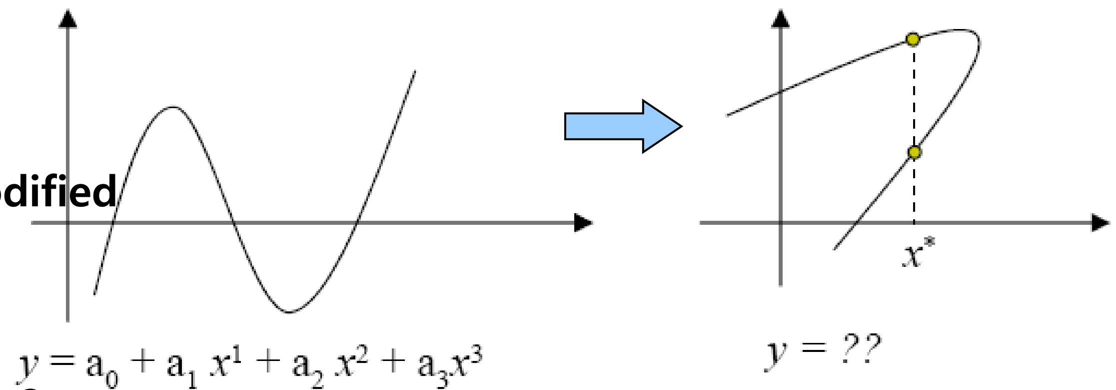
$$\frac{dx}{dt} = -r\sin t, \quad \frac{dy}{dt} = r\cos t$$

Characteristics of parametric function(2)

☑ **Explicit function of y=f(x)**

- It is difficult to express as explicit function* again after original explicit function is modified through rotation, move, etc

  *dimensional extension



☑ **implicit function of f(x,y)=0**

- Points on the curve can not be calculated in order
- Dimensional extension is difficult

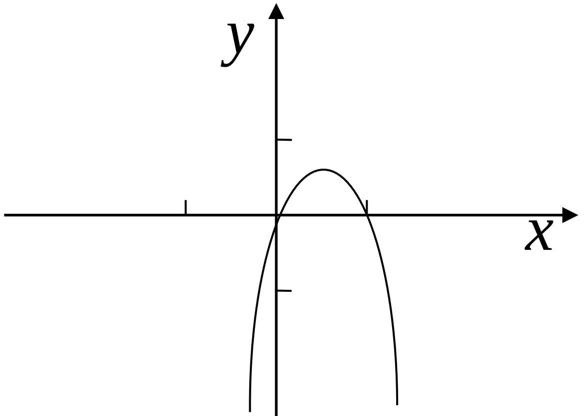
☑ **Parametric function of x = f(t), y = g(t)**

- Points on the curve can be easily calculated in order by varying the parameters
- Dimensional extension is easy.
- The reason why parametric function is commonly used for CAD systems

Characteristics of parametric function(3)

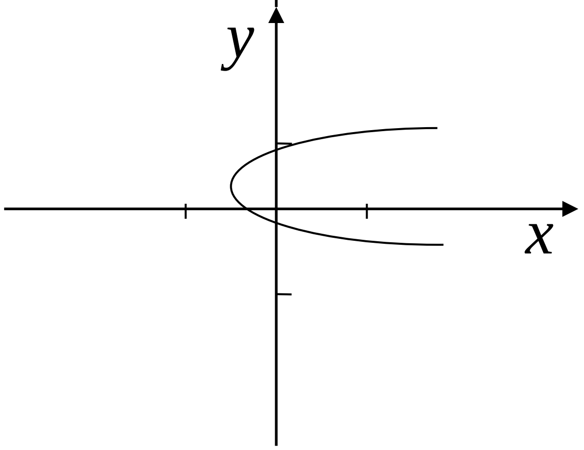
☑ **A curve is defined by the parametric functions as follows:**

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ 2t - 2t^2 \end{bmatrix}$$



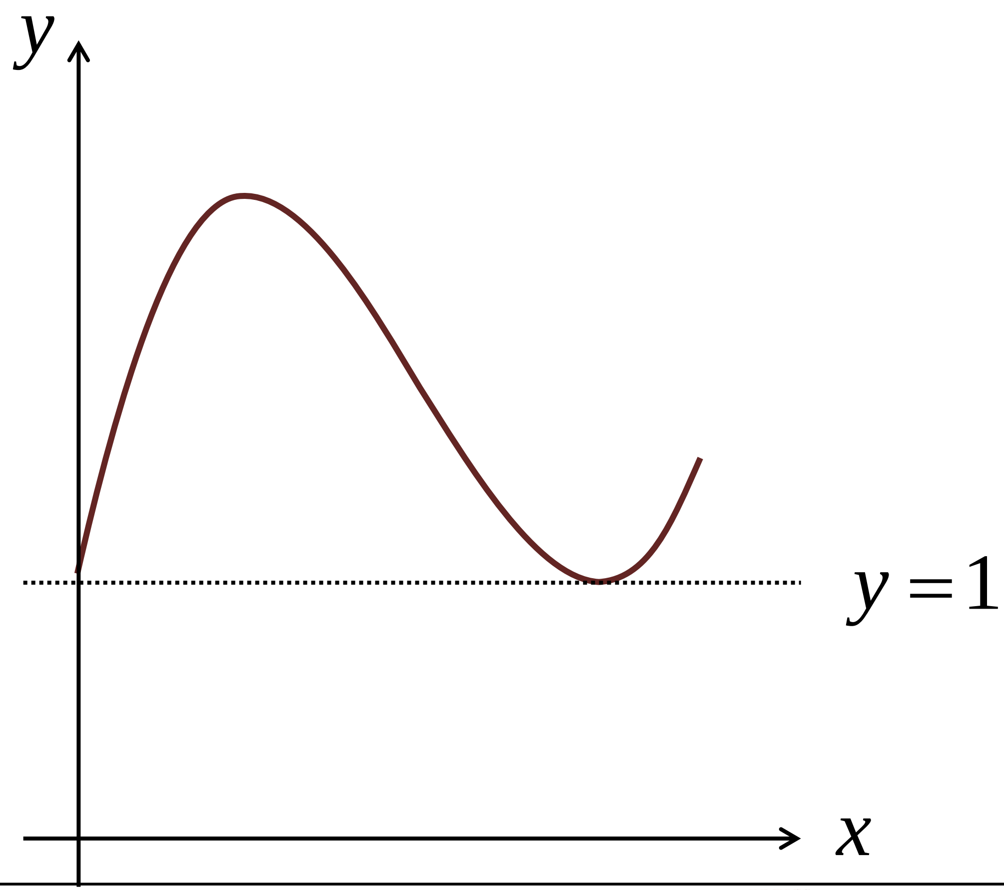
☑ **If the curve is rotated with angle of 90°, geometry('topology') is not changed, only its position vector are changed.**

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2t + 2t^2 \\ t \end{bmatrix}$$



3) Expression of general function by using parametric function(1)

Given: $y = 2x^3 - 4x^2 + 2x + 1$



Expression of general function by using parametric function(2)

$$y = 2x^3 - 4x^2 + 2x + 1$$

$$\mathbf{r}(t) = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} t \\ 2t^3 - 4t^2 + 2t + 1 \end{bmatrix}$$

- From this parametric function with coefficient 2, -4, 2, 1, it is not at all obvious what the function might look like.

- Alternatively, we can express the function in another way as follows:

$$\mathbf{r}(t) = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} (1-t)^3 x_0 + 3t(1-t)^2 x_1 + 3t^2(1-t)x_2 + t^3 x_3 \\ (1-t)^3 y_0 + 3t(1-t)^2 y_1 + 3t^2(1-t)y_2 + t^3 y_3 \end{bmatrix}$$

$$\begin{bmatrix} t \\ 2t^3 - 4t^2 + 2t + 1 \end{bmatrix} = \begin{bmatrix} (1-t)^3 x_0 + 3t(1-t)^2 x_1 + 3t^2(1-t)x_2 + t^3 x_3 \\ (1-t)^3 y_0 + 3t(1-t)^2 y_1 + 3t^2(1-t)y_2 + t^3 y_3 \end{bmatrix}$$

Expression of general function by using parametric function(3)

$$(1-t)^3 x_0 + 3t(1-t)^2 x_1 + 3t^2(1-t)x_2 + t^3 x_3 = t$$

Coefficient of constant:	$x_0 = 0$		$x_0 = 0$
Coefficient of *t* :	$-3x_0 + 3x_1 = 1$	$\Rightarrow$	$x_1 = 1/3$
Coefficient of *t*² :	$3x_0 - 6x_1 + 3x_2 = 0$		$x_2 = 2/3$
Coefficient of *t*³ :	$-x_0 + 3x_1 - 3x_2 + x_3 = 0$		$x_3 = 1$

$$b_{x_i}^0 = x_i = \frac{i}{n}$$

Linear Precision

Gerald E. Farin, The Essentials of CAGD, 2000, p. 29.

- *Linear precision*: If the control points b_1 and b_2 are evenly spaced on the straight line between b_0 and b_3, the cubic Bezier curve is the linear interpolant between b_0 and b_3.

Expression of general function by using parametric function(4)

$$(1-t)^3 y_0 + 3t(1-t)^2 y_1 + 3t^2(1-t)y_2 + t^3 y_3 = 2t^3 - 4t^2 + 2t + 1$$

Coefficient of constant : $y_0 = 1$

Coefficient of *t* : $-3y_0 + 3y_1 = 2$

Coefficient of *t*² : $3y_0 - 6y_1 + 3y_2 = -4$

Coefficient of *t*³ : $-y_0 + 3y_1 - 3y_2 + y_3 = 2$

$\Rightarrow$

$y_0 = 1$

$y_1 = 5/3$

$y_2 = 1$

$y_3 = 1$

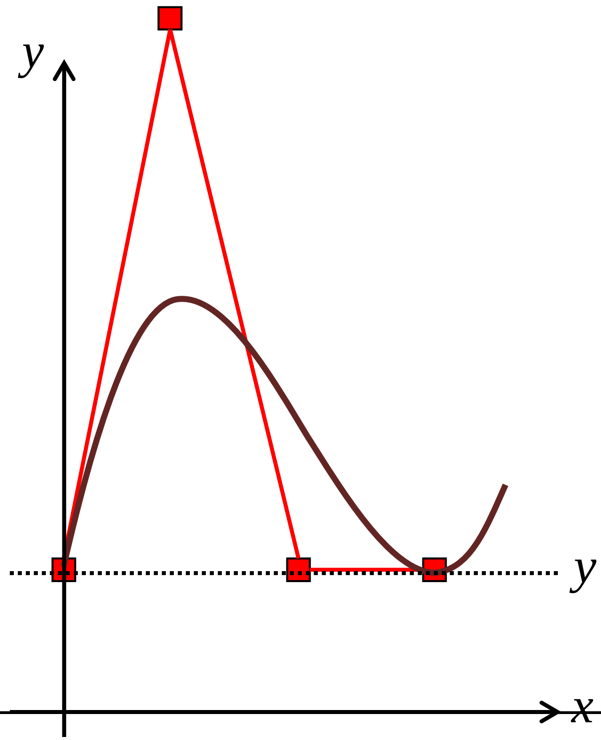
Expression of general function by using parametric function(5)

$$\mathbf{r}(t)=\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x(t)\\y(t)\end{bmatrix}=\begin{bmatrix}t\\2t^3-4t^2+2t+1\end{bmatrix}=\begin{bmatrix}(1-t)^3\cdot 0+3t(1-t)^2\cdot\frac{1}{3}+3t^2(1-t)\cdot\frac{2}{3}+t^3\cdot 1\\(1-t)^3\cdot 1+3t(1-t)^2\cdot\frac{5}{3}+3t^2(1-t)\cdot 1+t^3\cdot 1\end{bmatrix}$$

$$=(1-t)^3\begin{bmatrix}0\\1\end{bmatrix}+3t(1-t)^2\begin{bmatrix}1/3\\5/3\end{bmatrix}+3t^2(1-t)\begin{bmatrix}2/3\\1\end{bmatrix}+t^3\begin{bmatrix}1\\1\end{bmatrix}$$

$$=B_0^3(t)\boxed{\begin{bmatrix}0\\1\end{bmatrix}}+B_1^3(t)\boxed{\begin{bmatrix}1/3\\5/3\end{bmatrix}}+B_2^3(t)\boxed{\begin{bmatrix}2/3\\1\end{bmatrix}}+B_3^3(t)\boxed{\begin{bmatrix}1\\1\end{bmatrix}}$$

$$=B_0^3\mathbf{b}_0+B_1^3\mathbf{b}_1+B_2^3\mathbf{b}_2+B_3^3\mathbf{b}_3$$



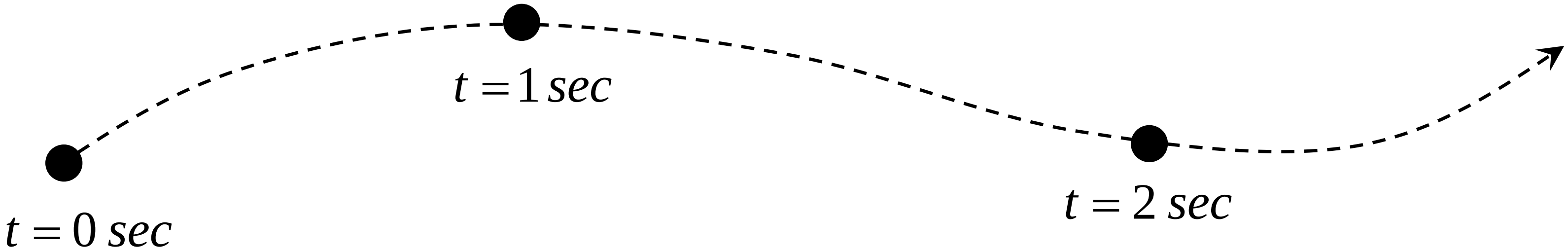
When the points are connected to the line, it shows a similar appearance for the original curve.

Expression of general function by using parametric function(6)

$$\mathbf{r}(t)=\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} x(t) \\ y(t) \end{bmatrix}=\begin{bmatrix} t \\ 2t^3-4t^2+2t+1 \end{bmatrix}=\begin{bmatrix} (1-t)^3\cdot 0+3t(1-t)^2\cdot\frac{1}{3}+3t^2(1-t)\cdot\frac{2}{3}+t^3\cdot 1 \\ (1-t)^3\cdot 1+3t(1-t)^2\cdot\frac{5}{3}+3t^2(1-t)\cdot 1+t^3\cdot 1 \end{bmatrix}$$

- **If the parameter 't' is time, then r(t) can be regarded as the moving trajectory of a rigid body**
- **In explicit or implicit function, it is only possible to express the moving trajectory of a rigid body, whereas the parametric function can express the detail of the position r(t) in particular time 't' as well as the moving trajectory of a rigid body**

$\mathrm{r}(t)$: Position of body, $\dot{\mathrm{r}}(t)$: Velocity of body, $\ddot{\mathrm{r}}(t)$: Acceleration of body

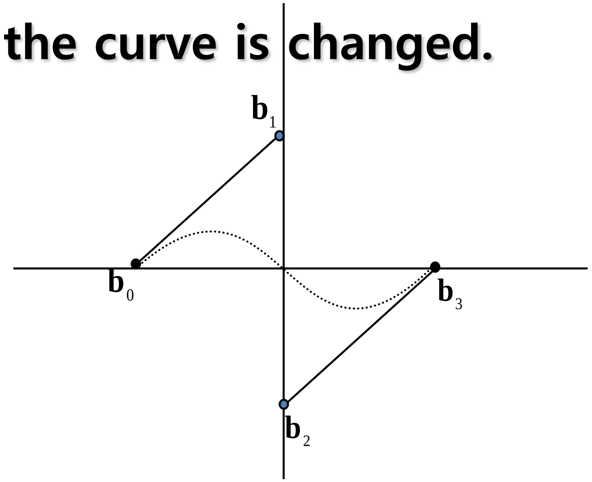


4) "Blending" the points in space and parametric functions

Curves can be represented by "blending" the points in space and parametric functions

If these points are moved, then the shape of the curve is changed.

So, these points are called "control points"



$$
\begin{aligned}
\mathrm{r}(t) = \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -(1-t)^3 + t^3 \\ 3(1-t)^2 t - 3(1-t)t^2 \end{bmatrix} \\
&= (1-t)^3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + 3(1-t)^2 t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 3(1-t)t^2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + t^3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= B_0^3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + B_1^3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + B_2^3 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + B_3^3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\
&= B_0^3 \mathbf{b}_0 + B_1^3 \mathbf{b}_1 + B_2^3 \mathbf{b}_2 + B_3^3 \mathbf{b}_3
\end{aligned}
$$

→French engineer P. Bezier at Renault formulated it in 1971

2.2 Bezier Curves

1) Definition of cubic "Bezier" curves

☑ **Cubic Bezier curve is defined by**

$$\text{if} \quad c_1 B_0^3(t) + c_2 B_1^3(t) + c_3 B_2^3(t) + c_4 B_3^3(t) = 0,$$
$$\text{then} \quad c_1 = c_2 = c_3 = c_4 = 0$$

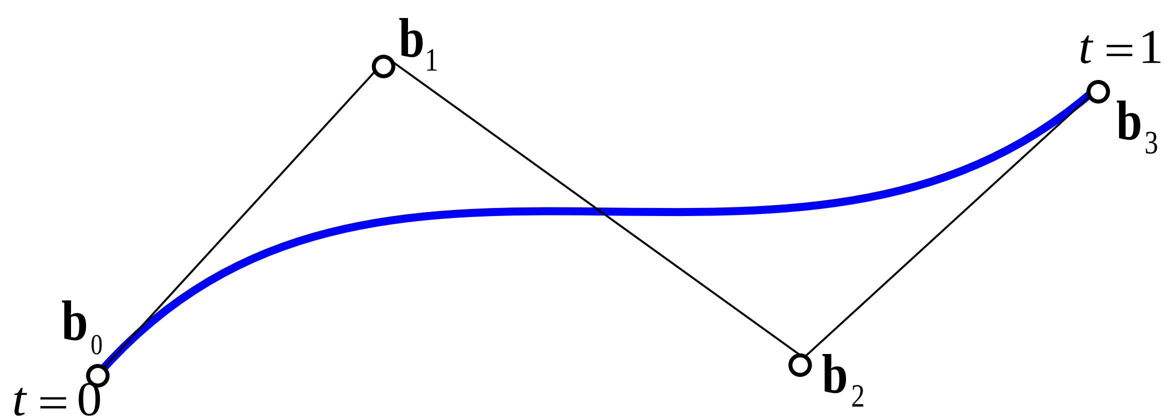
$$\mathbf{r}(t) = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \end{bmatrix} \text{ or } \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \\ \mathbf{z}(t) \end{bmatrix} = (1-t)^3 \mathbf{b}_0 + 3(1-t)^2 t\, \mathbf{b}_1 + 3(1-t) t^2\, \mathbf{b}_2 + t^3\, \mathbf{b}_3$$
$$= \underline{B_0^3(t)}\, \mathbf{b}_0 + \underline{B_1^3(t)}\, \mathbf{b}_1 + \underline{B_2^3(t)}\, \mathbf{b}_2 + \underline{B_3^3(t)}\, \mathbf{b}_3$$

linearly independent

where, $\mathbf{b}_i$: **Bezier control points** (b_{ix}, b_{iy}) **or** (b_{ix}, b_{iy}, b_{iz})

$B_i^3(t)$: **cubic Bernstein polynomial or Bernstein basis function** $\quad \sum_{i=0}^{3} B_i^3(t) = 1, \quad B_i^3(t) \geq 0$

$0 \leq t \leq 1$: **Bezier curve parameter**

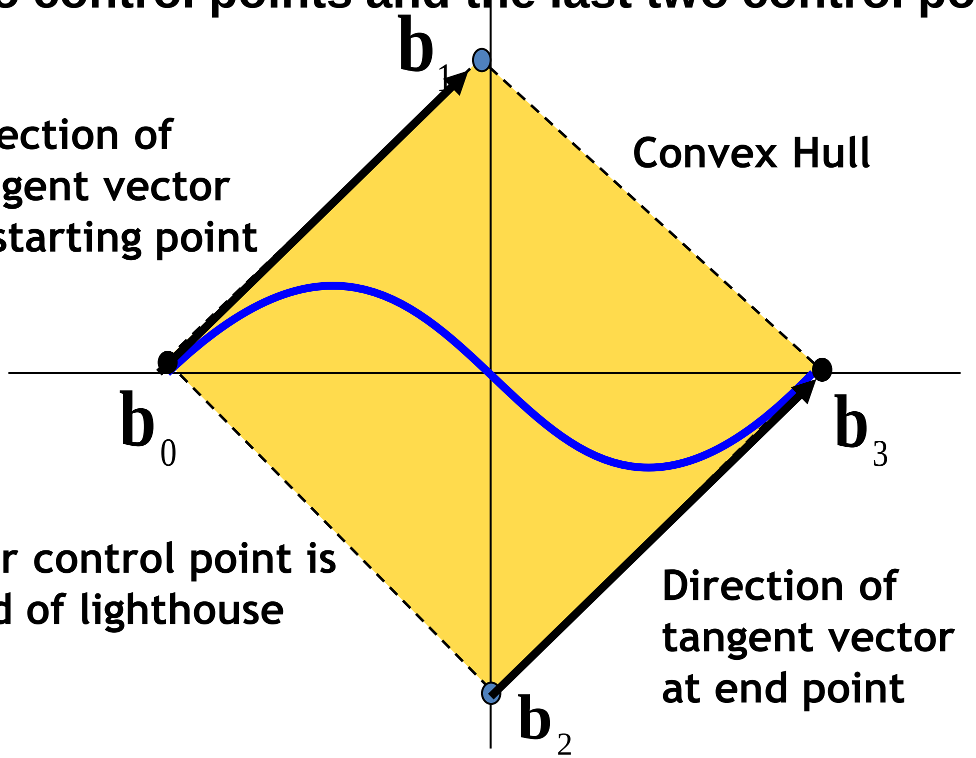


2) Characteristics of Bezier Curves (1)

- **Bezier curves are represented in a *convex hull* which is composed of the outer control points [1)]**

$$\left(\because \sum_{i=0}^{3} B_i^3(t) = 1\right)$$

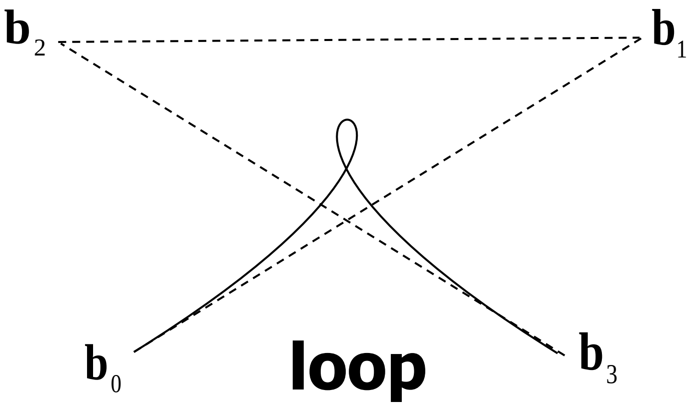
- **The Direction of tangent vector at the start and end points can be obtained from the first two control points and the last two control points**



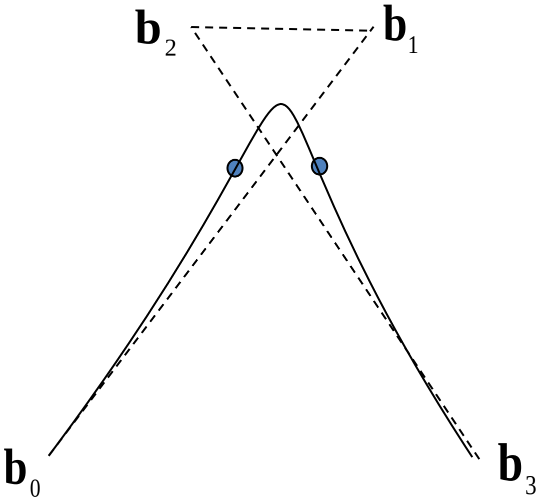
1) *Convex Hull Property*: For all *t*, the curve r(*t*) is in the convex hull of the control polygon.

Characteristics of Bezier Curves (2)

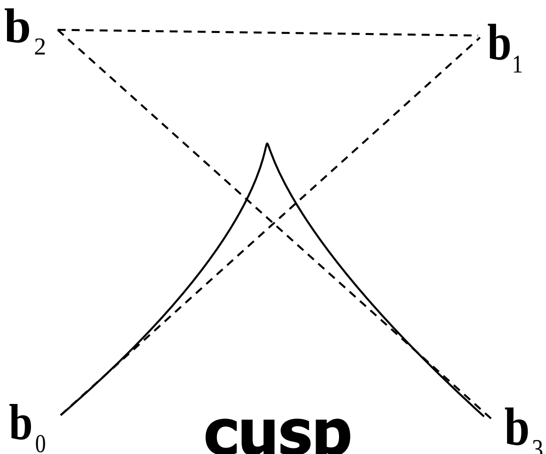
If the control points are moved, then shape of the curve is changed.



loop



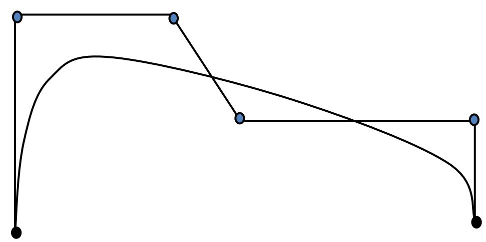
Two inflection points



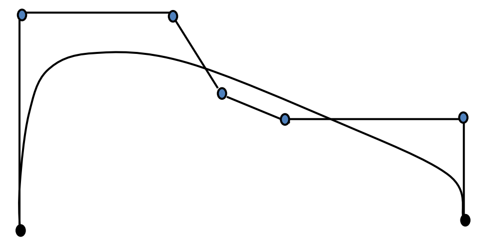
cusp

3) Higher order Bezier Curves

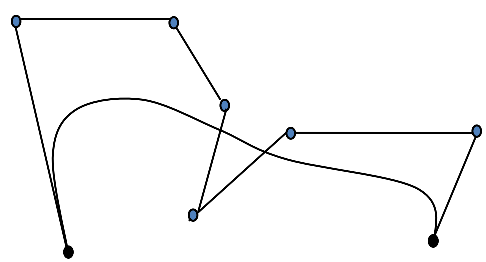
5th-degree Bezier curve



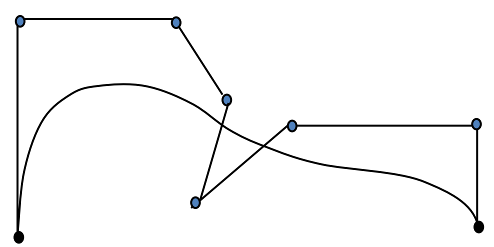
6th-degree Bezier curve



7th-degree Bezier curve



7th-degree Bezier curve



4) Derivatives of Cubic Bezier Curves (1)

$$\mathbf{r}(t) = (1-t)^3\mathbf{b}_0 + 3(1-t)^2 t\mathbf{b}_1 + 3(1-t)t^2\mathbf{b}_2 + t^3\mathbf{b}_3$$

☑ **First derivatives: Tangent vector of the curve**

: "Velocity of body at time = t"

$$\frac{d\mathbf{r}(t)}{dt} = -3(1-t)^2\mathbf{b}_0 + [3(1-t)^2 - 6(1-t)t]\mathbf{b}_1 + [6(1-t)t - 3t^2]\mathbf{b}_2 + 3t^2\mathbf{b}_3$$

$$= 3[\mathbf{b}_1 - \mathbf{b}_0](1-t)^2 + 6[\mathbf{b}_2 - \mathbf{b}_1](1-t)t + 3[\mathbf{b}_3 - \mathbf{b}_2]t^2$$

$$= 3\Delta\mathbf{b}_0(1-t)^2 + 6\Delta\mathbf{b}_1(1-t)t + 3\Delta\mathbf{b}_2 t^2$$

$$= 3(\Delta\mathbf{b}_0 B_0^2 + \Delta\mathbf{b}_1 B_1^2 + \Delta\mathbf{b}_2 B_2^2)$$

where, $\Delta\mathbf{b}_i = \mathbf{b}_{i+1} - \mathbf{b}_i$ **: forward differences**

1st Derivatives of Cubic Bezier Curves(2)

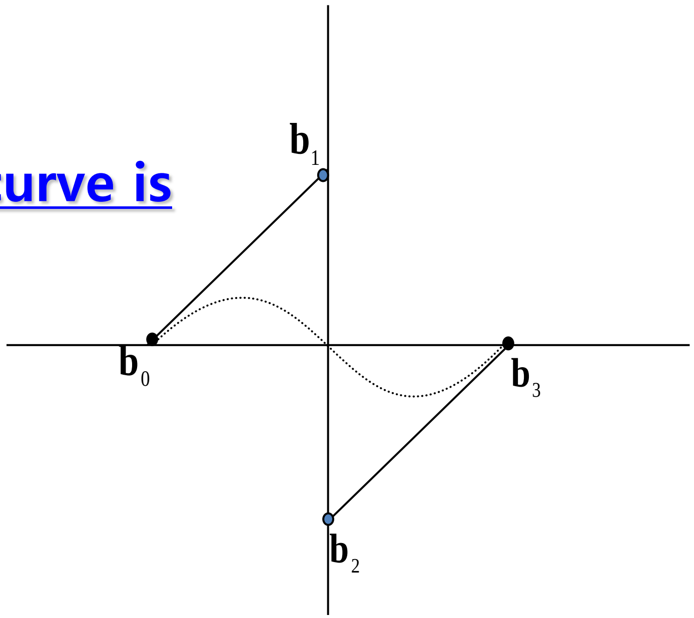
☑ **The derivative of the cubic curve is quadratic curve.**

$$\dot{\mathbf{r}}(t) = \frac{d\mathbf{r}(t)}{dt} = 3(\Delta\mathbf{b}_0 B_0^2 + \Delta\mathbf{b}_1 B_1^2 + \Delta\mathbf{b}_2 B_2^2). \; = 3\Delta\mathbf{b}_0(1-t)^2 + 6\Delta\mathbf{b}_1(1-t)t + 3\Delta\mathbf{b}_2 t^2$$

- **where,** B_i^2 **: quadratic Bernstein basis function.**

☑ **Most important tangent vectors at the curve is Endpoints tangent vectors:**

$$\dot{\mathbf{r}}(0) = 3\Delta\mathbf{b}_0 = 3(\mathbf{b_1} - \mathbf{b_o}),$$
$$\dot{\mathbf{r}}(1) = 3\Delta\mathbf{b}_2 = 3(\mathbf{b}_3 - \mathbf{b}_2)$$



Higher order Bezier Curves (1)

☑ **A Bezier Curve of degree n can be defined by;**

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^n(t) + \mathbf{b}_1 B_1^n(t) + \ldots\ldots + \mathbf{b}_n B_n^n(t).$$

☑ **where,** $B_i^n(t)$ **: Bernstein Polynomial Function.**

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i},$$

$$\binom{n}{i} = {}_nC_i = \begin{cases} \dfrac{n!}{i!(n-i)!} & \textbf{if } 0 \le i \le n \\ \mathbf{0} & \textbf{else} \end{cases}$$

$$B_i^n(t) = tB_{i-1}^{n-1}(t) + (1-t)B_i^{n-1}(t) \quad \text{with } B_0^0(t) \equiv 1$$

☑ **For cubic case, the Bezier curve is:**

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^3(t) + \mathbf{b}_1 B_1^3(t) + \mathbf{b}_2 B_2^3(t) + \mathbf{b}_3 B_3^3(t).$$

Higher order Bezier Curves (2)

- **Bernstein Polynomial Function:**

$$
\begin{aligned}
[(1-t)+t]^2 &= (1-t)^2 + 2(1-t)t + t^2 \\
&= B_0^2(t) + B_1^2(t) + B_2^2(t),
\end{aligned}
$$

$$
\begin{aligned}
[(1-t)+t]^3 &= [(1-t)+t]^2[(1-t)+t] \\
&= (1-t)^3 + 3(1-t)^2 t + 3(1-t)t^2 + t^3 \\
&= B_0^3(t) + B_1^3(t) + B_2^3(t) + B_3^3(t),
\end{aligned}
$$

$$
\begin{aligned}
[(1-t)+t]^4 &= [(1-t)+t]^3[(1-t)+t] \\
&= (1-t)^4 + 4(1-t)^3 t + 6(1-t)^2 t^2 + 4(1-t)t^3 + t^4 \\
&= B_0^4(t) + B_1^4(t) + B_2^4(t) + B_3^4(t) + B_4^4(t)
\end{aligned}
$$

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & \mathbf{1} & & \mathbf{1} & & & \\
 & & \mathbf{1} & & \mathbf{2} & & \mathbf{1} & & \\
 & \mathbf{1} & & \mathbf{3} & & \mathbf{3} & & \mathbf{1} & \\
\mathbf{1} & & \mathbf{4} & & \mathbf{6} & & \mathbf{4} & & \mathbf{1}
\end{array}
$$

Pascal's triangle

Derivatives of Higher Order Bezier Curves (1)

☑ **For Cubic Case (*n* = 3),**

$$\dot{\mathbf{r}}(t) = 3[\Delta\mathbf{b}_0 B_0^2 + \Delta\mathbf{b}_1 B_1^2 + \Delta\mathbf{b}_2 B_2^2].$$

☑ **For degree = *n*,**

$$\dot{\mathbf{r}}(t) = n[\Delta\mathbf{b}_0 B_0^{n-1} + \Delta\mathbf{b}_1 B_1^{n-1} + \ldots\ldots\ldots + \Delta\mathbf{b}_{n-1} B_{n-1}^{n-1}].$$

where $\Delta\mathbf{b}_i = \mathbf{b}_{i+1} - \mathbf{b}_i$ **: forward difference.**

☑ **Bezier Curve→ differentiated by more than one by parameter '*t*'.**

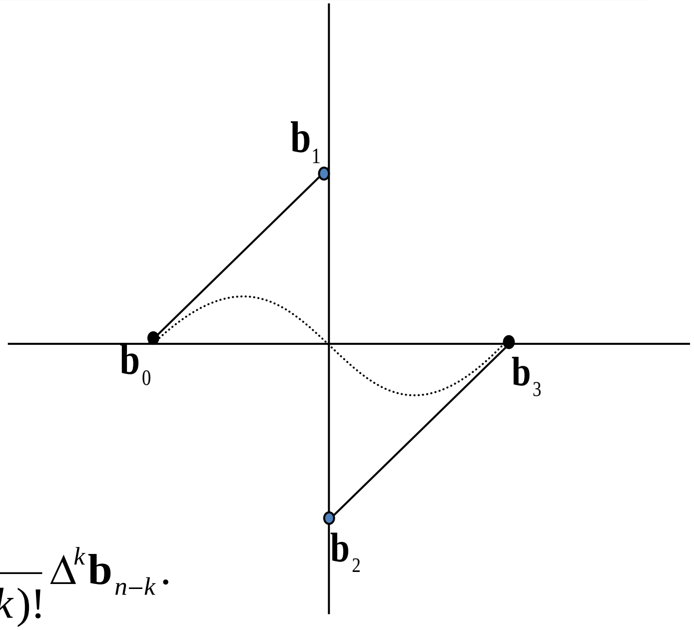
☑ **For the *k***

$$\frac{d^k\mathbf{r}(t)}{dt^k} = \frac{n!}{(n-k)!}[\Delta^k\mathbf{b}_0 B_0^{n-k}(t) + \Delta^k\mathbf{b}_1 B_1^{n-k}(t)\ldots\ldots + \Delta^k\mathbf{b}_{n-k} B_{n-k}^{n-k}(t)].$$

Derivatives of Higher Order Bezier Curves (2)

- ☑ where, Δ^k :forward operator.
- ☑ we can get $\Delta^k \mathbf{b}_i = \Delta^{k-1}\mathbf{b}_{i+1} - \Delta^{k-1}\mathbf{b}_i$.
  where, $\Delta^0 \mathbf{b}_i = \mathbf{b}_i$.
- ☑ for k=2 : $\mathbf{b}_{i+2} - 2\mathbf{b}_{i+1} + \mathbf{b}_i$.
- ☑ for k=3 : $\mathbf{b}_{i+3} - 3\mathbf{b}_{i+2} + 3\mathbf{b}_{i+1} - \mathbf{b}_i$.
- ☑ for k=4 : $\mathbf{b}_{i+4} - 4\mathbf{b}_{i+3} + 6\mathbf{b}_{i+2} - 4\mathbf{b}_{i+1} + \mathbf{b}_i$.
- ☑ the k^{th} derivative of $\mathbf{r}(0)$ and $\mathbf{r}(1)$;

$$\mathbf{r}^k(0) = \frac{n!}{(n-k)!}\Delta^k \mathbf{b}_0 \quad \text{and} \quad \mathbf{r}^k(1) = \frac{n!}{(n-k)!}\Delta^k \mathbf{b}_{n-k}.$$



■ **For n=3, k=2;**

$$\mathbf{r}^2(0) = \frac{3!}{(3-2)!}\Delta^2 \mathbf{b}_0$$
$$= 6(\Delta^1\mathbf{b}_1 - \Delta^1\mathbf{b}_0)$$
$$= 6((\Delta^0\mathbf{b}_2 - \Delta^0\mathbf{b}_1) - (\Delta^0\mathbf{b}_1 - \Delta^0\mathbf{b}_0))$$
$$= 6(\Delta^0\mathbf{b}_2 - 2\Delta^0\mathbf{b}_1 + \Delta^0\mathbf{b}_0)$$
$$= 6(\mathbf{b}_2 - 2\mathbf{b}_1 + \mathbf{b}_0)$$

$$\mathbf{r}^2(1) = \frac{3!}{(3-2)!}\Delta^2 \mathbf{b}_1$$
$$= 6(\Delta^1\mathbf{b}_2 - \Delta^1\mathbf{b}_1)$$
$$= 6((\Delta^0\mathbf{b}_3 - \Delta^0\mathbf{b}_2) - (\Delta^0\mathbf{b}_2 - \Delta^0\mathbf{b}_1))$$
$$= 6(\Delta^0\mathbf{b}_3 - 2\Delta^0\mathbf{b}_2 + \Delta^0\mathbf{b}_1)$$
$$= 6(\mathbf{b}_3 - 2\mathbf{b}_2 + \mathbf{b}_1)$$

5) Matrix form of Bezier curves(1)

☑ Cubic Bezier Curve

$$\mathbf{r}(t) = (1-t)^3\mathbf{b}_0 + 3(1-t)^2 t\mathbf{b}_1 + 3(1-t)t^2\mathbf{b}_2 + t^3\mathbf{b}_3$$

☑ applying the dot product to above equation;

$$\mathbf{r}(t) = \begin{bmatrix} \mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3} \end{bmatrix} \begin{bmatrix} (1-t)^3 \\ 3(1-t)^2 t \\ 3(1-t)\,t^2 \\ t^3 \end{bmatrix}$$

Matrix form of Bezier curves(2)

☑ **The Matrix form of Bezier Curve is**

$$\mathbf{r}(t)=\begin{bmatrix}\mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3}\end{bmatrix}\begin{bmatrix}(1-t)^3\\ 3(1-t)^2 t\\ 3(1-t)\,t^2\\ t^3\end{bmatrix}=\begin{bmatrix}\mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3}\end{bmatrix}\begin{bmatrix}B_0^3(\mathbf{t})\\ B_1^3(\mathbf{t})\\ B_2^3(\mathbf{t})\\ B_3^3(\mathbf{t})\end{bmatrix}$$

Conversion to the monomial form: $\mathbf{r}(t)=\mathbf{a}_0+\mathbf{a}_1 t+\mathbf{a}_2 t^2+\mathbf{a}_3 t^3$

$$\mathbf{r}(t)=\begin{bmatrix}\mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3}\end{bmatrix}\begin{bmatrix}(1-t)^3\\ 3(1-t)^2 t\\ 3(1-t)\,t^2\\ t^3\end{bmatrix}=\begin{bmatrix}\mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3}\end{bmatrix}\begin{bmatrix}1 & -3 & 3 & -1\\ 0 & 3 & -6 & 3\\ 0 & 0 & 3 & -3\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\ t\\ t^2\\ t^3\end{bmatrix}$$

$$=\begin{bmatrix}\mathbf{a_0} & \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3}\end{bmatrix}\begin{bmatrix}1\\ t\\ t^2\\ t^3\end{bmatrix}$$

Matrix form of Bezier curves(3)

☑ **The Matrix form of Monomial Curve is**

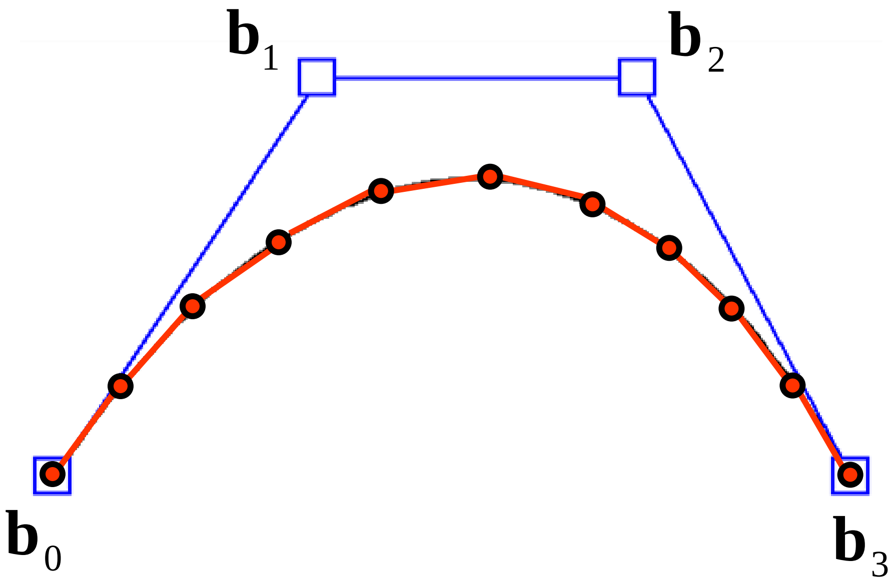
$$\mathbf{r}(t) = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{a}_2 t^2 + \mathbf{a}_3 t^3 = \begin{bmatrix} \mathbf{a_0} & \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix}$$

Transformation to the Bezier form:

$$\begin{aligned} \mathbf{r}(t) &= (1-t)^3 \mathbf{b}_0 + 3(1-t)^2 t \mathbf{b}_1 + 3(1-t)t^2 \mathbf{b}_2 + t^3 \mathbf{b}_3 \\ &= B_0^3(t)\,\mathbf{b}_0 + B_1^3(t)\,\mathbf{b}_1 + B_2^3(t)\,\mathbf{b}_2 + B_3^3(t)\,\mathbf{b}_3 \\ &= \begin{bmatrix} \mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3} \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \end{bmatrix} \end{aligned}$$

$$\therefore \begin{bmatrix} \mathbf{b_0} & \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3} \end{bmatrix} = \begin{bmatrix} \mathbf{a_0} & \mathbf{a_1} & \mathbf{a_2} & \mathbf{a_3} \end{bmatrix} \begin{bmatrix} 1 & -3 & 3 & -1 \\ 0 & 3 & -6 & 3 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}$$

6) Programming for Bezier Curve class



$$\mathbf{r}(t) = \boxed{(1-t)^3}\mathbf{b}_0 + \boxed{3(1-t)^2 t}\mathbf{b}_1 + \boxed{3(1-t)t^2}\mathbf{b}_2 + \boxed{t^3}\mathbf{b}_3$$

1) Bezier Curve is defined by

- Degree
- Control Point

<u>Member Variables of Bezier Curve Class</u>
int n: degree of Bezier Curve
Vector* m_ControlPoint: Control Point
int m_nControlPoint: the number of Control Point

2) Calculation of Bernstein Polynomial

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i},$$

$$\binom{n}{i} = {}_nC_i = \begin{cases} \dfrac{n!}{i!(n-i)!} & \textbf{if } 0 \le i \le n \\ \mathbf{0} & \textbf{else} \end{cases}$$

3) Bezier Curve construction

- Construct the curve divided by line segment
- After divide a parameter t(0~1) into n equal parts, find the points on the curve at the each point to be divided.
- Visualize the curve by connecting points with straight lines

Sample code for Bezier Curve class(1)

```
numberifndef __BezierCurve_h__
numberdefine __BezierCurve_h__

numberinclude “vector.h”

class BezierCurve {
public:
    int n;  // degree of Bezier Curve
    Vector* m_ControlPoint;   int m_nControlPoint;
    BezierCurve();
    ~BezierCurve();

    void SetDegree(int degree);
    void SetControlPoint(Vector* pControlPoint, int nControlPoint);
    Vector CalcPoint(double t);
    double B (int i, double t);                    // Bernstein Polynomial
};
numberendif
```

<u>Member Variables</u>
int n: degree of Bezier Curve
Vector* m_ControlPoint: Control Point
int m_nControlPoint: the number of Control Point

Sample code for Bezier Curve class(2)

```
BezierCurve::BezierCurve () {
  m_ControlPoint = 0;  n= 0;
  m_nControlPoint = 0;
}
BezierCurve::~BezierCurve () {
  if(m_ControlPoint) delete[] m_ControlPoint;
}
void BezierCurve::SetControlPoint(Vector* pControlPoint, int nControlPoint) {
  SetDegree( nControlPoint-1 );
  if(m_ControlPoint) delete[] m_ControlPoint;
  m_ControlPoint = new Vector[nControlPoint];
  for(int i=0; i < nControlPoint; i++) {
    m_ControlPoint[i] = pControlPoint[i];
  }
}
void BezierCurve::SetDegree(int degree){
  n = degree;
}
```

Sample code for Bezier Curve class(3)

```
Vector BezierCurve:: CalcPoint(double t) {
    Vector PointOnCurve(0,0,0);
    if ( t < 0.0 || t > 1.0 ) {
        return PointOnCurve;
    }
    for(int i = 0; i < m_nControlPoint; i++){
        PointOnCurve = PointOnCurve + m_ControlPoint[i] * B(i,t);
    }
    return PointOnCurve;
}

double BezierCurve:: B (int i, double t) {
    double result = 0;
    // Calculate ith Berstein Polynomial at parameter t
    result = comb(n, i) * pow(t, i) * pow(1.0 - t, n-i);
    return result;
}
```

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^n(t) + \mathbf{b}_1 B_1^n(t) + \dots\dots + \mathbf{b}_n B_n^n(t).$$

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i},$$

$$\binom{n}{i} = {}_nC_i = \begin{cases} \dfrac{n!}{i!(n-i)!} & \textbf{if } 0 \le i \le n \\ \mathbf{0} & \textbf{else} \end{cases}$$

2.3 Degree Elevation / Reduction of Bezier Curves

1) Degree Elevation (1)

☑ Objective

- To connect curves with different degree, we have to change the degree of the curves to be same.

  Ex) 3rd-degree Bezier curve + 4th-degree Bezier curve

  → 4th-degree Bezier curve + 4th-degree Bezier curve

- Free curve design by using more control points

  (Number of Bezier control point = degree+1)

Degree Elevation (2)

☑ 2^{nd}-degree Bézier curve->3^{rd}-degree Bézier curve

$$\mathbf{r}(t) = (1-t)^2\mathbf{b}_0 + 2(1-t)t\mathbf{b}_1 + t^2\mathbf{b}_2 \qquad \times\, [t + (1-t)]$$

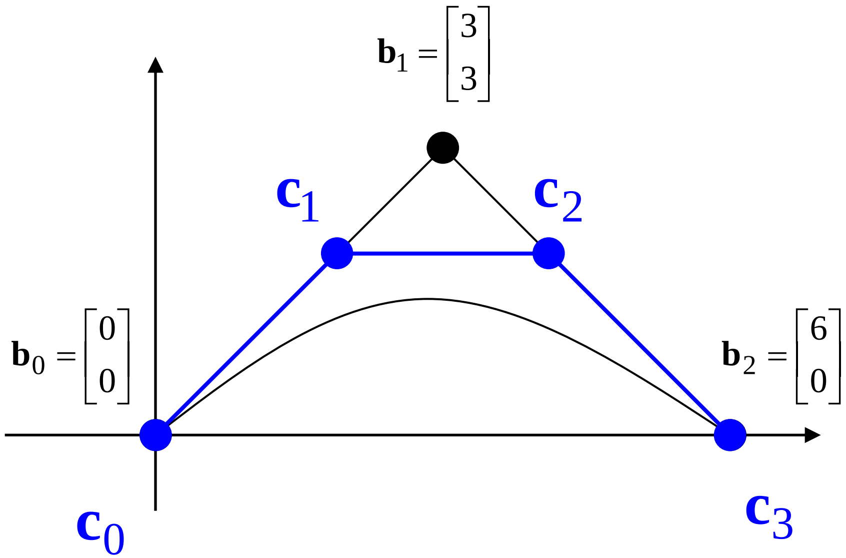
$$\mathbf{r}(t) = [t(1-t)^2 + (1-t)^3]\mathbf{b}_0 + 2[t^2(1-t) + (1-t)^2 t]\mathbf{b}_1 + [t^3 + t^2(1-t)]\mathbf{b}_2$$

$$\mathbf{r}(t) = (1-t)^3\boxed{\mathbf{b}_0} + 3(1-t)^2 t\boxed{[\frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_1]} + 3(1-t)t^2\boxed{[\frac{2}{3}\mathbf{b}_1 + \frac{1}{3}\mathbf{b}_2]} + t^3\boxed{\mathbf{b}_2}$$

$\boxed{}$: New control point

⇨ **Thus the original** 2^{nd}-degree Bézier curve **may also be written as a** 3^{rd}-degree Bézier curve **with new control points**

Degree Elevation (3)



$$\mathbf{c}_0 = \mathbf{b}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\mathbf{c}_1 = [\frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_1] = \begin{bmatrix} 2 \\ 2 \end{bmatrix},$$

$$\mathbf{c}_2 = [\frac{2}{3}\mathbf{b}_1 + \frac{1}{3}\mathbf{b}_2] = \begin{bmatrix} 4 \\ 2 \end{bmatrix},$$

$$\mathbf{c}_3 = \mathbf{b}_2 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

$$\mathbf{r}(t) = (1-t)^2\mathbf{b}_0 + 2(1-t)t\mathbf{b}_1 + t^2\mathbf{b}_2$$

$$\mathbf{r}(t) = (1-t)^3\mathbf{b}_0 + 3(1-t)^2 t[\frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_1] + 3(1-t)t^2[\frac{2}{3}\mathbf{b}_1 + \frac{1}{3}\mathbf{b}_2] + t^3\mathbf{b}_2$$

Degree Elevation (4)

☑ **Degree elevation of a degree n Bézier curve with control point $\mathbf{b}_0, \ldots, \mathbf{b}_n$ to n+1 degree**

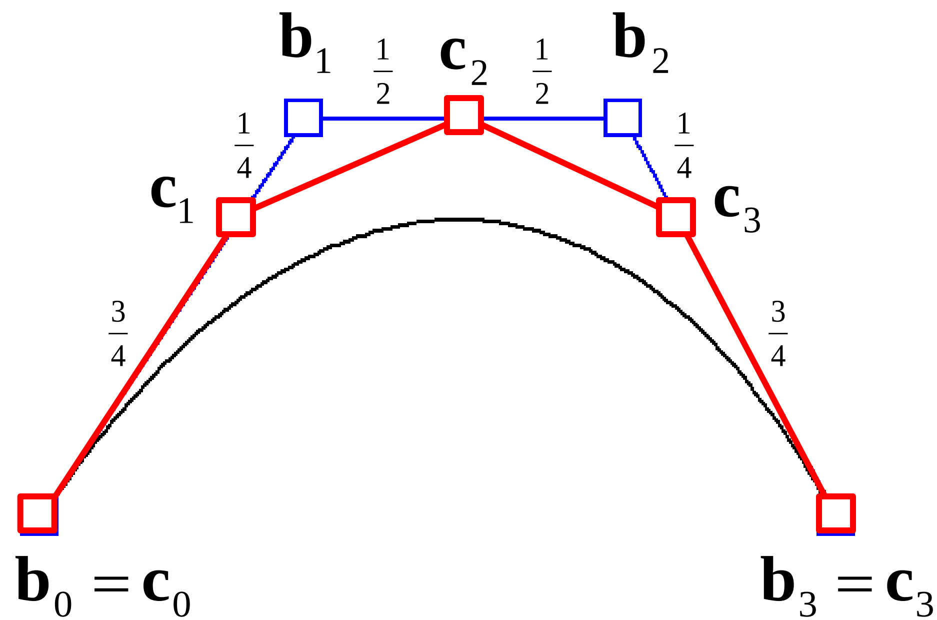
$$\mathbf{c}_0 = \mathbf{b}_0,$$

$$\vdots$$

$$\mathbf{c}_i = \frac{i}{n+1}\mathbf{b}_{i-1} + \left(1 - \frac{i}{n+1}\right)\mathbf{b}_i,$$

$$\vdots$$

$$\mathbf{c}_{n+1} = \mathbf{b}_n$$



Degree Elevation: 3rd-degree → 4th-degree

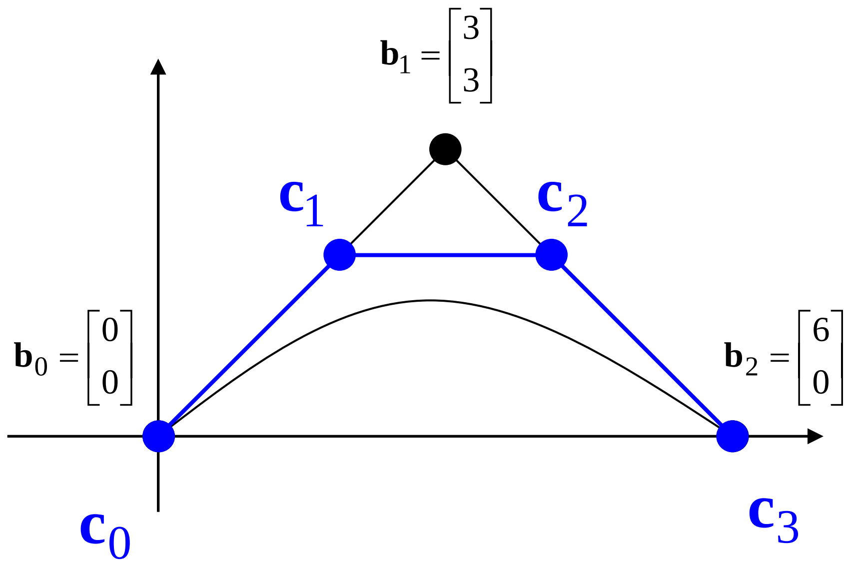
Degree Elevation (5)

$$
\begin{aligned}
\mathbf{c}_0 &= \mathbf{b}_0, \\
&\vdots \\
\mathbf{c}_i &= \frac{i}{n+1}\mathbf{b}_{i-1} + \left(1 - \frac{i}{n+1}\right)\mathbf{b}_i, \\
&\vdots \\
\mathbf{c}_{n+1} &= \mathbf{b}_n
\end{aligned}
$$

$$
\begin{array}{cc}
 & \overbrace{\hphantom{1 \quad * \quad * \quad * \quad * \quad *}}^{n+1 \ \ colums} \\
n+2 \ \ rows \left\{ \vphantom{\begin{matrix} 1 \\ * \\ * \\ \vdots \\ * \\ 1 \end{matrix}} \right. &
\begin{bmatrix}
1 & & & & & \\
* & * & & & & \\
 & * & * & & & \\
 & & & \vdots & \vdots & \\
 & & & & * & * \\
 & & & & & 1
\end{bmatrix}
\end{array}
\begin{bmatrix} \mathbf{b}_0 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}
=
\begin{bmatrix} \mathbf{c}_0 \\ \vdots \\ \mathbf{c}_{n+1} \end{bmatrix}
$$

$$\mathbf{DB} = \mathbf{C}$$

Degree Elevation (6)



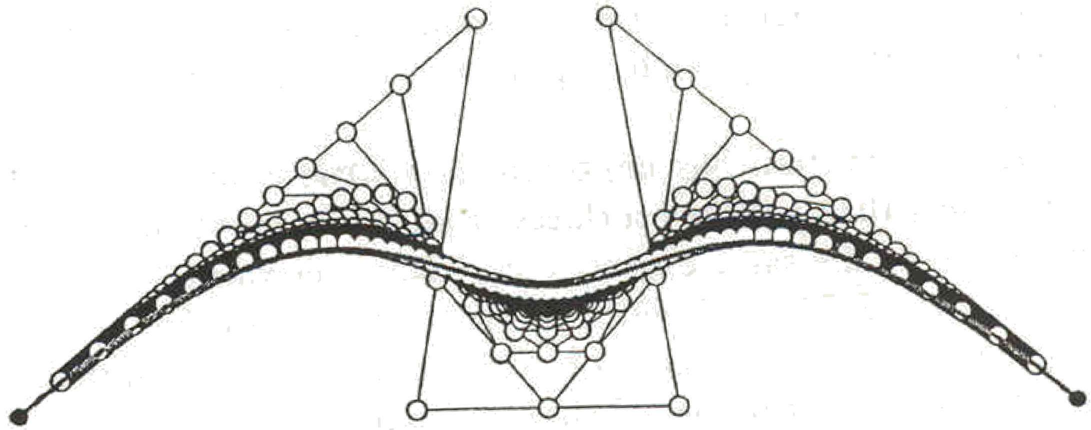
$$\mathbf{c}_0 = \mathbf{b}_0,$$
$$\mathbf{c}_1 = [\frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_1],$$
$$\mathbf{c}_2 = [\frac{2}{3}\mathbf{b}_1 + \frac{1}{3}\mathbf{b}_2],$$
$$\mathbf{c}_3 = \mathbf{b}_2$$

$$\mathbf{DB} = \mathbf{C}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3 & 3 \\ 6 & 0 \end{bmatrix} = \mathbf{C}$$

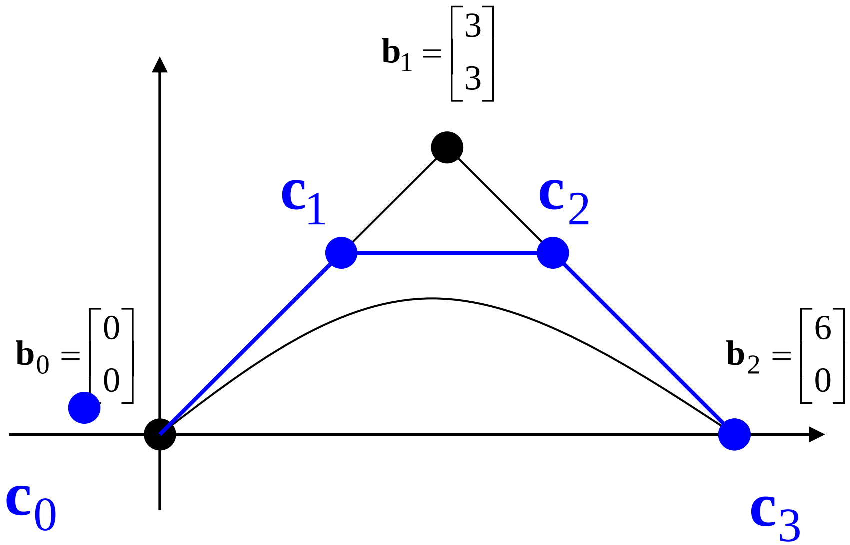
$$\therefore \mathbf{C} = \begin{bmatrix} 0 & 0 \\ 2 & 2 \\ 4 & 2 \\ 6 & 0 \end{bmatrix}$$

Repeated Degree Elevation



Repeated degree elevation
: a sequence of polygons approaching the curve

2) Degree Reduction



$$\mathbf{c}_0 = \mathbf{b}_0,$$
$$\mathbf{c}_1 = [\frac{1}{3}\mathbf{b}_0 + \frac{2}{3}\mathbf{b}_1],$$
$$\mathbf{c}_2 = [\frac{2}{3}\mathbf{b}_1 + \frac{1}{3}\mathbf{b}_2],$$
$$\mathbf{c}_3 = \mathbf{b}_2$$

$$\mathbf{DB} = \mathbf{C}$$

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 2/3 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 \\ 2 & 2 \\ 4 & 2 \\ 6 & 0 \end{bmatrix}$$

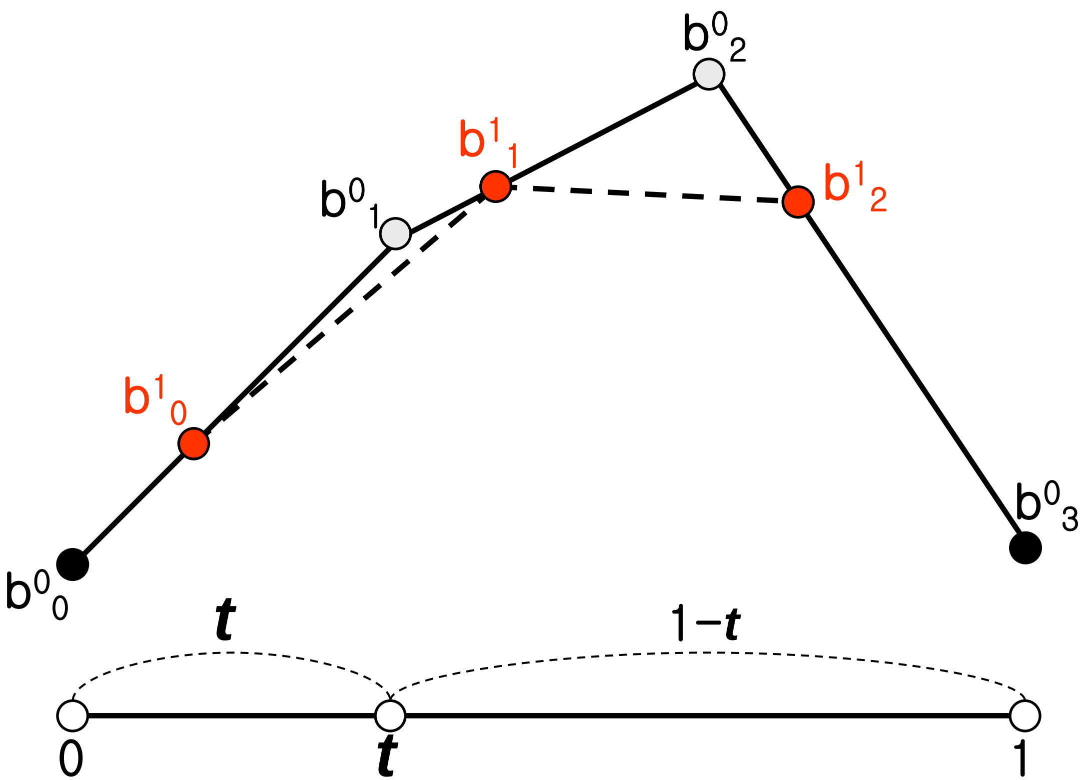
$$\mathbf{D}^T\mathbf{DB} = \mathbf{D}^T\mathbf{C}$$

$$\therefore \mathbf{B} = \left(\mathbf{D}^T\mathbf{D}\right)^{-1}\mathbf{D}^T\mathbf{C}$$

$$\mathbf{D}^T\mathbf{D} = \frac{1}{9}\begin{bmatrix} 10 & 2 & 0 \\ 2 & 8 & 2 \\ 0 & 2 & 10 \end{bmatrix}, \quad \mathbf{D}^T\mathbf{C} = \frac{1}{3}\begin{bmatrix} 2 & 2 \\ 12 & 8 \\ 22 & 2 \end{bmatrix}, \quad \therefore \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 3 & 3 \\ 6 & 0 \end{bmatrix}$$

2.4 de Casteljau algorithm

1) de Casteljau algorithm & Bezier curves (1)



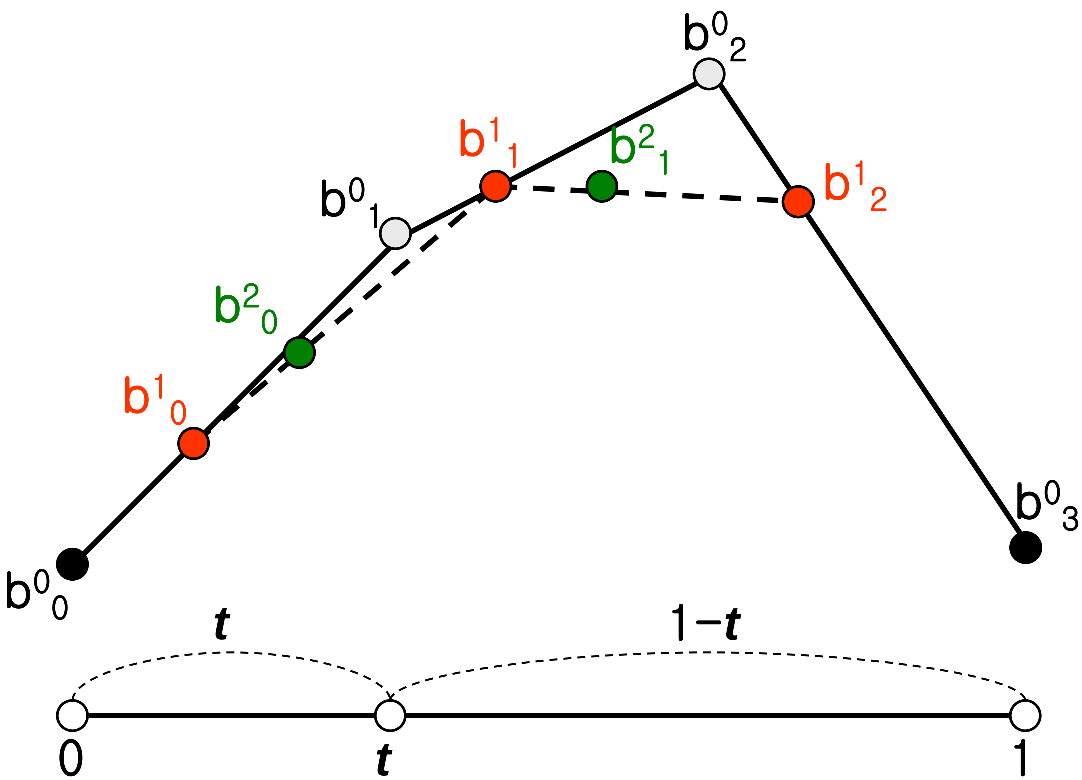
Linear interpolation

$$\mathbf{b}_0^1(t) = (1-t)\mathbf{b}_0^0 + t\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(t) = (1-t)\mathbf{b}_1^0 + t\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(t) = (1-t)\mathbf{b}_2^0 + t\mathbf{b}_3^0$$

de Casteljau algorithm & Bezier curves (2)



Linear interpolation

$$\mathbf{b}_0^1(t) = (1-t)\mathbf{b}_0^0 + t\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(t) = (1-t)\mathbf{b}_1^0 + t\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(t) = (1-t)\mathbf{b}_2^0 + t\mathbf{b}_3^0$$

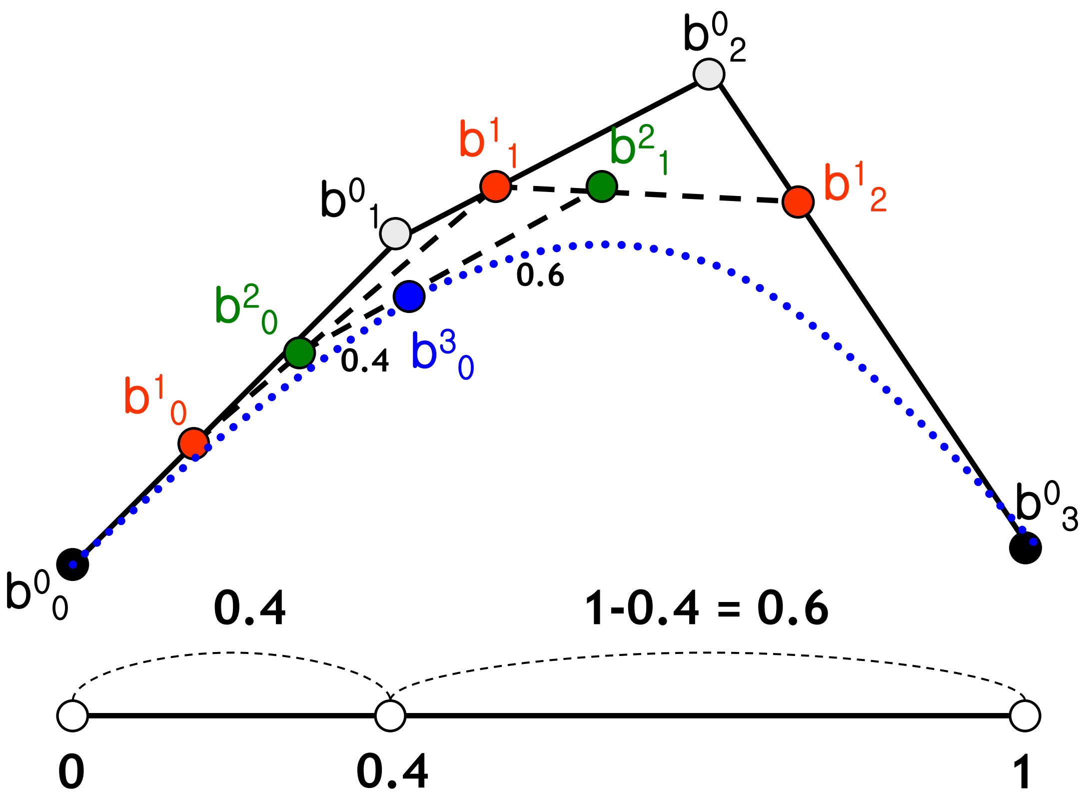
$$\mathbf{b}_0^2(t) = (1-t)\mathbf{b}_0^1 + t\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(t) = (1-t)\mathbf{b}_1^1 + t\mathbf{b}_2^1$$

Example of de Casteljau algorithm (3)

- de Casteljau algorithm at t = 0.4

$t = 0.4$



Linear interpolation

$$\mathbf{b}_0^1(0.4) = (1-0.4)\mathbf{b}_0^0 + 0.4\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(0.4) = (1-0.4)\mathbf{b}_1^0 + 0.4\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(0.4) = (1-0.4)\mathbf{b}_2^0 + 0.4\mathbf{b}_3^0$$

$$\mathbf{b}_0^2(0.4) = (1-0.4)\mathbf{b}_0^1 + 0.4\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(0.4) = (1-0.4)\mathbf{b}_1^1 + 0.4\mathbf{b}_2^1$$

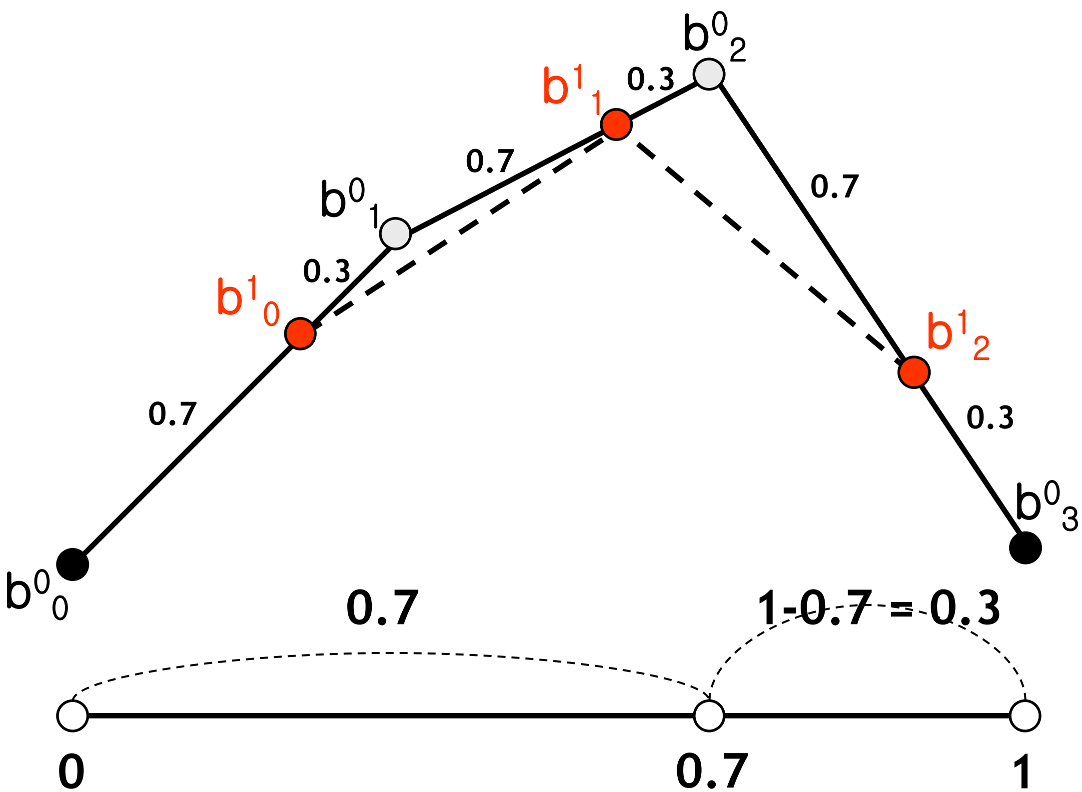
$$\mathbf{b}_0^3(0.4) = (1-0.4)\,\mathbf{b}_0^2 + 0.4\mathbf{b}_1^2$$

$$\mathbf{b}_0^3(0.4) = (1-0.4)^3\mathbf{b}_0^0 + 3\cdot 0.4(1-0.4)^2\mathbf{b}_1^0 + 3\cdot 0.4^2(1-0.4)\mathbf{b}_2^0 + 0.4^3\mathbf{b}_3^0$$

Example of de Casteljau algorithm (1)

- de Casteljau algorithm at t = 0.7

$t = 0.7$



Linear interpolation

$$\mathbf{b}_0^1(0.7) = (1-0.7)\mathbf{b}_0^0 + 0.7\mathbf{b}_1^0$$

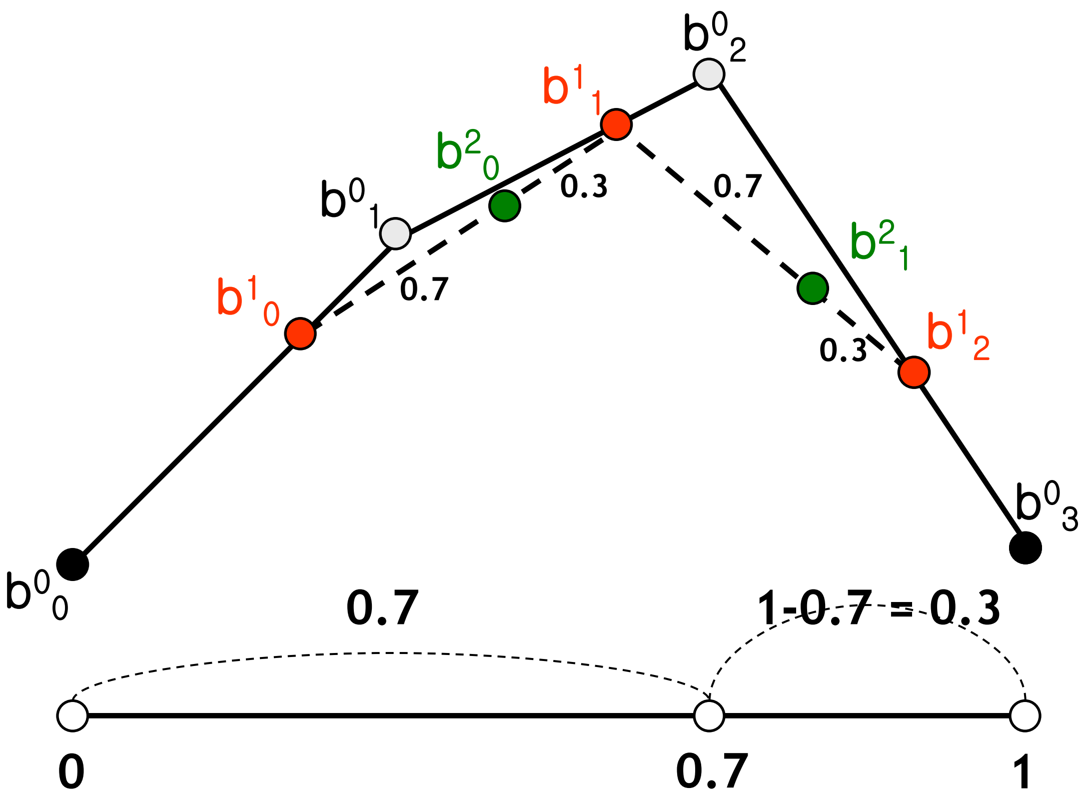
$$\mathbf{b}_1^1(0.7) = (1-0.7)\mathbf{b}_1^0 + 0.7\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(0.7) = (1-0.7)\mathbf{b}_2^0 + 0.7\mathbf{b}_3^0$$

Example of de Casteljau algorithm (2)

- de Casteljau algorithm at t = 0.7

$t = 0.7$



Linear interpolation

$$\mathbf{b}_0^1(0.7) = (1-0.7)\mathbf{b}_0^0 + 0.7\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(0.7) = (1-0.7)\mathbf{b}_1^0 + 0.7\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(0.7) = (1-0.7)\mathbf{b}_2^0 + 0.7\mathbf{b}_3^0$$

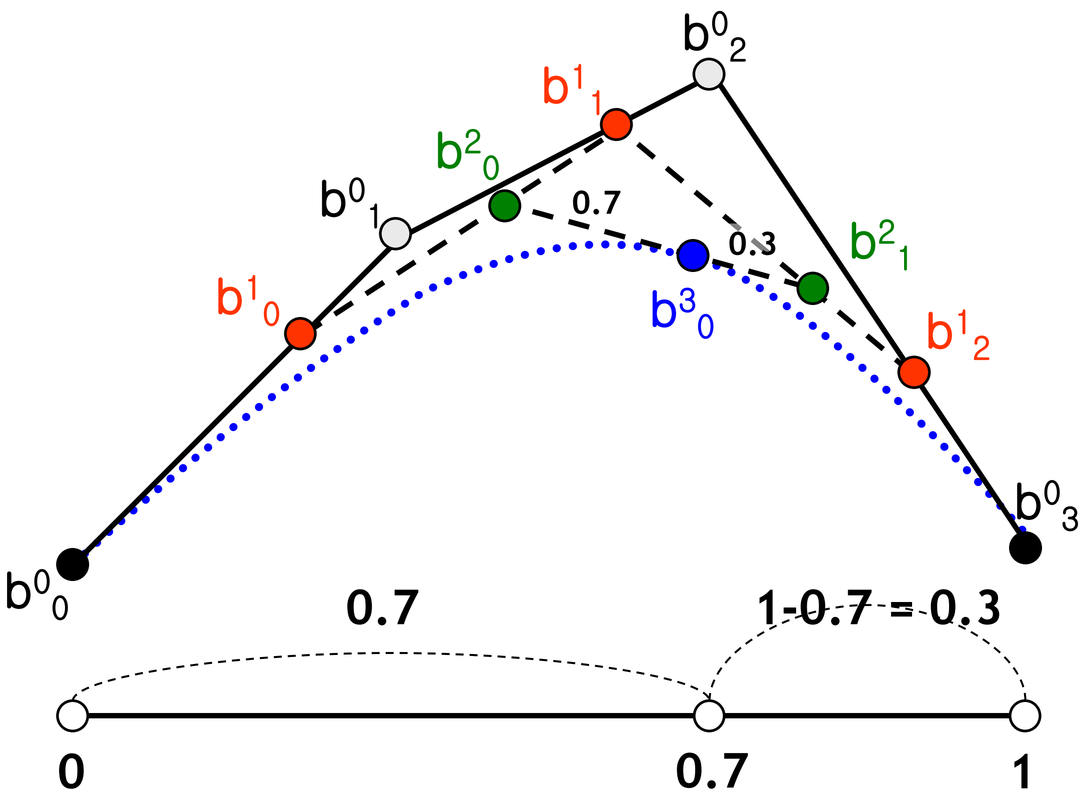
$$\mathbf{b}_0^2(0.7) = (1-0.7)\,\mathbf{b}_0^1 + 0.7\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(0.7) = (1-0.7)\,\mathbf{b}_1^1 + 0.7\mathbf{b}_2^1$$

Example of de Casteljau algorithm (3)

- de Casteljau algorithm at t = 0.7

$t = 0.7$



Linear interpolation

$$\mathbf{b}_0^1(0.7) = (1-0.7)\mathbf{b}_0^0 + 0.7\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(0.7) = (1-0.7)\mathbf{b}_1^0 + 0.7\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(0.7) = (1-0.7)\mathbf{b}_2^0 + 0.7\mathbf{b}_3^0$$

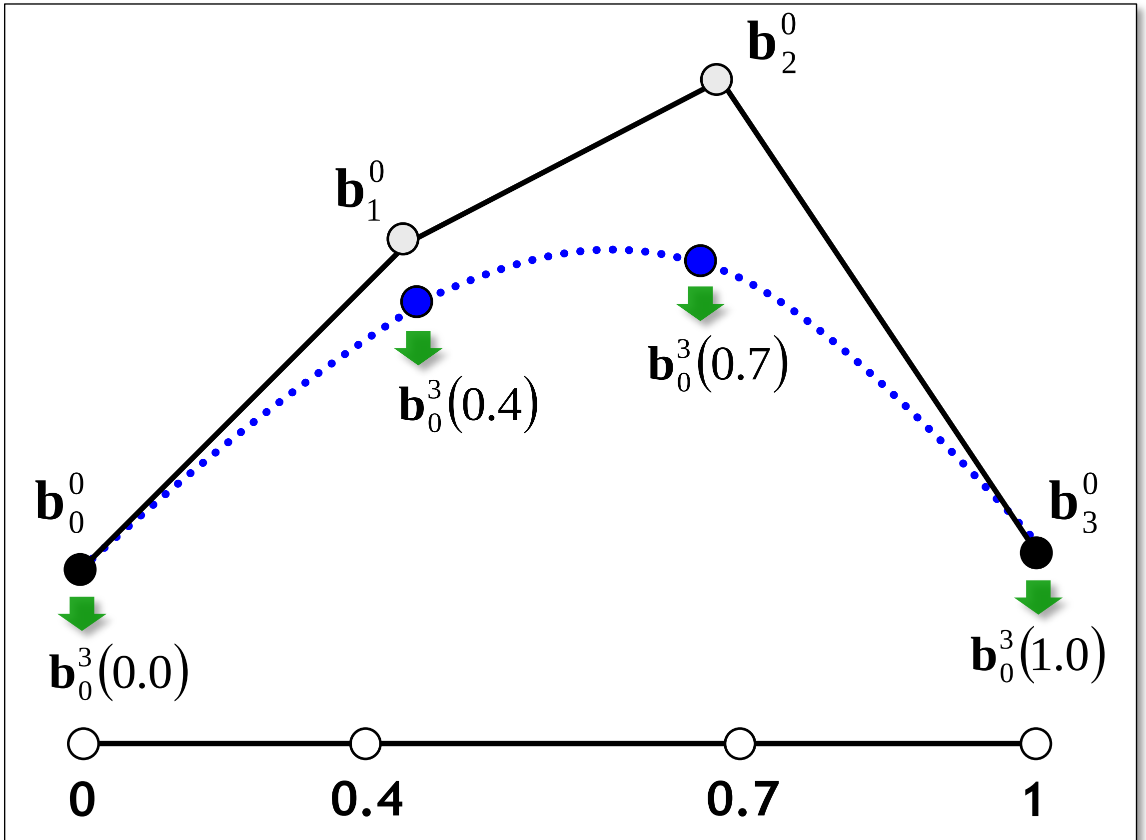
$$\mathbf{b}_0^2(0.7) = (1-0.7)\,\mathbf{b}_0^1 + 0.7\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(0.7) = (1-0.7)\,\mathbf{b}_1^1 + 0.7\mathbf{b}_2^1$$

$$\mathbf{b}_0^3(0.7) = (1-0.7)\,\mathbf{b}_0^2 + 0.7\mathbf{b}_1^2$$

$$\mathbf{b}_0^3(0.7) = (1-0.7)^3\mathbf{b}_0^0 + 3\cdot 0.7(1-0.7)^2\mathbf{b}_1^0 + 3\cdot 0.7^2(1-0.7)\mathbf{b}_2^0 + 0.7^3\mathbf{b}_3^0$$

Example of de Casteljau algorithm (4)



Given

$\mathbf{b}_0^0,\ \mathbf{b}_1^0,\ \mathbf{b}_2^0,\ \mathbf{b}_3^0$

Find

Points on bezier curve at t = 0.0, 0.4, 0.7, 1.0

$$\mathbf{b}_0^3(0.0) = (1-0.0)^3\mathbf{b}_0^0 + 3\cdot 0.0(1-0.0)^2\mathbf{b}_1^0 + 3\cdot 0.0^2(1-0.0)\mathbf{b}_2^0 + 0.0^3\mathbf{b}_3^0 = \mathbf{b}_0^0$$

$$\mathbf{b}_0^3(0.4) = (1-0.4)^3\mathbf{b}_0^0 + 3\cdot 0.4(1-0.4)^2\mathbf{b}_1^0 + 3\cdot 0.4^2(1-0.4)\mathbf{b}_2^0 + 0.4^3\mathbf{b}_3^0$$

$$\mathbf{b}_0^3(0.7) = (1-0.7)^3\mathbf{b}_0^0 + 3\cdot 0.7(1-0.7)^2\mathbf{b}_1^0 + 3\cdot 0.7^2(1-0.7)\mathbf{b}_2^0 + 0.7^3\mathbf{b}_3^0$$

$$\mathbf{b}_0^3(0.0) = (1-1.0)^3\mathbf{b}_0^0 + 3\cdot 1.0(1-1.0)^2\mathbf{b}_1^0 + 3\cdot 1.0^2(1-1.0)\mathbf{b}_2^0 + 1.0^3\mathbf{b}_3^0 = \mathbf{b}_3^0$$

2) Sample code of de Casteljau algorithm (1)

```
numberifndef __BezierCurve_h__
numberdefine __BezierCurve_h__

numberinclude “vector.h”

class BezierCurve {
public:
    int m_nDegree;
    Vector* m_ControlPoint;   int m_nControlPoint;
    BezierCurve();
    ~BezierCurve();

    void SetDegree(int nDegree);
    void SetControlPoint(Vector* pControlPoint, int nControlPoint);
    Vector CalcPoint(double t);
    Vector deCasteljau(double t);                // CalcPoint by de Casteljau algorithm
    double B (int i, double t);
};
numberendif
```

Sample code of de Casteljau algorithm (2)

```
Vector BezierCurve:: deCasteljau (double t) {
    Vector* TmpControlPoint = new Vector [m_nControlPoint];
    for(int i = 0; i < m_nControlPoints; i++) TmpControlPoint[i] = m_ControlPoint[i];

    for(i = 1; i < m_nControlPoint; i++){
       for(int j = 0; j < m_nDegree - i; j++){
          TmpControlPoint[j] = (1-t)*TmpControlPoint[j]  + t*TmpControlPoint[j+1];
          //        bj^i                        bj^i-1                        bj+1^i-1
       }
    }
    Vector result = TmpControlPoint[0];  // b0^3
    delete[] TmpControlPoint;
   return result;
}
```

$$
\begin{array}{cccc}
\mathbf{b}_0^0 & \mathbf{b}_0^1 & \mathbf{b}_0^2 & \mathbf{b}_0^3 \\
\mathbf{b}_1^0 & \mathbf{b}_1^1 & \mathbf{b}_1^2 & \\
\mathbf{b}_2^0 & \mathbf{b}_2^1 & & \\
\mathbf{b}_3^0 & & &
\end{array}
$$

3) Comparison between the de Casteljau algorithm & Bezier curves

☑ **de Casteljau algorithm: “Constructive Approach”**

Input: b_i (Bezier control points)

Processor: Sequentially n-times ‘linear interpolation’

Output : Point on the n^{th}-degree Bezier curve

☑ **Bezier curve : “Bernstein Function evaluation Approach”**

Input: b_i (Bezier control points)

Processor: Curve by "blending" the control points(b_i) and Bernstein Basis functions

4) Parameter Transformation

☑ The affine map for the interval of $t \in [0,1] \rightarrow u \in [a,b]$,

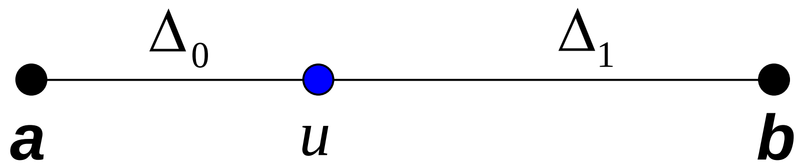
☑ Change the interval of [a, b] to the interval of [0, 1]

$$t = \frac{u-a}{b-a}. \qquad \text{and} \qquad 1-t = \frac{b-u}{b-a}.$$

☑ u → global parameter, t → local parameter

☑ the process of changing interval is called parameter transformation.

5) Linear Interpolation on [a, b]



$$u-a : b-u = \Delta_0 : \Delta_1$$

$$\Delta_0(b-u) = \Delta_1(u-a)$$

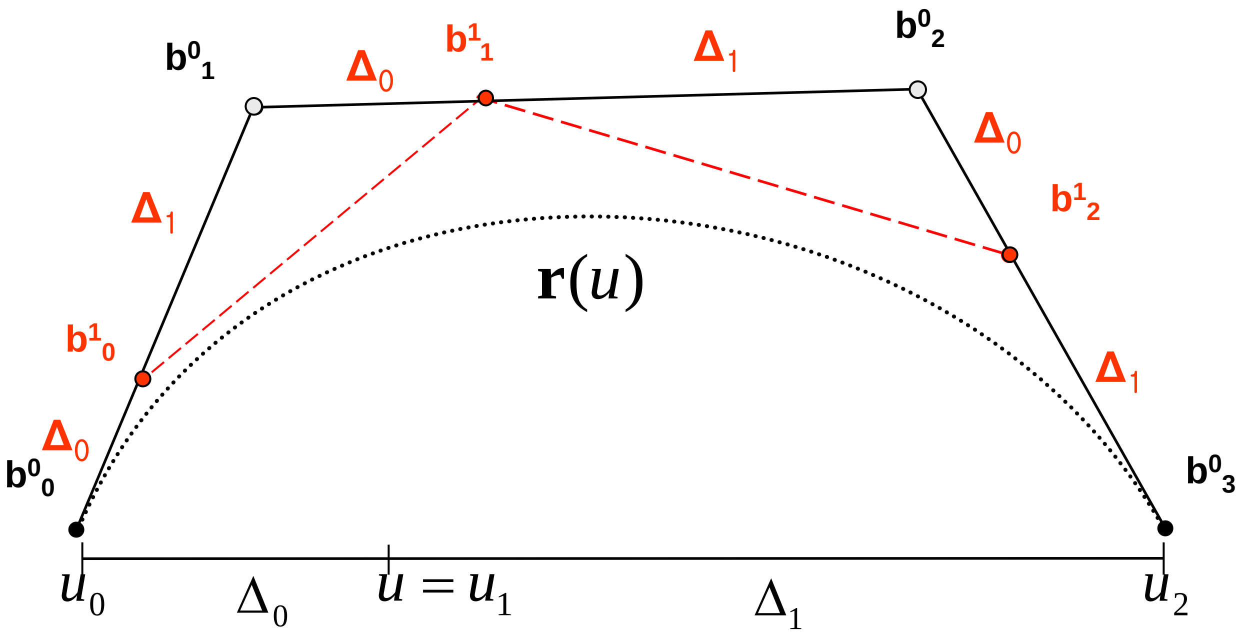
$$(\Delta_0+\Delta_1)u = \Delta_1 a + \Delta_0 b$$

$$u = \frac{\Delta_1 a + \Delta_0 b}{\Delta_0 + \Delta_1}$$

$$\therefore u = \frac{\Delta_1}{\Delta_0+\Delta_1} a + \frac{\Delta_0}{\Delta_0+\Delta_1} b$$

$$ratio(a,u,b) = \frac{\Delta_0}{\Delta_1}$$

6) Interval of the parameter u is given by [u_0, u_2]. For given four control points, construct the point on the curve at $u=u_1$ by using de Casteljau Algorithm

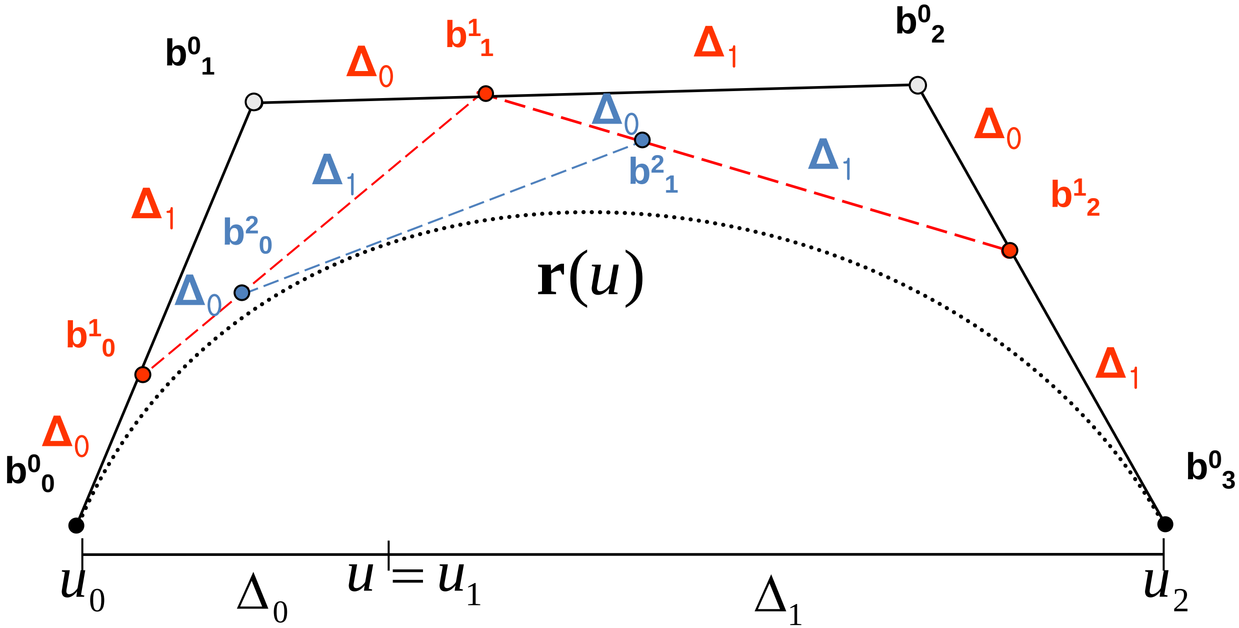


$$\mathbf{b}_1^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_1^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_2^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_3^0$$

$$\Delta = u_2 - u_0, \quad \Delta_1 = u_2 - u_1, \quad \Delta_0 = u_1 - u_0, \quad \Delta = \Delta_0 + \Delta_1$$

$$\text{ratio}(b_0^2, b_0^3, b_1^2) = \frac{u - u_0}{u_2 - u} = \frac{\Delta_0}{\Delta_1}$$



$$\mathrm{b}_0^1(u) = \frac{\Delta_1}{\Delta}\mathrm{b}_0^0 + \frac{\Delta_0}{\Delta}\mathrm{b}_1^0$$

$$\mathbf{b}_1^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_1^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_2^0$$

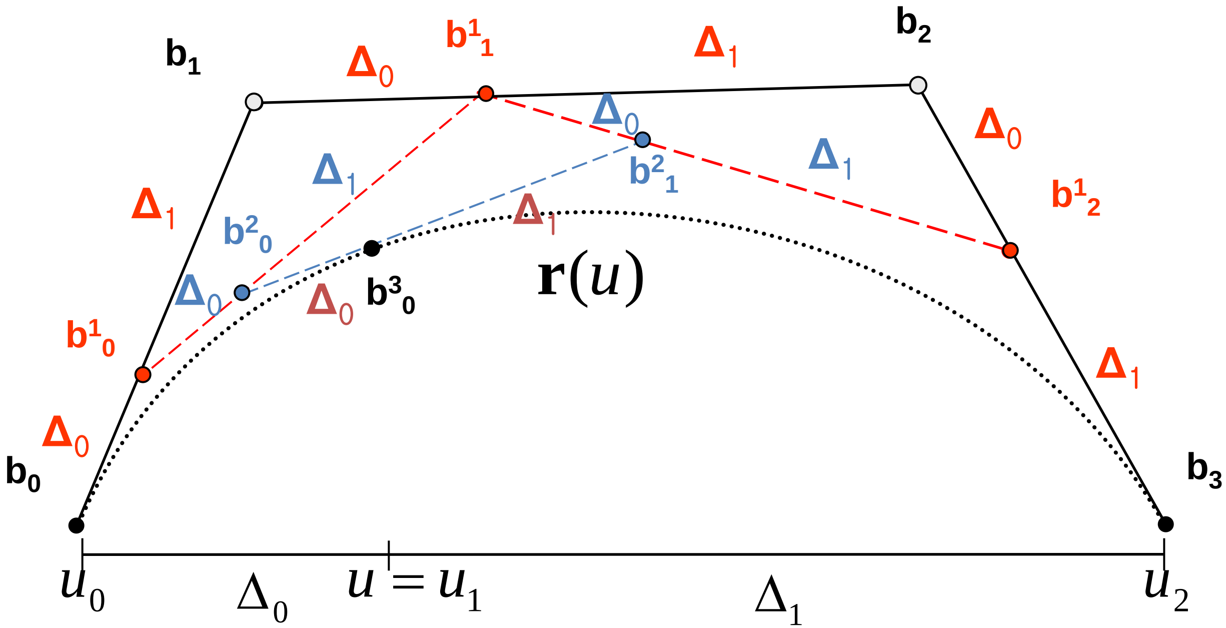
$$\mathbf{b}_2^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_2^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_3^0$$

$$\mathbf{b}_0^2(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_0^1 + \frac{\Delta_0}{\Delta}\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_1^1 + \frac{\Delta_0}{\Delta}\mathbf{b}_2^1$$

$$\Delta = u_2 - u_0, \quad \Delta_1 = u_2 - u_1, \quad \Delta_0 = u_1 - u_0, \quad \Delta = \Delta_0 + \Delta_1$$

$$\mathrm{ratio}(b_0^2, b_0^3, b_1^2) = \frac{u - u_0}{u_2 - u} = \frac{\Delta_0}{\Delta_1}$$



$$\Delta = u_2 - u_0, \quad \Delta_1 = u_2 - u_1, \quad \Delta_0 = u_1 - u_0, \quad \Delta = \Delta_0 + \Delta_1$$

$$\text{ratio}(b_0^2, b_0^3, b_1^2) = \frac{u - u_0}{u_2 - u} = \frac{\Delta_0}{\Delta_1}$$

$$\mathbf{b}_0^1(u) = \frac{u_2 - u}{u_2 - u_0}\mathbf{b}_0^0 + \frac{u - u_0}{u_2 - u_0}\mathbf{b}_1^0$$

$$= \frac{\Delta_1}{\Delta}\mathbf{b}_0^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_1^0$$

$$\mathbf{b}_1^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_1^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_2^0$$

$$\mathbf{b}_2^1(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_2^0 + \frac{\Delta_0}{\Delta}\mathbf{b}_3^0$$

$$\mathbf{b}_0^2(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_0^1 + \frac{\Delta_0}{\Delta}\mathbf{b}_1^1$$

$$\mathbf{b}_1^2(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_1^1 + \frac{\Delta_0}{\Delta}\mathbf{b}_2^1$$

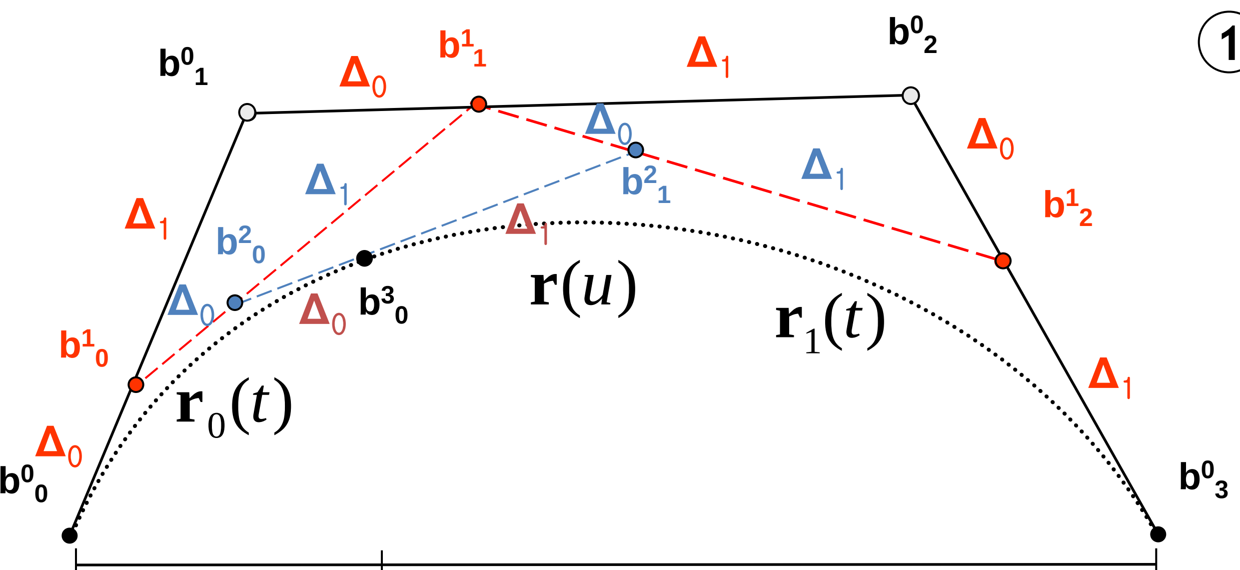
$$\mathbf{b}_0^3(u) = \frac{\Delta_1}{\Delta}\mathbf{b}_0^2 + \frac{\Delta_0}{\Delta}\mathbf{b}_1^2$$

Let $t = \frac{u - u_0}{u_2 - u_0}$,

Identical with 3rd degree Bezier curves!!!

$$\mathbf{b}_0^3(u) = \mathbf{r}(t) = (1-t)^3\mathbf{b}_0 + 3(1-t)^2 t\mathbf{b}_1 + 3(1-t)t^2\mathbf{b}_2 + t^3\mathbf{b}_3$$

7) Point on the Bezier curve-> divided into two Bezier curves at the point



$\Delta = u_2 - u_0,\quad \Delta_1 = u_2 - u_1,\quad \Delta_0 = u_1 - u_0,\quad \Delta = \Delta_0 + \Delta_1$

$$\mathrm{ratio}(b_0^2, b_0^3, b_1^2) = \frac{u - u_0}{u_2 - u} = \frac{\Delta_0}{\Delta_1}$$

1

1. Evaluation of a point on the Bezier curve at $u=u_1$ generates two sets of Bezier control points:
Bezier control points $b^0{}_0$, $b^1{}_0$, $b^2{}_0$, $b^3{}_0$, with parameter interval of Δ_0 represent the left curve $r_0(t)$, and Bezier control points $b^3{}_0$, $b^2{}_1$, $b^1{}_2$, $b^0{}_3$ with parameter interval of Δ_1 represent the right curve $r_1(t)$

2

$$\mathbf{r}(t) = (1-t)^3\mathbf{b}_0^0 + 3(1-t)^2 t\mathbf{b}_1^0 + 3(1-t)t^2\mathbf{b}_2^0 + t^3\mathbf{b}_3^0,\ t = \frac{u-u_0}{u_2-u_0}$$

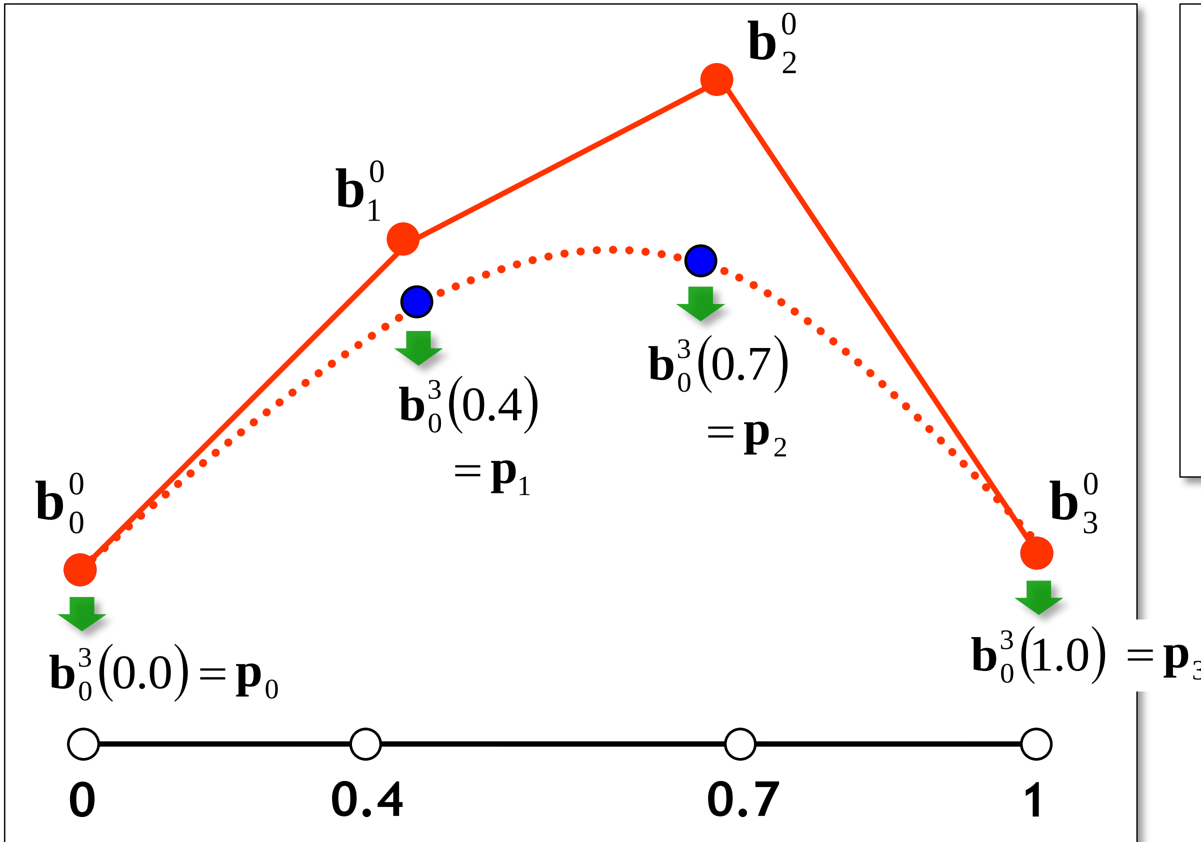
$$\mathbf{r}_0(t) = (1-t)^3\mathbf{b}_0^0 + 3(1-t)^2 t\mathbf{b}_0^1 + 3(1-t)t^2\mathbf{b}_0^2 + t^3\mathbf{b}_0^3,\ t = \frac{u-u_0}{u_1-u_0}$$

$$\mathbf{r}_1(t) = (1-t)^3\mathbf{b}_0^3 + 3(1-t)^2 t\mathbf{b}_1^2 + 3(1-t)t^2\mathbf{b}_2^1 + t^3\mathbf{b}_3^0,\ t = \frac{u-u_1}{u_2-u_1}$$

2.5 Bezier Curve Interpolation / Approximation

1) Introduction to Curve Interpolation
2) Cubic Bezier curve Interpolation
3) Bezier curve Interpolation beyond Cubics
4) Bezier curve Approximation
5) Finding the right parameters
6) Sample code of Bezier curve Interpolation

Points on the Cubic Bezier Curve at Parameter *t*



Given

$$\mathbf{b}_0^0,\ \mathbf{b}_1^0,\ \mathbf{b}_2^0,\ \mathbf{b}_3^0$$

Find

Points on the Bezier curve at *t* = 0.0, 0.4, 0.7, 1.0

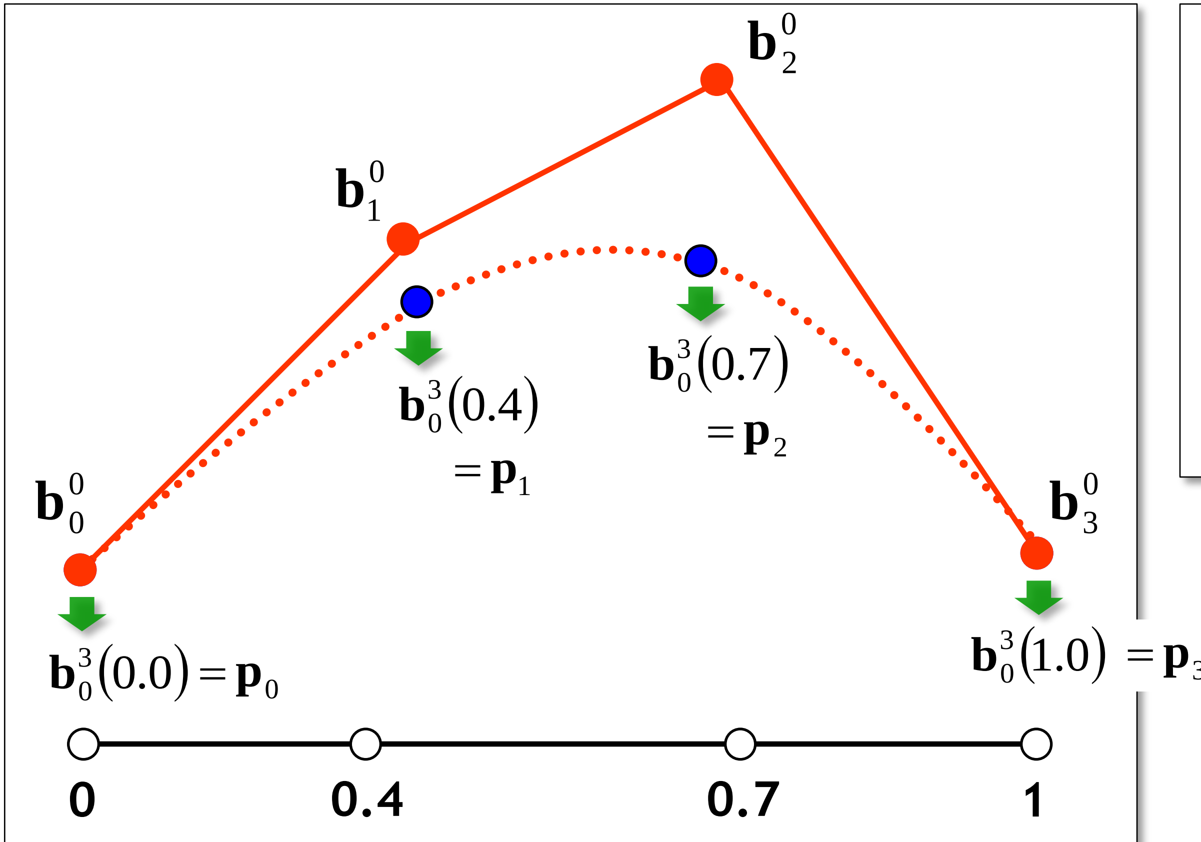
$$\mathbf{b}_0^3(0.0) = (1-0.0)^3\mathbf{b}_0^0 + 3\cdot 0.0(1-0.0)^2\mathbf{b}_1^0 + 3\cdot 0.0^2(1-0.0)\mathbf{b}_2^0 + 0.0^3\mathbf{b}_3^0 = \mathbf{b}_0^0$$

$$\mathbf{b}_0^3(0.4) = (1-0.4)^3\mathbf{b}_0^0 + 3\cdot 0.4(1-0.4)^2\mathbf{b}_1^0 + 3\cdot 0.4^2(1-0.4)\mathbf{b}_2^0 + 0.4^3\mathbf{b}_3^0$$

$$\mathbf{b}_0^3(0.7) = (1-0.7)^3\mathbf{b}_0^0 + 3\cdot 0.7(1-0.7)^2\mathbf{b}_1^0 + 3\cdot 0.7^2(1-0.7)\mathbf{b}_2^0 + 0.7^3\mathbf{b}_3^0$$

$$\mathbf{b}_0^3(0.0) = (1-1.0)^3\mathbf{b}_0^0 + 3\cdot 1.0(1-1.0)^2\mathbf{b}_1^0 + 3\cdot 1.0^2(1-1.0)\mathbf{b}_2^0 + 1.0^3\mathbf{b}_3^0 = \mathbf{b}_3^0$$

1) Curve Interpolation (1)



<u>Given</u>

Points on the Bezier curve at t = 0.0, 0.4, 0.7, 1.0 $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$

<u>Find</u>: Cubic Bezier Curve

$\mathbf{b}_0^0,\ \mathbf{b}_1^0,\ \mathbf{b}_2^0,\ \mathbf{b}_3^0$

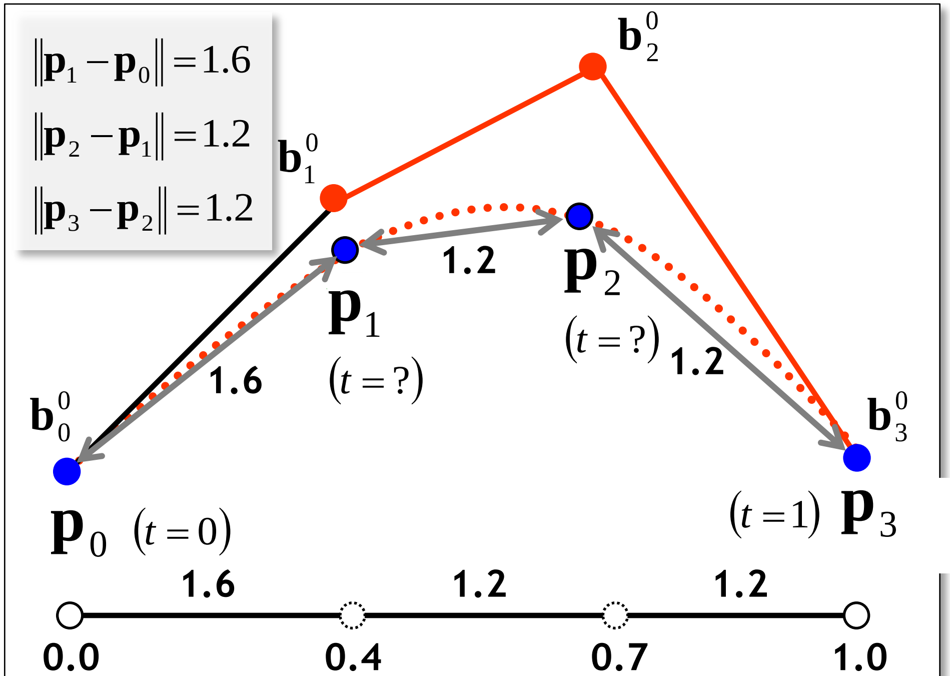
☑ If we are given fitting points $\mathbf{P}_i$ and we wish to pass a curve through them, called "curve interpolation".

$$\mathbf{b}_0^3(0.0) = (1-0.0)^3\mathbf{b}_0^0 + 3\cdot 0.0(1-0.0)^2\mathbf{b}_1^0 + 3\cdot 0.0^2(1-0.0)\mathbf{b}_2^0 + 0.0^3\mathbf{b}_3^0 = \mathbf{b}_0^0$$
$$\mathbf{b}_0^3(0.4) = (1-0.4)^3\mathbf{b}_0^0 + 3\cdot 0.4(1-0.4)^2\mathbf{b}_1^0 + 3\cdot 0.4^2(1-0.4)\mathbf{b}_2^0 + 0.4^3\mathbf{b}_3^0$$
$$\mathbf{b}_0^3(0.7) = (1-0.7)^3\mathbf{b}_0^0 + 3\cdot 0.7(1-0.7)^2\mathbf{b}_1^0 + 3\cdot 0.7^2(1-0.7)\mathbf{b}_2^0 + 0.7^3\mathbf{b}_3^0$$
$$\mathbf{b}_0^3(0.0) = (1-1.0)^3\mathbf{b}_0^0 + 3\cdot 1.0(1-1.0)^2\mathbf{b}_1^0 + 3\cdot 1.0^2(1-1.0)\mathbf{b}_2^0 + 1.0^3\mathbf{b}_3^0 = \mathbf{b}_3^0$$

- We may choose among many kinds of curves; If we use a cubic Bezier curve as an interpolation curve.

→ "cubic Bezier curve interpolation"

Set of parameter using chord length



$$\mathbf{r}(t)=(1-t)^3\mathbf{b}_0+3(1-t)^2t\mathbf{b}_1+3(1-t)t^2\mathbf{b}_2+t^3\mathbf{b}_3$$

Given

Points on the Bezier curve
at t = 0.0, 0.4, 0.7, 1.0
($\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$)

Find

$$\mathbf{b}_0^0,\ \mathbf{b}_1^0,\ \mathbf{b}_2^0,\ \mathbf{b}_3^0$$

- Every point on a Bezier curve has a parameter value t; in order to solve interpolation problem, we have to assign a parameter value t_i to every point P_i.

$$0=t_0<t_1<t_2<t_3=1$$

- A natural choice is to associate the ratio of distances between each P_i.

$$\left.\begin{array}{l} t_0=0.0,\ t_3=1.0 \\ t_1=\frac{1.6}{1.6+1.2+1.2}=0.4 \\ t_2=\frac{1.6+1.2}{1.4+1.0+1.6}=0.7 \end{array}\right\}$$

Set parameter t using chord length

- Then, we want a cubic Bezier curve such that:

$$\mathbf{r}(t_i)=\mathbf{p}_i;\quad i=0,1,2,3$$

2) Cubic Bezier curve interpolation (1)

☑ **The cubic Bezier curve of the form:**

$$\mathbf{r}(t) = B_0^3(t)\mathbf{b}_0 + B_1^3(t)\mathbf{b}_1 + B_2^3(t)\mathbf{b}_2 + B_3^3(t)\mathbf{b}_3.$$

☑ **All interpolation conditions are:**

$$\mathbf{p}_0 = B_0^3(t_0)\mathbf{b}_0 + B_1^3(t_0)\mathbf{b}_1 + B_2^3(t_0)\mathbf{b}_2 + B_3^3(t_0)\mathbf{b}_3,$$
$$\mathbf{p}_1 = B_0^3(t_1)\mathbf{b}_0 + B_1^3(t_1)\mathbf{b}_1 + B_2^3(t_1)\mathbf{b}_2 + B_3^3(t_1)\mathbf{b}_3,$$
$$\mathbf{p}_2 = B_0^3(t_2)\mathbf{b}_0 + B_1^3(t_2)\mathbf{b}_1 + B_2^3(t_2)\mathbf{b}_2 + B_3^3(t_2)\mathbf{b}_3,$$
$$\mathbf{p}_3 = B_0^3(t_3)\mathbf{b}_0 + B_1^3(t_3)\mathbf{b}_1 + B_2^3(t_3)\mathbf{b}_2 + B_3^3(t_3)\mathbf{b}_3,$$

4 Unknown Vectors, 4 Vector Equations

Cubic Bezier curve interpolation (2)

☑ To find the solution of these four equations for four unknowns, we can write in matrix form:

$$\begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \begin{bmatrix} B_0^3(t_0) & B_1^3(t_0) & B_2^3(t_0) & B_3^3(t_0) \\ B_0^3(t_1) & B_1^3(t_1) & B_2^3(t_1) & B_3^3(t_1) \\ B_0^3(t_2) & B_1^3(t_2) & B_2^3(t_2) & B_3^3(t_2) \\ B_0^3(t_3) & B_1^3(t_3) & B_2^3(t_3) & B_3^3(t_3) \end{bmatrix} \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}.$$

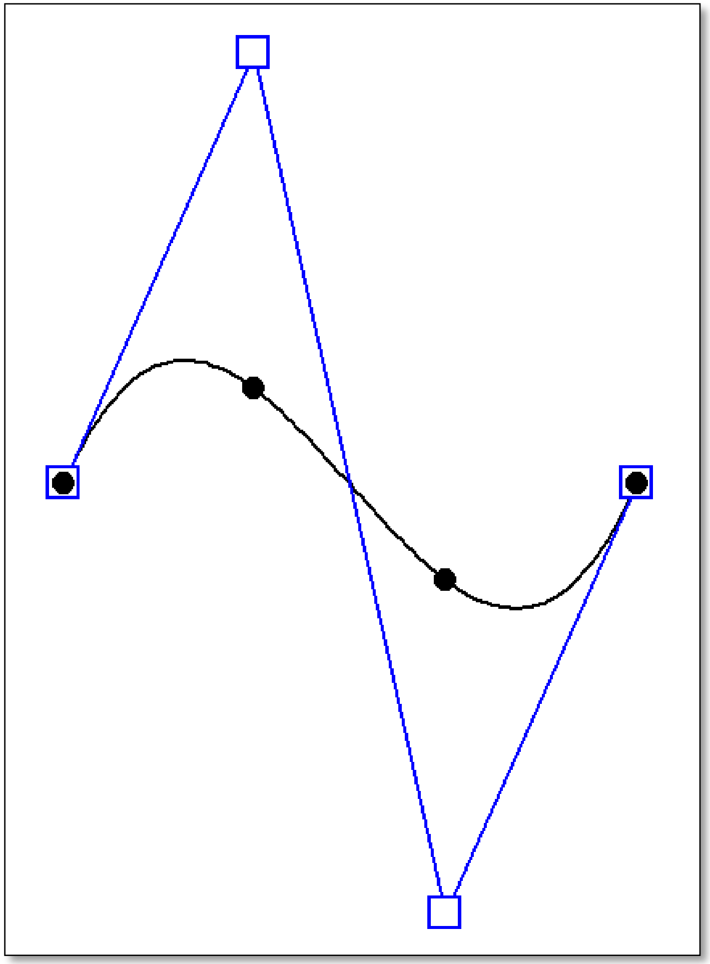
☑ To abbreviate the above form as: $\mathbf{P} = \mathbf{MB}.$

☑ The solution is: $\mathbf{B} = \mathbf{M}^{-1}\mathbf{P}.$

☑ Although it looks like the solution to one linear system but it is the two or three systems depending on the dimensionality of the $\mathbf{p}_i$.

$$ex)\ \mathbf{p}_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}^T \quad or \quad \begin{bmatrix} x_0 & y_0 & z_0 \end{bmatrix}^T$$

Cubic Bezier curve interpolation (3)



Cubic Bezier interpolation.

3) Bezier curve interpolation beyond Cubics (1)

☑ Polynomial interpolation can also works for more than four data points.

☑ Given: points $\mathbf{p}_0,\ldots.\mathbf{p}_m$ and
corresponding parameter values $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = 1.$

☑ If we choose a Bezier curve of degree n for interpolation,
we have "m+1 vector equations" for "n+1 unknown vectors".

☑ $n > m$: underdetermined system,
We need *additional conditions* to solve the interpolation problem

☑ $n = m$: determinate linear system → "Interpolation problem"

☑ $n < m$: overdetermined system → "Approximation problem"

Bezier curve interpolation beyond Cubics (2)

☑ Given: points $\mathbf{p}_0, \ldots, \mathbf{p}_m$ and corresponding parameter values $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = 1.$

☑ If we use a Bezier curve of degree n (=m), we have a linear system: $\mathbf{P} = \mathbf{MB}.$

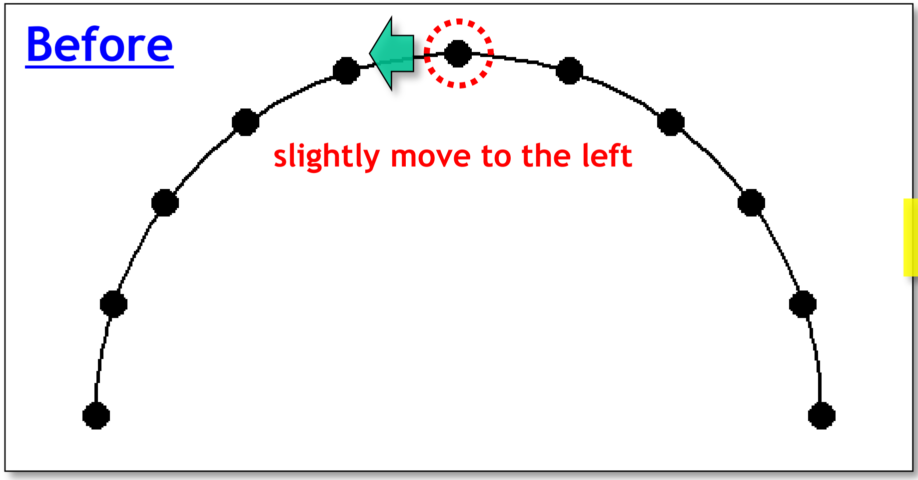
☑ $\mathbf{M}$ is an $(m+1) \times (m+1)$ matrix with elements;

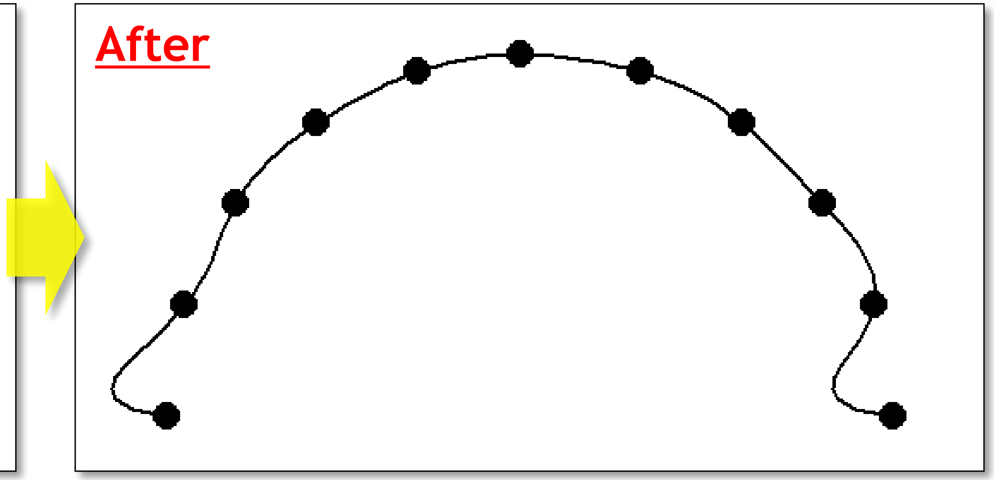
$$e_{ij} = B_j^m(t_i)$$

☑ It can be solved with any linear solver.

☑ Polynomial interpolation does not provide satisfied result for higher degrees. Figure in the next slide should be convincing enough.

Bezier curve interpolation beyond Cubics (3)





Top: Data from a circle; Bottom: one point is slightly modified.

- The processes of a small change in data can lead large change in the interpolating curve is called ill-conditioned.
- Different polynomial forms will give the identical result.

4) Bezier curve approximation (1)

☑ One is given more data points than should be interpolated by a polynomial curve (i.e. number of data points more than degree of curve)

→ We can solve the problem
by interpolating with a higher degree Bezier curve,
but higher degree interpolation becomes ill-conditioned.

☑ In such cases, an approximating curve will be needed, which does not pass through the data points exactly; rather it passes near them.

- the best technique to find such curves
- →'least squares approximation'.

Bezier curve approximation (2)

☑ **Given: points** $\mathbf{p}_0, \ldots\ldots \mathbf{p}_m$ **and**

corresponding parameter values $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = 1.$

☑ **We wish to find a polynomial curve** $\boldsymbol{r(t)}$ **of a given degree** $\boldsymbol{n}$ **(**$< \boldsymbol{m}$**) such that**

$$\sum_{i=1}^{m} \left\| \mathbf{p}_i - \mathrm{r}(t_i) \right\| \to minimize \quad (or) \quad \mathbf{p}_i = \mathrm{r}(t_i); \quad i = 0, 1, \ldots\ldots, m$$

☑ **Polynomial curve is of the Bezier form:**

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^n(t) + \mathbf{b}_1 B_1^n(t) + \ldots\ldots\ldots + \mathbf{b}_n B_n^n(t).$$

Bezier curve approximation (3)

☑ **We would like the following to hold:**

$$
\begin{aligned}
\mathbf{p}_0 &= \mathbf{b}_0 B_0^n(t_0) + \dots\dots\dots + \mathbf{b}_n B_n^n(t_0) \\
\mathbf{p}_1 &= \mathbf{b}_0 B_0^n(t_1) + \dots\dots\dots + \mathbf{b}_n B_n^n(t_1) \\
\vdots \quad & \quad \vdots \qquad\qquad\qquad\qquad \vdots \\
\mathbf{p}_m &= \mathbf{b}_0 B_0^n(t_m) + \dots\dots\dots + \mathbf{b}_n B_n^n(t_m)
\end{aligned}
\;\Rightarrow\;
\begin{bmatrix} B_0^n(t_0) & \cdots & B_n^n(t_0) \\ \vdots & & \vdots \\ B_0^n(t_m) & & B_n^n(t_m) \end{bmatrix}
\begin{bmatrix} \mathbf{b}_0 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}
=
\begin{bmatrix} \mathbf{p}_0 \\ \vdots \\ \mathbf{p}_m \end{bmatrix}
$$

$$\mathbf{MB} = \mathbf{P}$$

(*n*+1)*(2 or 3) Unknowns < (m+1)*(2 or 3) Equations

Bezier curve approximation (4)

☑ Multiply both sides by : $\mathbf{M}^T$

$$\mathbf{M}^T\mathbf{M}\mathbf{B} = \mathbf{M}^T\mathbf{P}.$$

← Normal equation

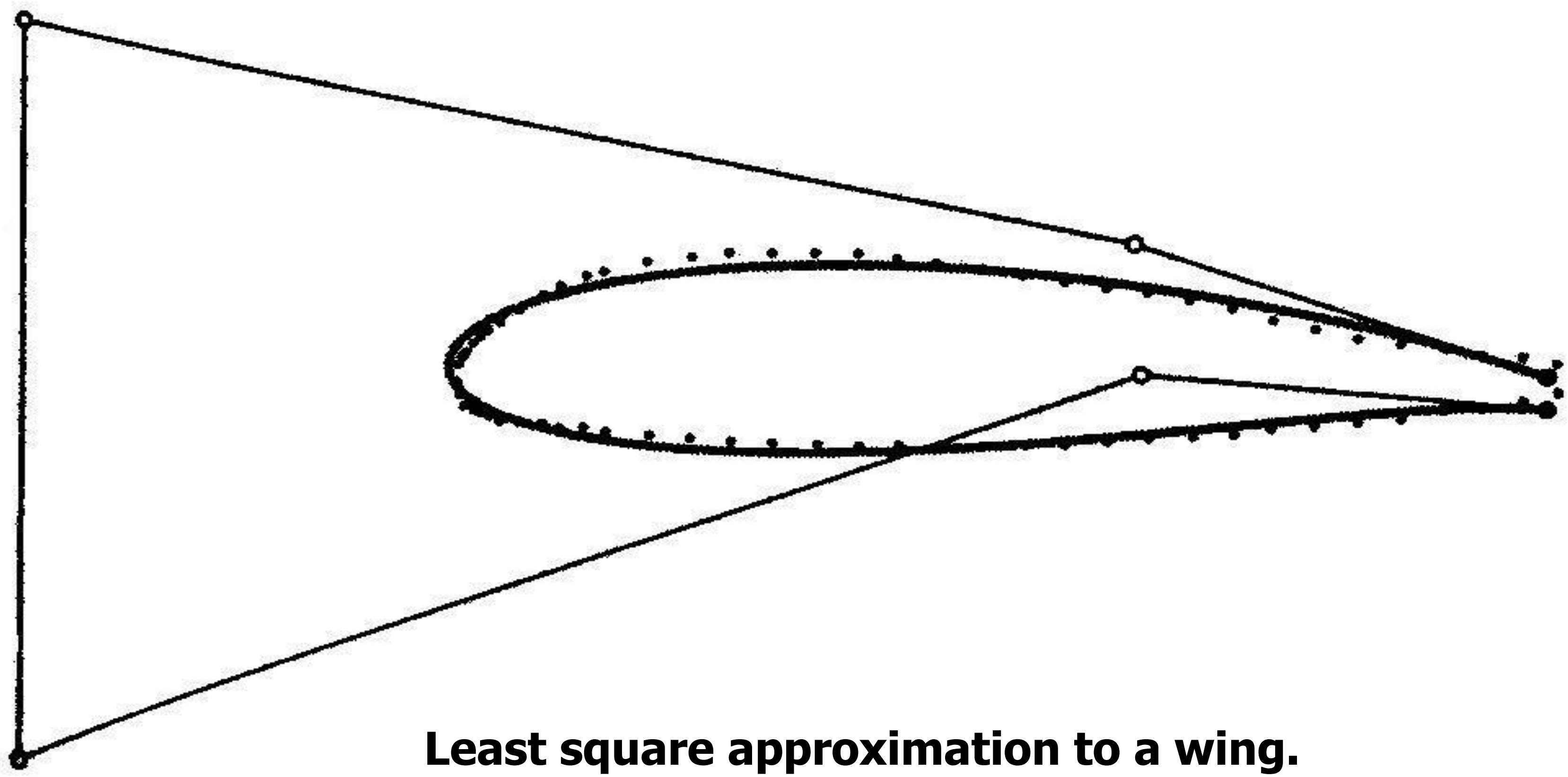
where $\mathbf{M}^T\mathbf{M}$ is a square and symmetric matrix, which is always invertible.

- The curve B minimizes the sum of the $\left\|\mathbf{p}_i - \mathbf{r}(t_i)\right\|, \quad i = 0,1,\ldots.,m$

$$\therefore \ \mathbf{B} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T\mathbf{P}.$$

note that any modification of the t_i would result in an entirely different solution.

Bezier curve approximation (5)



Least square approximation to a wing.
A quintic Bezier curve with chord length parameters assigned to the data.

5) chord length parameter

☑ In both interpolation & approximation curve,
in practice, the parameter value t_i are not normally given,
and have to be made up.

☑ There are two types to be made up:

(1) Uniform sets of parameters;

- If there are $(m+1)$ points $\mathbf{p}_i$,
- then set $t_i = i/l$.

(2) chord length parameters;

- if the distance between two points is relatively large, then their parameter values should also be fairly different.

$$t_0 = 0$$

$$t_1 = t_0 + \left\| \mathbf{p}_1 - \mathbf{p}_0 \right\|$$

$$\vdots$$

$$t_l = t_{l-1} + \left\| \mathbf{p}_l - \mathbf{p}_{l-1} \right\|$$

☑ If desired (it makes no difference to the interpolation or approximation result), the parameters may be normalized by scaling the parameters to live between zero and one:

$$t_i = \frac{t_i - t_0}{t_m - t_0}.$$

☑ In general, chord length parameterization method is superior to the uniform method, because it takes into account the geometry of the data.

6) Sample code of Interpolation/Approximation (1)

```
numberinclude "vector.h"

class BezierCurve {
public:
    int m_nDegree;
    Vector* m_ControlPoint;   int m_nControlPoint;
    ......
    void SetDegree(int nDegree);
    void SetControlPoint(Vector* pControlPoint, int nControlPoint);
    Vector CalcPoint(double t);
    double B (int i, double t);
    int Approximation(int nDegree, int nType, Vector* FittingPoint, int nPoint);
    int Interpolation(int nType, Vector* FittingPoint, int nPoint);
    void Parameterization(int nType, Vector* FittingPoint, int nPoint, double* t);
};
```

Sample code of Interpolation/Approximation (2)

```
void BezierCurve:: Parameterization (int nType, Vector* FittingPoint, int nPoint, double* t){
    // assume t is allocated out of function
    if( nType == 1) {   // Uniform Set
        for (int i = 0; i < nPoint; i++)
                            t[i] = 1./(nPoint-1);
    } else if ( nType == 2) { // Chord length
        t[0] = 0.;
        for (int i=0; i < nPoint-1; i++)
            t[i+1] = t[i] + (FittingPoint[i+1] - FittingPoint[i]).Magnitude();
        double t0 = t[0], tm = t[nPoint-1];
        for (int i=0; i < nPoint; i++)
            t[i] = (t[i] -  t0)/(tm - t0);      // Normalize
    }
}
```

Sample code of Interpolation/Approximation (3)

```
int BezierCurve:: Approximation(int nDegree, int nType, Vector* FittingPoint, int nPoint){
    m_nDegree = nDegree;
    m_nControlPoint = m_nDegree+1;
    if(m_ControlPoint) = delete[] m_ControlPoint;
    m_ControlPoint = new Vector[m_nControlPoint];

    double* t = new double[nPoint];
    Parameterization(nType, FittingPoint, nPoint, t);

    // Solve normal equation
    ....
   delete[] t;
}
```

Sample code of Interpolation/Approximation (4)

```
int BezierCurve:: Interpolation(int nType, Vector* FittingPoint, int nPoint){
    …

    double** M = new double*[nNumOfPoint];
    for (i=0; i<nNumOfPoint; i++) M[i] = new double[nNumOfPoint];

    for (i=0; i<nNumOfPoint; i++)    {
        for (j=0; j<nNumOfPoint; j++)    {
            M[i][j] = B(j, t[i]);
        }
    }

    // Solve MB = P
    GaussElimination(nNumOfPoint, M, p_x, b_x);
    GaussElimination(nNumOfPoint, M, p_y, b_y);
    GaussElimination(nNumOfPoint, M, p_z, b_z);
    ….

}
```

$$\begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \boxed{\begin{bmatrix} B_0^3(t_0) & B_1^3(t_0) & B_2^3(t_0) & B_3^3(t_0) \\ B_0^3(t_1) & B_1^3(t_1) & B_2^3(t_1) & B_3^3(t_1) \\ B_0^3(t_2) & B_1^3(t_2) & B_2^3(t_2) & B_3^3(t_2) \\ B_0^3(t_3) & B_1^3(t_3) & B_2^3(t_3) & B_3^3(t_3) \end{bmatrix}} \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix}$$

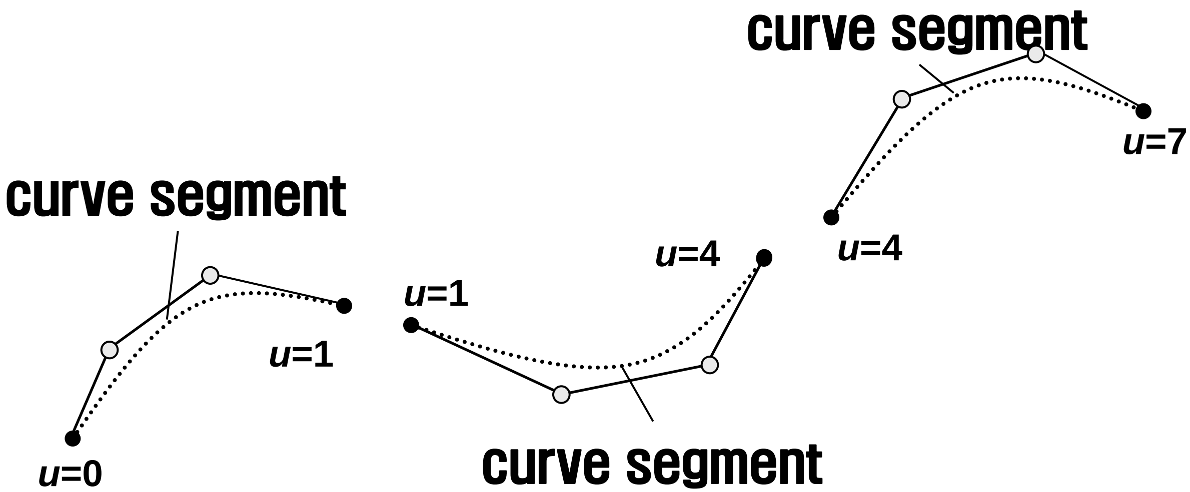
$$\mathbf{P} = \mathbf{MB}$$

$$\boxed{\mathbf{B} = \mathbf{M}^{-1}\mathbf{P}}$$

Chapter 3. B[asis]-spline Curves

3.1 Introduction to B-spline Curves

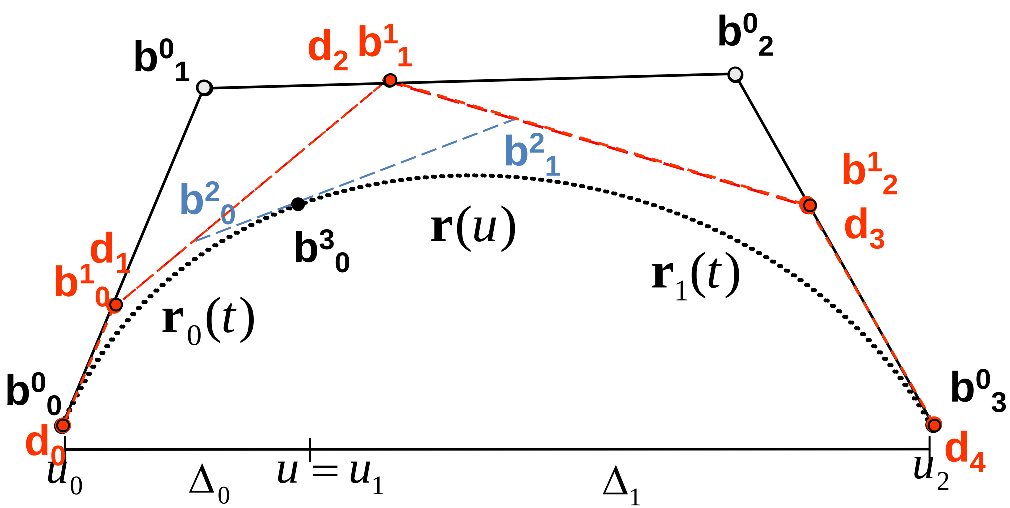
1) 'Smooth' connection of separate curve segments at knots : Spline curve



Knot = {..., 0, 1, 4, 7, ...}

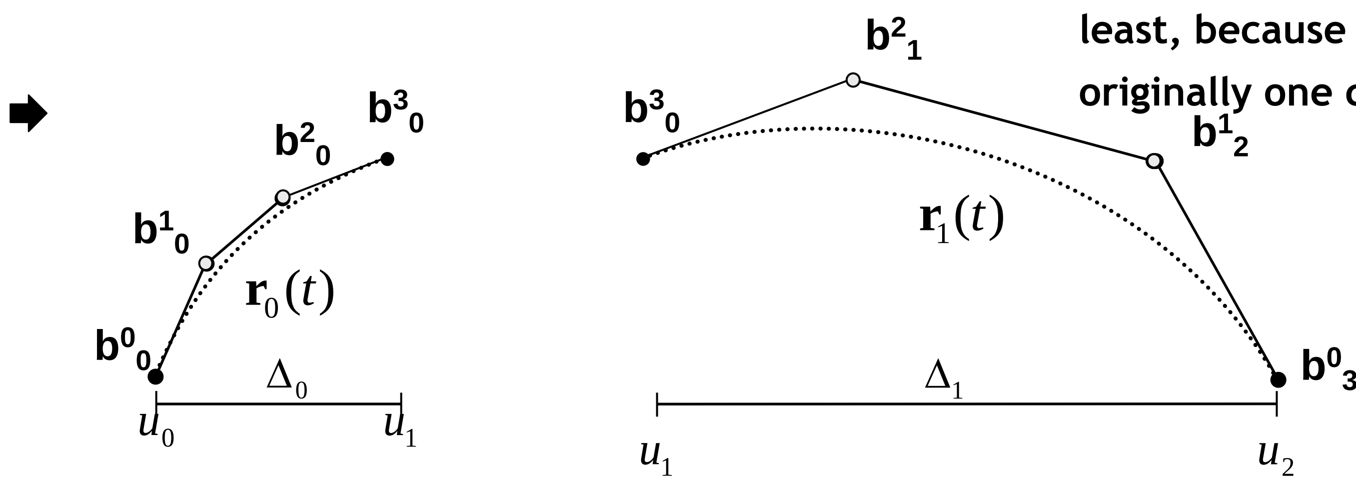
- Curve is "smoothly" connected with curve segments : spline curve
- Curve segments are tied by knots : knot

Point on the Bezier curve-> connected two Bezier curves at the 'knot'



In contrast with dividing into two Bezier curves, we can imagine as if two Bezier curves $r_0(t)$ and $r_1(t)$ were connected at $u=u_1$. Here, u1 is called 'knot', that means the knot ties the curves .

At $u=u_1$, the curves obviously satisfy C_0, C_1, C_2 continuity conditions at least, because the curve r(u) was originally one curve.



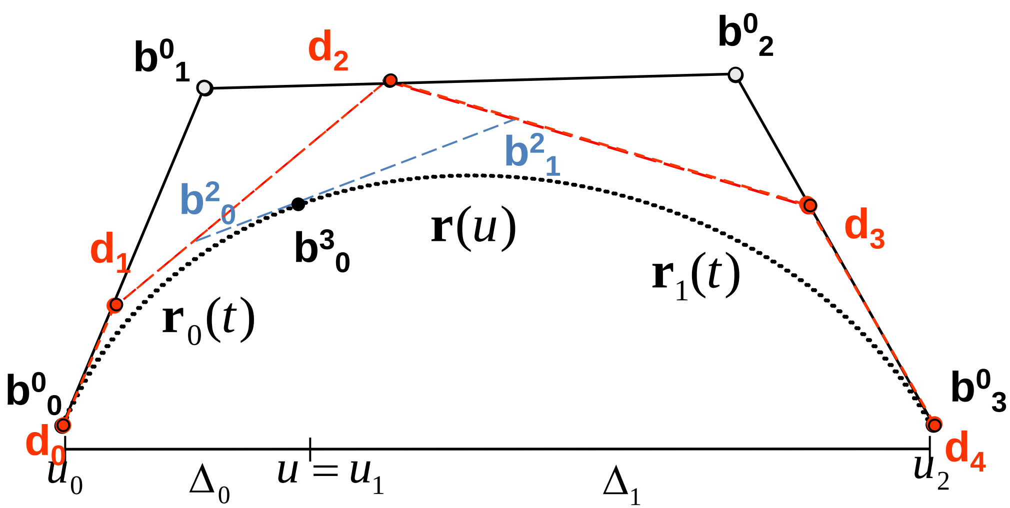
$$\mathbf{r}_0(t) = (1-t)^3\mathbf{b}_0^0 + 3(1-t)^2 t\mathbf{b}_0^1 + 3(1-t)t^2\mathbf{b}_0^2 + t^3\mathbf{b}_0^3,$$

$$t = \frac{u-u_0}{u_1-u_0}$$

$$\mathbf{r}_1(t) = (1-t)^3\mathbf{b}_0^3 + 3(1-t)^2 t\mathbf{b}_1^2 + 3(1-t)t^2\mathbf{b}_2^1 + t^3\mathbf{b}_3^0,$$

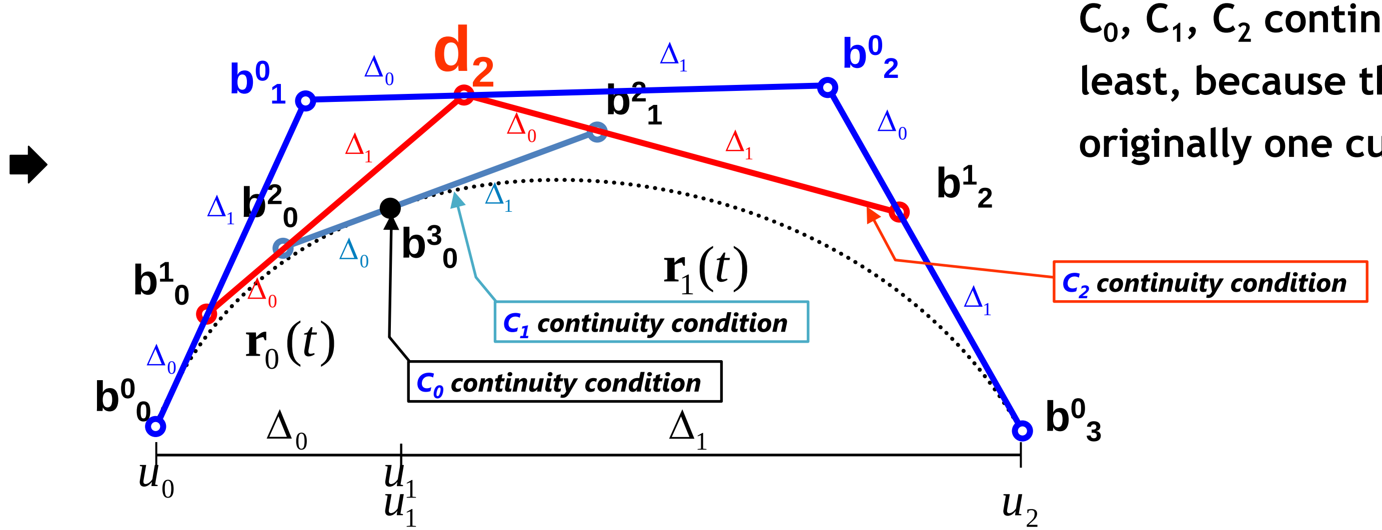
$$t = \frac{u-u_1}{u_2-u_1}$$

Point on the Bezier curve-> connected two Bezier curves at the 'knot'



In contrast with dividing into two Bezier curves, we can imagine as if two Bezier curves r_0(t) and r_1(t) were connected at u=u_1. Here, u1 is called 'knot', that means the knot ties the curves .

At u=u_1 , the curves obviously satisfy C_0, C_1, C_2 continuity conditions at least, because the curve r(u) was originally one curve.

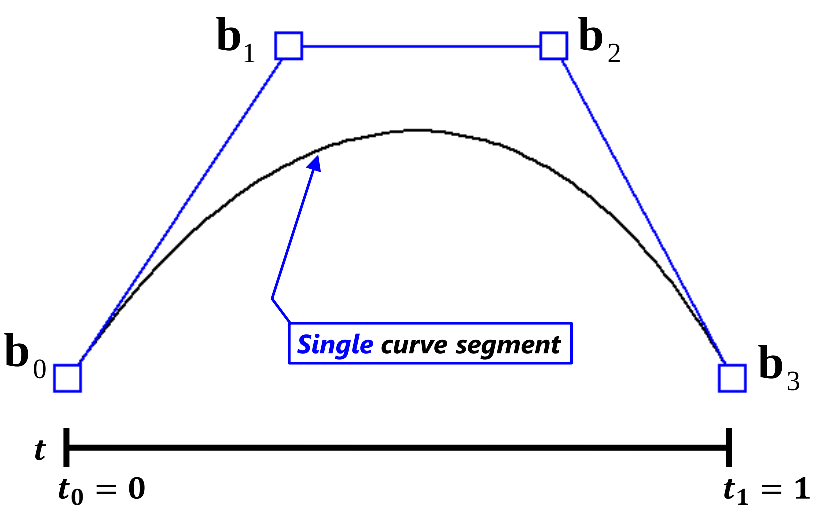


$$\mathbf{r}_0(t) = (1-t)^3\mathbf{b}_0^0 + 3(1-t)^2 t\mathbf{b}_0^1 + 3(1-t)t^2\mathbf{b}_0^2 + t^3\mathbf{b}_0^3,\qquad t = \frac{u-u_0}{u_1-u_0}$$

$$\mathbf{r}_1(t) = (1-t)^3\mathbf{b}_0^3 + 3(1-t)^2 t\mathbf{b}_1^2 + 3(1-t)t^2\mathbf{b}_2^1 + t^3\mathbf{b}_3^0,\qquad t = \frac{u-u_1}{u_2-u_1}$$

2) Definition of B-spline curves

Ex) Cubic Bezier Curve



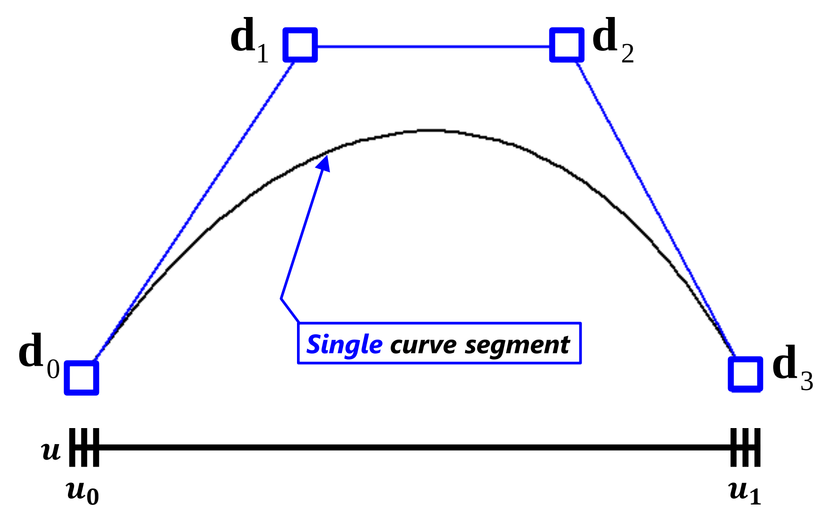
Given: $\mathbf{b}_0$, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, t

Find: points on the curve at parameter t

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^3(t) + \mathbf{b}_1 B_1^3(t) + \mathbf{b}_2 B_2^3(t) + \mathbf{b}_3 B_3^3(t)$$

Bernstein Polynomial Function

Ex) Cubic B-spline Curve



Given: $\mathbf{d}_0$, $\mathbf{d}_1$, $\mathbf{d}_2$, $\mathbf{d}_3$, u

Find: points on the curve at parameter u

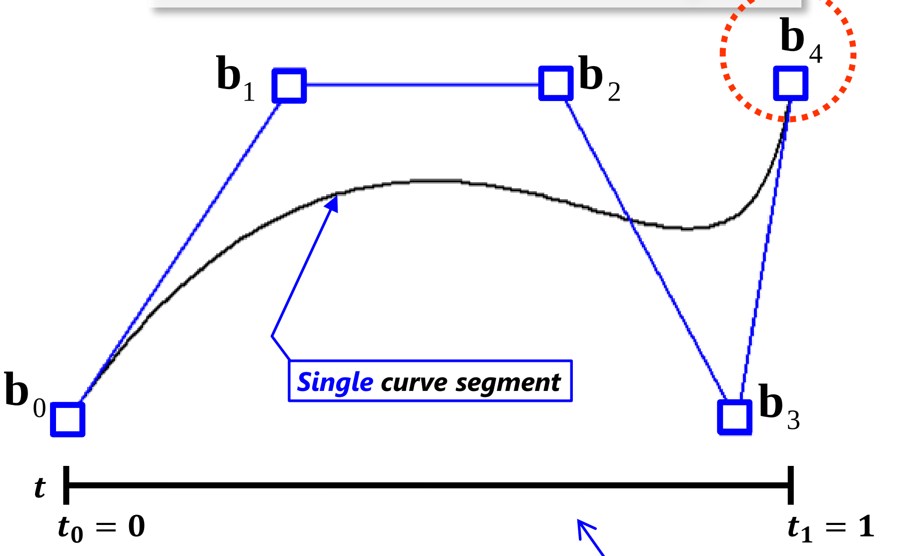
$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$

B-spline Basis Function
(Cox-de Boor Recursive Formula)

What happens if one more control point is included?

Property of Bezier curves and B-spline curves in increasing(changing) the number of the control points

Bezier Curve of 4th degree



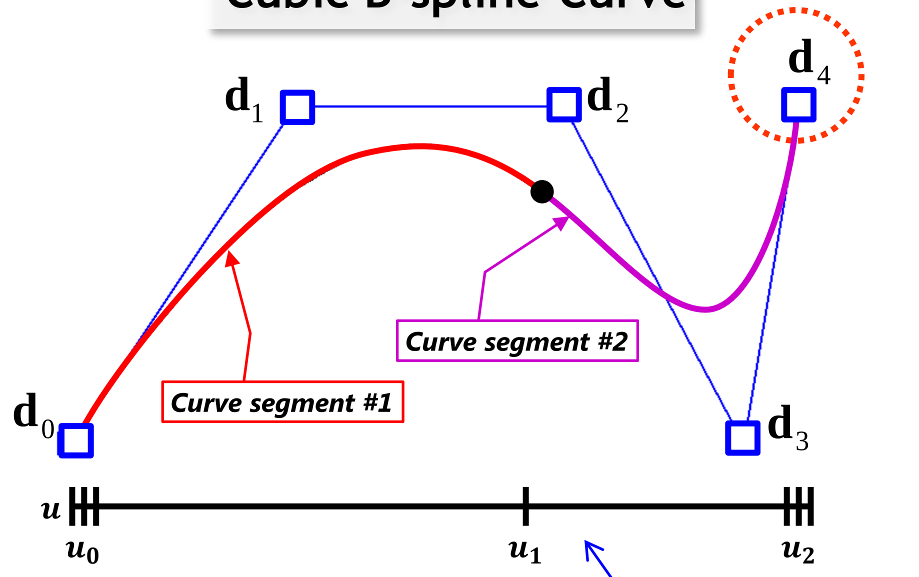
Given: $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4, t$

Find: points on the curve at parameter t

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^4(t) + \mathbf{b}_1 B_1^4(t) + \mathbf{b}_2 B_2^4(t) + \mathbf{b}_3 B_3^4(t) + \boxed{\mathbf{b}_4 B_4^4(t)}$$

For the Bezier curve, if the number of the control points increases, the **degree** of the Bezier curve will also **increase.**

Cubic B-spline Curve



Given: $\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, u$

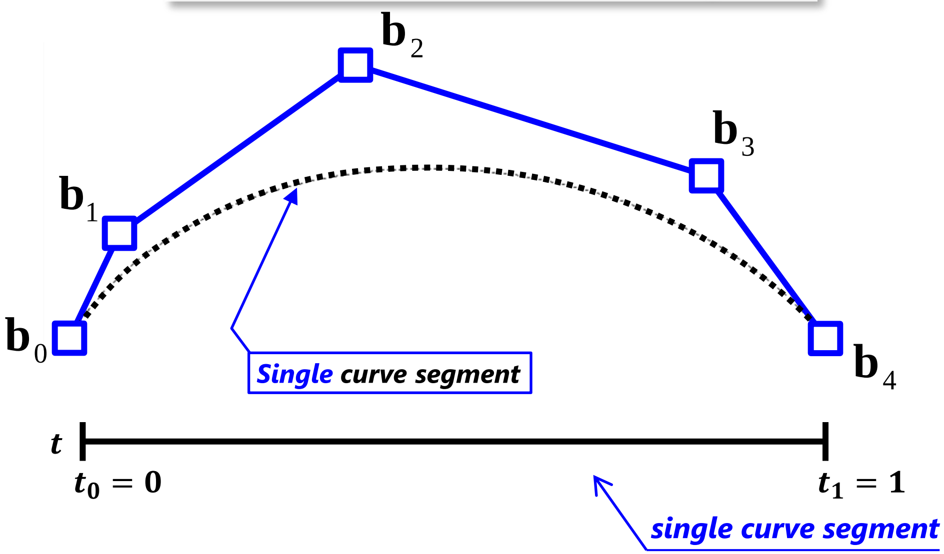
Find: points on the curve at parameter u

$$\mathrm{r}(u) = \mathrm{d}_0 N_0^3(u) + \mathrm{d}_1 N_1^3(u) + \mathrm{d}_2 N_2^3(u) + \mathrm{d}_3 N_3^3(u) + \boxed{\mathrm{d}_4 N_4^3(u)}$$

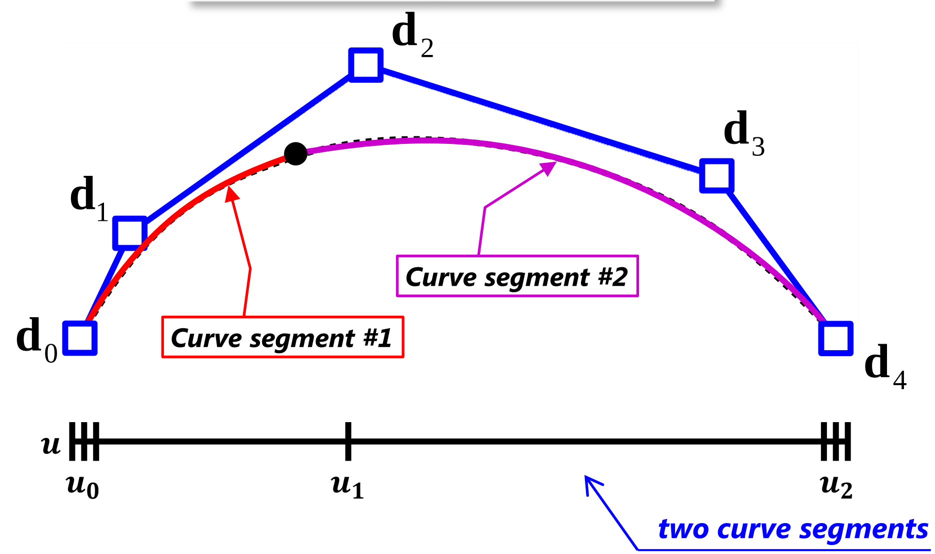
For the B-spline curve, if the number of the control points increases,
the degree of the curve does not change
but **additional** one **Bezier curve of 3rd degree** is generated.

Properties of Bezier curves and B-spline curves with same control points

Bezier Curve of 4th degree



Cubic B-spline Curve



Given: $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4, t$

Find: points on the curve at parameter t

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^4(t) + \mathbf{b}_1 B_1^4(t) + \mathbf{b}_2 B_2^4(t) + \mathbf{b}_3 B_3^4(t) + \mathbf{b}_4 B_4^4(t)$$

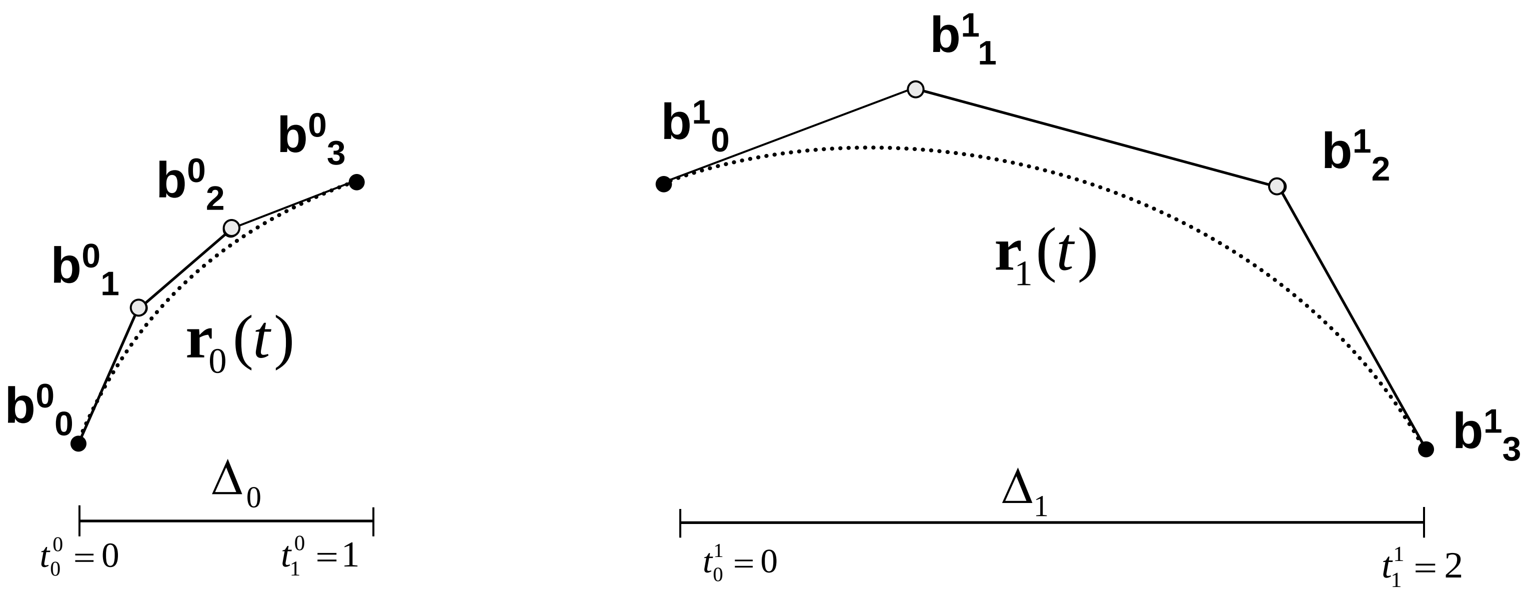
Given: $\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, u$

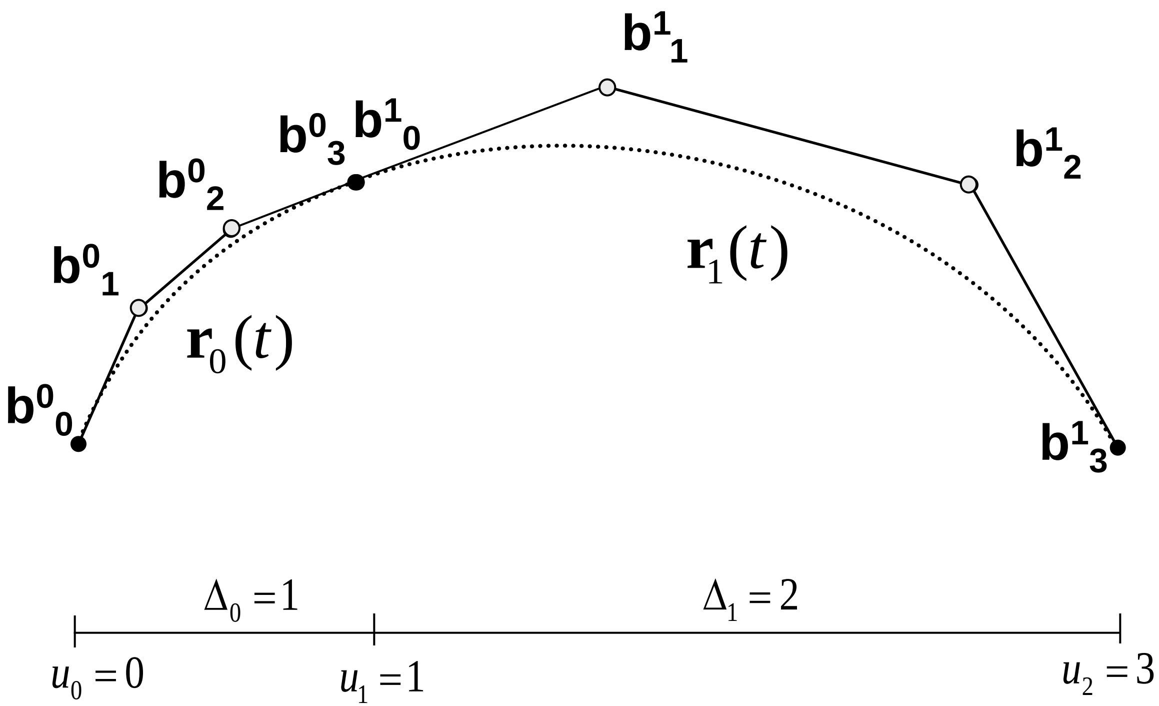
Find: points on the curve at parameter u

$$\mathrm{r}(u) = \mathrm{d}_0 N_0^3(u) + \mathrm{d}_1 N_1^3(u) + \mathrm{d}_2 N_2^3(u) + \mathrm{d}_3 N_3^3(u) + \mathrm{d}_4 N_4^3(u)$$

3) Geometric meanings of B-spline curve

☑ Ex) 'Cubic' B-spline curve is composed of several 'cubic' Bezier curves, which are connected with the C^2 continuity condition (condition of continuous 2^{nd} derivative)



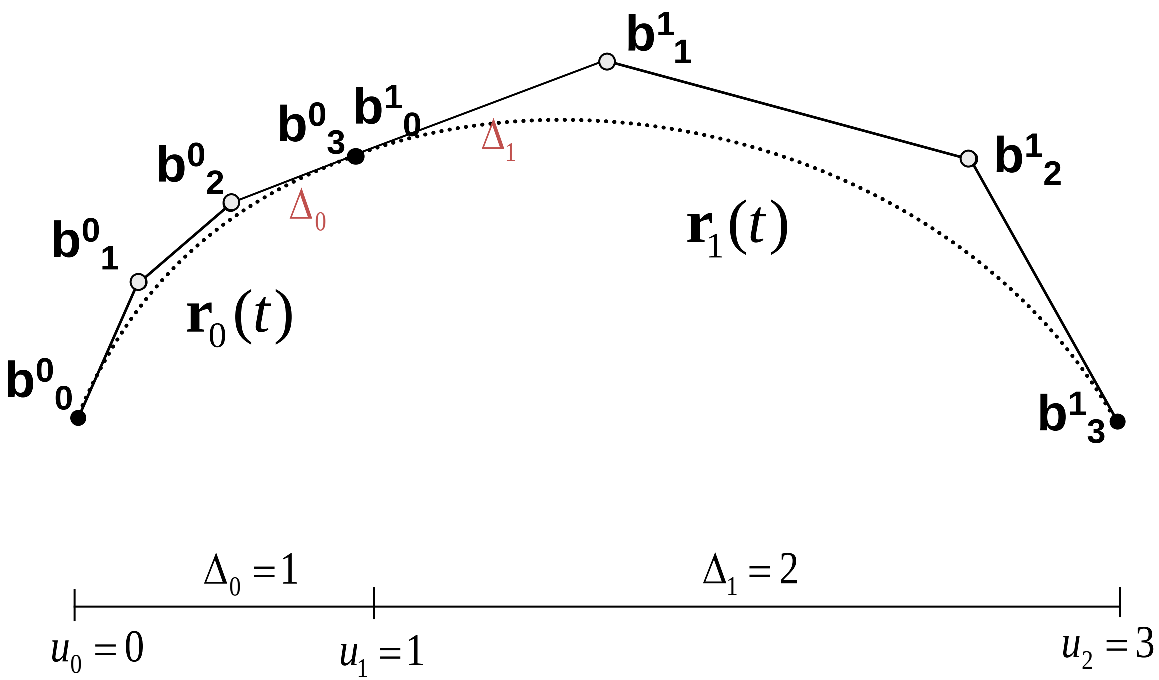


Assign new global parameter *u* to the connected curves

C^0 continuity condition (coincident position)

$$\mathbf{b}_3^0 = \mathbf{b}_0^1$$

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



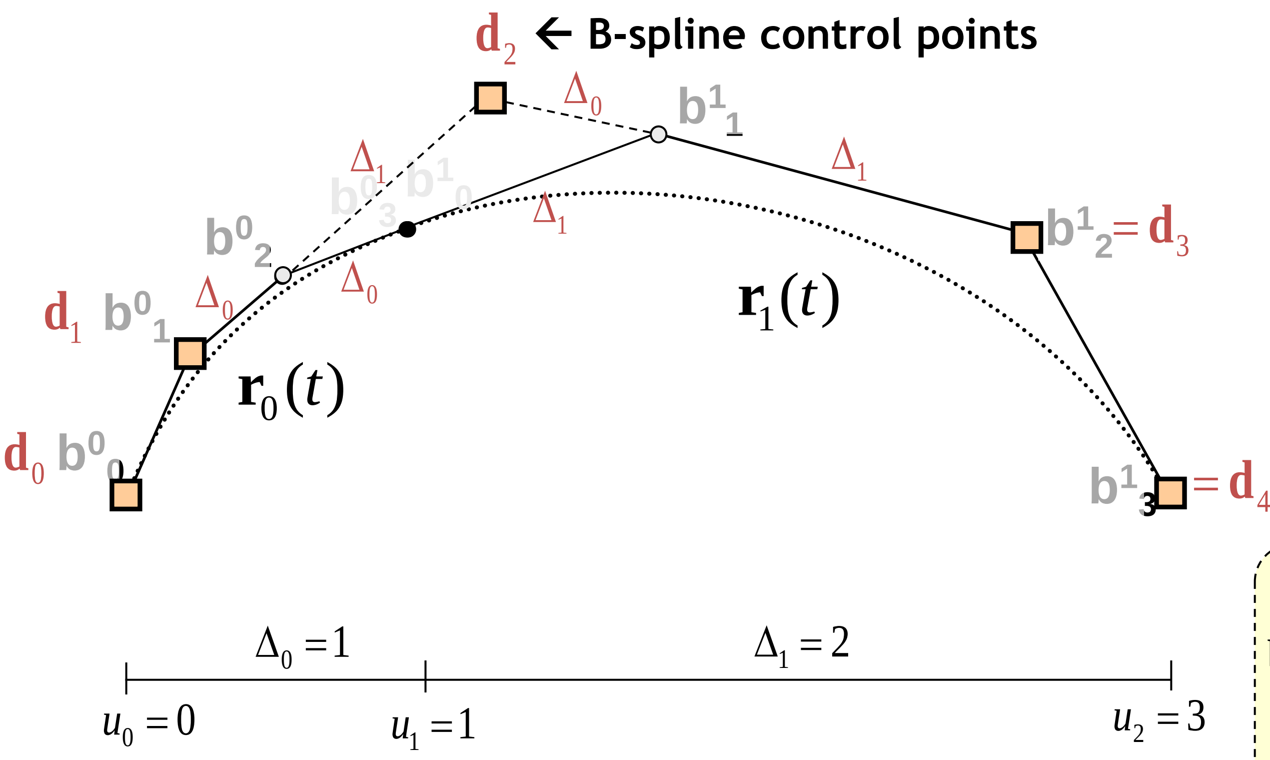
Assign new global parameter *u* to jointed curve

C^1 continuity condition

$$\mathbf{b}_3^0 = \mathbf{b}_0^1$$

$$\mathbf{b}_3^0 = \mathbf{b}_0^1 = \frac{\Delta_1}{\Delta_0 + \Delta_1}\mathbf{b}_2^0 + \frac{\Delta_0}{\Delta_0 + \Delta_1}\mathbf{b}_1^1$$

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



Assign new global parameter *u* to jointed curve

C² continuity condition

$$\mathbf{b}_3^0 = \mathbf{b}_0^1$$

$$\mathbf{b}_3^0 = \mathbf{b}_0^1 = \frac{\Delta_1}{\Delta_0 + \Delta_1}\mathbf{b}_2^0 + \frac{\Delta_0}{\Delta_0 + \Delta_1}\mathbf{b}_1^1$$

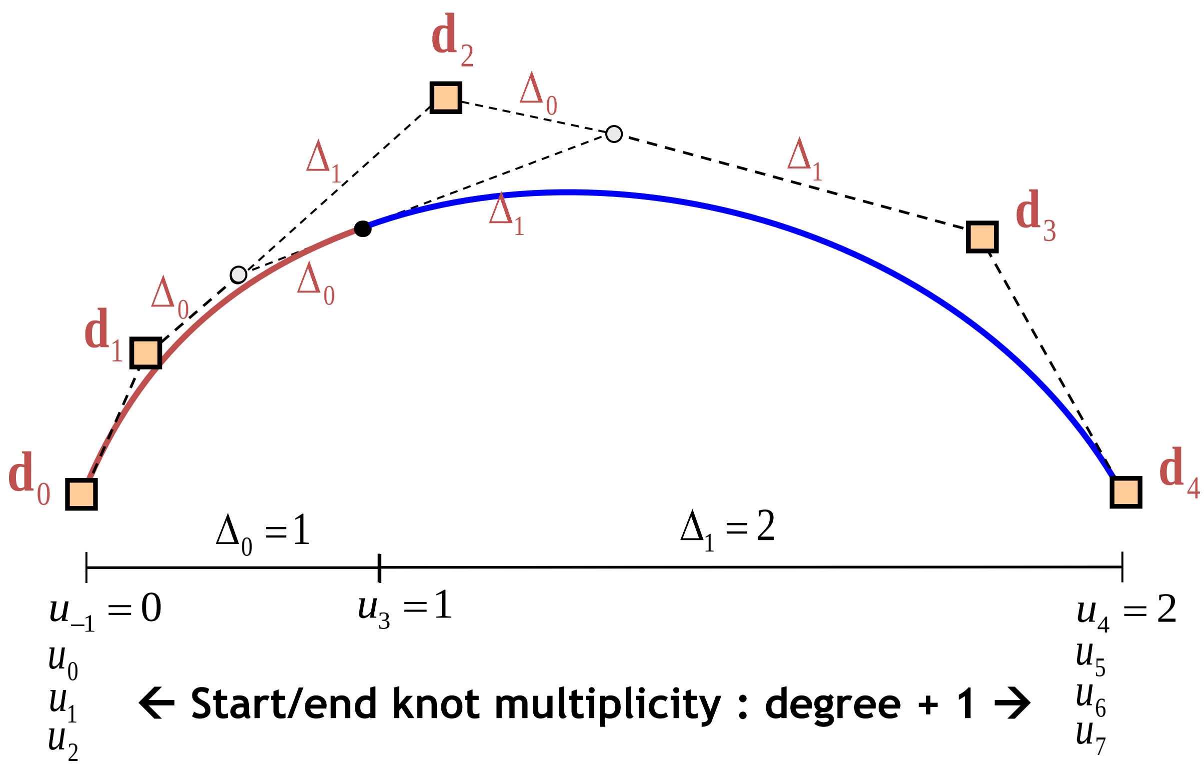
$$\mathbf{b}_2^0 = \{\frac{\Delta_1}{\Delta_0 + \Delta_1}\}\mathbf{d}_1 + \{\frac{\Delta_0}{\Delta_0 + \Delta_1}\}\mathbf{d}_2$$

$$\mathbf{b}_1^1 = \{\frac{\Delta_1}{\Delta_0 + \Delta_1}\}\mathbf{d}_2 + \{\frac{\Delta_0}{\Delta_0 + \Delta_1}\}\mathbf{d}_3$$

Definition of B-spline curve: degree, control point, knot, curve segments

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

- ☑ **n: degree**
- ☑ **S: number of Bezier curve segments**
- ☑ **number of knot = (S-1) + 2(n+1)**
- ☑ **number of control points = (n+1) + (S-1)**



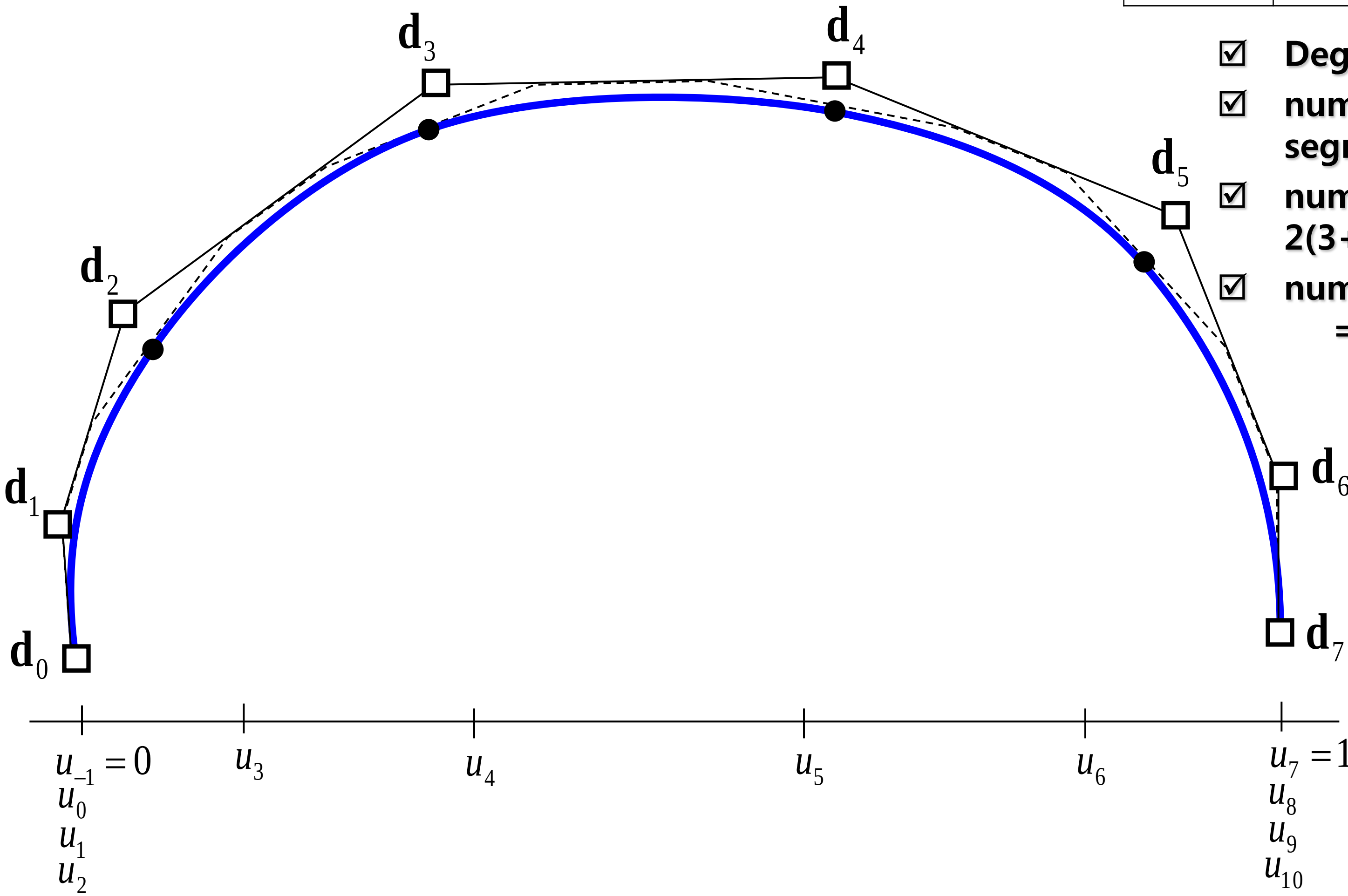
← Start/end knot multiplicity : degree + 1 →

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u)$$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u) \qquad N_i^0(u) = \begin{cases} 1 & \textbf{if } u_{i-1} \le u < u_i \\ 0 & \textbf{else} \end{cases}, \quad \sum_{i=0}^{D-1} N_i^n(u) = 1$$

Example of cubic B-spline curve with eight control points

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



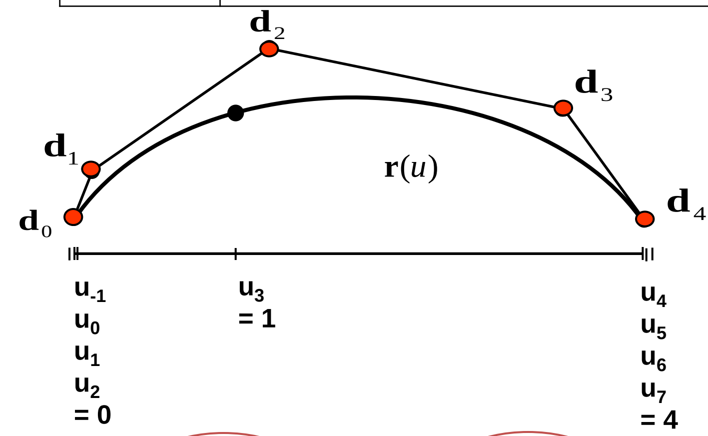
- ☑ **Degree :3**
- ☑ **number of Bezier curve segments :5**
- ☑ **number of knot = (5-1) + 2(3+1)**
- ☑ **number of control points = 4 + (5-1)**

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_6^3(u) + \mathbf{d}_7 N_7^3(u)$$

3.2 B-spline Basis Function
(Cox-de Boor recurrence formula)

1) Cox-de Boor Recurrence Formula (B-spline function) (1)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



☑ Example: Cubic B-spline curve

$$\mathbf{r}(u)=\begin{bmatrix} x(u) \\ y(u) \\ z(u) \end{bmatrix}=\sum_{i=0}^{D-1}\mathbf{d}_i N_i^n(u)$$

$$=\mathbf{d}_0 N_0^3(u)+\mathbf{d}_1 N_1^3(u)+\mathbf{d}_2 N_2^3(u)+\mathbf{d}_3 N_3^3(u)+\mathbf{d}_4 N_4^3(u)$$

- Cox-de Boor Recurrence Formula (B-spline function)

$$N_i^n(u)=\frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}}N_i^{n-1}(u)+\frac{u_{i+n}-u}{u_{i+n}-u_i}N_{i+1}^{n-1}(u)$$

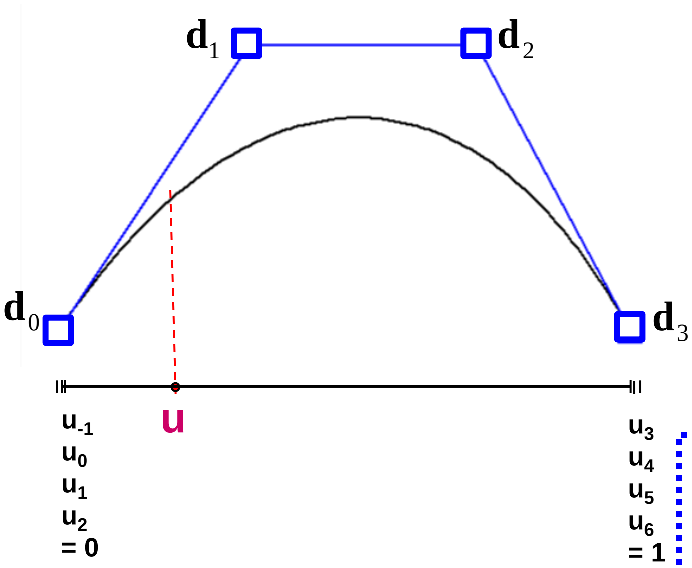
$$N_i^0(u)=\begin{cases}1 & \text{if} \quad u_{i-1}\le u<u_i \\ 0 & \text{else}\end{cases}$$

B-spline function (2)

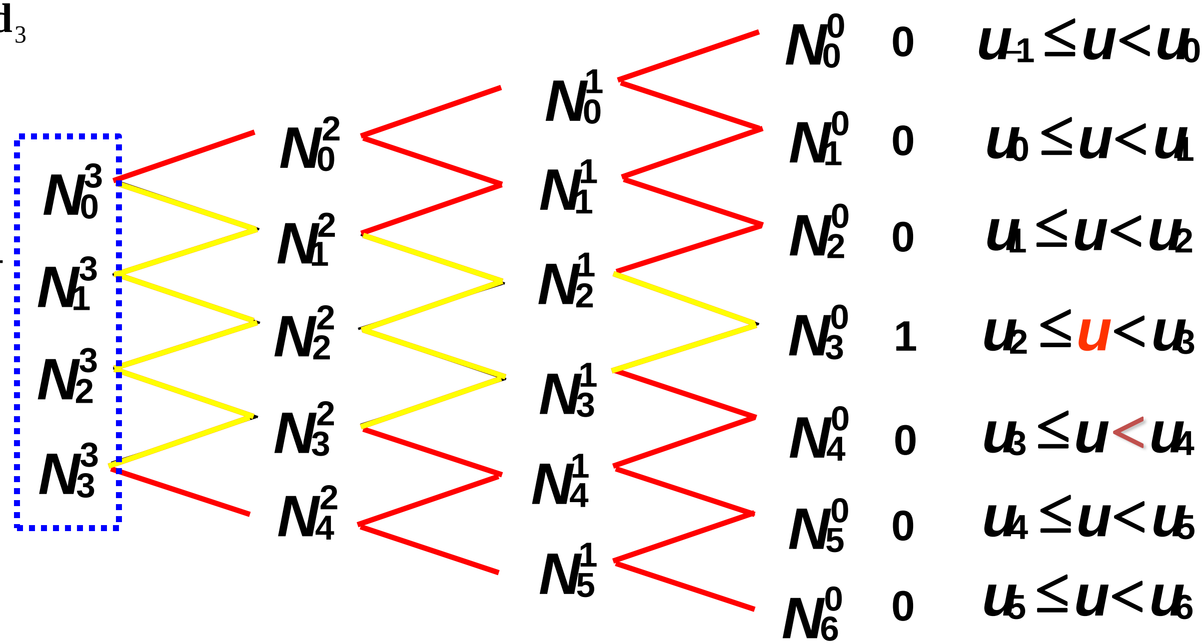
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\boxed{\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)}$$

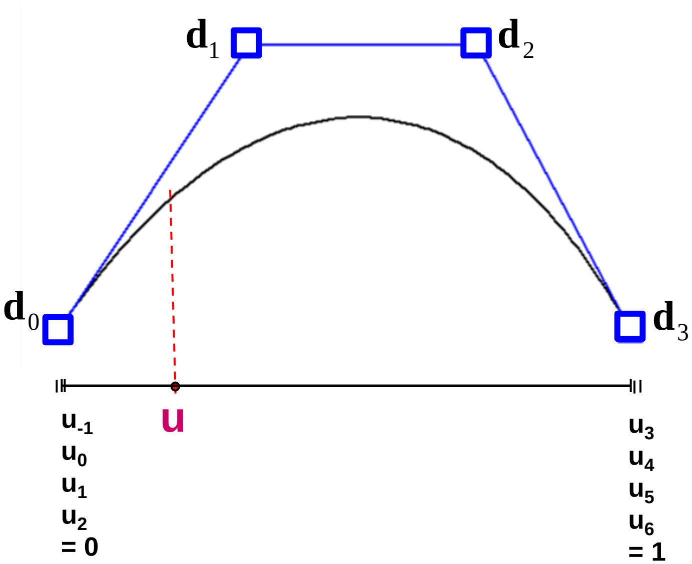


B-spline function (3)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$

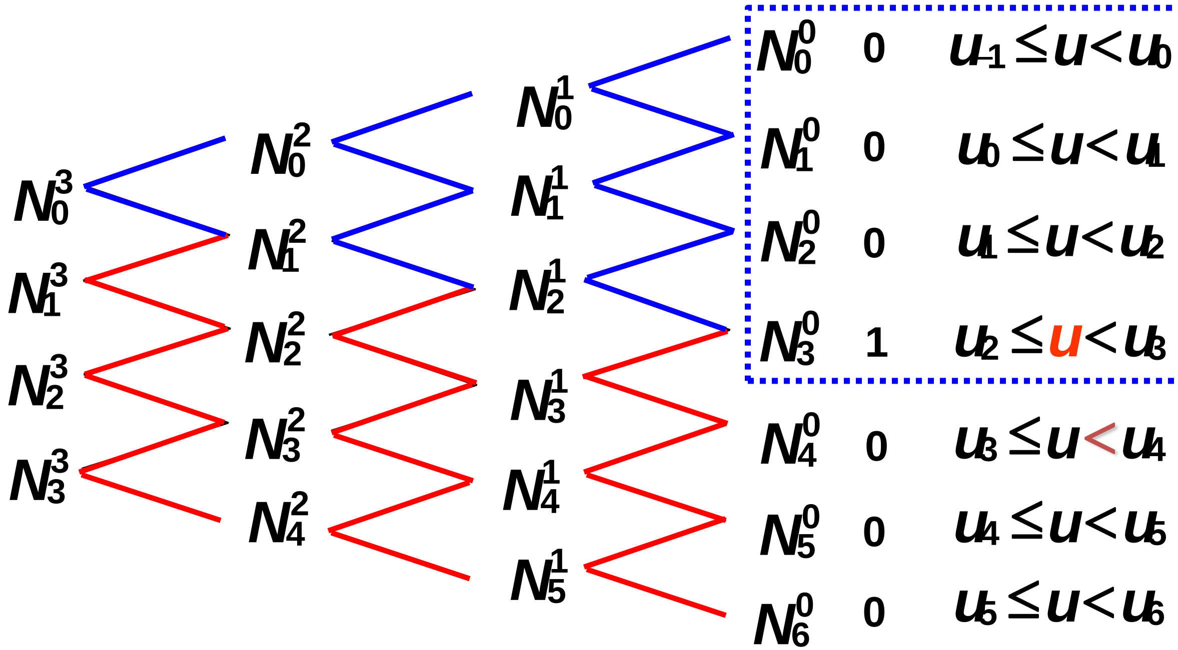
From $u_2 \le u < u_3$,

we can get $N_0^0(u) = 0$

$$N_1^0(u) = 0$$

$$N_2^0(u) = 0$$

$$N_3^0(u) = 1$$

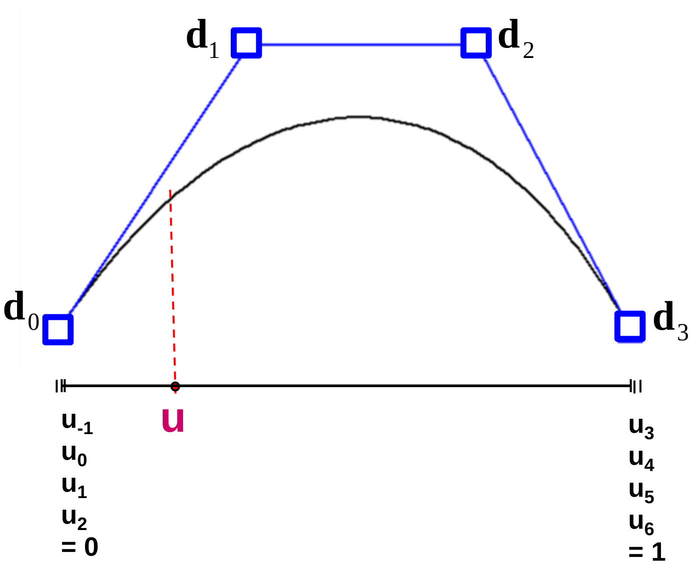


B-spline function (4)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

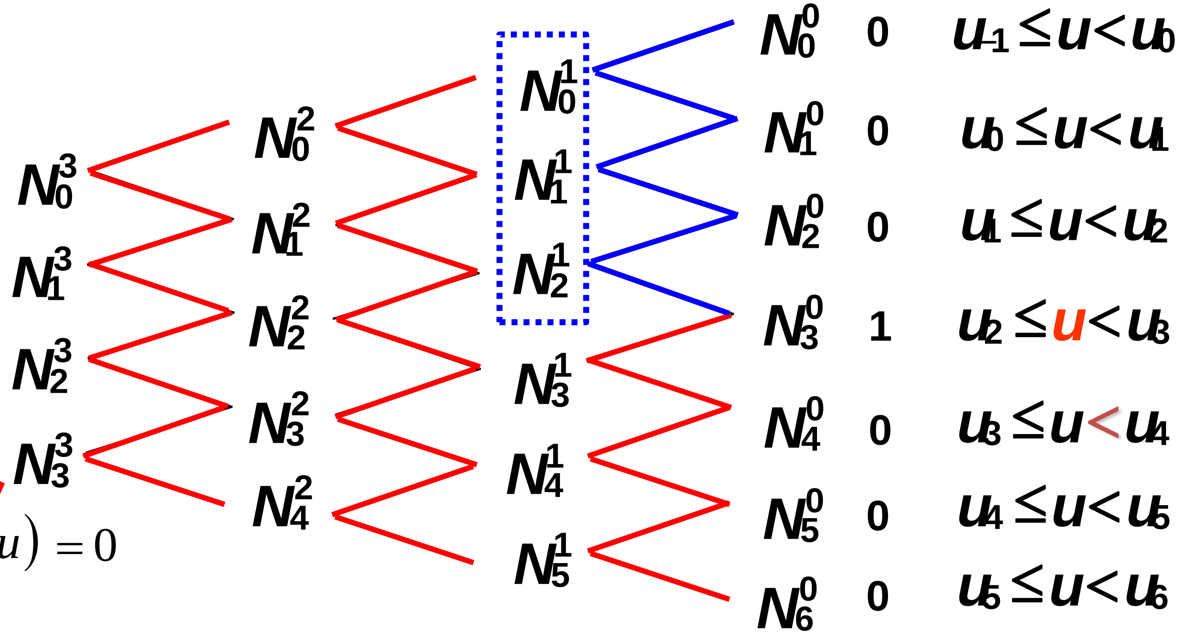
$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\boxed{\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)}$$

$$N_0^0(u) = 0, \quad N_1^0(u) = 0$$

$$N_2^0(u) = 0, \quad N_3^0(u) = 1$$



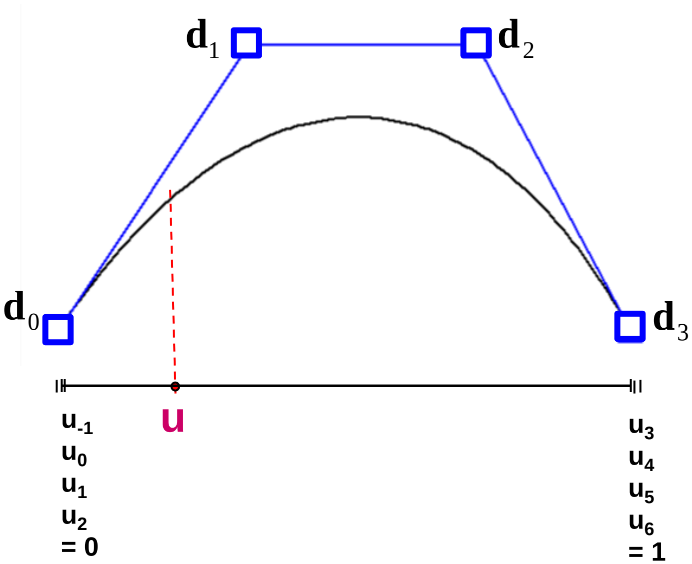
$$N_0^1(u) = \frac{u - u_{-1}}{u_0 - u_{-1}} \cancel{N_0^0(u)} + \frac{u_1 - u}{u_1 - u_0} \cancel{N_0^0(u)} = 0$$

$$N_1^1(u) = \frac{u - u_0}{u_1 - u_0} \cancel{N_1^0(u)} + \frac{u_2 - u}{u_2 - u_1} \cancel{N_2^0(u)} = 0$$

$$N_2^1(u) = \frac{u - u_1}{u_2 - u_1} \cancel{N_2^0(u)} + \frac{u_3 - u}{u_3 - u_2} N_3^0(u) = \frac{u_3 - u}{u_3 - u_2} = 1 - u$$

B-spline function (5)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

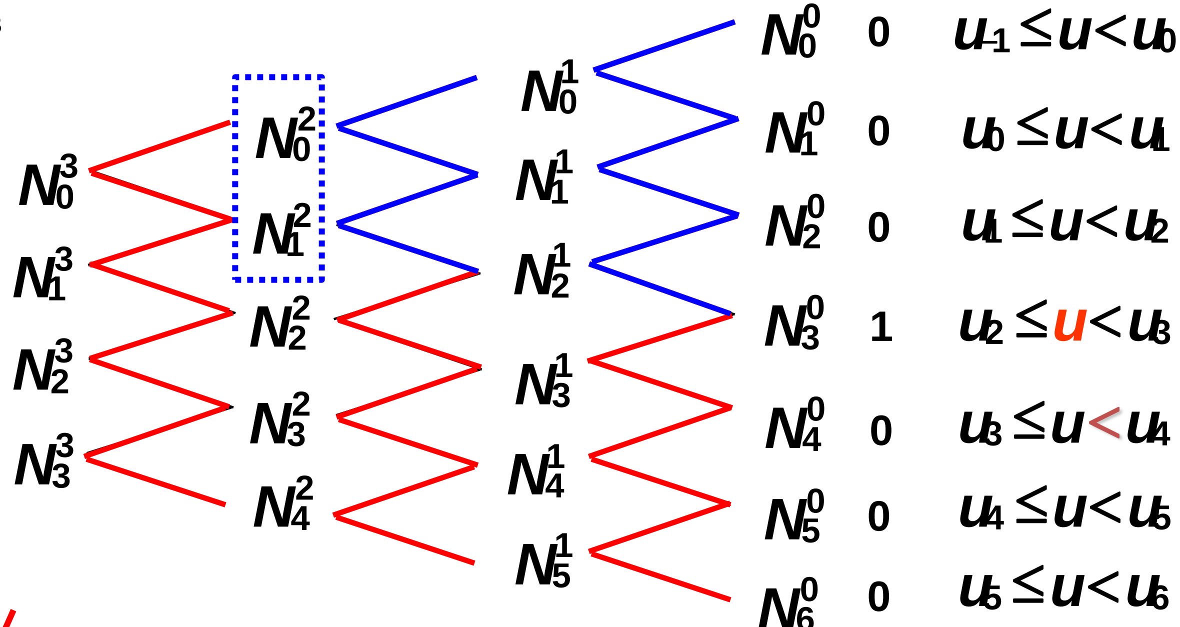
$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

$$\boxed{\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)}$$

$N_0^0(u) = 0, \quad N_1^0(u) = 0$

$N_2^0(u) = 0, \quad N_3^0(u) = 1$

$N_0^1(u) = 0, \quad N_1^1(u) = 0, \quad N_2^1(u) = 1 - u$



$$N_0^2(u) = \frac{u - u_{-1}}{u_1 - u_{-1}} \cancel{N_0^1(u)} + \frac{u_2 - u}{u_2 - u_0} \cancel{N_1^1(u)} = 0$$

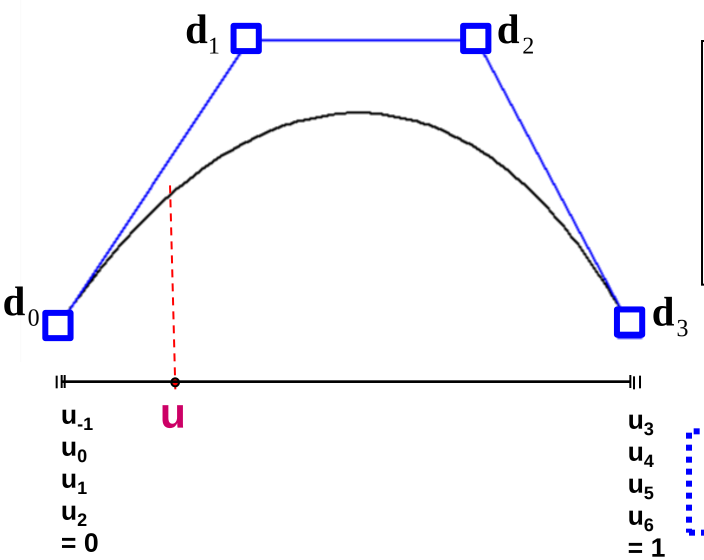
$$N_1^2(u) = \frac{u - u_0}{u_2 - u_0} \cancel{N_1^1(u)} + \frac{u_3 - u}{u_3 - u_1} N_2^1(u) = \frac{u_3 - u}{u_3 - u_1}(1 - u) = (1 - u)^2$$

B-spline function (6)

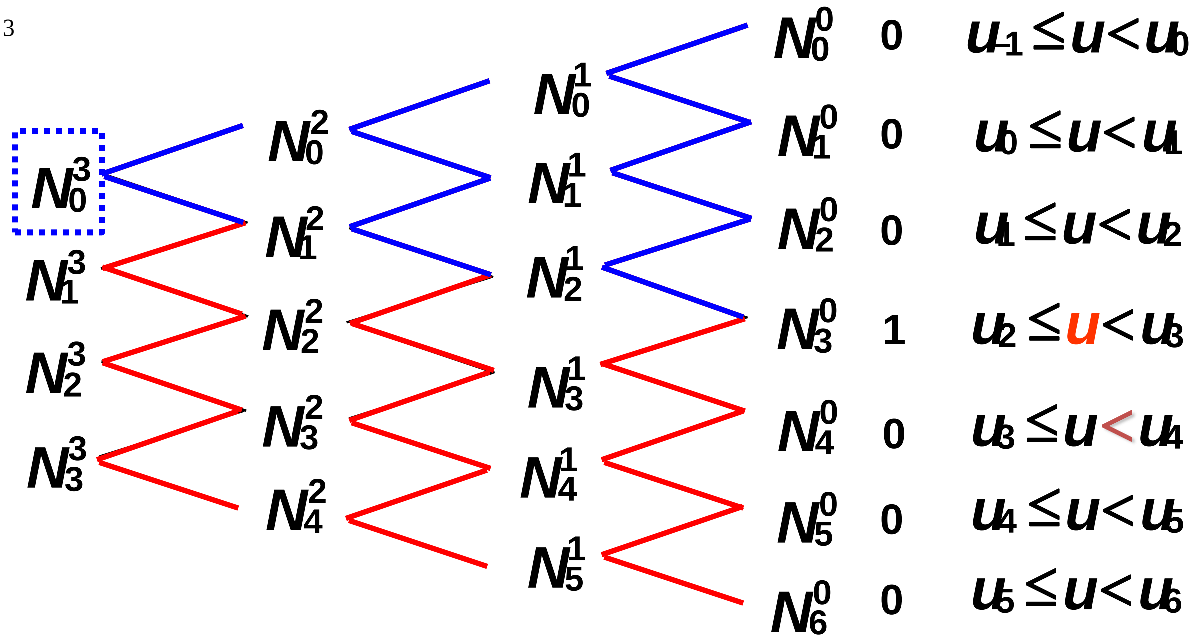
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$



$$N_0^0(u) = 0, \quad N_1^0(u) = 0$$

$$N_2^0(u) = 0, \quad N_3^0(u) = 1$$

$$N_0^1(u) = 0, \quad N_1^1(u) = 0, \quad N_2^1(u) = 1 - u$$

$$N_0^2(u) = 0, \; N_1^2(u) = (1-u)^2$$

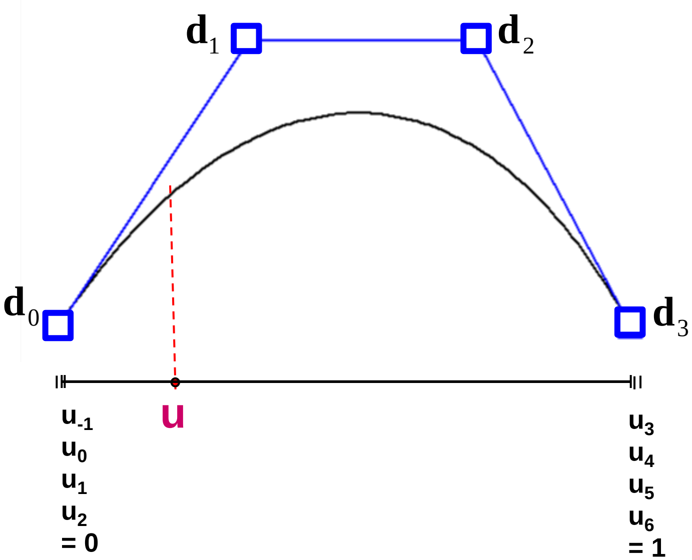
$$N_0^3(u) = \frac{u - u_{-1}}{u_2 - u_{-1}} \cancel{N_0^2(u)} + \frac{u_3 - u}{u_3 - u_0} N_1^2(u) = \frac{u_3 - u}{u_3 - u_0}(1-u)^2 = \boxed{(1-u)^3}$$

B-spline function (7)

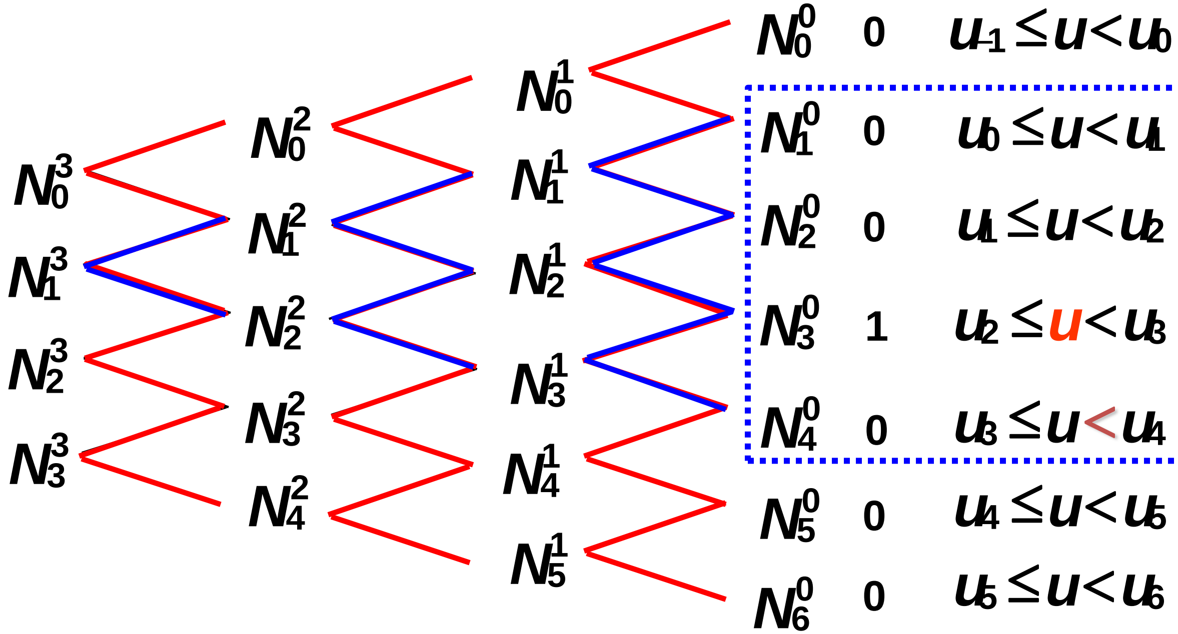
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \ u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$



From $u_2 \le u < u_3$,

we can get $N_1^0(u) = 0$

$$N_2^0(u) = 0$$

$$N_3^0(u) = 1$$

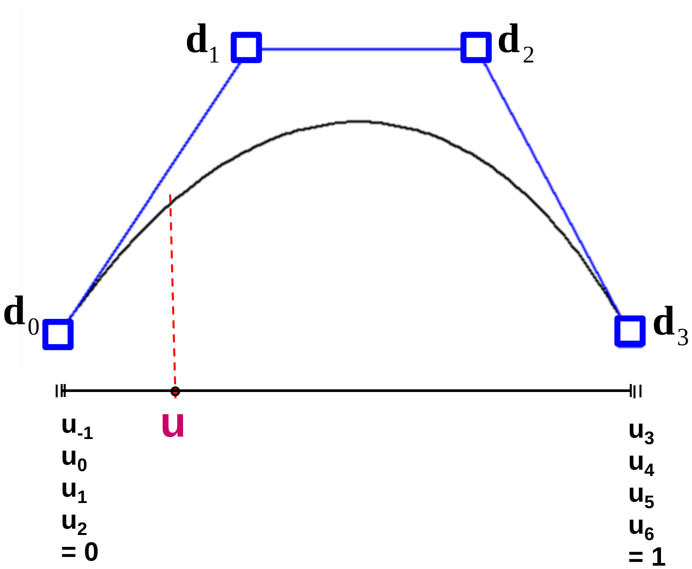
$$N_4^0(u) = 0$$

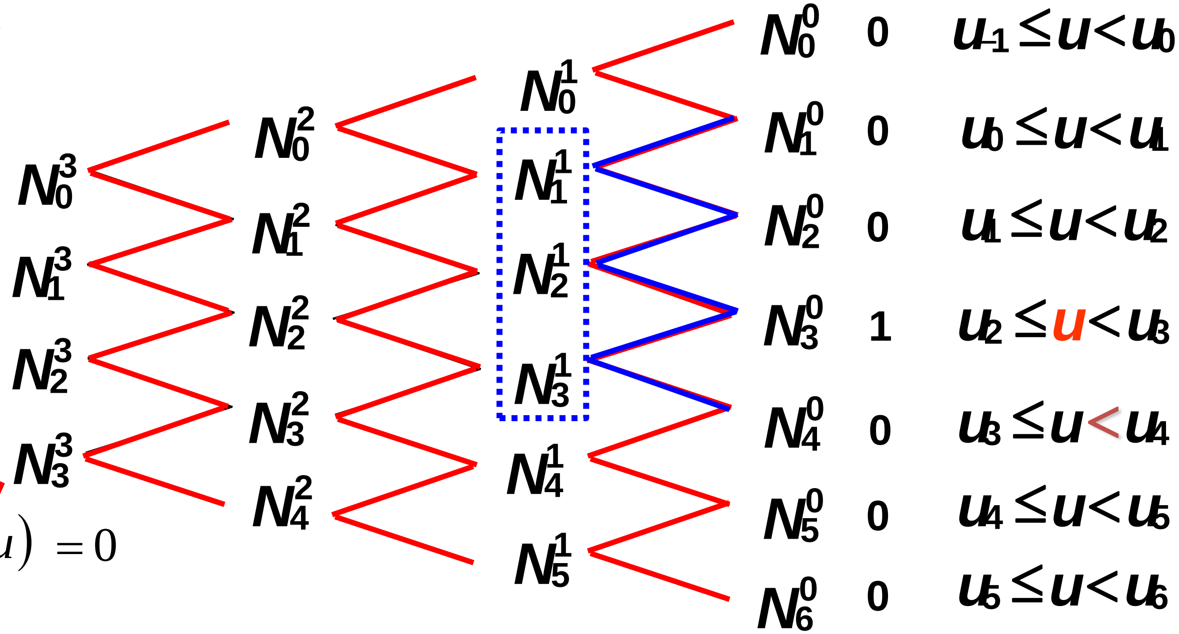
B-spline function (8)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$





$$\boxed{\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)}$$

$$N_1^0(u) = 0, \quad N_2^0(u) = 0$$

$$N_3^0(u) = 1, \quad N_4^0(u) = 0$$

$$N_1^1(u) = \frac{u - u_0}{u_1 - u_0} \cancel{N_1^0(u)} + \frac{u_2 - u}{u_2 - u_0} \cancel{N_2^0(u)} = 0$$

$$N_2^1(u) = \frac{u - u_1}{u_2 - u_1} \cancel{N_2^0(u)} + \frac{u_3 - u}{u_3 - u_2} N_3^0(u) = \frac{u_3 - u}{u_3 - u_2} = 1 - u$$

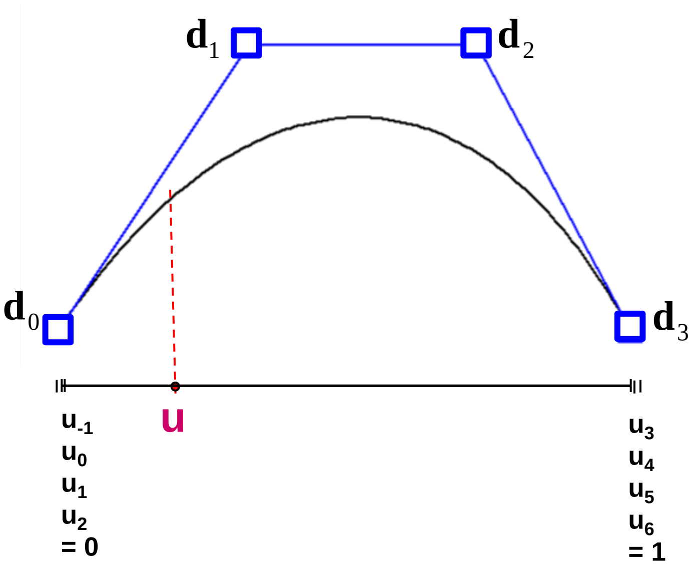
$$N_3^1(u) = \frac{u - u_1}{u_3 - u_2} N_3^0(u) + \frac{u_4 - u}{u_4 - u_3} \cancel{N_4^0(u)} = \frac{u - u_1}{u_3 - u_2} = u$$

B-spline function (9)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n}-u}{u_{i+n}-u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

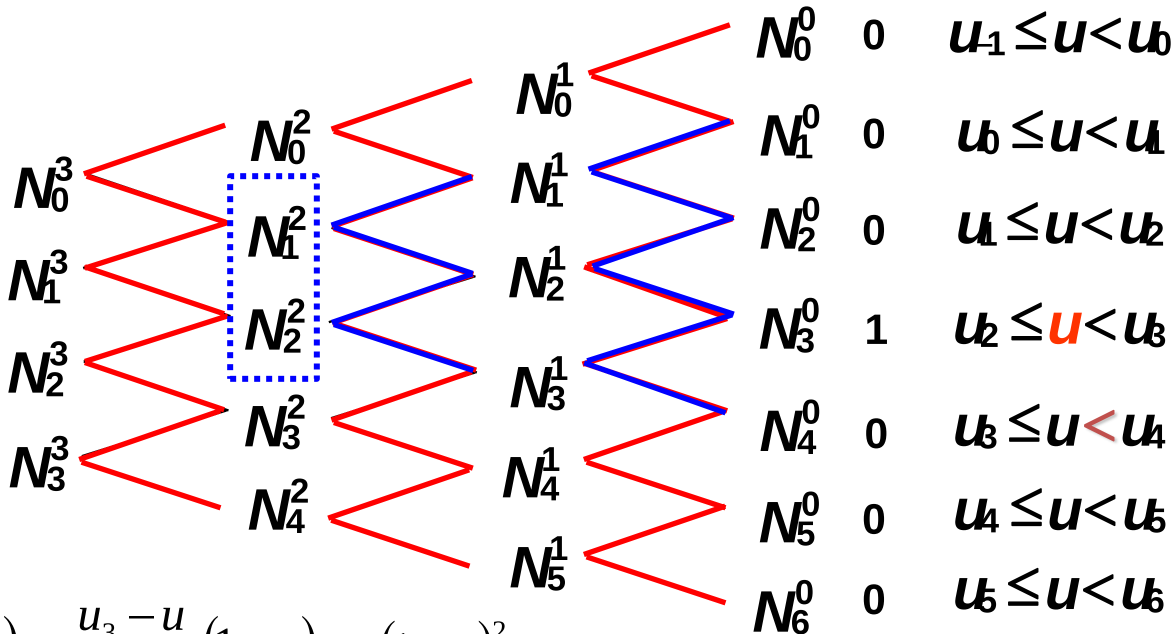


$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$

$$N_1^0(u) = 0, \quad N_2^0(u) = 0$$

$$N_3^0(u) = 1, \quad N_4^0(u) = 0$$

$$N_1^1(u) = 0, \quad N_2^1(u) = 1-u, \quad N_3^1(u) = u$$



$$N_1^2(u) = \frac{u-u_0}{u_2-u_0} \cancel{N_1^1(u)} + \frac{u_3-u}{u_3-u_1} N_2^1(u) = \frac{u_3-u}{u_3-u_1}(1-u) = (1-u)^2$$

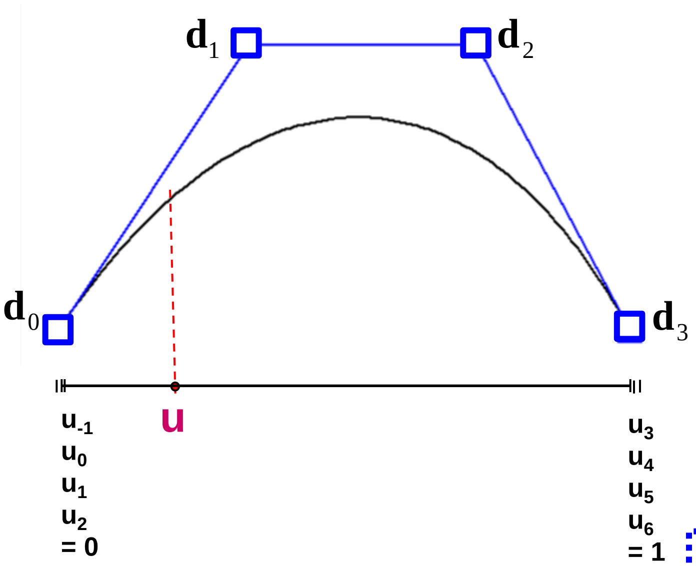
$$N_2^2(u) = \frac{u-u_1}{u_3-u_1} N_2^1(u) + \frac{u_4-u}{u_4-u_2} N_3^1(u) = \frac{u-u_1}{u_3-u_1}(1-u) + \frac{u_4-u}{u_4-u_2} u = 2u(1-u)$$

B-spline function (10)

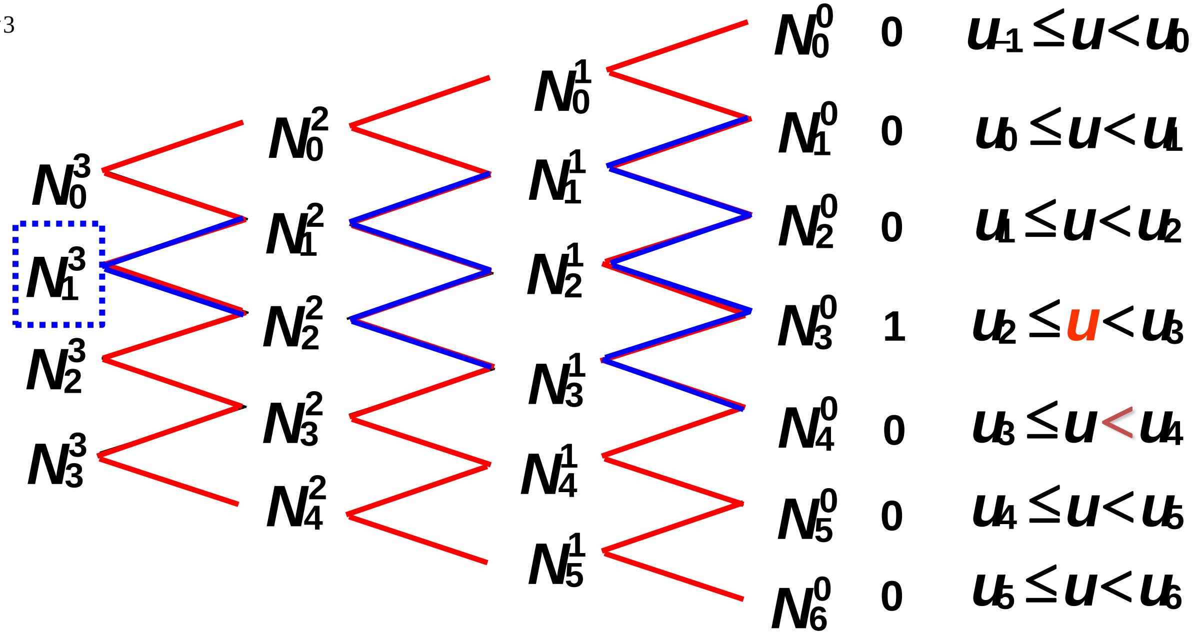
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$\boxed{\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)}$$



$$N_1^0(u) = 0, \quad N_2^0(u) = 0$$

$$N_3^0(u) = 1, \quad N_4^0(u) = 0$$

$$N_0^1(u) = 0, \quad N_1^1(u) = 0, \quad N_2^1(u) = 1 - u$$

$$N_1^2(u) = (1-u)^2, \quad N_2^2(u) = 2u(1-u)$$

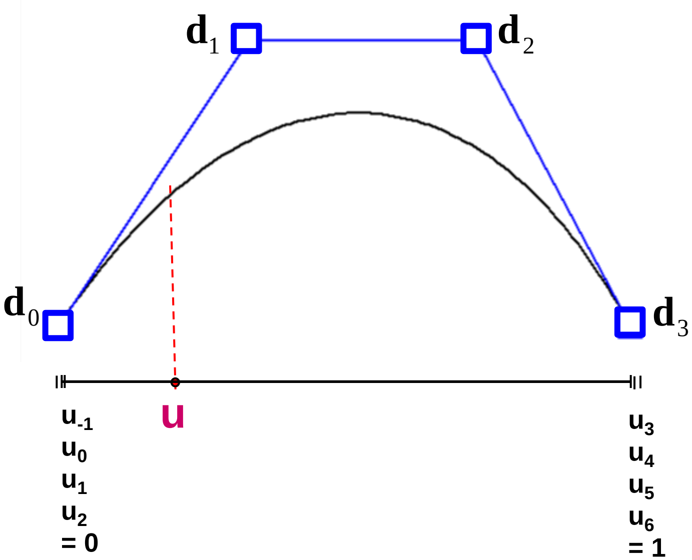
$$N_1^3(u) = \frac{u - u_0}{u_3 - u_0} N_1^2(u) + \frac{u_4 - u}{u_4 - u_1} N_2^2(u) = \frac{u - u_0}{u_3 - u_0}(1-u)^2 + \frac{u_4 - u}{u_4 - u_1} \cdot 2u(1-u) = 3u(1-u)^2$$

B-spline function (11)

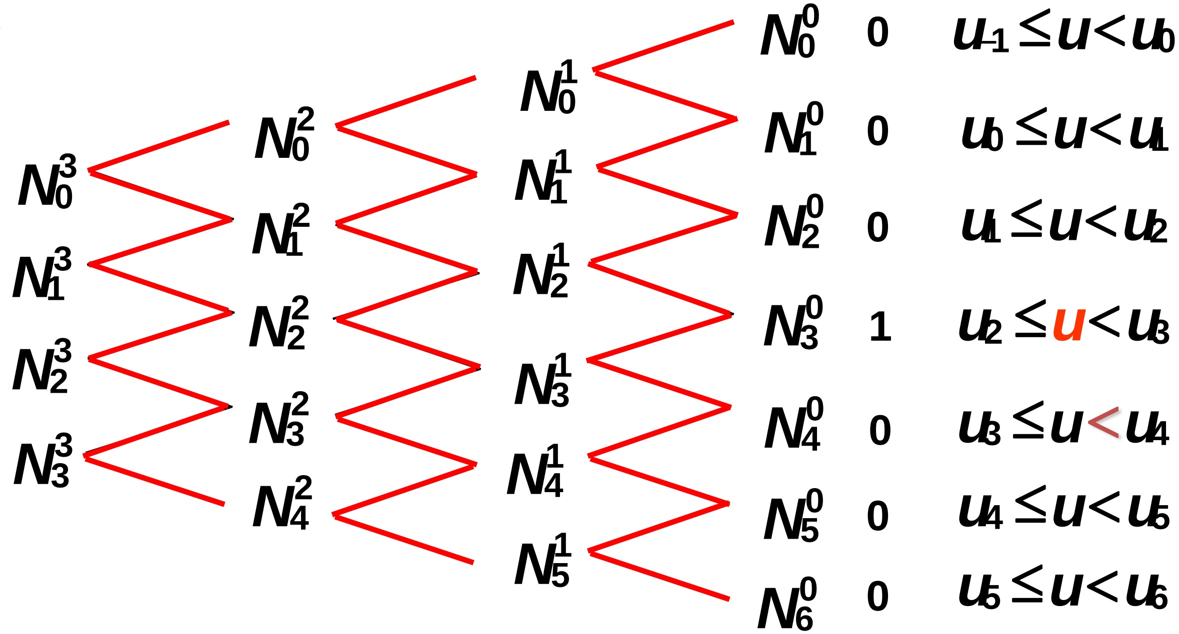
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$



Calculate $N_2^3(u), N_2^3(u)$
In the same manner

$$N_0^3(u) = (1-u)^3$$

$$N_1^3(u) = 3u(1-u)^2$$

$$N_2^3(u) = 3u^2(1-u)$$

$$N_3^3(u) = 3u^3$$

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$

$$\mathbf{r}(u) = (1-u)^3 \mathbf{d}_0 + 3u(1-u)^2 \mathbf{d}_1 + 3u^2(1-u)\mathbf{d}_2 + u^3 \mathbf{d}_3$$

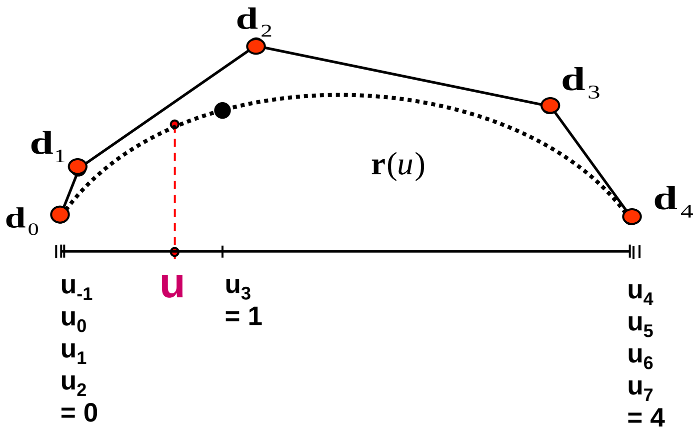
➔The B-spline curve is identical with 3rd-degree Bezier curves four given control points

B-spline function (12)

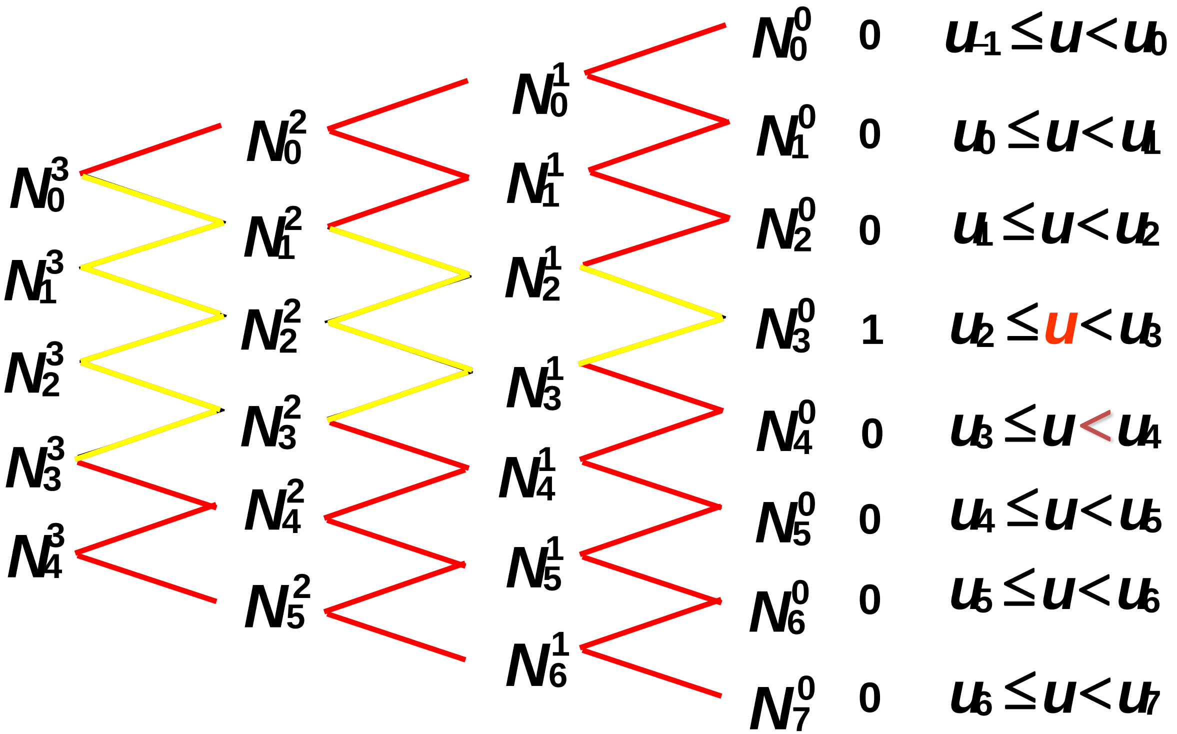
Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \; u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

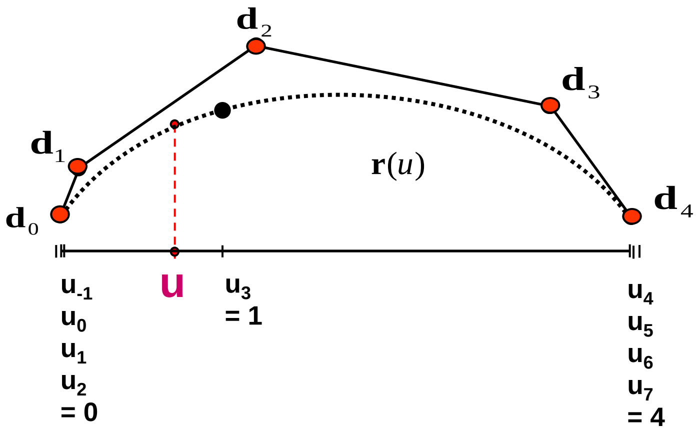


$$\mathbf{r}(u) = \sum_{i=0}^{D-1} \mathbf{d}_i N_i^n(u)$$



B-spline function (13)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$

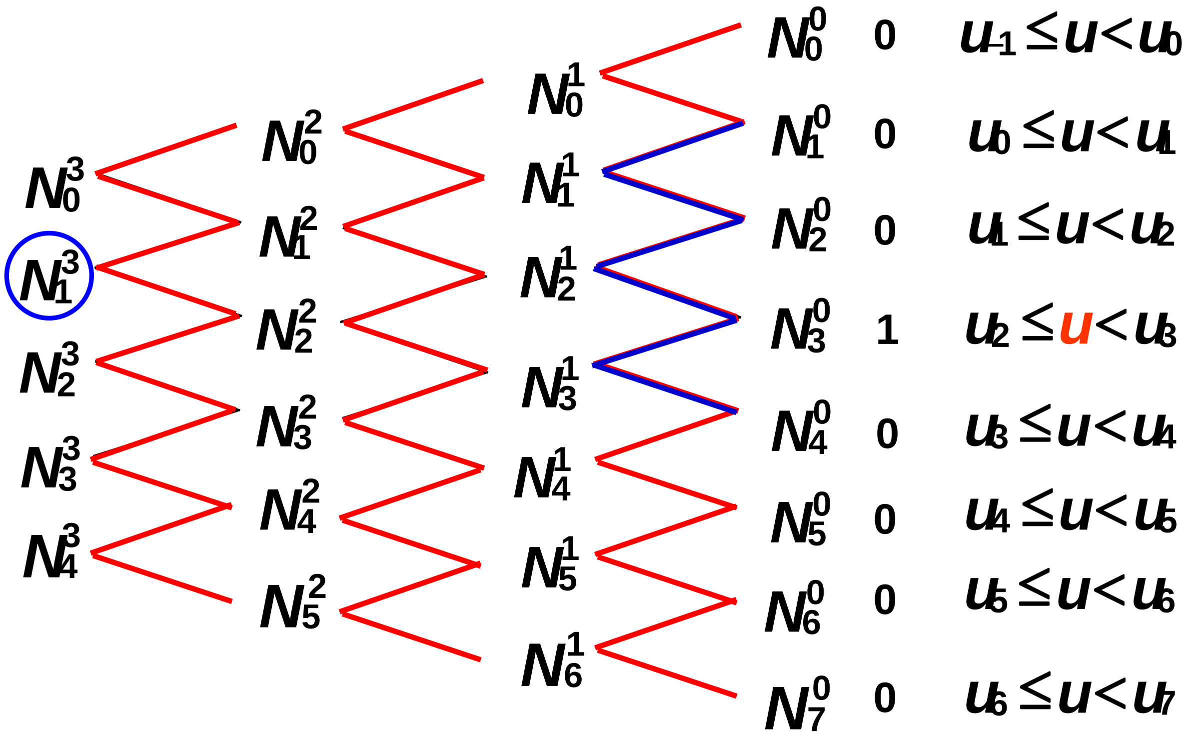


$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

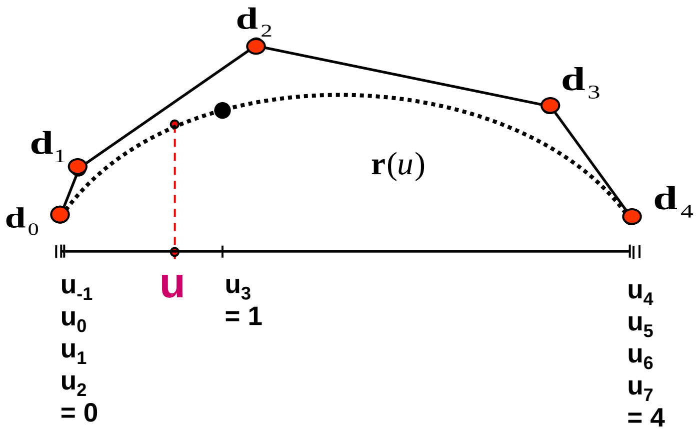
$$N_1^1 = 0 \qquad N_2^1 = \frac{u_3 - u}{u_3 - u_2} \qquad N_3^1 = \frac{u - u_2}{u_3 - u_2}$$

$$\mathbf{r}(u) = \sum_{i=0}^{D-1} \mathbf{d}_i N_i^n(u)$$



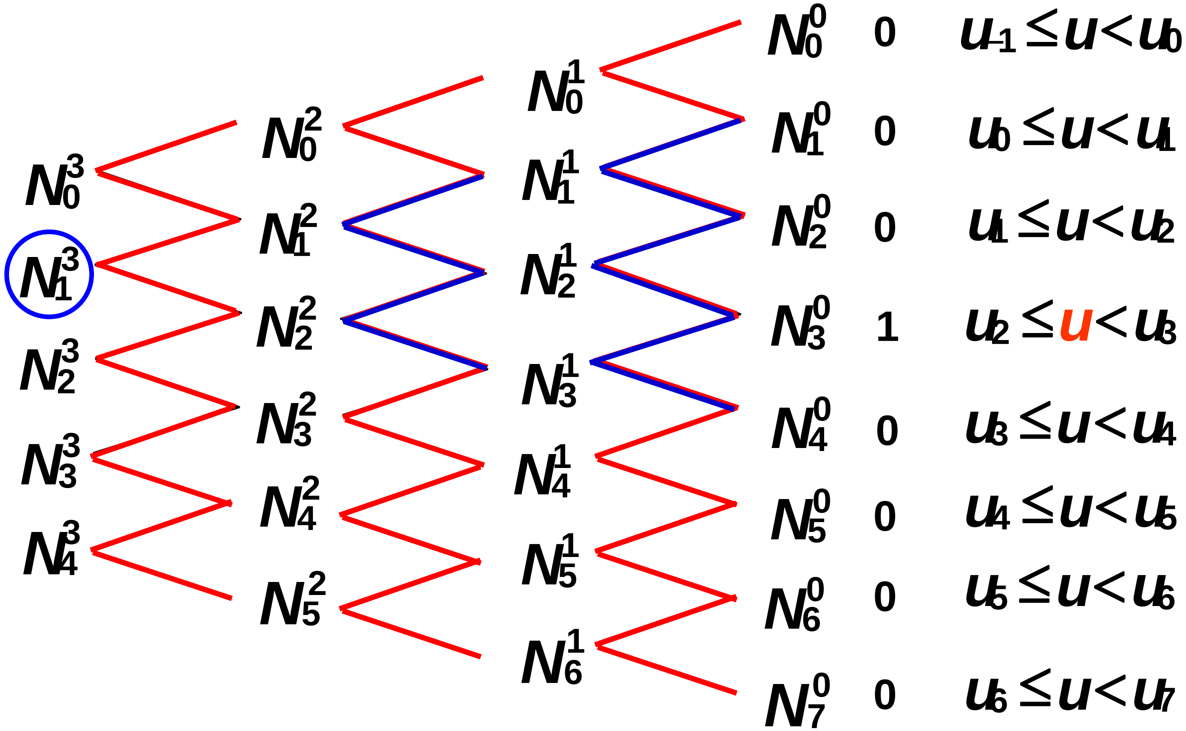
B-spline function (14)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \; u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$N_1^1 = 0 \qquad N_2^1 = \frac{u_3 - u}{u_3 - u_2} \qquad N_3^1 = \frac{u - u_2}{u_3 - u_2}$$

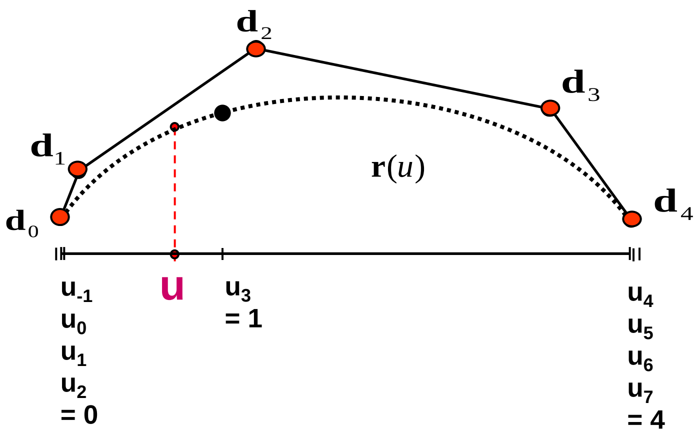
$$N_1^2 = \frac{u_3 - u}{u_3 - u_1} N_2^1 = \frac{u_3 - u}{u_3 - u_1} \cdot \frac{u_3 - u}{u_3 - u_2}$$

$$N_2^2 = \frac{u - u_1}{u_3 - u_1} N_2^1 + \frac{u_4 - u}{u_4 - u_2} N_3^1$$

$$= \frac{u - u_1}{u_3 - u_1} \cdot \frac{u_3 - u}{u_3 - u_2} + \frac{u_4 - u}{u_4 - u_2} \cdot \frac{u - u_2}{u_3 - u_2}$$

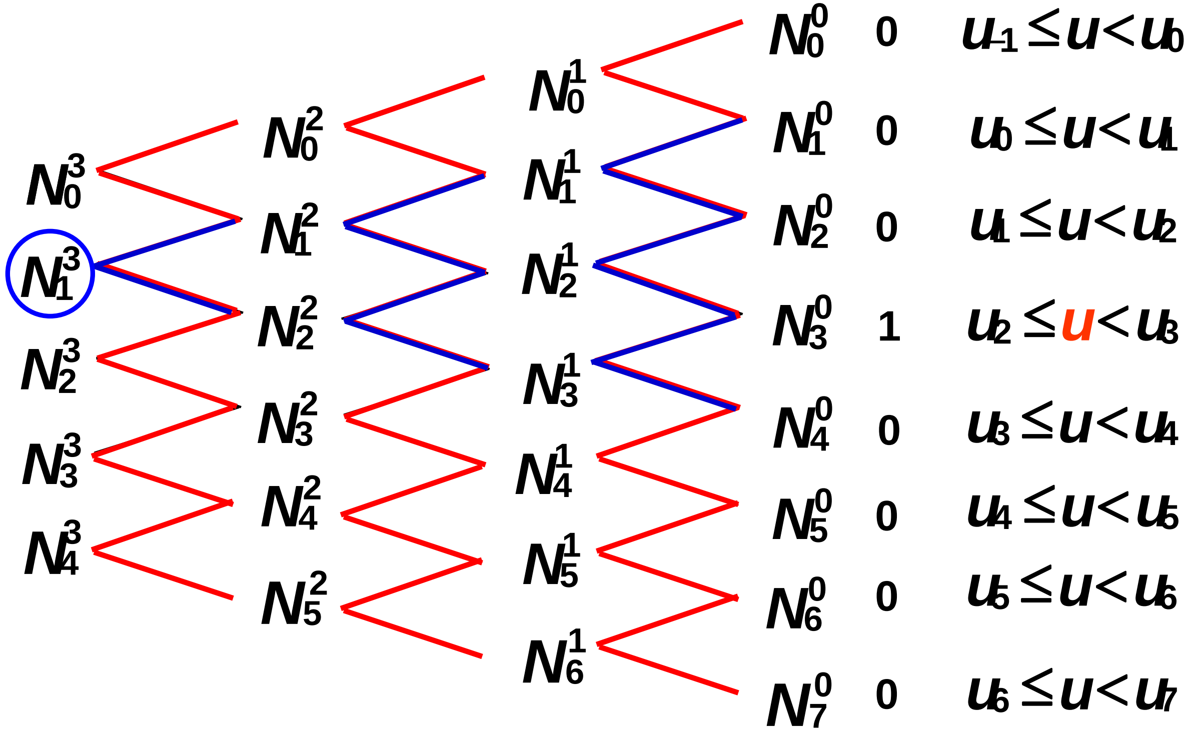
B-spline function (15)

Given	B-spline Control Point $\mathbf{d}_i$ Parameter u B-spline Basis Func. $N_i^n(u)$
Find	B-spline Curve $\mathbf{r}(u)$



$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$



$$N_1^1 = 0 \qquad N_2^1 = \frac{u_2 - u}{u_3 - u_2} \qquad N_3^1 = \frac{u - u_2}{u_3 - u_2}$$

$$N_1^2 = \frac{u_3 - u}{u_3 - u_1} N_2^1 = \frac{u_3 - u}{u_3 - u_1} \cdot \frac{u_2 - u}{u_3 - u_2}$$

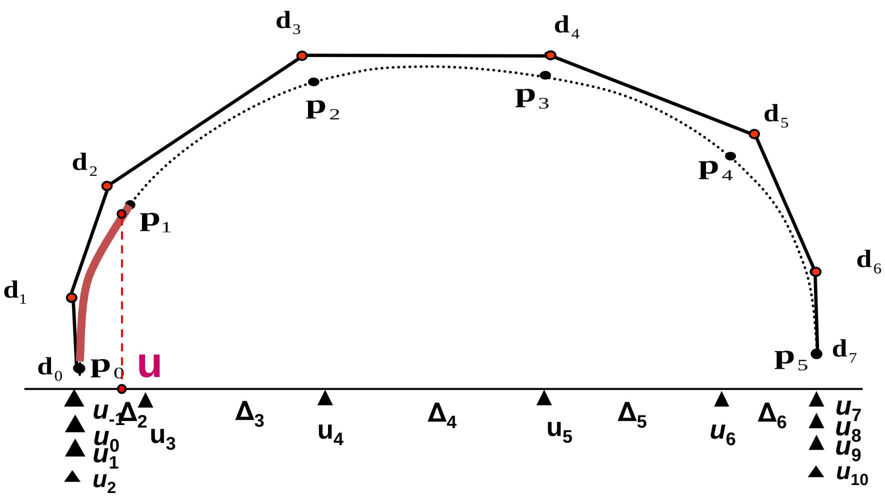
$$N_2^2 = \frac{u - u_1}{u_3 - u_1} N_2^1 + \frac{u_4 - u}{u_4 - u_2} N_3^1$$

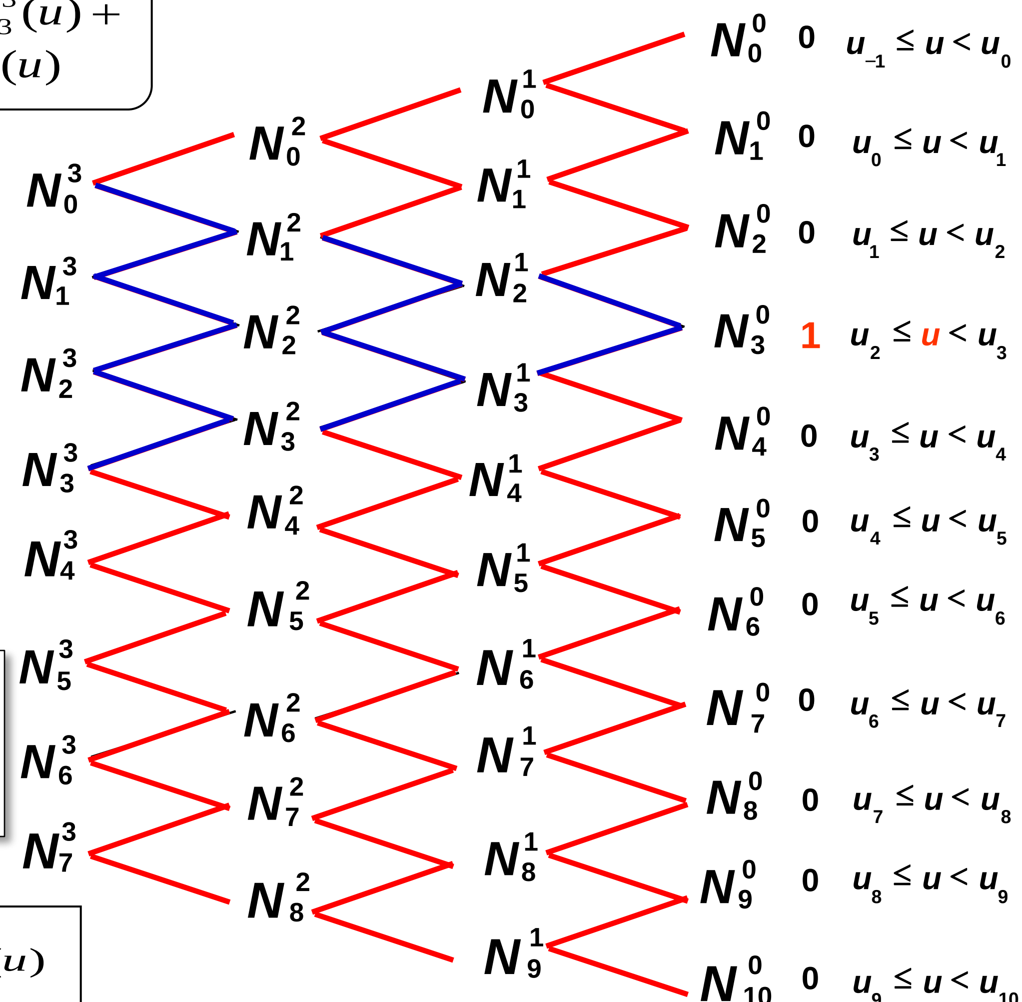
$$= \frac{u - u_1}{u_3 - u_1} \cdot \frac{u_2 - u}{u_3 - u_2} + \frac{u_4 - u}{u_4 - u_2} \cdot \frac{u - u_2}{u_3 - u_2}$$

$$N_1^3 = \frac{u - u_0}{u_3 - u_2} N_1^2 + \frac{u_4 - u}{u_4 - u_1} N_2^2 = \frac{u - u_0}{u_3 - u_2} \cdot \frac{u_3 - u}{u_3 - u_1} \cdot \frac{u_2 - u}{u_3 - u_2} + \frac{u_4 - u}{u_4 - u_1} \cdot \frac{u - u_1}{u_3 - u_1} \cdot \frac{u_2 - u}{u_3 - u_2} + \frac{u_4 - u}{u_4 - u_1} \cdot \frac{u_4 - u}{u_4 - u_2} \cdot \frac{u - u_2}{u_3 - u_2}$$

3.3 B-spline curves (1)

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_6^3(u) + \mathbf{d}_7 N_7^3(u)$$





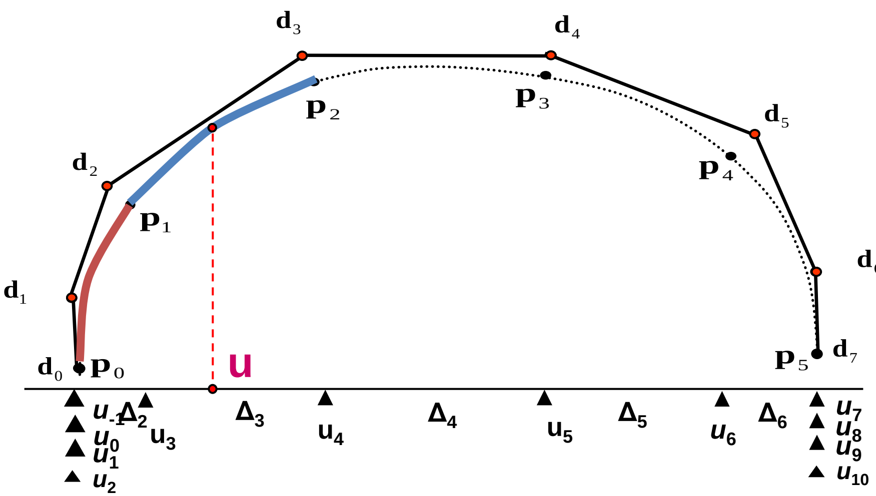
$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_4 \cancel{N_4^3}(u) + \mathbf{d}_5 \cancel{N_5^3}(u) + \mathbf{d}_6 \cancel{N_6^3}(u) + \mathbf{d}_7 \cancel{N_7^3}(u)$$

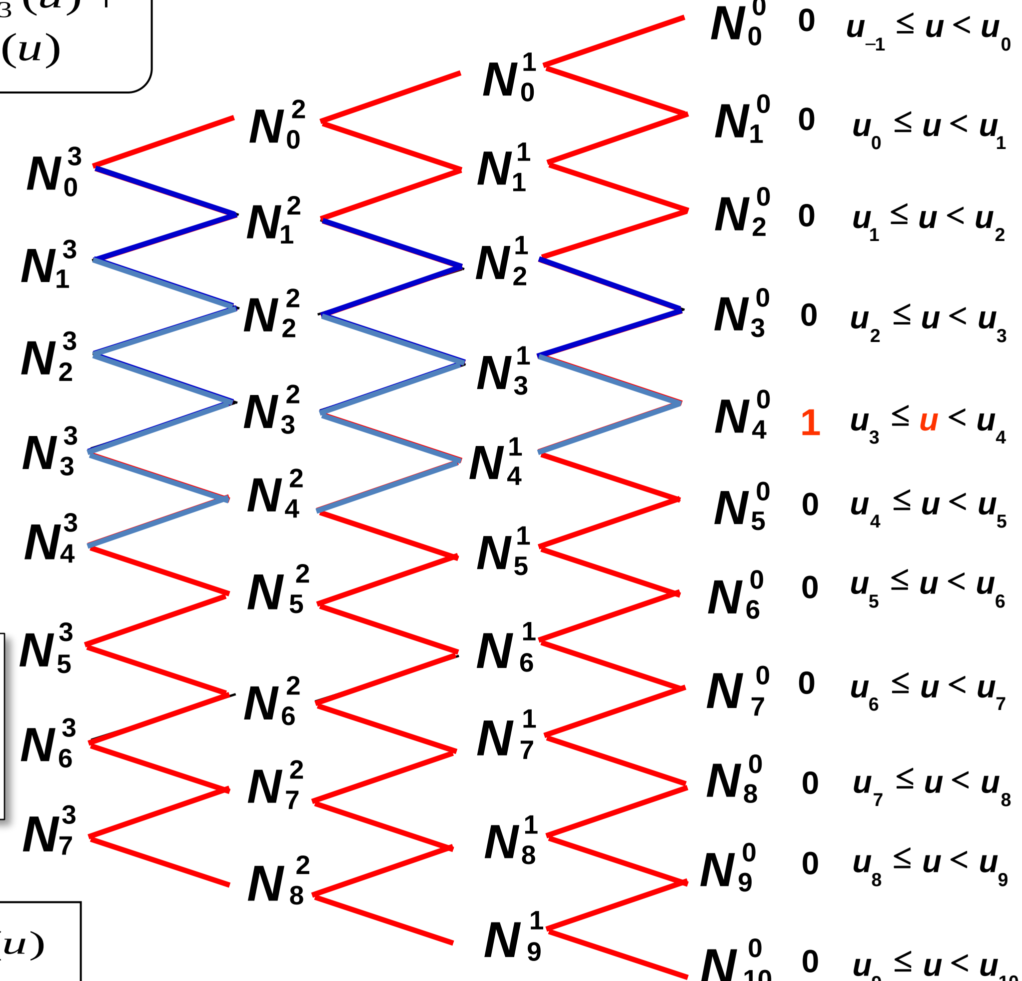
$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

B-spline curves (2)

$$\mathbf{r}(u)=\mathbf{d}_0N_0^3(u)+\mathbf{d}_1N_1^3(u)+\mathbf{d}_2N_2^3(u)+\mathbf{d}_3N_3^3(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_5N_5^3(u)+\mathbf{d}_6N_6^3(u)+\mathbf{d}_7N_7^3(u)$$





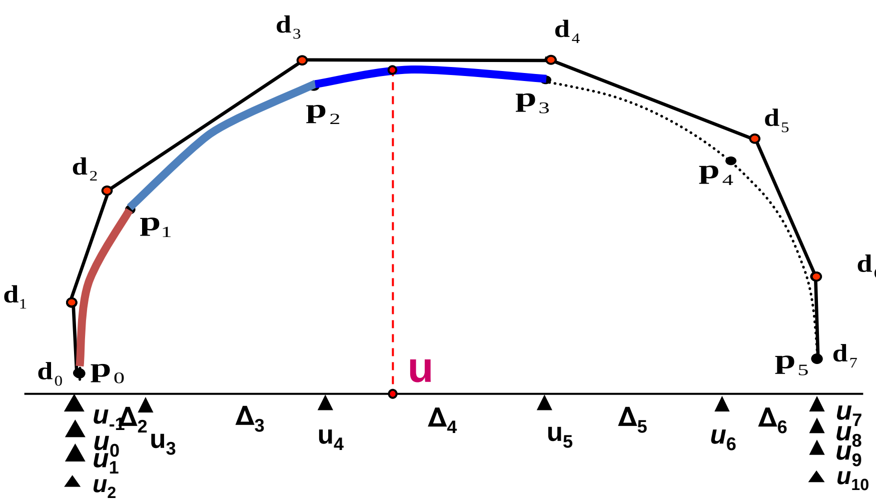
$$\mathbf{r}(u)=\mathbf{d}_0\cancel{N_0^3}(u)+\mathbf{d}_1N_1^3(u)+\mathbf{d}_2N_2^3(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_5\cancel{N_5^3}(u)+\mathbf{d}_6\cancel{N_6^3}(u)+\mathbf{d}_7\cancel{N_7^3}(u)$$

$$N_i^n(u)=\frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}}N_i^{n-1}(u)+\frac{u_{i+n}-u}{u_{i+n}-u_i}N_{i+1}^{n-1}(u)$$

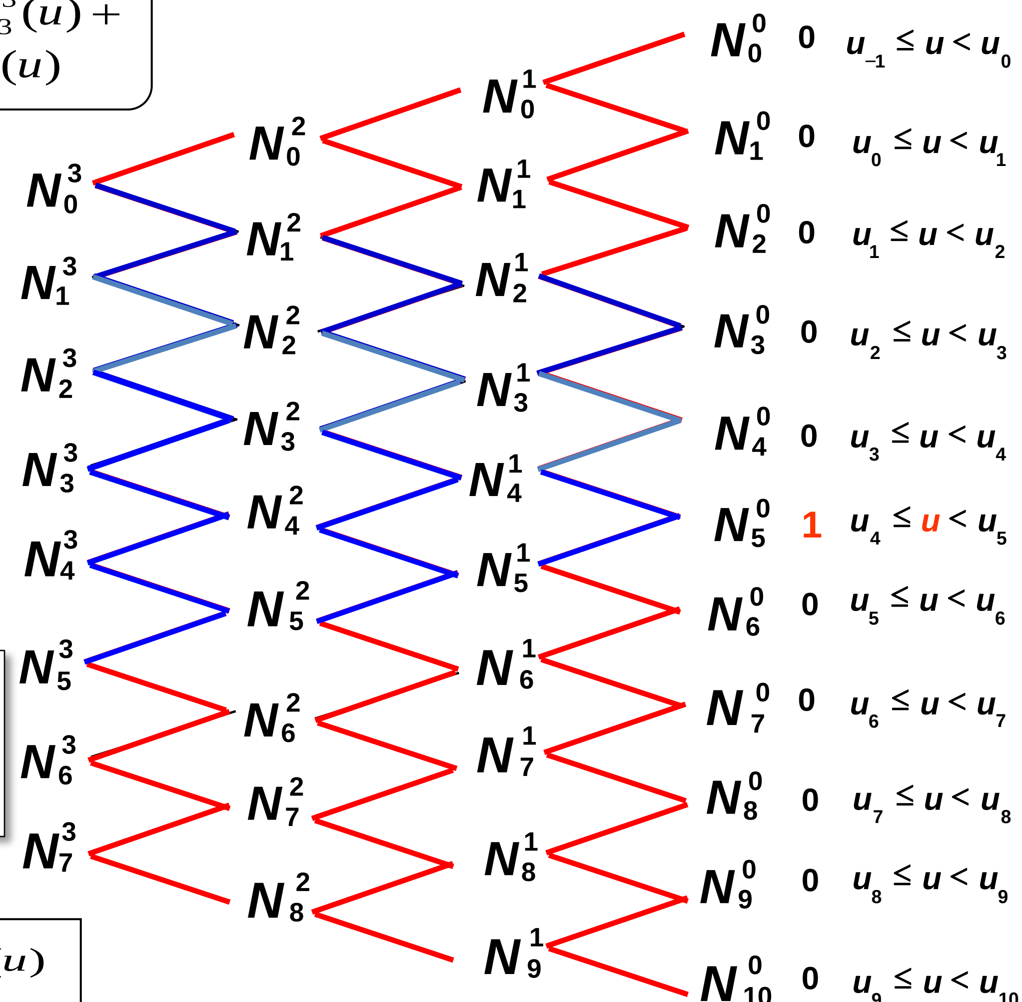
$$N_i^0(u)=\begin{cases}1 & \text{if} \quad u_{i-1}\le u<u_i\\ 0 & \text{else}\end{cases}$$

B-spline curves (3)

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_6^3(u) + \mathbf{d}_7 N_7^3(u)$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_{\cancel{0}}^{3}(u) + \mathbf{d}_1 N_{\cancel{1}}^{3}(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_{\cancel{6}}^{3}(u) + \mathbf{d}_7 N_{\cancel{7}}^{3}(u)$$

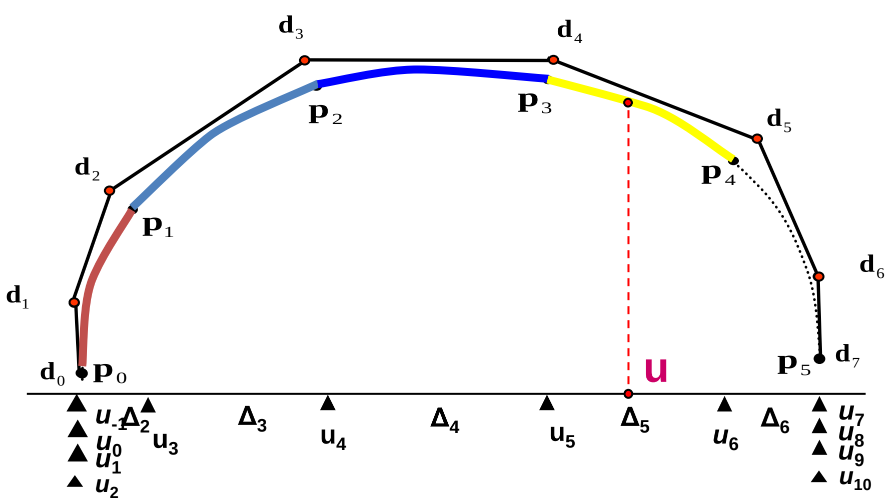


$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

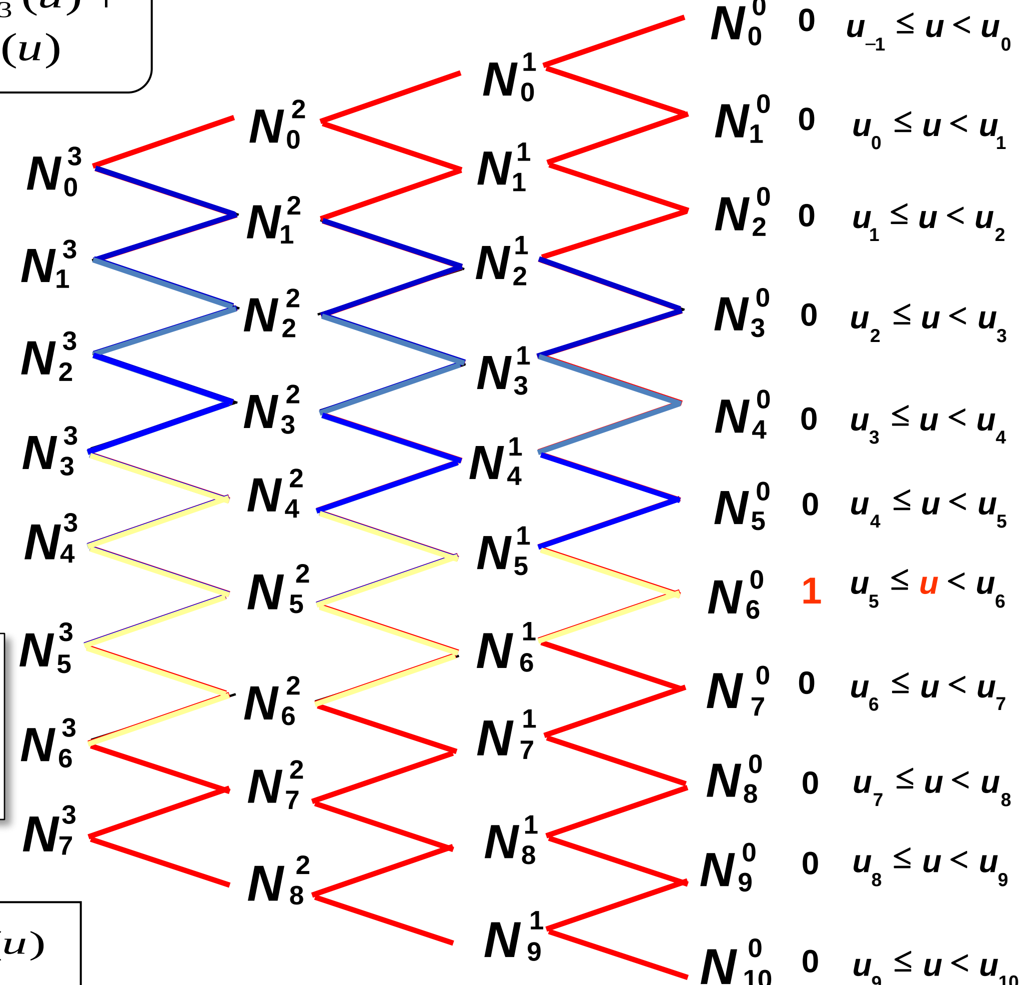
$$N_i^0(u) = \begin{cases} 1 & \text{if} \quad u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

B-spline curves (4)

$$\mathbf{r}(u)=\mathbf{d}_0N_0^3(u)+\mathbf{d}_1N_1^3(u)+\mathbf{d}_2N_2^3(u)+\mathbf{d}_3N_3^3(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_5N_5^3(u)+\mathbf{d}_6N_6^3(u)+\mathbf{d}_7N_7^3(u)$$



$$\mathbf{r}(u)=\mathbf{d}_0\cancel{N_0^3}(u)+\mathbf{d}_1\cancel{N_1^3}(u)+\mathbf{d}_2\cancel{N_2^3}(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_4N_4^3(u)+\mathbf{d}_5N_5^3(u)+\mathbf{d}_6N_6^3(u)+\mathbf{d}_7\cancel{N_7^3}(u)$$

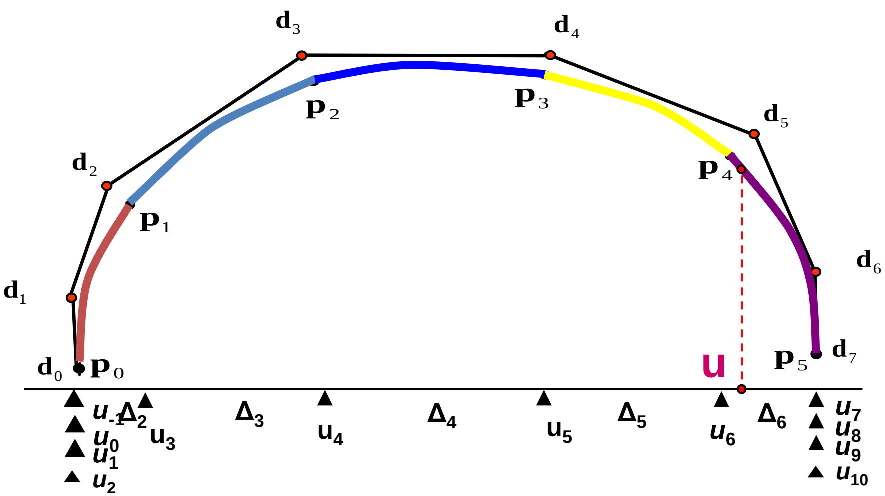


$$N_i^n(u)=\frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}}N_i^{n-1}(u)+\frac{u_{i+n}-u}{u_{i+n}-u_i}N_{i+1}^{n-1}(u)$$

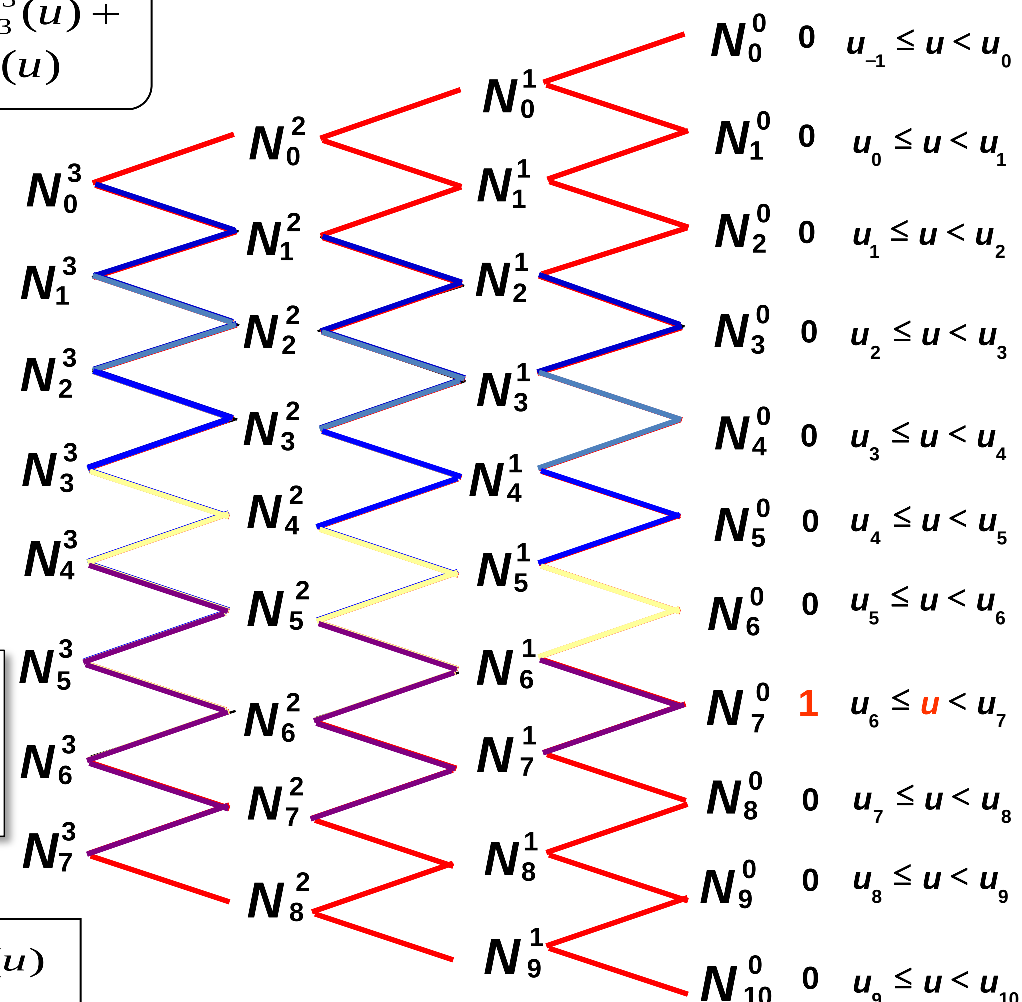
$$N_i^0(u)=\begin{cases}1 & \text{if } \ u_{i-1}\le u<u_i\\ 0 & \text{else}\end{cases}$$

B-spline curves (5)

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_6^3(u) + \mathbf{d}_7 N_7^3(u)$$



$$\mathbf{r}(u) = \mathbf{d}_0 \cancel{N_0^3}(u) + \mathbf{d}_1 \cancel{N_1^3}(u) + \mathbf{d}_2 \cancel{N_2^3}(u) + \mathbf{d}_4 \cancel{N_4^3}(u) + \mathbf{d}_4 N_4^3(u) + \mathbf{d}_5 N_5^3(u) + \mathbf{d}_6 N_6^3(u) + \mathbf{d}_7 N_7^3(u)$$



$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

Cubic Bezier Curve

Given: $\mathbf{b}_0$, $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, t

Find

$$\mathbf{r}(t) = \mathbf{b}_0 B_0^3(t) + \mathbf{b}_1 B_1^3(t) + \mathbf{b}_2 B_2^3(t) + \mathbf{b}_3 B_3^3(t)$$

Bernstein polynomial function

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i},$$

$$\binom{n}{i} = {}_nC_i = \begin{cases} \dfrac{n!}{i!(n-i)!} & \textbf{if } 0 \le i \le n \\ \mathbf{0} & \textbf{else} \end{cases}$$

☑ Cubic B-spline curves

- **Given:** $\mathbf{d}_i$, u_j
- **Find:** $\mathbf{r}(u)$ **(Points on curve at parameter** u**)**

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_{D-1} N_{D-1}^3(u)$$

$\mathbf{d}_i$: control points(de Boor points), i = 0, 1, … , D-1

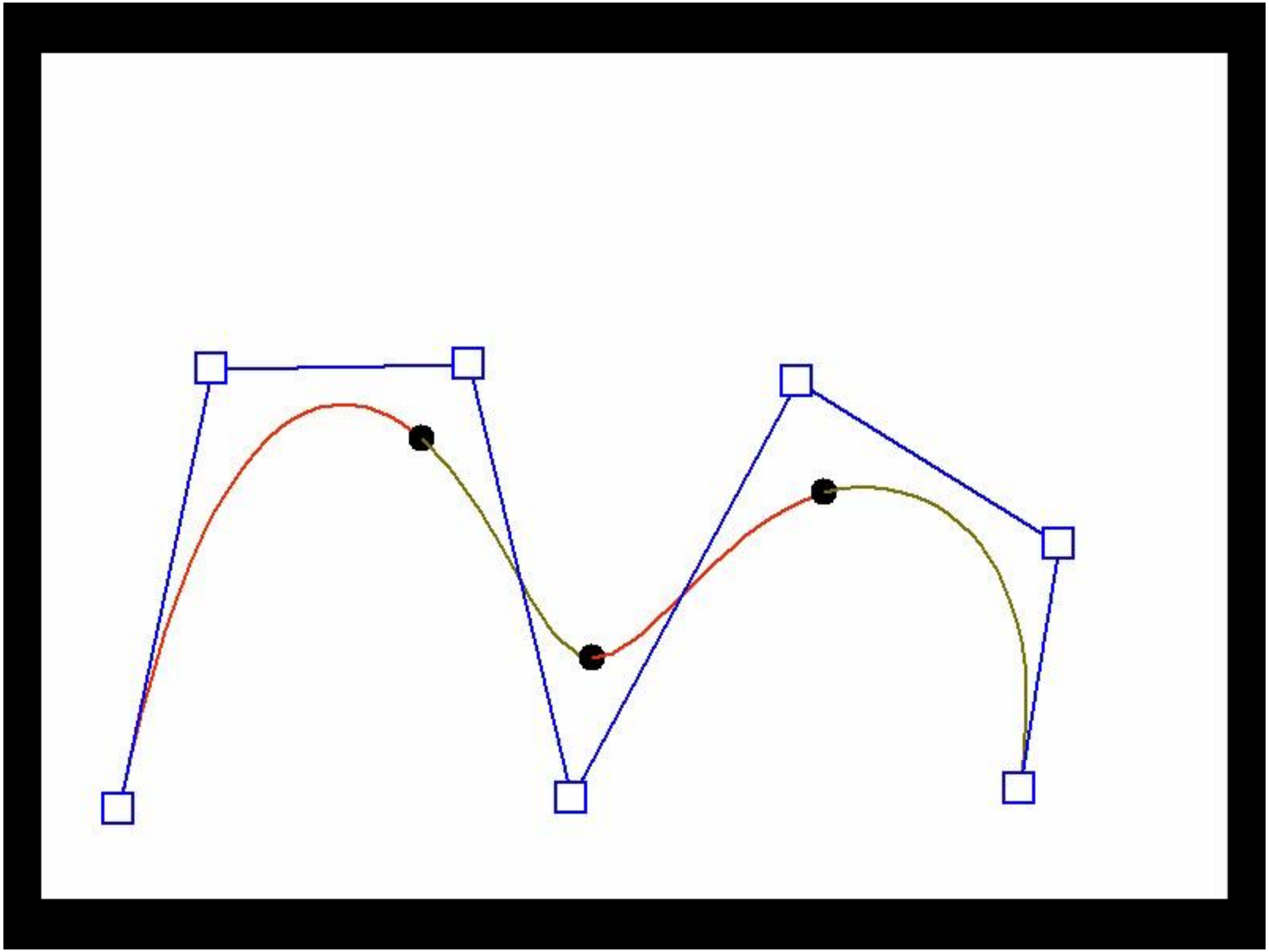
$N_i^n(u)$: B-spline basis function of degree n(=3)

u_j : knots, j = 0, 1, …, K-1

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

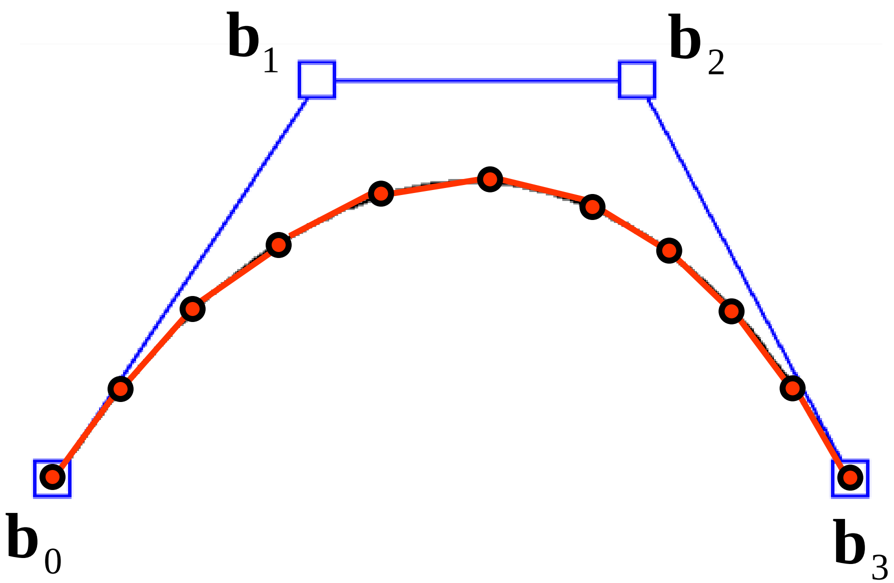
$$N_i^0(u) = \begin{cases} 1 & \textbf{if } u_{i-1} \le u < u_i \\ 0 & \textbf{else} \end{cases} \quad , \sum_{i=0}^{D-1} N_i^n(u) = 1$$

Computer Implementation of B-spline Curve



3.4 Programming for B-spline Curve class

Cubic B-spline 예시



$$\mathbf{r}(u) = \mathbf{d}_0 \boxed{N_0^3(u)} + \mathbf{d}_1 \boxed{N_1^3(u)} + \mathbf{d}_2 \boxed{N_2^3(u)} + \mathbf{d}_3 \boxed{N_3^3(u)}$$

1) Definition of B-spline Curve

- Degree
- Control Point

<u>Member Variables of B-spline Curve Class</u>
int n: degree of B-spline Curve
Vector* m_ControlPoint: Control Point
int m_nControlPoint: the number of Control Point

2) Calculation of B-spline Basis Function

(Cox-de Boor Recurrence Formula)

$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \; u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

3) B-spline curve construction

- Divide the parameter u(u_{min}~u_{max}) into n equal parts
- Find the points on the curve at the each divided parameter
- Represent curve by connecting points with straight lines

Sample code of Cubic B-spline Curve (1)

Member Variables of B-spline Curve Class
int m_nDegree : degree of B-spline Curve
Vector* m_ControlPoint: Control Point
int m_nControlPoint: the number of Control Point

```
#include "vector.h“

class CubicBsplineCurve {
public:
    Vector* m_ControlPoint;   int m_nControlPoint;
    double* m_Knot; int m_nKnot;
    int m_nDegree;

    CubicBsplineCurve();
    ~CubicBsplineCurve();

    void SetControlPoint(Vector* pControlPoint, int nControlPoint);
    void SetKnot(double* pKnot, int nKnot);
    Vector CalcPoint(double u);
    double N(int d, int i, double u);                    // B-spline basis function
};
```

Sample code of Cubic B-spline Curve (2)

```
CubicBsplineCurve::CubicBsplineCurve () {
  m_ControlPoint = 0;        m_Knot = 0;
  m_nControlPoint = 0;       m_nKnot = 0;       int m_nDegree =3;
}
CubicBsplineCurve::~CubicBsplineCurve () {
  if(m_ControlPoint)  delete[] m_ControlPoint;
  if(m_Knot)  delete[] m_Knot;
}
void CubicBsplineCurve::SetControlPoint(Vector* pControlPoint, int nControlPoint) {
  m_ControlPoint = new Vector[nControlPoint];
  for(int i=0; i < nControlPoint; i++) {
    m_ControlPoint[i] = pControlPoint[i];
  }
}
void CubicBsplineCurve::SetKnot(double* pKnot, int nKnot){
  m_Knot = new double[nKnot];
  for(int i=0; i < nKnot; i++) {
    m_Knot[i] = pKnot[i];
  }
}
```

Sample code of Cubic B-spline Curve (3)

```
Vector CubicBsplineCurve::CalcPoint(double u)
{
    Vector PointOnCurve(0,0,0);
    if ( t < m_Knot[0] || t > m_Knot[m_nKnot-1] ) {
        return PointOnCurve;
    }
    for(int i = 0; i < m_nControlPoint; i++){
        PointOnCurve = PointOnCurve + m_ControlPoint[i] * N(m_nDegree, i, u);
    }
    return PointOnCurve;
}
```

Calculate points on the curve at parameter *u*

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \mathbf{d}_3 N_3^3(u)$$

Sample code of Cubic B-spline Curve (4)

```
double CubicBsplineCurve:: N(int d, int i,
    // Find Span k
    // U i-1 <= U < U i  → k = i

    if( d == 0 ) {
        // return 0 or 1;
    } else {
        // return Cox de-Boor recurrence f
    }
}
```

Calculation of B-spline Basis Function

(Cox-de Boor Recurrence Formula)

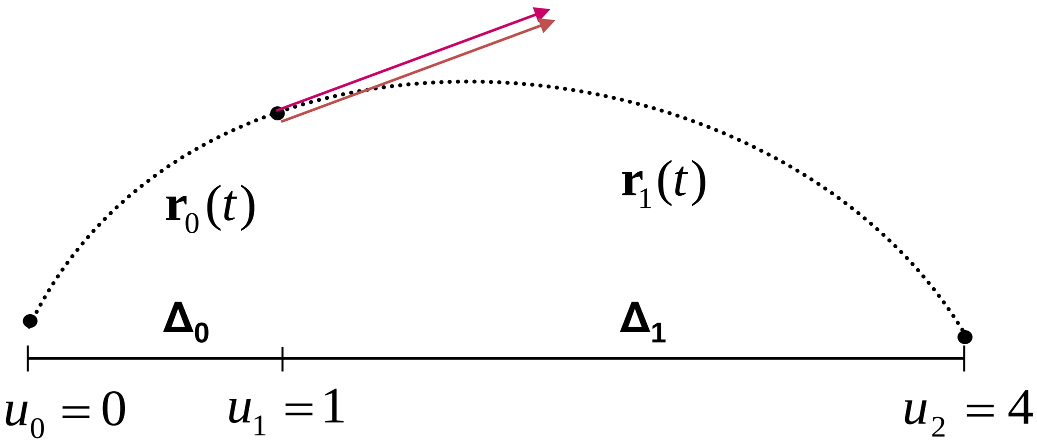
$$N_i^n(u) = \frac{u - u_{i-1}}{u_{i+n-1} - u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n} - u}{u_{i+n} - u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } \ u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

3.3 C^1 and C^2 Continuity Condition

1) 1^{st} Derivatives of Cubic Bezier Curves at Junction point
2) C^1 continuity condition of composite curves
3) 2^{nd} Derivatives of Cubic Bezier Curves
4) C^2 continuity condition of composite curves

1) 1st Derivatives of Cubic Bezier Curves at Junction point



$t = \dfrac{u-u_i}{u_{i+1}-u_i} = \dfrac{u-u_i}{\Delta_i}$ **t is local parameter in [0,1] section**

$$\frac{d\mathbf{r}(u(t))}{du} = \frac{d\mathbf{r}_i(t)}{dt}\frac{dt}{du} = \frac{1}{\Delta_i}\frac{d\mathbf{r}_i(t)}{dt}$$

The derivative value of $\dfrac{d\mathbf{r}\,(u)}{du}$ **in** $u_0 \le u \le u_1$

$t = \dfrac{u-u_0}{u_1-u_0} = \dfrac{u-u_0}{\Delta_0}$ **t is local parameter in [0,1] section**

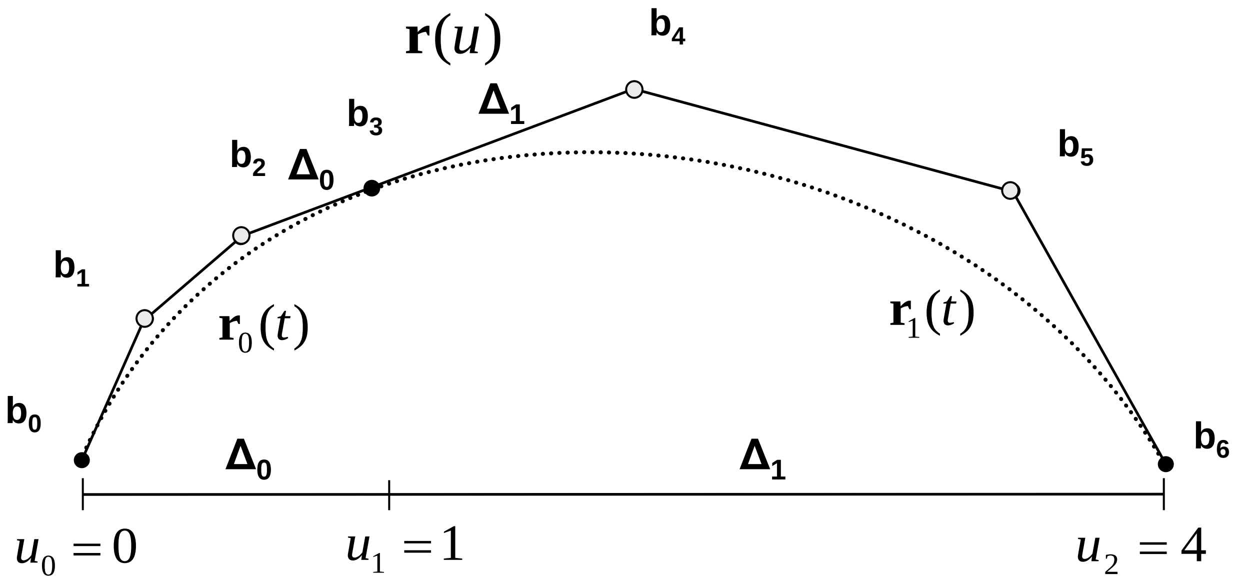
$$\frac{d\mathbf{r}\,(u)}{du} = \frac{d\,\mathbf{r}_0(u(t))}{dt}\frac{dt}{du} = \frac{1}{\Delta_0}\frac{d\,\mathbf{r}_0(t)}{dt}$$

The derivative value of $\dfrac{d\mathbf{r}\,(u)}{du}$ **in** $u_1 \le u \le u_2$

$t = \dfrac{u-u_1}{u_2-u_1} = \dfrac{u-u_1}{\Delta_1}$ **t is local parameter in [0,1] section**

$$\frac{d\mathbf{r}\,(u)}{du} = \frac{d\,\mathbf{r}_1(t)}{dt}\frac{dt}{du} = \frac{1}{\Delta_1}\frac{d\,\mathbf{r}_1(t)}{dt}$$

2) C^1 continuity condition of composite curves



r(*u* = *u*$_1$)= r$_0$(*t* = 1) = r$_1$(*t* = 0) . **C^1 condition must satisfy on connecting point**

$$\left.\frac{d\mathbf{r}(u)}{du}\right|_{u_1=1} = \frac{1}{\Delta_0}\left.\frac{d\mathbf{r}_0(t)}{dt}\right|_{t=1} = \frac{1}{\Delta_0}3(\mathbf{b}_3-\mathbf{b}_2)$$

$$= \frac{1}{\Delta_1}\left.\frac{d\mathbf{r}_1(t)}{dt}\right|_{t=0} = \frac{1}{\Delta_1}\cdot 3(\mathbf{b}_4-\mathbf{b}_3)$$

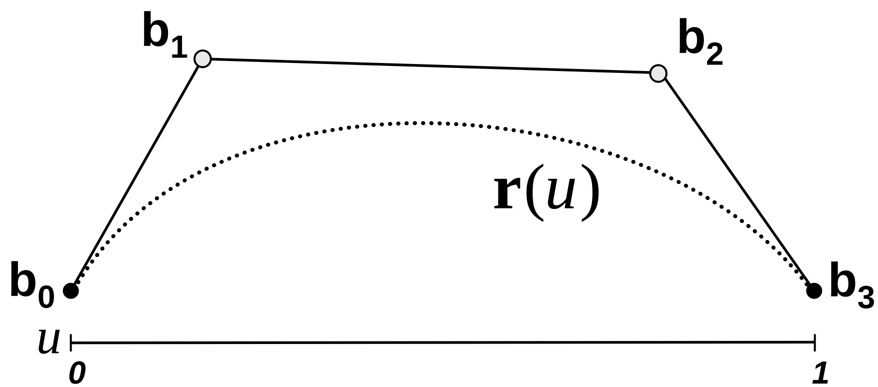
$$(\mathbf{b}_3-\mathbf{b}_2):(\mathbf{b}_4-\mathbf{b}_3) = \Delta_0 : \Delta_1$$

$$\Downarrow$$

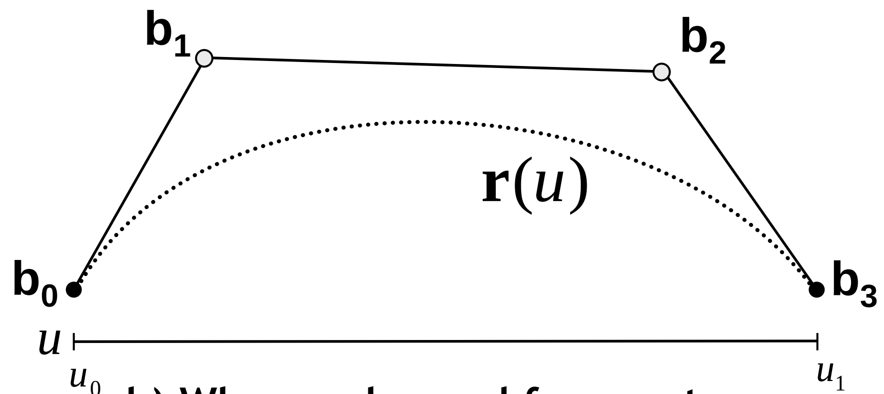
$$\mathbf{b}_3 = \frac{\Delta_1}{\Delta}\mathbf{b}_2 + \frac{\Delta_0}{\Delta}\mathbf{b}_4$$

Suppose the parameter u is time, then, 1st derivative is velocity of the point which passing through the curve.
If 1st derivative of the curve is continuous on the connecting point b$_3$, , then the velocity must be continuous.
Accordingly, if the time interval is changed from $\triangle_0$ to $\triangle_1$. the distance must be changed proportionally , because the velocity is continuous on the connecting point.

3) 2nd Derivatives of Cubic Bezier Curves



a) When u changed from 0 to 1



b) When u changed form u_0 to u_1

The second derivative of n^{th}-degree Bezier curve

$$\frac{d^2\mathbf{r}(u)}{du^2} = n(n-1)\sum_{i=0}^{n-2}(\mathbf{b}_{i+2} - 2\mathbf{b}_{i+1} + \mathbf{b}_i)B_i^{n-2}$$

The second derivative of 3^{rd}-degree Bezier curve

$$\frac{d^2\mathbf{r}(u)}{du^2} = 3(3-1)\sum_{i=0}^{1}(\mathbf{b}_{i+2} - 2\mathbf{b}_{i+1} + \mathbf{b}_i)B_i^1(u)$$

When $u = 1$

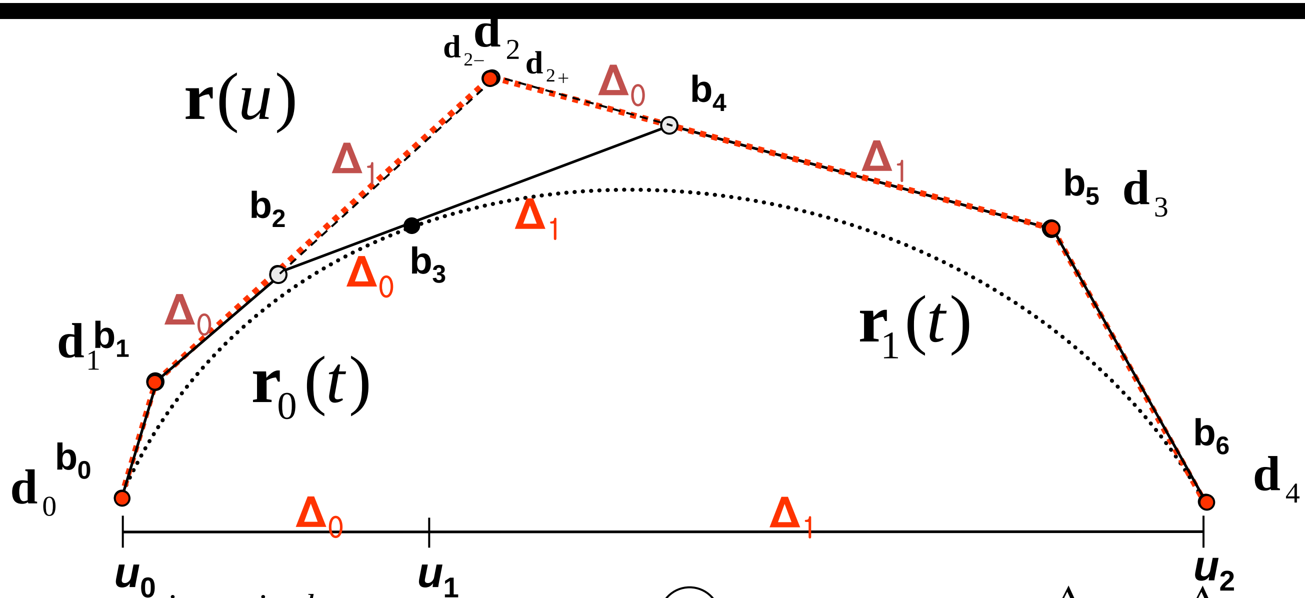
$$\frac{d^2\mathbf{r}(1)}{du^2} = 3(3-1)(\mathbf{b}_3 - 2\mathbf{b}_2 + \mathbf{b}_1)$$

$$\frac{d^2\mathbf{r}(u(t))}{du^2} = \frac{1}{(\Delta)^2}\frac{d^2\mathbf{r}(t)}{dt^2} \qquad (\Delta = u_1 - u_0)$$

When $u = u_1$

$$\frac{d^2\mathbf{r}(u_1)}{du^2} = \frac{1}{(\Delta)^2}\frac{d^2\mathbf{r}(1)}{dt^2} = \frac{1}{(\Delta)^2}3(3-1)(\mathbf{b}_3 - 2\mathbf{b}_2 + \mathbf{b}_1)$$

4) C² continuity condition of composite curves



① C^2 condition on connecting point b_3

$$\frac{d^2\mathbf{r}(u_{1-})}{du^2}=\frac{1}{(\Delta_0)^2}\frac{d^2\mathbf{r}_0(1)}{dt^2}=\frac{1}{(\Delta_0)^2}3(3-1)(\mathbf{b}_3-2\mathbf{b}_2+\mathbf{b}_1)$$

$$\frac{d^2\mathbf{r}(u_{1+})}{du^2}=\frac{1}{(\Delta_1)^2}\frac{d^2\mathbf{r}_1(0)}{dt^2}=\frac{1}{(\Delta_1)^2}3(3-1)(\mathbf{b}_5-2\mathbf{b}_4+\mathbf{b}_3)$$

② since $\dfrac{d^2\mathbf{r}(u_{1-})}{du^2}=\dfrac{d^2\mathbf{r}(u_{1+})}{du^2}$

$$\frac{6}{(\Delta_0)^2}(\mathbf{b}_3-2\mathbf{b}_2+\mathbf{b}_1)=\frac{6}{(\Delta_0)^2}(\mathbf{b}_5-2\mathbf{b}_4+\mathbf{b}_3)$$

③ And substitute C¹ condition($\mathbf{b}_3=\dfrac{\Delta_1}{\Delta}\mathbf{b}_2+\dfrac{\Delta_0}{\Delta}\mathbf{b}_4$)

To summarize

$$\Rightarrow -\frac{\Delta_1}{\Delta_0}\mathbf{b}_1+\frac{\Delta}{\Delta_0}\mathbf{b}_2=\frac{\Delta}{\Delta_1}\mathbf{b}_4-\frac{\Delta_0}{\Delta_1}\mathbf{b}_5$$

④ If LHS is $\mathbf{d}_{2-}=-\dfrac{\Delta_1}{\Delta_0}\mathbf{b}_1+\dfrac{\Delta}{\Delta_0}\mathbf{b}_2$

$$\mathbf{b}_2=\frac{\Delta_1}{\Delta}\mathbf{b}_1+\frac{\Delta_0}{\Delta}\mathbf{d}_{2-}$$

⑤ If RHS is $\mathbf{d}_{2+}=\dfrac{\Delta}{\Delta_1}\mathbf{b}_4-\dfrac{\Delta_0}{\Delta_1}\mathbf{b}_5$

$$\mathbf{b}_4=\frac{\Delta_1}{\Delta}\mathbf{d}_{2+}+\frac{\Delta_1}{\Delta}\mathbf{b}_5$$

⑥ Thus, if the points $\mathbf{d}_{2-}=\mathbf{d}_{2+}=\mathbf{d}_2$ are present, C² condition is satisfied

⑦ C² condition on connecting point

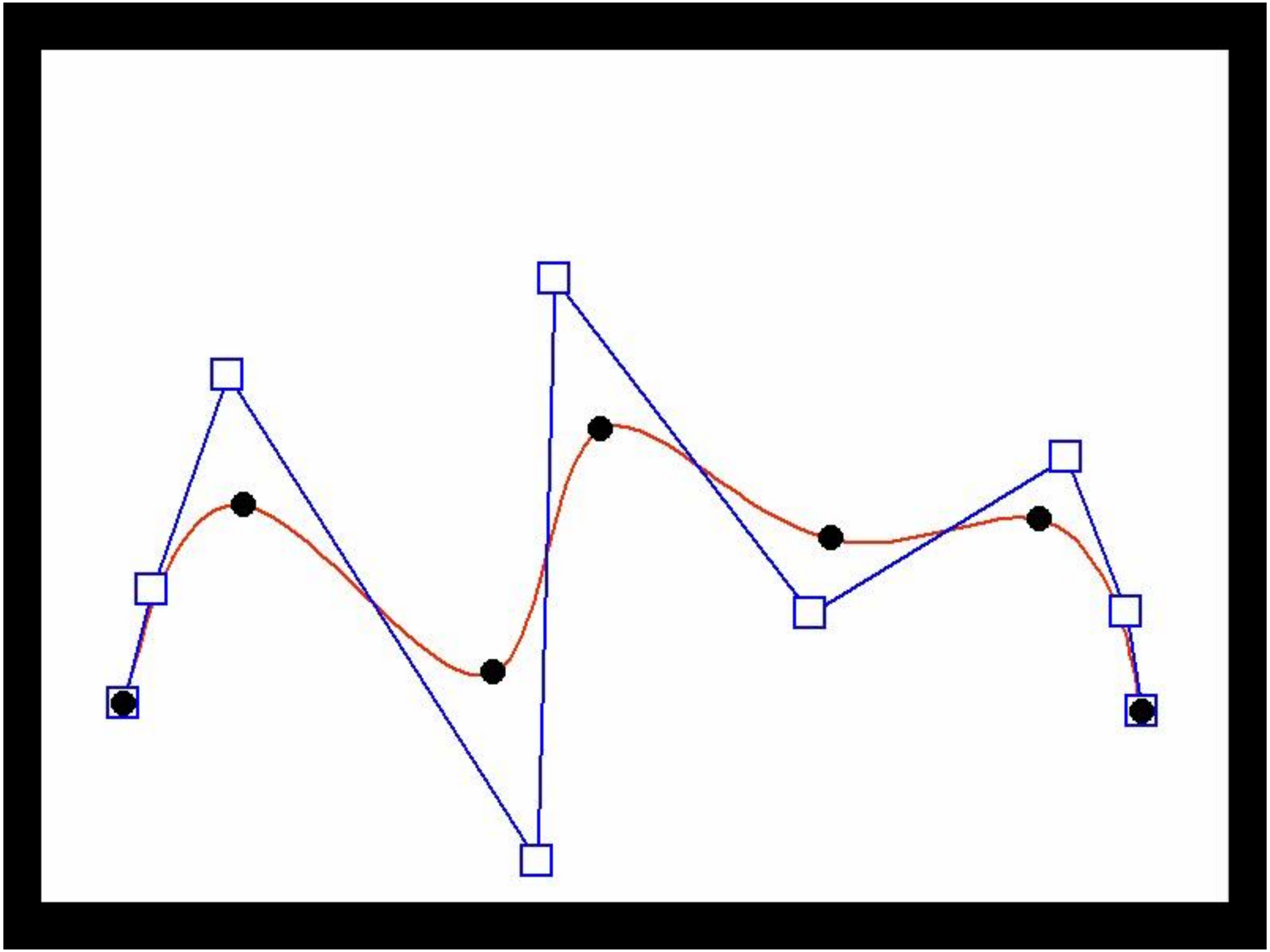
$$-\frac{\Delta_1}{\Delta_0}\mathbf{b}_1+\frac{\Delta}{\Delta_0}\mathbf{b}_2=\frac{\Delta}{\Delta_1}\mathbf{b}_4-\frac{\Delta_0}{\Delta_1}\mathbf{b}_5$$

$$ratio(\mathbf{b}_1,\mathbf{b}_2,d_2)=ratio(d_2,\mathbf{b}_4,\mathbf{b}_5)=\frac{\Delta_0}{\Delta_1}$$

3.4 B-spline Curve Interpolation

1) Determine the number of curve segments & knots values
2) Problem definition of B-spline curve interpolation
3) Determine Bezier end control points by end tangent vectors
4) Determine Bezier control points satisfying C1 continuity condition
5) Determine B-spline control points satisfying C2 continuity condition
6) Calculate B-spline control points by using tri-diagonal matrix solution
7) Bessel end condition
8) Sample code of cubic B-spline curve interpolation

Example of B-spline Interpolation



1) Determine the number of Bezier curve segments & knot values

☑ **Given: fitting points $\mathbf{P}_i$ and corresponding parameter t_i**

1. Determine the number of Bezier curve segment to be (number of fitting point -1)

2. We can determine knots to be same as the parameters t_i

3. How can we determine the B-spline control points ?

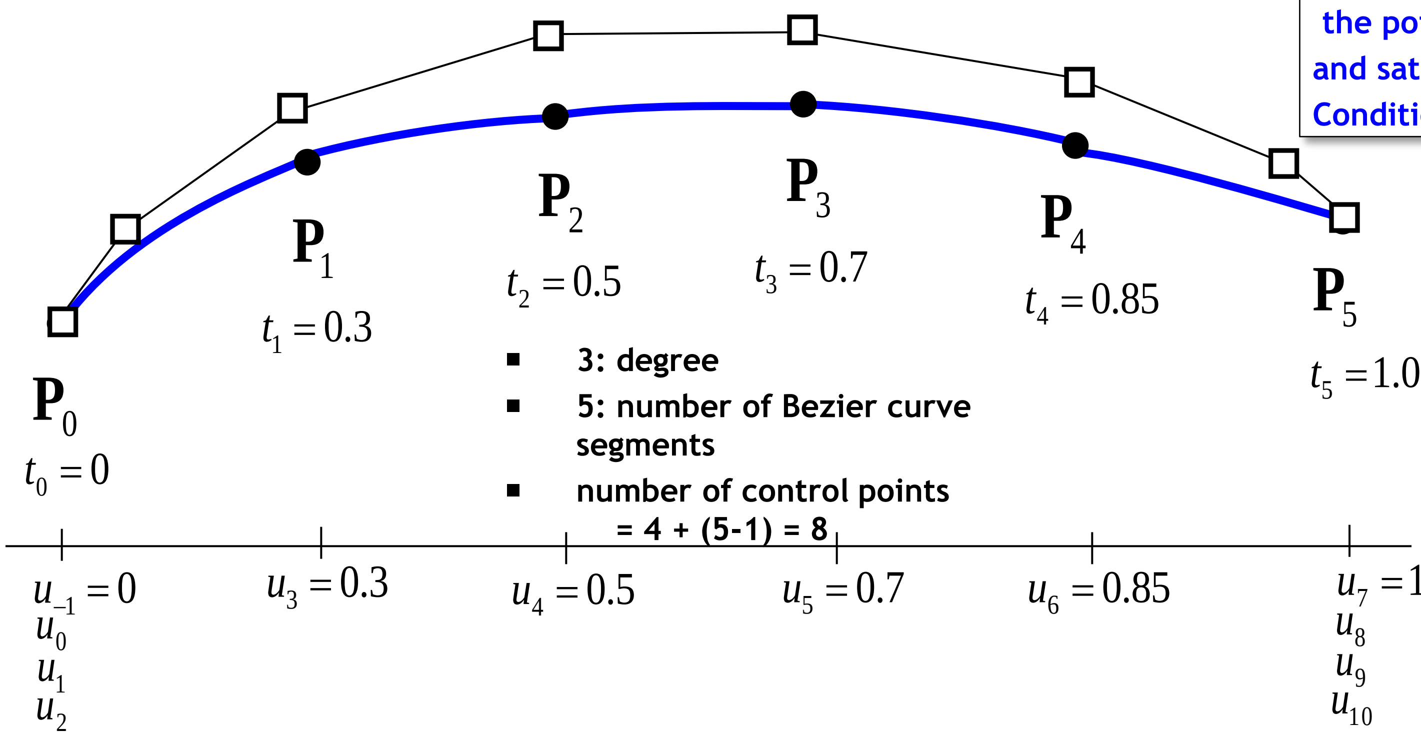
Given:

- Points p_i on the curve
- Knots u_j of the given points on the curve
- Tangent vectors t_0, t_1 at both ends

Find:

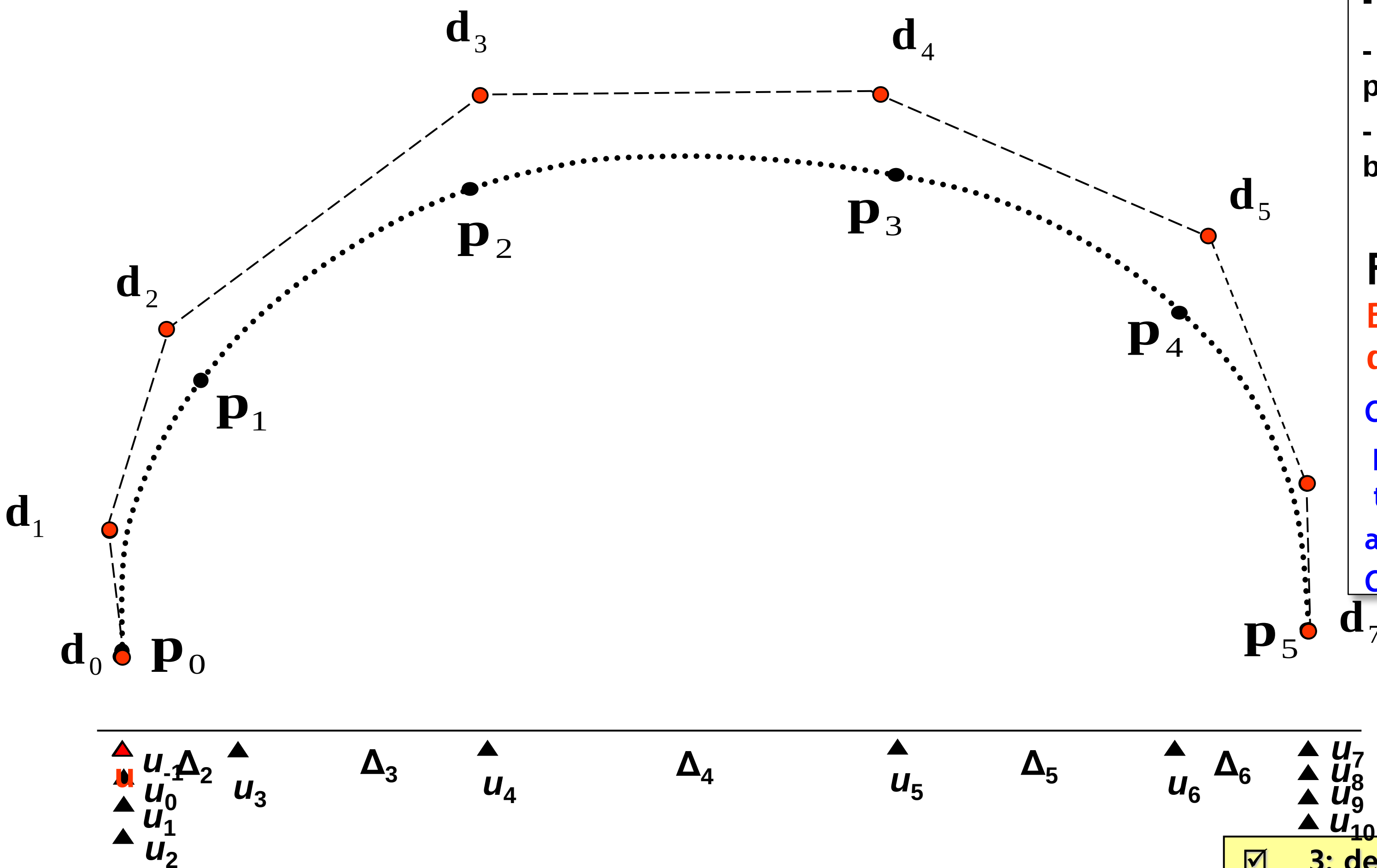
B-spline control point d_i

Cubic B-spline curve r(u) passing through the points p_i on the curve and satisfying C^2 continuity Condition:



- 3: degree
- 5: number of Bezier curve segments
- number of control points = 4 + (5-1) = 8

2) Problem Statement of cubic B-spline curve interpolation



Given:

- Points p_i on the curve
- Knots u_j of the given points on the curve
- Tangent vectors t_0, t_1 at both ends

Find:

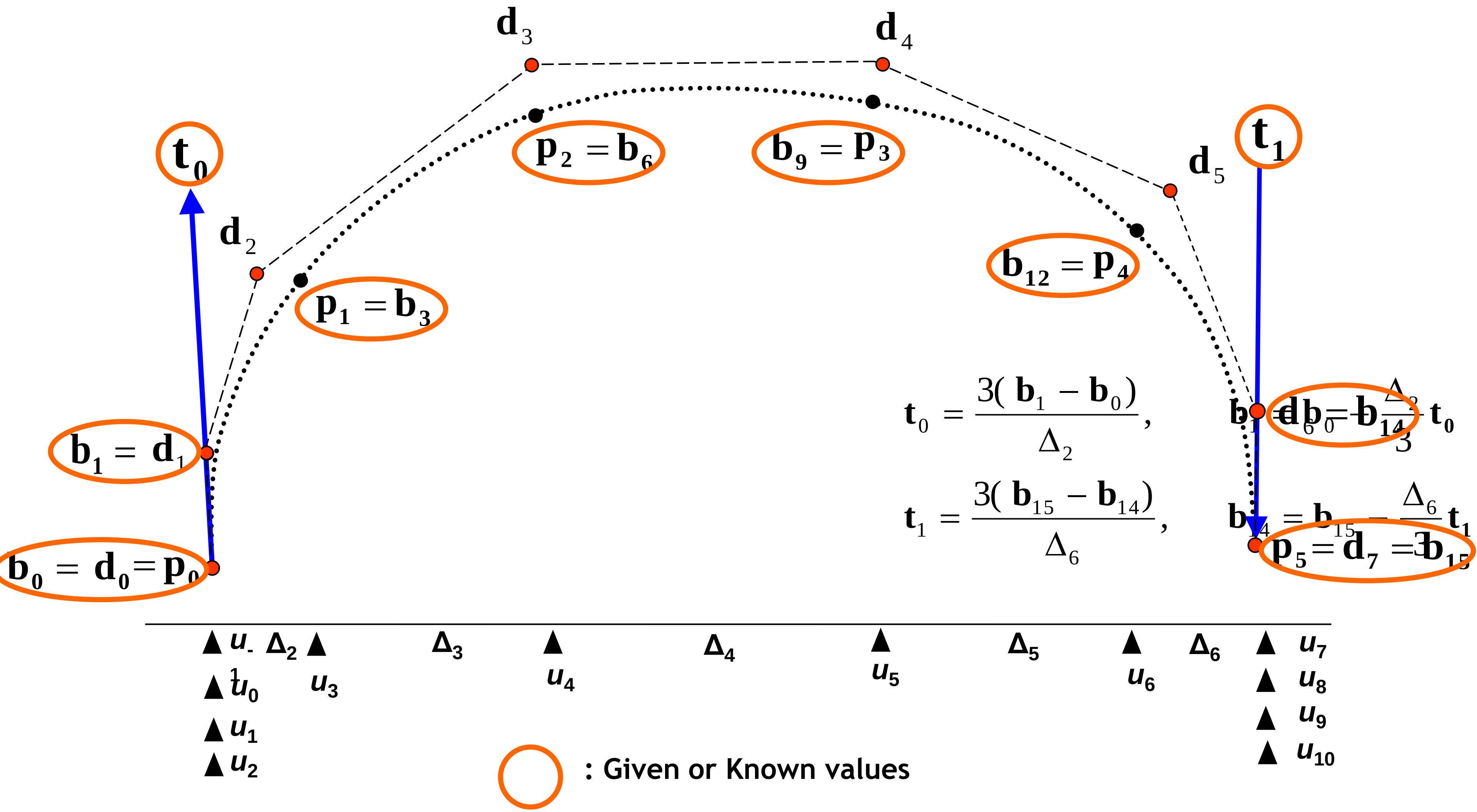
B-spline control point d_i

Cubic B-spline curve r(u) passing through the points p_i on the curve and satisfying C^2 continuity Condition:

Assumption of basis function:
Each curve segment is 3rd -degree Bezier curve
Constraints :
satisfying C0, C^1 ,C^2 continuity conditions at knots

- ☑ 3: degree
- ☑ 5: number of Bezier curve segments
- ☑ number of knot = (5-1) + 2(3+1)
- ☑ number of control points = 4 + (5-1) = (3+1) + (5-1)

3) Determine Bezier end control points by using end tangent vectors

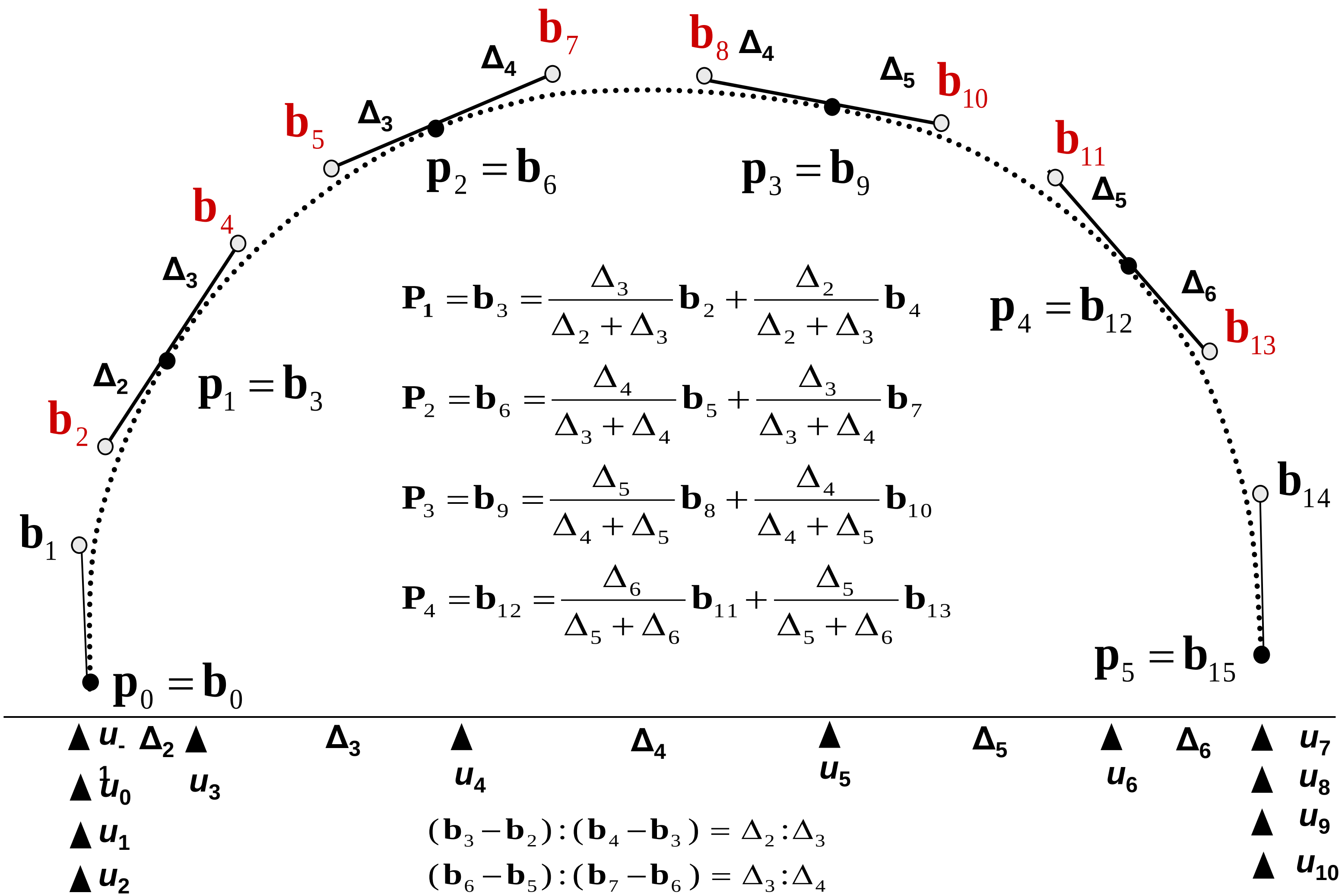


$$\mathbf{t}_0 = \frac{3(\mathbf{b}_1 - \mathbf{b}_0)}{\Delta_2}, \qquad \mathbf{b}_1 = \mathbf{b}_0 + \frac{\Delta_2}{3}\mathbf{t}_0$$

$$\mathbf{t}_1 = \frac{3(\mathbf{b}_{15} - \mathbf{b}_{14})}{\Delta_6}, \qquad \mathbf{b}_{14} = \mathbf{b}_{15} - \frac{\Delta_6}{3}\mathbf{t}_1$$

: Given or Known values

4) Determine **Bezier control points** satisfying C0, C^1 continuity conditions



$$\mathbf{P}_1 = \mathbf{b}_3 = \frac{\Delta_3}{\Delta_2+\Delta_3}\mathbf{b}_2 + \frac{\Delta_2}{\Delta_2+\Delta_3}\mathbf{b}_4$$

$$\mathbf{P}_2 = \mathbf{b}_6 = \frac{\Delta_4}{\Delta_3+\Delta_4}\mathbf{b}_5 + \frac{\Delta_3}{\Delta_3+\Delta_4}\mathbf{b}_7$$

$$\mathbf{P}_3 = \mathbf{b}_9 = \frac{\Delta_5}{\Delta_4+\Delta_5}\mathbf{b}_8 + \frac{\Delta_4}{\Delta_4+\Delta_5}\mathbf{b}_{10}$$

$$\mathbf{P}_4 = \mathbf{b}_{12} = \frac{\Delta_6}{\Delta_5+\Delta_6}\mathbf{b}_{11} + \frac{\Delta_5}{\Delta_5+\Delta_6}\mathbf{b}_{13}$$

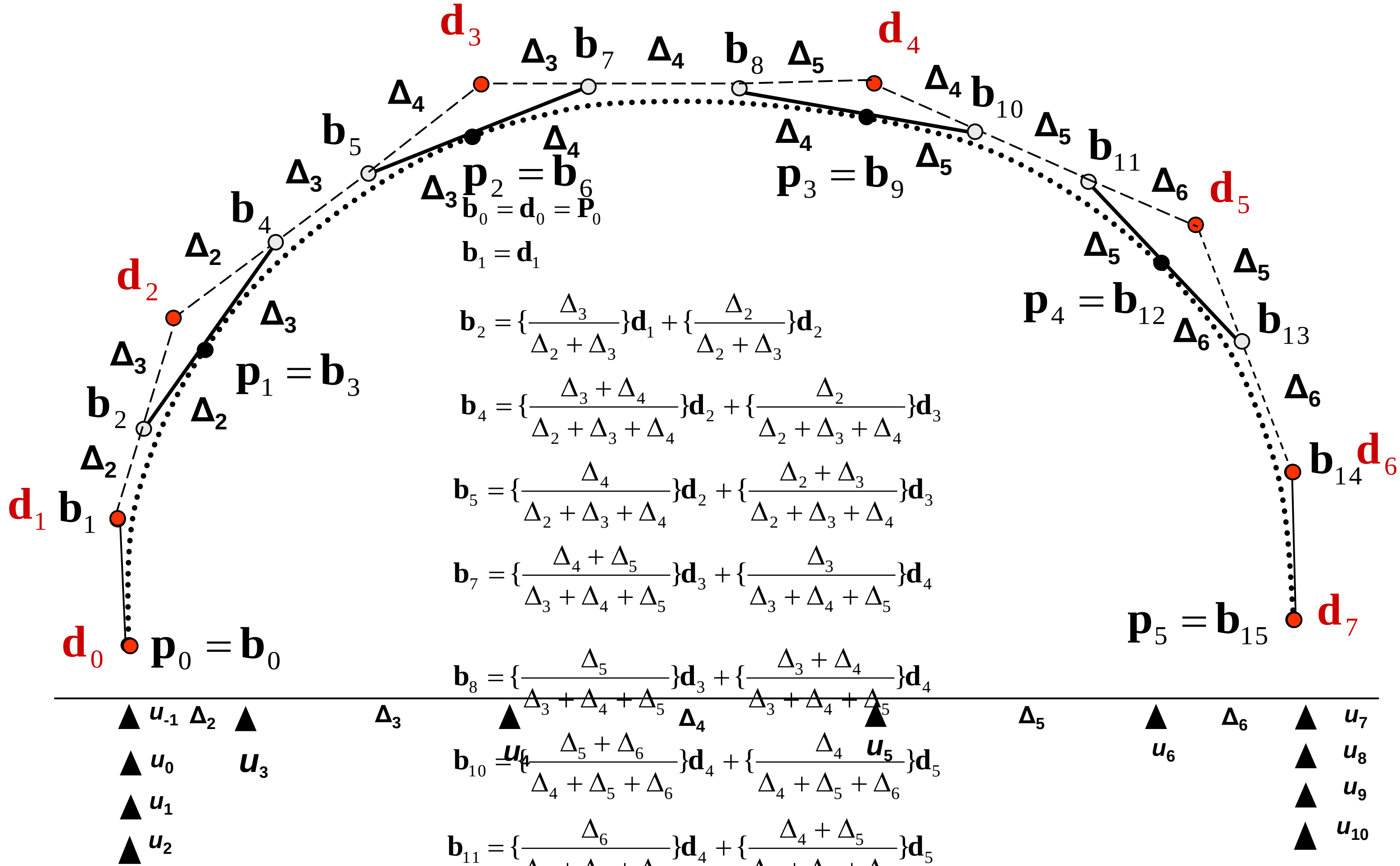
$$(\mathbf{b}_3-\mathbf{b}_2):(\mathbf{b}_4-\mathbf{b}_3) = \Delta_2:\Delta_3$$

$$(\mathbf{b}_6-\mathbf{b}_5):(\mathbf{b}_7-\mathbf{b}_6) = \Delta_3:\Delta_4$$

$$(\mathbf{b}_9-\mathbf{b}_8):(\mathbf{b}_{10}-\mathbf{b}_9) = \Delta_4:\Delta_5$$

$$(\mathbf{b}_{12}-\mathbf{b}_{11}):(\mathbf{b}_{13}-\mathbf{b}_{12}) = \Delta_5:\Delta_6$$

5) Determine B-spline control points satisfying C^2 continuity condition



$$\mathbf{b}_0 = \mathbf{d}_0 = \mathbf{P}_0$$

$$\mathbf{b}_1 = \mathbf{d}_1$$

$$\mathbf{b}_2 = \{\frac{\Delta_3}{\Delta_2+\Delta_3}\}\mathbf{d}_1 + \{\frac{\Delta_2}{\Delta_2+\Delta_3}\}\mathbf{d}_2$$

$$\mathbf{b}_4 = \{\frac{\Delta_3+\Delta_4}{\Delta_2+\Delta_3+\Delta_4}\}\mathbf{d}_2 + \{\frac{\Delta_2}{\Delta_2+\Delta_3+\Delta_4}\}\mathbf{d}_3$$

$$\mathbf{b}_5 = \{\frac{\Delta_4}{\Delta_2+\Delta_3+\Delta_4}\}\mathbf{d}_2 + \{\frac{\Delta_2+\Delta_3}{\Delta_2+\Delta_3+\Delta_4}\}\mathbf{d}_3$$

$$\mathbf{b}_7 = \{\frac{\Delta_4+\Delta_5}{\Delta_3+\Delta_4+\Delta_5}\}\mathbf{d}_3 + \{\frac{\Delta_3}{\Delta_3+\Delta_4+\Delta_5}\}\mathbf{d}_4$$

$$\mathbf{b}_8 = \{\frac{\Delta_5}{\Delta_3+\Delta_4+\Delta_5}\}\mathbf{d}_3 + \{\frac{\Delta_3+\Delta_4}{\Delta_3+\Delta_4+\Delta_5}\}\mathbf{d}_4$$

$$\mathbf{b}_{10} = \{\frac{\Delta_5+\Delta_6}{\Delta_4+\Delta_5+\Delta_6}\}\mathbf{d}_4 + \{\frac{\Delta_4}{\Delta_4+\Delta_5+\Delta_6}\}\mathbf{d}_5$$

$$\mathbf{b}_{11} = \{\frac{\Delta_6}{\Delta_4+\Delta_5+\Delta_6}\}\mathbf{d}_4 + \{\frac{\Delta_4+\Delta_5}{\Delta_4+\Delta_5+\Delta_6}\}\mathbf{d}_5$$

$$\mathbf{b}_{13} = \{\frac{\Delta_6}{\Delta_5+\Delta_6}\}\mathbf{d}_5 + \{\frac{\Delta_5}{\Delta_5+\Delta_6}\}\mathbf{d}_6$$

$$\mathbf{b}_{14} = \mathbf{d}_6$$

$$\mathbf{b}_{15} = \mathbf{d}_7 = \mathbf{p}_5$$

Construct equations for d_i satisfying C0, C^1, C^2 conditions

<u>C0, C^1 Conditions</u>

$$\mathbf{P}_1 = \mathbf{b}_3 = \frac{\Delta_3}{\Delta_2 + \Delta_3}\mathbf{b}_2 + \frac{\Delta_2}{\Delta_2 + \Delta_3}\mathbf{b}_4$$

$$\mathbf{P}_2 = \mathbf{b}_6 = \frac{\Delta_4}{\Delta_3 + \Delta_4}\mathbf{b}_5 + \frac{\Delta_3}{\Delta_3 + \Delta_4}\mathbf{b}_7$$

$$\mathbf{P}_3 = \mathbf{b}_9 = \frac{\Delta_5}{\Delta_4 + \Delta_5}\mathbf{b}_8 + \frac{\Delta_4}{\Delta_4 + \Delta_5}\mathbf{b}_{10}$$

$$\mathbf{P}_4 = \mathbf{b}_{12} = \frac{\Delta_6}{\Delta_5 + \Delta_6}\mathbf{b}_{11} + \frac{\Delta_5}{\Delta_5 + \Delta_6}\mathbf{b}_{13}$$

<u>C^2 Condition</u>

$$\mathbf{b}_0 = \mathbf{d}_0 = \mathbf{P}_0$$

$$\mathbf{b}_1 = \mathbf{d}_1$$

$$\mathbf{b}_2 = \{\frac{\Delta_3}{\Delta_2 + \Delta_3}\}\mathbf{d}_1 + \{\frac{\Delta_2}{\Delta_2 + \Delta_3}\}\mathbf{d}_2$$

$$\mathbf{b}_4 = \{\frac{\Delta_3 + \Delta_4}{\Delta_2 + \Delta_3 + \Delta_4}\}\mathbf{d}_2 + \{\frac{\Delta_2}{\Delta_2 + \Delta_3 + \Delta_4}\}\mathbf{d}_3$$

$$\mathbf{b}_5 = \{\frac{\Delta_4}{\Delta_2 + \Delta_3 + \Delta_4}\}\mathbf{d}_2 + \{\frac{\Delta_2 + \Delta_3}{\Delta_2 + \Delta_3 + \Delta_4}\}\mathbf{d}_3$$

$$\mathbf{b}_7 = \{\frac{\Delta_4 + \Delta_5}{\Delta_3 + \Delta_4 + \Delta_5}\}\mathbf{d}_3 + \{\frac{\Delta_3}{\Delta_3 + \Delta_4 + \Delta_5}\}\mathbf{d}_4$$

$$\mathbf{b}_8 = \{\frac{\Delta_5}{\Delta_3 + \Delta_4 + \Delta_5}\}\mathbf{d}_3 + \{\frac{\Delta_3 + \Delta_4}{\Delta_3 + \Delta_4 + \Delta_5}\}\mathbf{d}_4$$

$$\mathbf{b}_{10} = \{\frac{\Delta_5 + \Delta_6}{\Delta_4 + \Delta_5 + \Delta_6}\}\mathbf{d}_4 + \{\frac{\Delta_4}{\Delta_4 + \Delta_5 + \Delta_6}\}\mathbf{d}_5$$

$$\mathbf{b}_{11} = \{\frac{\Delta_6}{\Delta_4 + \Delta_5 + \Delta_6}\}\mathbf{d}_4 + \{\frac{\Delta_4 + \Delta_5}{\Delta_4 + \Delta_5 + \Delta_6}\}\mathbf{d}_5$$

$$\mathbf{b}_{13} = \{\frac{\Delta_6}{\Delta_5 + \Delta_6}\}\mathbf{d}_5 + \{\frac{\Delta_5}{\Delta_5 + \Delta_6}\}\mathbf{d}_6$$

$$\mathbf{b}_{14} = \mathbf{d}_6$$

$$\mathbf{b}_{15} = \mathbf{d}_7 = \mathbf{p}_5$$

$$\mathbf{P}_1 = \frac{1}{(\Delta_2+\Delta_3)(\Delta_2+\Delta_3+\Delta_4)}[(\Delta_3)^2(\Delta_2+\Delta_3+\Delta_4)/(\Delta_2+\Delta_3)\mathbf{d}_1$$
$$+\{\Delta_2\Delta_3(\Delta_2+\Delta_3+\Delta_4)+\Delta_2(\Delta_2+\Delta_3)(\Delta_3+\Delta_4)\}/(\Delta_2+\Delta_3)\mathbf{d}_2+(\Delta_2)^2\mathbf{d}_3]$$
$$=\alpha_1\mathbf{d}_1+\beta_1\mathbf{d}_2+\gamma_1\mathbf{d}_3$$

$$\alpha_i = \frac{(\Delta_{i+2})^2}{(\Delta_i+\Delta_{i+1}+\Delta_{i+2})(\Delta_{i+1}+\Delta_{i+2})}$$
$$\beta_i = \{\frac{\Delta_{i+2}(\Delta_i+\Delta_{i+1})}{(\Delta_i+\Delta_{i+1}+\Delta_{i+2})}+\frac{\Delta_{i+1}(\Delta_{i+2}+\Delta_{i+3})}{(\Delta_{i+1}+\Delta_{i+2}+\Delta_{i+3})}\}/(\Delta_{i+1}+\Delta_{i+2})$$
$$\gamma_i = \frac{(\Delta_{i+1})^2}{(\Delta_{i+1}+\Delta_{i+2}+\Delta_{i+3})(\Delta_{i+1}+\Delta_{i+2})}$$

$$\mathbf{P}_2 = \frac{1}{(\Delta_3+\Delta_4)(\Delta_3+\Delta_4+\Delta_5)}[(\Delta_4)^2\mathbf{d}_2+\{\Delta_4(\Delta_2+\Delta_3)+$$
$$\Delta_3(\Delta_4+\Delta_5)\}\mathbf{d}_3+(\Delta_3)^2\mathbf{d}_4]=\alpha_2\mathbf{d}_2+\beta_2\mathbf{d}_3+\gamma_2\mathbf{d}_4$$

$$\mathbf{P}_3 = \frac{1}{(\Delta_4+\Delta_5)(\Delta_3+\Delta_4+\Delta_5)}[(\Delta_5)^2\mathbf{d}_3+\{\Delta_5(\Delta_3+\Delta_4)(\Delta_4+\Delta_5+\Delta_6)$$
$$+\Delta_4(\Delta_5+\Delta_6)(\Delta_3+\Delta_4+\Delta_5)\}/(\Delta_4+\Delta_5+\Delta_6)\mathbf{d}_4+(\Delta_4)^2(\Delta_3+\Delta_4+\Delta_5)$$
$$/(\Delta_4+\Delta_5+\Delta_6)\mathbf{d}_5]=\alpha_3\mathbf{d}_3+\beta_3\mathbf{d}_4+\gamma_3\mathbf{d}_5$$

$$\mathbf{P}_4 = \frac{1}{(\Delta_5+\Delta_6)(\Delta_4+\Delta_5+\Delta_6)}[(\Delta_6)^2\mathbf{d}_4+$$
$$\{\Delta_6(\Delta_4+\Delta_5)+\Delta_5\Delta_6(\Delta_4+\Delta_5+\Delta_6)\}\mathbf{d}_5$$
$$+(\Delta_5)^2(\Delta_4+\Delta_5+\Delta_6)\mathbf{d}_6]=\alpha_4\mathbf{d}_4+\beta_4\mathbf{d}_5+\gamma_4\mathbf{d}_6$$

Known **Unknown**

$$\begin{vmatrix}\mathbf{p}_0\\ \mathbf{t}_0\\ \mathbf{p}_1\\ \mathbf{p}_2\\ \mathbf{p}_3\\ \mathbf{p}_4\\ \mathbf{t}_1\\ \mathbf{p}_5\end{vmatrix} = \begin{vmatrix}1&0&0&0&0&0&0&0\\ \frac{-3}{\Delta_2}&\frac{3}{\Delta_2}&0&0&0&0&0&0\\ 0&\alpha_1&\beta_1&\gamma_1&0&0&0&0\\ 0&0&\alpha_2&\beta_2&\gamma_2&0&0&0\\ 0&0&0&\alpha_3&\beta_3&\gamma_3&0&0\\ 0&0&0&0&\alpha_4&\beta_4&\gamma_4&0\\ 0&0&0&0&0&0&\frac{-3}{\Delta_6}&\frac{3}{\Delta_6}\\ 0&0&0&0&0&0&0&1\end{vmatrix}\begin{vmatrix}\mathbf{d}_0\\ \mathbf{d}_1\\ \mathbf{d}_2\\ \mathbf{d}_3\\ \mathbf{d}_4\\ \mathbf{d}_5\\ \mathbf{d}_6\\ \mathbf{d}_7\end{vmatrix}$$

6) Calculate B-spline control points(d_i) by using Tri-diagonal matrix solution

$$\underbrace{\begin{bmatrix} \mathbf{p}_0 \\ \mathbf{t}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \mathbf{p}_4 \\ \mathbf{t}_1 \\ \mathbf{p}_5 \end{bmatrix}}_{=\mathbf{D}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-3}{\Delta_2} & \frac{3}{\Delta_2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_1 & \beta_1 & \gamma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_2 & \beta_2 & \gamma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_3 & \beta_3 & \gamma_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_4 & \beta_4 & \gamma_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{-3}{\Delta_6} & \frac{3}{\Delta_6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{d}_3 \\ \mathbf{d}_4 \\ \mathbf{d}_5 \\ \mathbf{d}_6 \\ \mathbf{d}_7 \end{bmatrix}}_{=\mathbf{X}}$$

Known Known Unknown

$$\mathbf{D} = \mathbf{A}\mathbf{X}$$

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{D}$$

Because Matrix A is tri-diagonal matrix, inverse matrix A^{-1} is easy to obtain.

Tridiagonal matrix

A tridiagonal matrix has nonzero elements only in the main diagonal, the first diagonal below the main diagonal, and the first diagonal above the main diagonal. ➡ Tri + Diagonal

$$
\begin{bmatrix}
b_0 & c_0 & 0 & & & & \\
a_1 & b_1 & c_1 & 0 & & & \\
0 & a_2 & b_2 & c_2 & 0 & & \\
 & & & \ddots & & & \\
 & & & & \ddots & & \\
 & & & 0 & a_{n-1} & b_{n-1} & c_{n-1} \\
 & & & & 0 & a_n & b_n
\end{bmatrix}
\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}
=
\begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}
$$

Diagonal elements

$$\mathbf{A} \qquad \mathbf{x} = \mathbf{d}$$

Determine x for known A and d

$\mathbf{Ax} = \mathbf{d}$

① $\mathbf{A} = \mathbf{LU}$

A is decomposed into multiplication of L and U

$\mathbf{LUx} = \mathbf{d}$

Substitute y for Ux

③ $\mathbf{Ux} = \mathbf{y}$

② $\mathbf{Ly} = \mathbf{d}$

① **A is decomposed into multiplication of L and U**

② **Solve the equation Ly = d for y**

③ **Solve the equation Ux = y for x, which is the solution of the equation of Ax=d**

① $\mathbf{A} = \mathbf{LU}$

$$\begin{bmatrix} b_0 & c_0 & 0 & & & & \\ a_1 & b_1 & c_1 & 0 & & & \\ 0 & a_2 & b_2 & c_2 & 0 & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & 0 & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & 0 & a_n & b_n \end{bmatrix} = \begin{bmatrix} \beta_0 & 0 & & & & & \\ \alpha_1 & \beta_1 & 0 & & & & \\ 0 & \alpha_2 & \beta_2 & 0 & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & 0 & \alpha_{n-1} & \beta_{n-1} & 0 \\ & & & & 0 & \alpha_n & \beta_n \end{bmatrix} \begin{bmatrix} 1 & \gamma_1 & 0 & & & & \\ 0 & 1 & \gamma_2 & 0 & & & \\ & 0 & 1 & \gamma_3 & 0 & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & 0 & 1 & \gamma_n \\ & & & & & 0 & 1 \end{bmatrix}$$

Diagonal elements

$$\mathbf{A} = \mathbf{L} \quad \mathbf{U}$$

$$\begin{array}{lll} & b_0 = \beta_0 & c_0 = \beta_0 \gamma_1 \\ a_1 = \alpha_1 & b_1 = \alpha_1 \gamma_1 + \beta_1 & c_1 = \beta_1 \gamma_2 \\ a_2 = \alpha_2 & b_2 = \alpha_2 \gamma_2 + \beta_2 & c_2 = \beta_2 \gamma_3 \\ \vdots & \vdots & \vdots \\ a_{n-1} = \alpha_{n-1} & b_{n-1} = \alpha_{n-1} \gamma_{n-1} + \beta_{n-1} & c_{n-1} = \beta_{n-1} \gamma_n \\ a_n = \alpha_n & b_n = \alpha_n \gamma_n + \beta_n & \end{array}$$

$$\alpha_i = a_i \qquad i = 1, \cdots, n$$

$$\gamma_{i+1} = \frac{c_i}{\beta_i} \qquad i = 0, \cdots, n-1$$

$$\beta_{i+1} = b_{i+1} - \alpha_{i+1} \gamma_{i+1} \qquad i = 0, \cdots, n-1$$

with $\beta_0 = b_0$

$\mathbf{Ax} = \mathbf{d}$?

① $\mathbf{A} = \mathbf{LU}$

$\mathbf{LUx} = \mathbf{d}$

③ $\mathbf{Ux} = \mathbf{y}$? !

② $\mathbf{Ly} = \mathbf{d}$!

② $\mathbf{L}\mathbf{y} = \mathbf{d}$

$$\underbrace{\begin{bmatrix} \beta_0 & 0 & & & & & \\ \alpha_1 & \beta_1 & 0 & & & & \\ 0 & \alpha_2 & \beta_2 & 0 & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & 0 & \alpha_{n-1} & \beta_{n-1} & 0 \\ & & & & 0 & \alpha_n & \beta_n \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}}_{\mathbf{d}}$$

Diagonal elements

$$\beta_0 y_0 = d_0$$
$$\alpha_1 y_0 + \beta_1 y_1 = d_1$$
$$\alpha_2 y_1 + \beta_2 y_2 = d_2$$
$$\vdots$$
$$\alpha_{n-1} y_{n-2} + \beta_{n-1} y_{n-1} = d_{n-1}$$
$$\alpha_n y_{n-1} + \beta_n y_n = d_n$$

Forward substitution

$$y_i = \frac{d_i - \alpha_i y_{i-1}}{\beta_i} \qquad i = 1, \cdots, n$$

$$\text{with } y_0 = \frac{d_0}{\beta_0}$$

$\mathbf{A}\mathbf{x} = \mathbf{d}$

① $\mathbf{A} = \mathbf{L}\mathbf{U}$

$\mathbf{L}\mathbf{U}\mathbf{x} = \mathbf{d}$

③ $\mathbf{U}\mathbf{x} = \mathbf{y}$

② $\mathbf{L}\mathbf{y} = \mathbf{d}$

③ $\mathbf{U}\mathbf{x} = \mathbf{y}$

$$\underbrace{\begin{bmatrix} 1 & \gamma_1 & 0 & & & & \\ 0 & 1 & \gamma_2 & 0 & & & \\ & 0 & 1 & \gamma_3 & 0 & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & 0 & 1 & \gamma_n \\ & & & & & 0 & 1 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}}_{\mathbf{y}}$$

$$x_0 + \gamma_1 x_1 = y_0$$

$$x_1 + \gamma_2 x_2 = y_1$$

$$x_2 + \gamma_3 x_3 = y_2$$

$$\vdots$$

$$x_{n-1} + \gamma_n x_n = y_{n-1}$$

$$x_n = y_n$$

Backward substitution

$$x_i = y_i - \gamma_{i+1} x_{i+1}$$

$$i = n-1, \cdots, 0$$

with $x_n = y_n$

$\mathbf{A}\mathbf{x} = \mathbf{d}$

① $\mathbf{A} = \mathbf{L}\mathbf{U}$

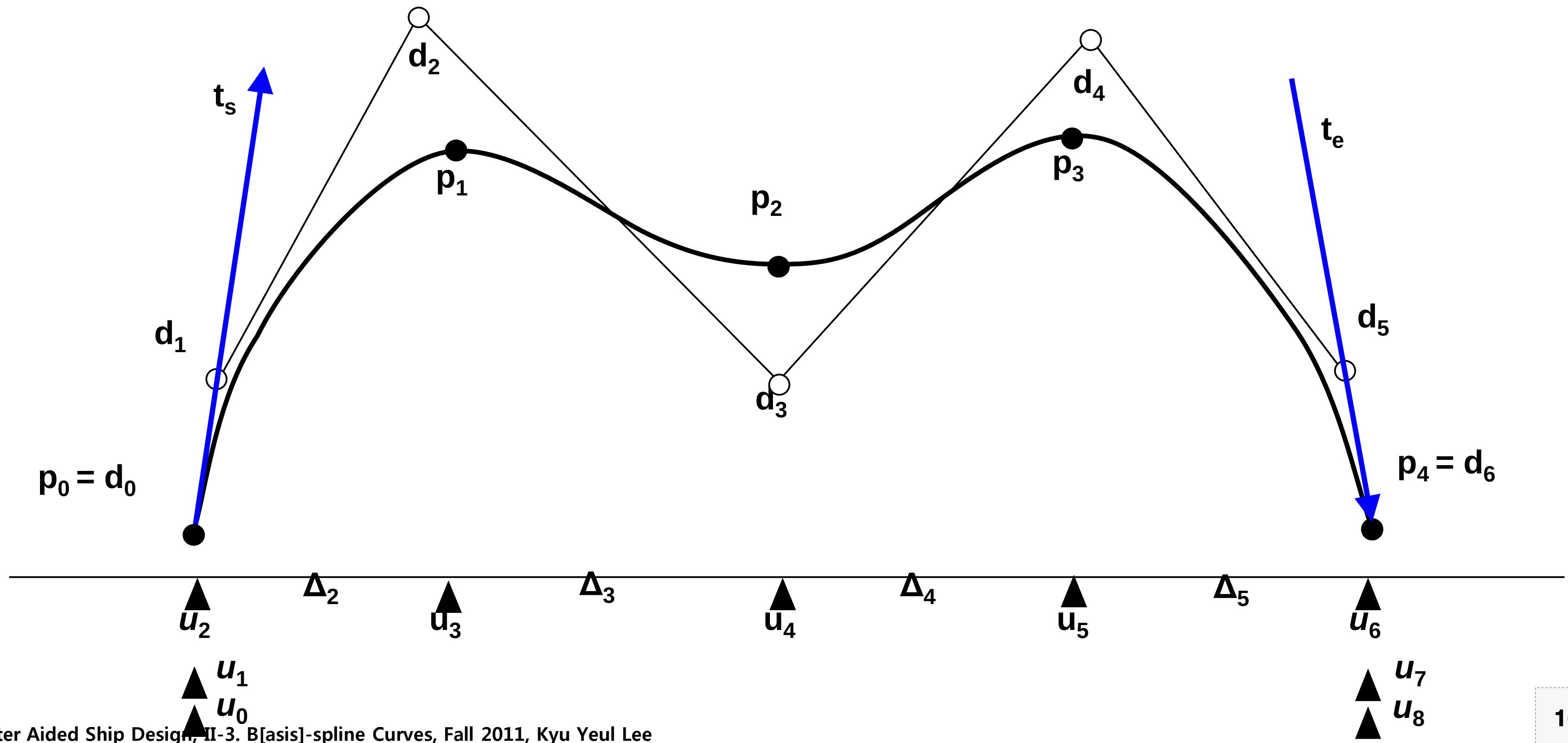
$\mathbf{L}\mathbf{U}\mathbf{x} = \mathbf{d}$

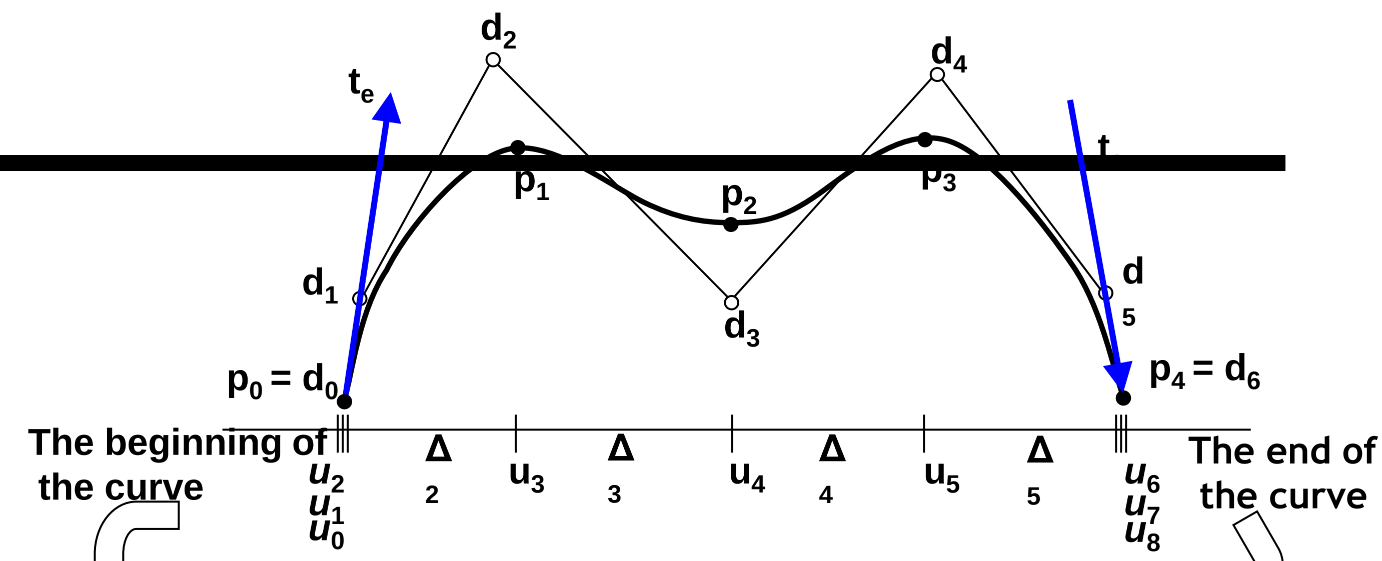
③ $\mathbf{U}\mathbf{x} = \mathbf{y}$

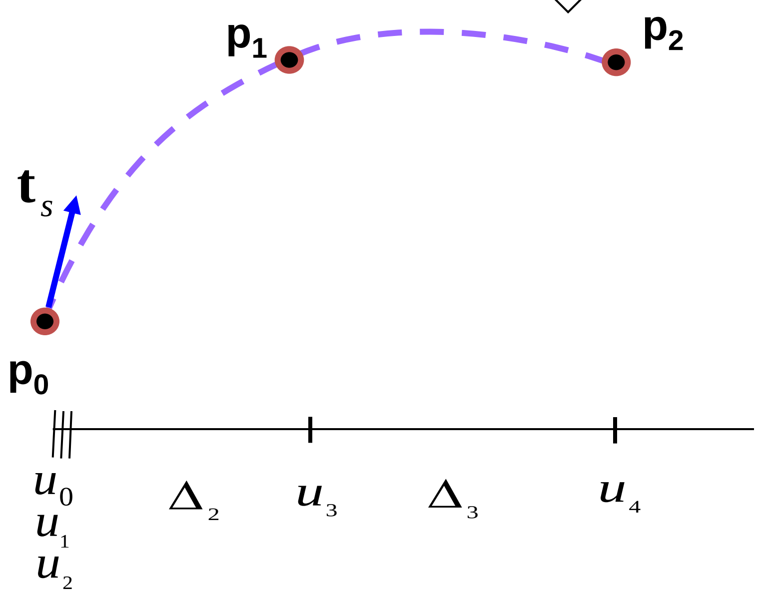
② $\mathbf{L}\mathbf{y} = \mathbf{d}$

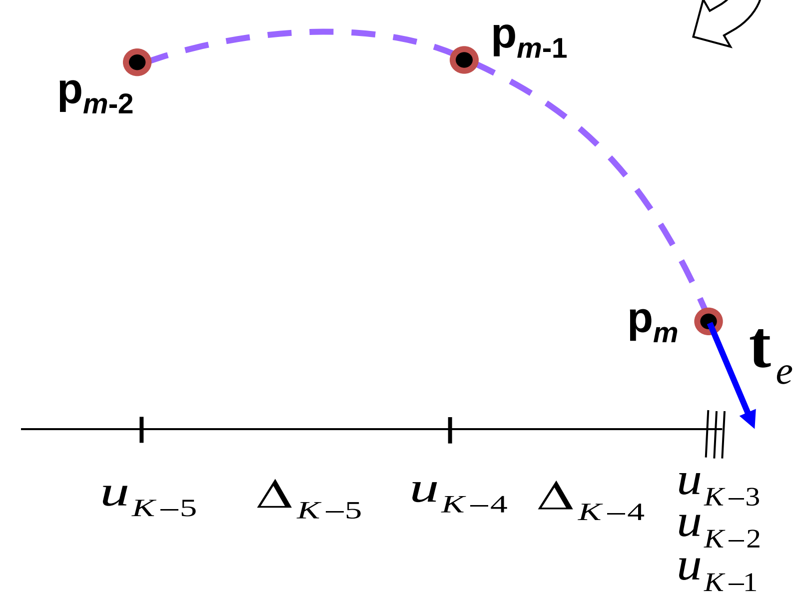
7) Bessel End Condition

☑ If the tangent vectors t_s, t_e at both end points are not given,
(1) Construct 2nd-degree curve(quadratic curve) from three consecutive points at both ends of the curve.
(2) And assume that the tangent vectors at each end point are the same with the value of the first derivatives of the constructed quadratic curves at each end point.





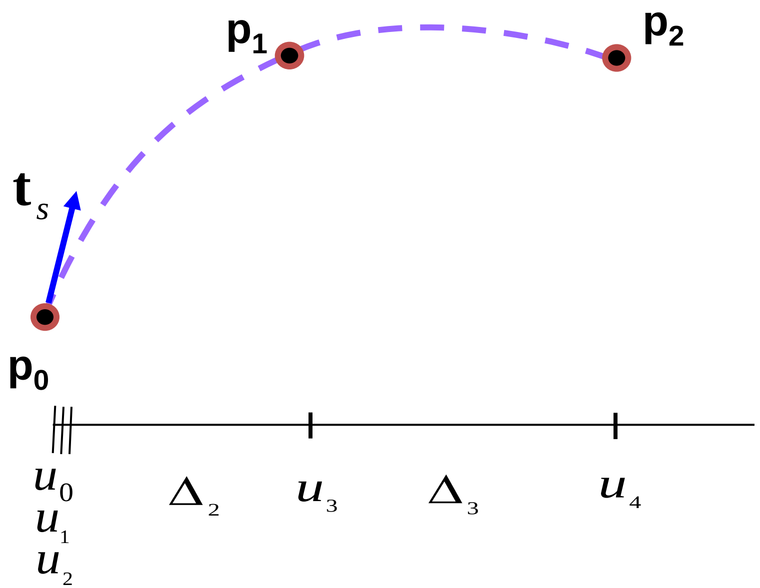




$$\mathbf{t}_s = \left(-\frac{2\Delta_2+\Delta_3}{\Delta_2(\Delta_2+\Delta_3)}\mathbf{p}_0 + \frac{(\Delta_2+\Delta_3)}{\Delta_2\Delta_3}\mathbf{p}_1 - \frac{\Delta_2}{\Delta_3(\Delta_2+\Delta_3)}\mathbf{p}_2\right)$$

$$\mathbf{t}_e = \left(\frac{\Delta_{K-4}}{\Delta_{K-5}(\Delta_{K-5}+\Delta_{K-4})}\mathbf{p}_{m-2} - \frac{(\Delta_{K-5}+\Delta_{K-4})}{\Delta_{K-5}\Delta_{K-4}}\mathbf{p}_{m-1} + \frac{(2\Delta_{K-4}+\Delta_{K-5})}{(\Delta_{K-5}+\Delta_{K-4})\Delta_{K-4}}\mathbf{p}_m\right)$$

Determination of t_s, t_e using Lagrange polynomial



$$\mathbf{x}(u) = L_0^2(u)\mathbf{p}_0 + L_1^2(u)\mathbf{p}_1 + L_2^2(u)\mathbf{p}_2$$

$$L_0^2(u) = \frac{(u-u_3)(u-u_4)}{(u_2-u_3)(u_2-u_4)} \qquad \dot{L}_0^2(u) = \frac{(u-u_4)+(u-u_3)}{(u_2-u_3)(u_2-u_4)}$$

$$L_1^2(u) = \frac{(u-u_2)(u-u_4)}{(u_3-u_2)(u_3-u_4)} \qquad \dot{L}_1^2(u) = \frac{(u-u_4)+(u-u_2)}{(u_3-u_2)(u_3-u_4)}$$

$$L_2^2(u) = \frac{(u-u_2)(u-u_3)}{(u_4-u_2)(u_4-u_3)} \qquad \dot{L}_2^2(u) = \frac{(u-u_3)+(u-u_2)}{(u_4-u_2)(u_4-u_3)}$$

$$\dot{\mathbf{x}}(u) = \dot{L}_0^2(u)\mathbf{p}_0 + \dot{L}_1^2(u)\mathbf{p}_1 + \dot{L}_2^2(u)\mathbf{p}_2$$

$$\dot{\mathbf{x}}(u_2) = \dot{L}_0^2(u_2)\mathbf{p}_0 + \dot{L}_1^2(u_2)\mathbf{p}_1 + \dot{L}_2^2(u_2)\mathbf{p}_2$$

$$= \frac{(u_2-u_4)+(u_2-u_3)}{(u_2-u_3)(u_2-u_4)}\mathbf{p}_0 + \frac{(u_2-u_4)+(u_2-u_2)}{(u_3-u_2)(u_3-u_4)}\mathbf{p}_1 + \frac{(u_2-u_3)+(u_2-u_2)}{(u_4-u_2)(u_4-u_3)}\mathbf{p}_2$$

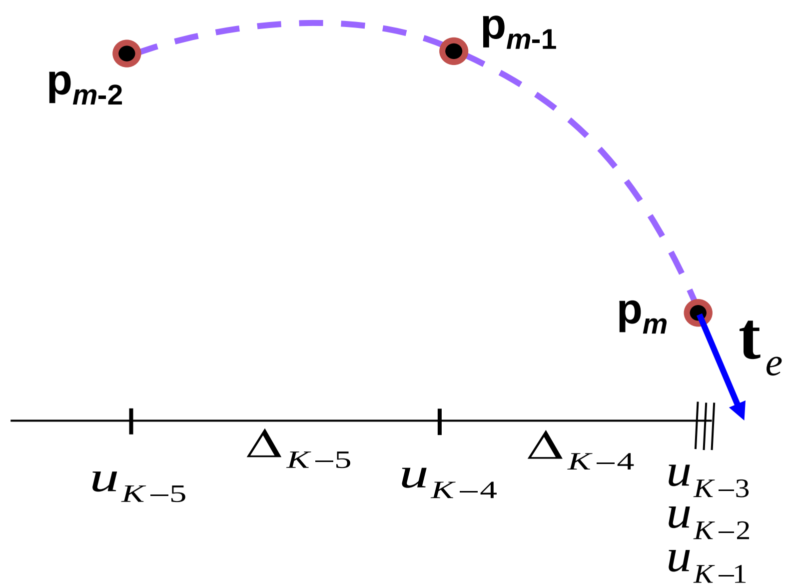
$$= \frac{-(\Delta_2+\Delta_3)-\Delta_2}{(-\Delta_2)(-(\Delta_2+\Delta_3))}\mathbf{p}_0 + \frac{-(\Delta_2+\Delta_3)}{\Delta_2(-\Delta_3)}\mathbf{p}_1 + \frac{-\Delta_2}{(\Delta_2+\Delta_3)(\Delta_2)}\mathbf{p}_2$$

$$= (-\frac{2\Delta_2+\Delta_3}{\Delta_2(\Delta_2+\Delta_3)}\mathbf{p}_0 + \frac{(\Delta_2+\Delta_3)}{\Delta_2\Delta_3}\mathbf{p}_1 - \frac{\Delta_2}{\Delta_3(\Delta_2+\Delta_3)}\mathbf{p}_2)$$

Therefore

$$\mathbf{t}_s = (-\frac{2\Delta_2+\Delta_3}{\Delta_2(\Delta_2+\Delta_3)}\mathbf{p}_0 + \frac{(\Delta_2+\Delta_3)}{\Delta_2\Delta_3}\mathbf{p}_1 - \frac{\Delta_2}{\Delta_3(\Delta_2+\Delta_3)}\mathbf{p}_2)$$

Determination of t_s, t_e using Lagrange polynomial



$$\mathbf{x}(u) = L_0^2(u)\mathbf{p}_{m-2} + L_1^2(u)\mathbf{p}_{m-1} + L_2^2(u)\mathbf{p}_m$$

$$L_0^2(u) = \frac{(u-u_{K-4})(u-u_{K-3})}{(u_{K-5}-u_{K-4})(u_{K-5}-u_{K-3})} \qquad \dot{L}_0^2(u) = \frac{(u-u_{K-3})+(u-u_{K-4})}{(u_{K-5}-u_{K-4})(u_{K-5}-u_{K-3})}$$

$$L_1^2(u) = \frac{(u-u_{K-5})(u-u_{K-3})}{(u_{K-4}-u_{K-5})(u_{K-4}-u_{K-3})} \qquad \dot{L}_1^2(u) = \frac{(u-u_{K-3})+(u-u_{K-5})}{(u_{K-4}-u_{K-5})(u_{K-4}-u_{K-3})}$$

$$L_2^2(u) = \frac{(u-u_{K-5})(u-u_{K-4})}{(u_{K-3}-u_{K-5})(u_{K-3}-u_{K-4})} \qquad \dot{L}_2^2(u) = \frac{(u-u_{K-4})+(u-u_{K-5})}{(u_{K-3}-u_{K-5})(u_{K-3}-u_{K-4})}$$

$$\dot{\mathbf{x}}(u) = \dot{L}_0^2(u)\mathbf{p}_{m-2} + \dot{L}_1^2(u)\mathbf{p}_{m-1} + \dot{L}_2^2(u)\mathbf{p}_m$$

$$\begin{aligned}
\dot{\mathbf{x}}(u_{K-3}) &= \dot{L}_0^2(u_{K-3})\mathbf{p}_{m-2} + \dot{L}_1^2(u_{K-3})\mathbf{p}_{m-1} + \dot{L}_2^2(u_{K-3})\mathbf{p}_m \\
&= \frac{(u_{K-3}-u_{K-3})+(u_{K-3}-u_{K-4})}{(u_{K-5}-u_{K-4})(u_{K-5}-u_{K-3})}\mathbf{p}_{m-2} + \frac{(u_{K-3}-u_{K-3})+(u_{K-3}-u_{K-5})}{(u_{K-4}-u_{K-5})(u_{K-4}-u_{K-3})}\mathbf{p}_{m-1} + \frac{(u_{K-3}-u_{K-4})+(u_{K-3}-u_{K-5})}{(u_{K-3}-u_{K-5})(u_{K-3}-u_{K-4})}\mathbf{p}_m \\
&= \frac{\Delta_{K-4}}{(-\Delta_{K-5})(-(\Delta_{K-4}+\Delta_{K-5}))}\mathbf{p}_{m-2} + \frac{-(\Delta_{K-4}+\Delta_{K-5})}{\Delta_{K-5}(-\Delta_{K-4})}\mathbf{p}_{m-1} + \frac{\Delta_{K-4}+(\Delta_{K-4}+\Delta_{K-5})}{(\Delta_{K-4}+\Delta_{K-5})(\Delta_{K-4})}\mathbf{p}_m \\
&= \left(\frac{\Delta_{K-4}}{\Delta_{K-5}(\Delta_{K-5}+\Delta_{K-4})}\mathbf{p}_{m-2} - \frac{(\Delta_{K-5}+\Delta_{K-4})}{\Delta_{K-5}\Delta_{K-4}}\mathbf{p}_{m-1} + \frac{(2\Delta_{K-4}+\Delta_{K-5})}{(\Delta_{K-5}+\Delta_{K-4})\Delta_{K-4}}\mathbf{p}_m\right)
\end{aligned}$$

Therefore

$$\mathbf{t}_e = \left(\frac{\Delta_{K-4}}{\Delta_{K-5}(\Delta_{K-5}+\Delta_{K-4})}\mathbf{p}_{m-2} - \frac{(\Delta_{K-5}+\Delta_{K-4})}{\Delta_{K-5}\Delta_{K-4}}\mathbf{p}_{m-1} + \frac{(2\Delta_{K-4}+\Delta_{K-5})}{(\Delta_{K-5}+\Delta_{K-4})\Delta_{K-4}}\mathbf{p}_m\right)$$

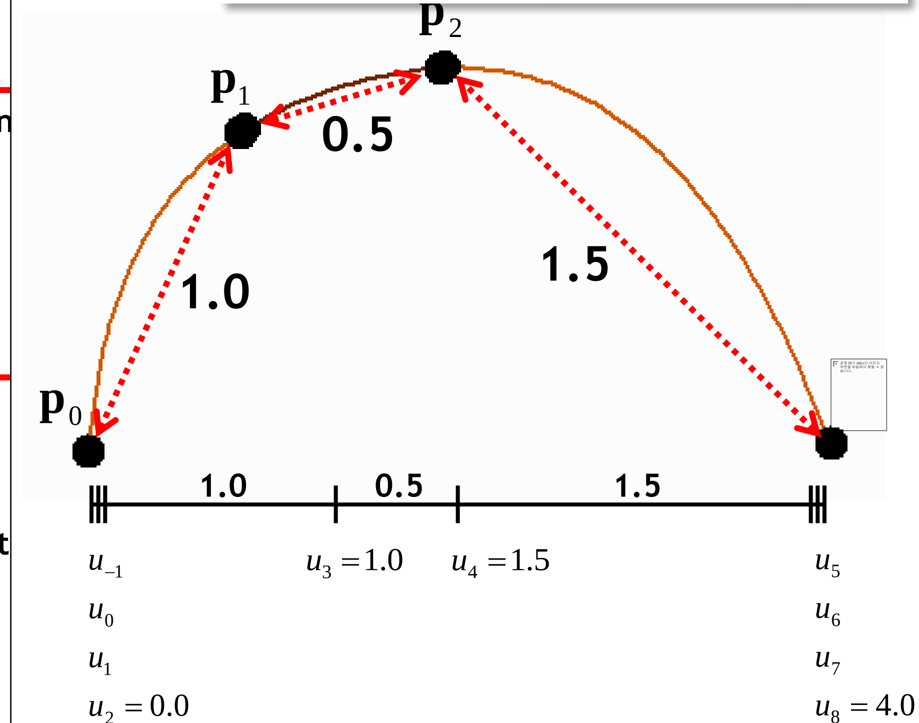
8) Sample code of Cubic B-spline Curve (1)

```
numberifndef __CubicBspline_h__
numberdefine __CubicBspline_h__

numberinclude "vector.h“

class CubicBsplineCurve {
public:
    Vector* m_ControlPoint;   int m_nControlPoint;
    double* m_Knot; int m_nKnot;  int m_nDegree;

    .........

    void SetControlPoint(Vector* pControlPoint, int nControlPoint);
    void SetKnot(double* pKnot, int nKnot);
    Vector CalcPoint(double u);
    double N(int d, int i, double u);
    void Interpolate(Vector *pFittingPoint, int nFittingPoint);
    void Parameterization(int nType, Vector* FittingPoint, int nPoint, double* t);
};
numberendif
```

Sample code of Cubic B-spline Curve (2)

```
void CubicBsplineCurve::Interpolate(Vector *pFittingPoint, in
{
  // Generate Knot
    if(m_Knot) delete[] m_Knot;
    m_nKnot = (m_nFittingPoint - 2) + 2*(3+1);
    m_Knot = new double [m_nKnot];
    // Use Chord length or Centripetal method
    ......

//----------------------------------------
// Generate Matrix : (L+1) * (L+1)
   int L = m_nFittingPoint + 1;           // (L+1)*(L+1) size Mat

  // Fill rhs
    Vector* rhs = new Vector[L+1];
    for(i = 1; i <= L-1 ; i++) rhs[i] = pFittingPoint[i-1];

  // Bessel End condition
    rhs[0] = rhs[1]; rhs[L] = rhs[L-1];
    rhs[1] = StartTangentByBesselEndCondition;   rhs[L-1] = EndTangentByBesselEndCondition;
```

Set knot using chord length



Sample code of Cubic B-spline Curve (3)

```
void CubicBsplineCurve::Interpolate(Vector *pFittingPoint, int nFittingPoint)
{
   // Generate Knot
      if(m_Knot) delete[] m_Knot;
      m_nKnot = (m_nFittingPoint - 2) + 2*(3+1);
      m_Knot = new double [m_nKnot];
      // Use Chord length or Centripetal method
      ......

//----------------------------------------
// Generate Matrix : (L+1) * (L+1)
    int L = m_nFittingPoint + 1;          // (L+1)*(L+1) size Matrix

   // Fill rhs
      Vector* rhs = new Vector[L+1];
      for(i = 1; i <= L-1 ; i++) rhs[i] = pFittingPoint[i-1];

   // Bessel End condition
      rhs[0] = rhs[1]; rhs[L] = rhs[L-1];
      rhs[1] = StartTangentByBesselEndCondition;   rhs[L-1] = EndTangentByBe
```

Bessel End Condition

$$\mathbf{t}_s = -\frac{2\Delta_2+\Delta_3}{\Delta_2(\Delta_2+\Delta_3)}\mathbf{p}_0 + \frac{(\Delta_2+\Delta_3)}{\Delta_2\Delta_3}\mathbf{p}_1 - \frac{\Delta_2}{\Delta_3(\Delta_2+\Delta_3)}\mathbf{p}_2$$

$$\mathbf{t}_e = \frac{\Delta_{K-4}}{\Delta_{K-5}(\Delta_{K-5}+\Delta_{K-4})}\mathbf{p}_{m-2} - \frac{(\Delta_{K-5}+\Delta_{K-4})}{\Delta_{K-5}\Delta_{K-4}}\mathbf{p}_{m-1} + \frac{(2\Delta_{K-4}+\Delta_{K-5})}{(\Delta_{K-5}+\Delta_{K-4})\Delta_{K-4}}\mathbf{p}_m$$

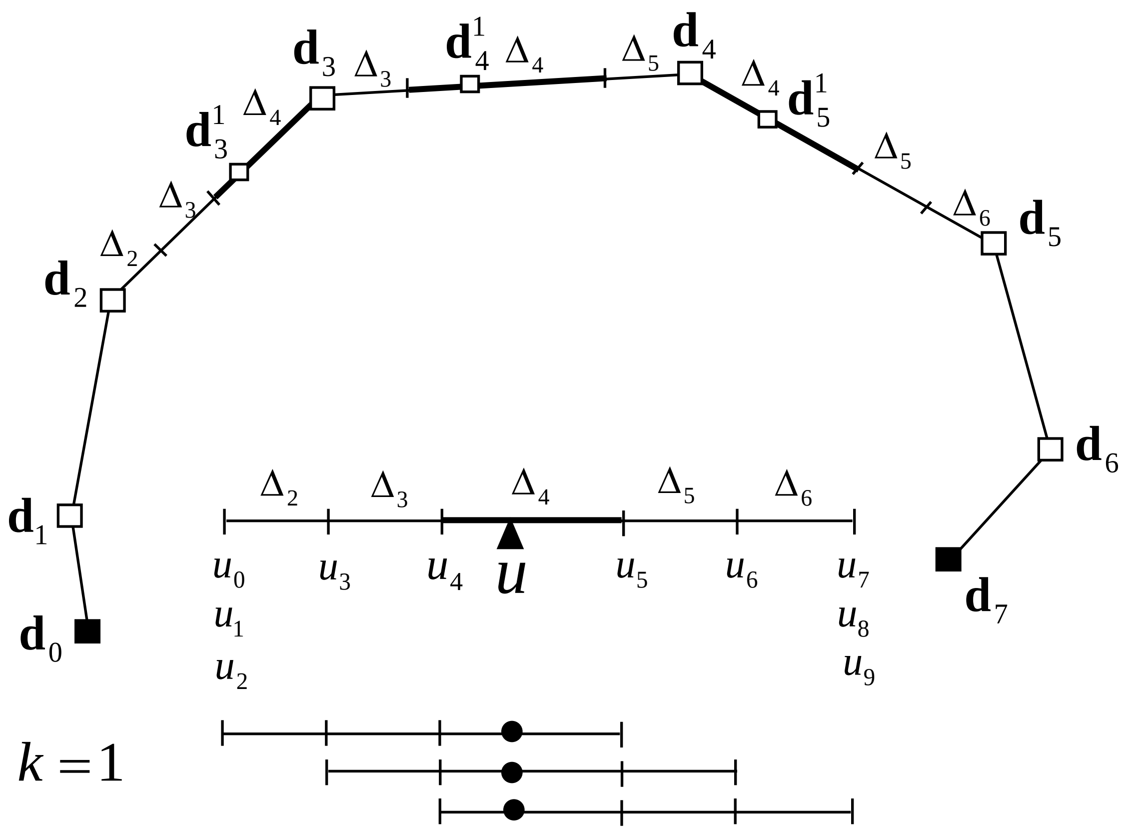
Sample code of Cubic B-spline Curve (4)

```
    double* alpha = new double[L+1];
    double* beta = new double[L+1];
    double* gamma = new double[L+1];
    double* up = new double[L+1];
    double* low = new double[L+1];
    if(m_ControlPoint) delete[] m_ControlPoint;
    m_nControlPoint = L+1;
    m_ControlPoint = new Vector[m_nControlPoint];
  // Fill alpha, beta, gamma
    ......
  // Solve LU system
    l_u_system(alpha, beta, gamma, L, up, low);
    solve_system(up, low, gamma, L, rhs, m_ControlPoint);

//-------------------------------
// Release memory
  delete[] rhs;   delete[] alpha;   delete[] beta;   delete[] gamma;   delete[] up;   delete[] low;
}
```

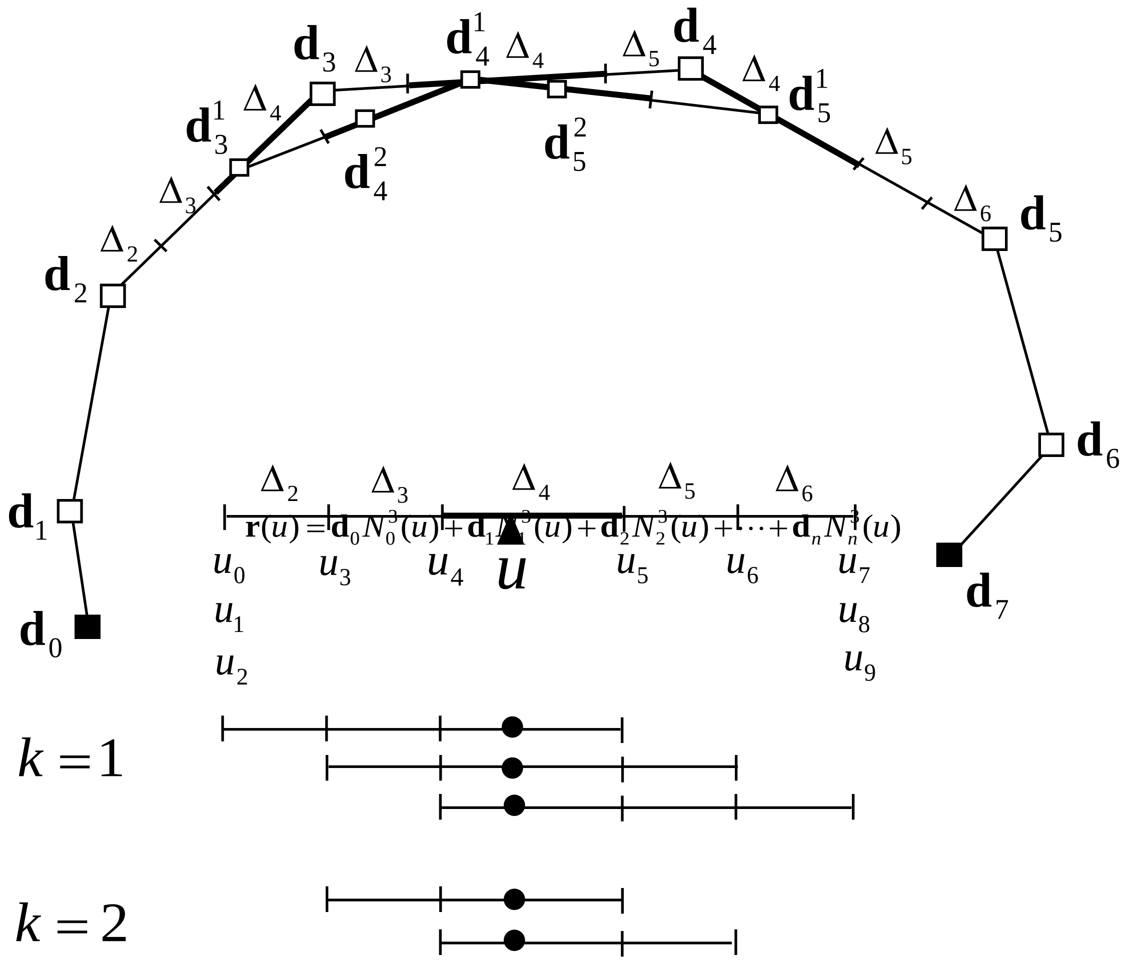
Calculate inverse matrix by using LU decomposition

3.5 de Boor Algorithm

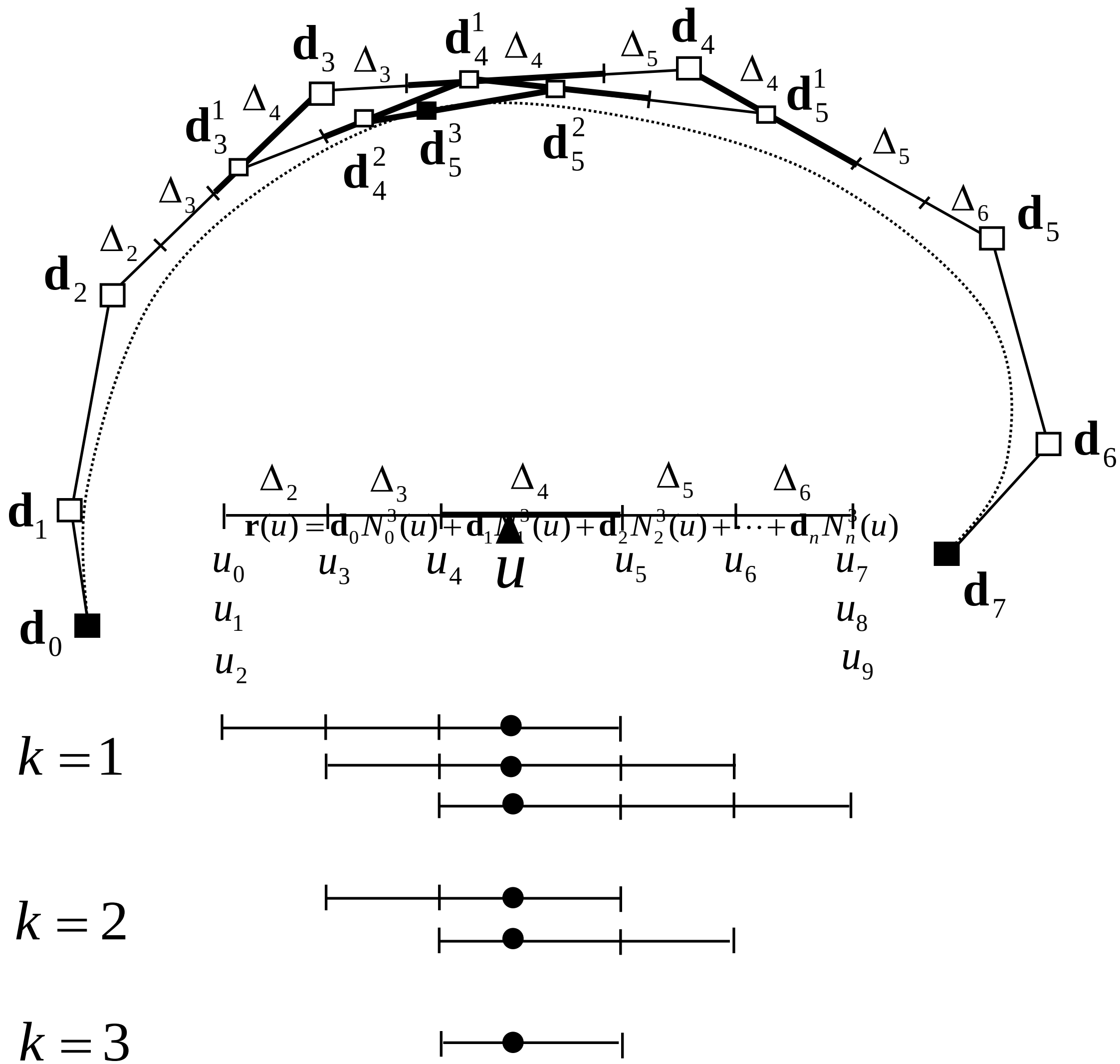
1) de Boor Algorithm



- The ratio of the linear interpolation used in the de Casteljau algorithm is constant.
- In contrast, the ratio of the linear interpolation used in the de Boor algorithm is not constant, since the intervals of the parameters of the Bezier curve segments, which B-spline curve is composed of, are different from each other.



$\mathbf{d}_0$
$\mathbf{d}_1$
$\mathbf{d}_2$
$\mathbf{d}_3$
$\mathbf{d}_4$
$\mathbf{d}_5$
$\mathbf{d}_6$
$\mathbf{d}_7$
$\mathbf{d}_3^1$
$\mathbf{d}_4^1$
$\mathbf{d}_5^1$
$\mathbf{d}_4^2$
$\mathbf{d}_5^2$
Δ_2 Δ_3 Δ_4 Δ_5 Δ_6
$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$
u_0 u_1 u_2 u_3 u_4 u u_5 u_6 u_7 u_8 u_9
$k = 1$
$k = 2$



$\mathbf{d}_0$
$\mathbf{d}_1$
$\mathbf{d}_2$
$\mathbf{d}_3$
$\mathbf{d}_4$
$\mathbf{d}_5$
$\mathbf{d}_6$
$\mathbf{d}_7$
$\mathbf{d}_3^1$
$\mathbf{d}_4^1$
$\mathbf{d}_5^1$
$\mathbf{d}_4^2$
$\mathbf{d}_5^2$
$\mathbf{d}_5^3$
Δ_2 Δ_3 Δ_4 Δ_5 Δ_6
$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$
u_0 u_1 u_2 u_3 u_4 u u_5 u_6 u_7 u_8 u_9
$k=1$
$k=2$
$k=3$

2) Relationship between de Boor algorithm & B-spline curves

☑ **de Boor Algorithm : "Constructive Approach"**

Input: $\mathbf{d}_i$ (de Boor Points)

Processor: Sequentially n-times 'linear interpolation' at $\mathbf{d}_i$ by section

Output : Point on n^{th}-degree curve

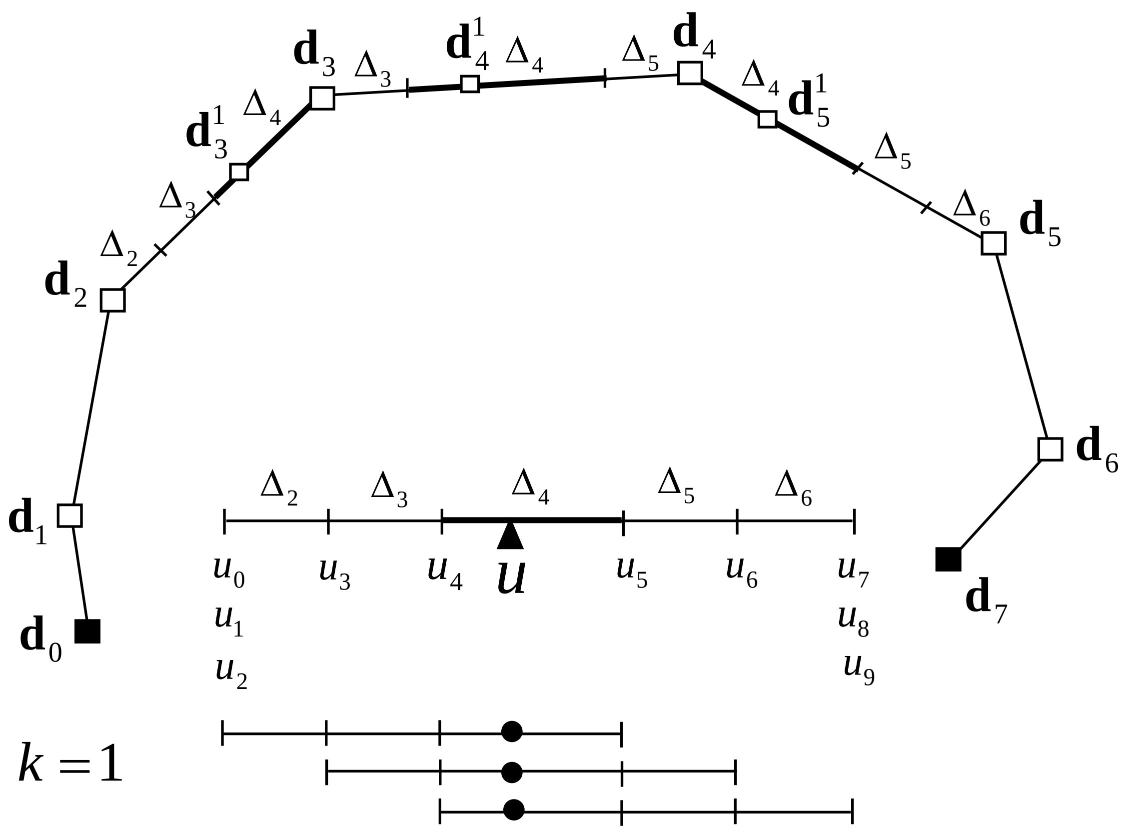
→Expressed 'B-spline function'

(Cox-de Boor recurrence formula)

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$$

$$\mathbf{d}_i^k(u) = \frac{u_{i+n-k} - u}{u_{i+n-k} - u_{i-1}} \mathbf{d}_{i-1}^{k-1}(u) + \frac{u - u_{i-1}}{u_{i+n-k} - u_{i-1}} \mathbf{d}_i^{k-1}(u)$$

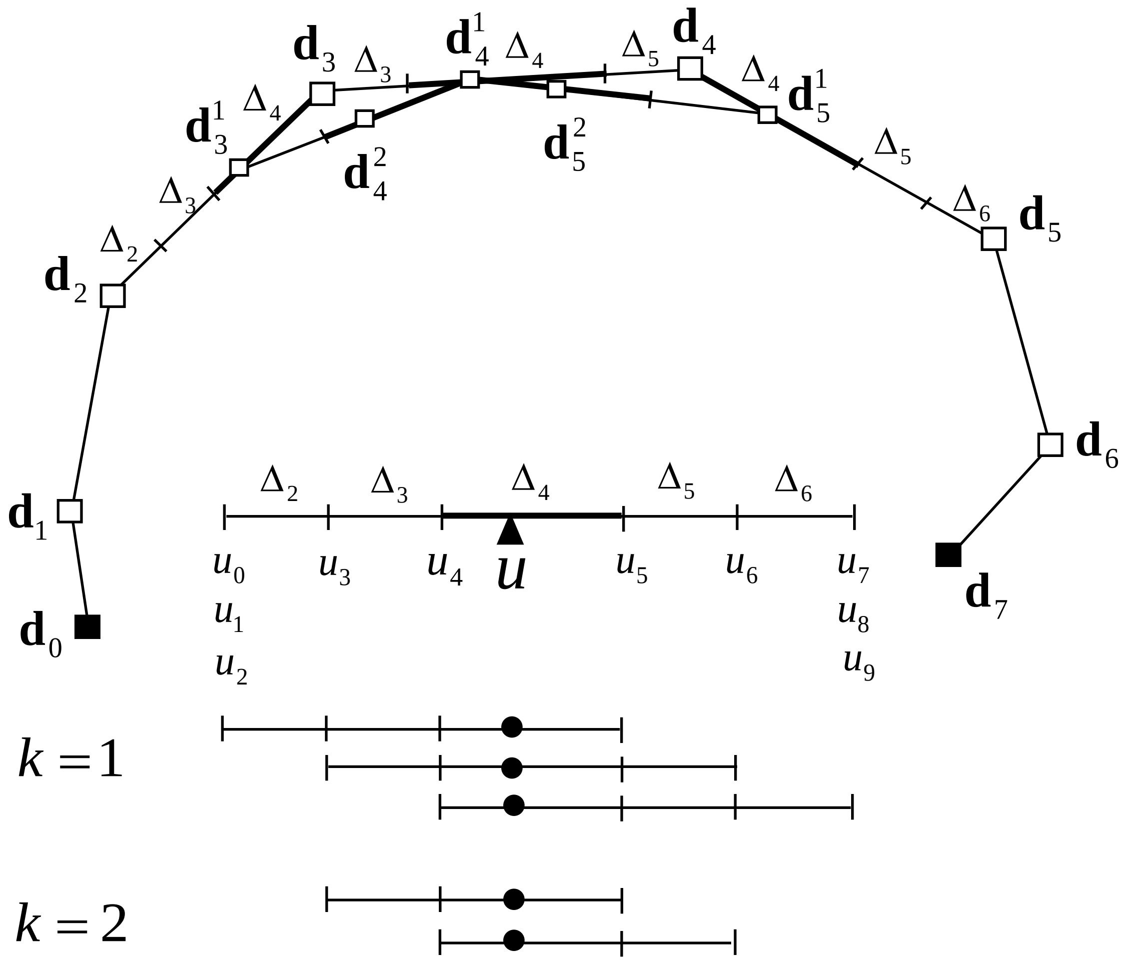
3) Geometrical Meaning of the de Boor Algorithm(1)



- Linear Interpolation 비율이 t:(1-t)로 일정했던 de Casteljau algorithm에 비하여 de Boor algorithm에서는 Linear Interpolation 비율이 변한다
- 이는 B-spline curve 를 구성하는 Bezier curve segment의 매개변수 간격이 서로 다르기 때문이다

Geometrical Meaning of the de Boor Algorithm(2)

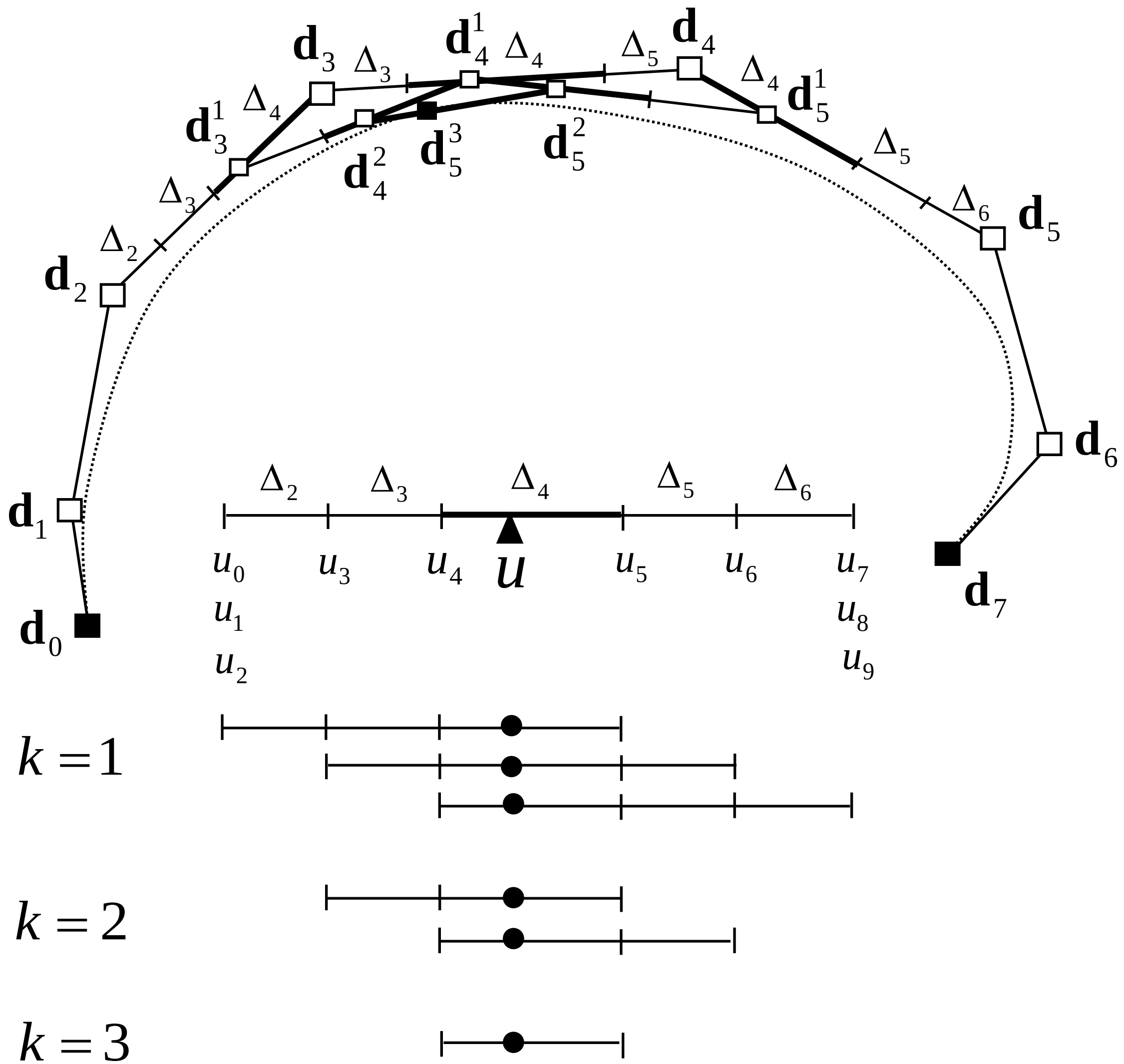
$$\mathbf{d}_i^k(u) = \frac{u_{i+n-k} - u}{u_{i+n-k} - u_{i-1}} \mathbf{d}_{i-1}^{k-1}(u) + \frac{u - u_{i-1}}{u_{i+n-k} - u_{i-1}} \mathbf{d}_i^{k-1}(u)$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$$

Geometrical Meaning of the de Boor Algorithm(3)

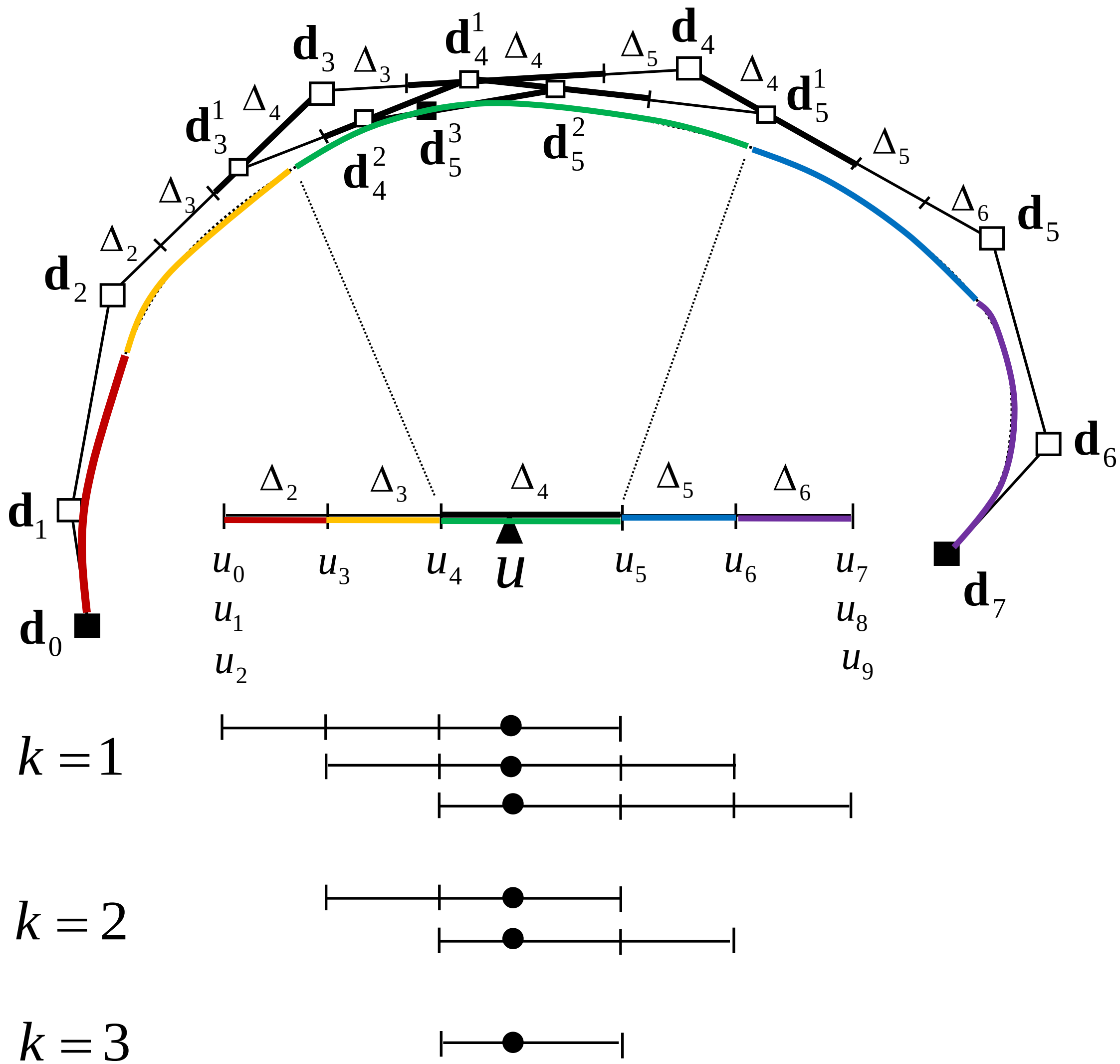
$$\mathbf{d}_i^k(u) = \frac{u_{i+n-k} - u}{u_{i+n-k} - u_{i-1}} \mathbf{d}_{i-1}^{k-1}(u) + \frac{u - u_{i-1}}{u_{i+n-k} - u_{i-1}} \mathbf{d}_i^{k-1}(u)$$



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$$

$$\mathbf{d}_i^k(u) = \frac{u_{i+n-k} - u}{u_{i+n-k} - u_{i-1}} \mathbf{d}_{i-1}^{k-1}(u) + \frac{u - u_{i-1}}{u_{i+n-k} - u_{i-1}} \mathbf{d}_i^{k-1}(u)$$

Geometrical Meaning of the de Boor Algorithm(3)



$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$$

4) Relationship between de Boor algorithm & B-spline curves

☑ **de Boor 알고리즘 : "Constructive Approach"**

Input: $\mathbf{d}_i$ (de Boor Points)

Processor: 구간별로 $\mathbf{d}_i$를 n번 순차적 'linear interpolation'

Output : n차 곡선상의 점
→ 'B-spline function'(Cox-de Boor recurrence formula) 형태로 표현 됨

$$\mathbf{r}(u) = \mathbf{d}_0 N_0^3(u) + \mathbf{d}_1 N_1^3(u) + \mathbf{d}_2 N_2^3(u) + \cdots + \mathbf{d}_n N_n^3(u)$$

☑ de Boor 알고리즘 : "Constructive Approach"

Input: $\mathbf{d}_i$ (de Boor Points)

Processor: 구간별로 $\mathbf{d}_i$를 n번 순차적 'linear interpolation'

Output : n차 곡선상의 점

→'B-spline function'(Cox-de Boor recurrence formula) 형태로 표현 됨

☑ B-spline 곡선식: "B-spline function evaluation Approach"

Input: $\mathbf{d}_i$ (de Boor Points)

Processor: 공간 상의 점 $\mathbf{d}_i$와 B-spline function을 "blending"하여 함수 값을 계산하면 곡선상의 점을 구할 수 있음

Output: B-spline function과 $\mathbf{d}_i$의 혼합 함수 형태로 표현

Relationship between de Boor algorithm & B-spline curves

☑ **de Boor Algorithm : "Constructive Approach"**

Input: $\mathbf{d}_i$ (de Boor Points)

Processor: Sequentially n-times 'linear interpolation' at $\mathbf{d}_i$ by section

Output : Point on n^{th}-degree curve

→ Expressed 'B-spline function'

(Cox-de Boor recurrence formula)

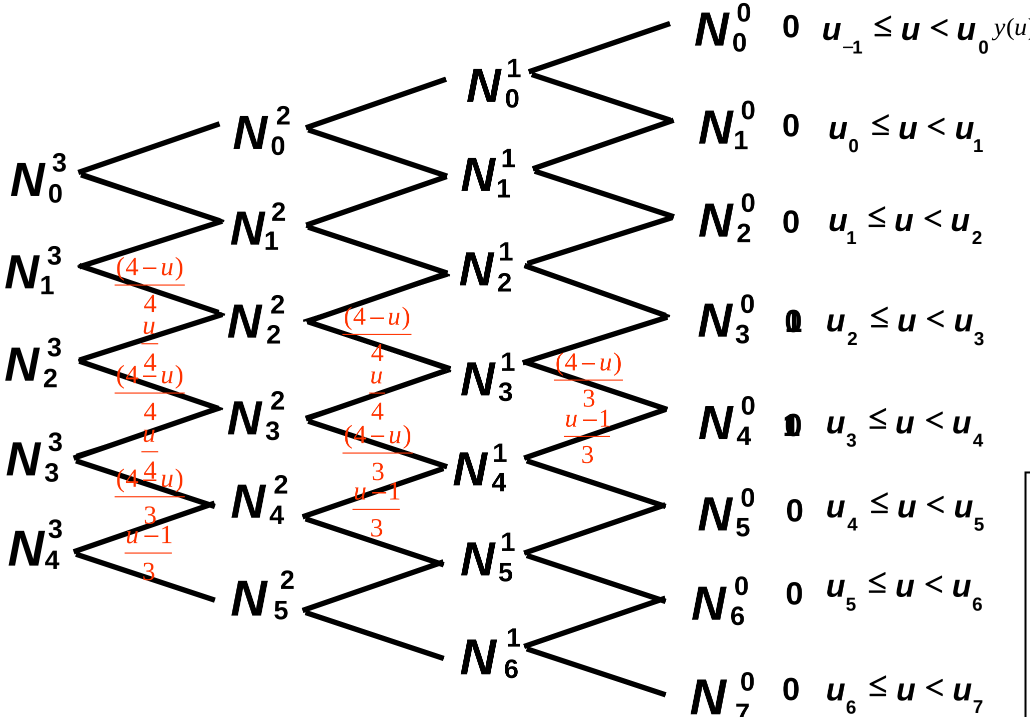
☑ **B-spline curve equation: "B-spline function evaluation Approach"**

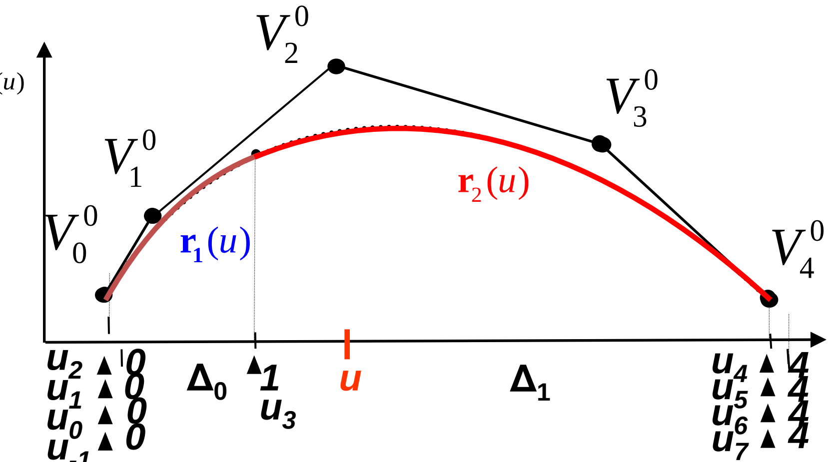
Input: $\mathbf{d}_i$ (de Boor Points)

Processor: Can be represented points on curve to "blending" points(d_i) in space and B-spline function

Output: Expressed mixture function of Bernstein basis function (polynomial) and $\mathbf{d}_i$

Cox–de Boor Algorithm:
Evaluation of Cox–de Boor basis function





$$\mathbf{r}(u) = V_0^0 N_0^3(u) + V_1^0 N_1^3(u) + V_2^0 N_2^3(u) + V_3^0 N_3^3(u) + V_4^0 N_4^3(u)$$

Cox-de Boor Algorithm

$$N_i^n(u) = \frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n}-u}{u_{i+n}-u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

$$\mathbf{r}_2(u) = \frac{(4-u)^3}{48} V_1^0 + (4-u)^2 \left(\frac{u}{24} + \frac{u-1}{36} \right) V_2^0 + (4-u) \left(\frac{u^2}{48} + \frac{u(u-1)}{36} + \frac{(u-1)^2}{27} \right) V_3^0 + \frac{(u-1)^3}{27} V_4^0$$

$$V_4^3 = \frac{(4-u)^3}{48} V_1^0 + (4-u)^2 \left(\frac{u}{24} + \frac{(u-1)}{36} \right) V_2^0 + (4-u) \left(\frac{u^2}{48} + \frac{u(u-1)}{36} + \frac{(u-1)^2}{27} \right) V_3^0 + \frac{(u-1)^3}{27} V_4^0$$

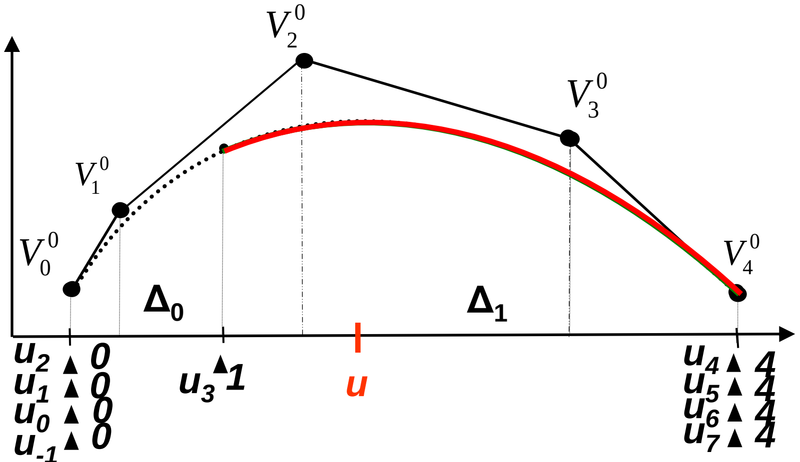
de Boor algorithm으로 구한 결과와 Cox- de Boor algorithm으로 구한 결과가 같다.

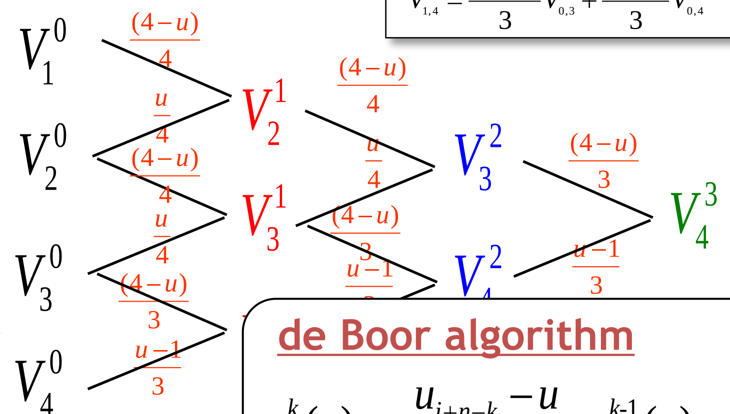
de Boor Algorithm & Cox-de Boor Algorithm

$$V_{1,2}^y = \frac{(4-u)}{4}V_{0,1}^y + \frac{u}{4}V_{0,2}^y \qquad V_{2,3}^y = \frac{(4-u)}{4}V_{1,2}^y + \frac{u}{4}V_{1,3}^y$$

$$V_{1,3}^y = \frac{(4-u)}{4}V_{0,2}^y + \frac{u}{4}V_{0,3}^y \qquad V_{2,4}^y = \frac{(4-u)}{3}V_{1,3}^y + \frac{(u-1)}{3}V_{1,4}^y$$

$$V_{1,4}^y = \frac{(4-u)}{3}V_{0,3}^y + \frac{(u-1)}{3}V_{0,4}^y \qquad V_{3,4}^y = \frac{(4-u)}{3}V_{2,3}^y + \frac{(u-1)}{3}V_{2,4}^y$$



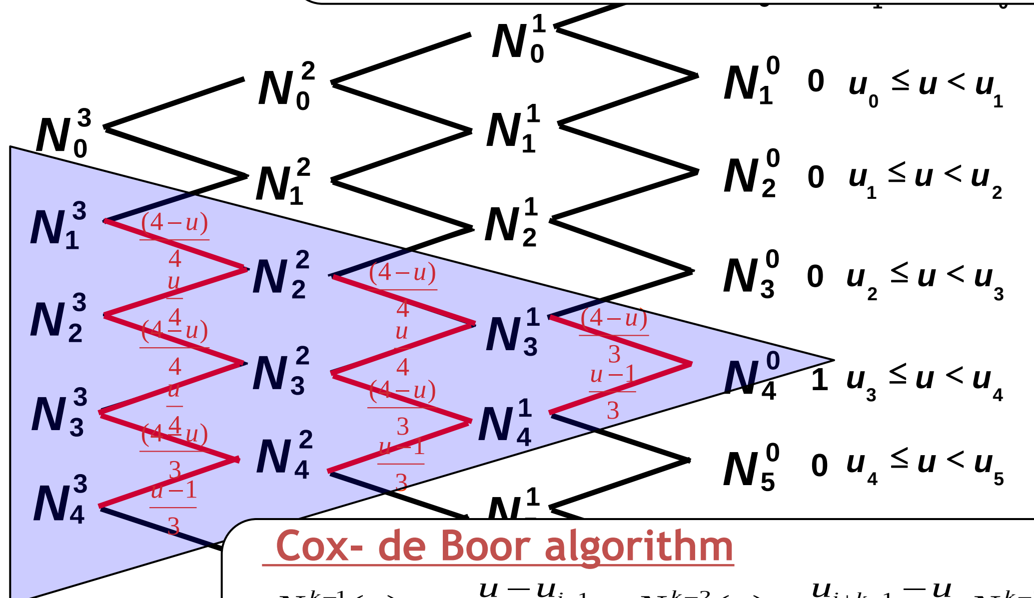


de Boor algorithm

$$v_i^k(u) = \frac{u_{i+n-k} - u}{u_{i+n-k} - u_{i-1}} v_{i-1}^{k-1}(u) + \frac{u - u_{i-1}}{u_{i+n-k} - u_{i-1}} v_i^{k-1}(u)$$

$$V_4^3 = \frac{(4-u)^3}{48}V_1^0 + (4-u)^2\left(\frac{u}{24} + \frac{(u-1)}{36}\right)V_2^0 +$$

$$(4-u)\left(\frac{u^2}{48} + \frac{u(u-1)}{36} + \frac{(u-1)^2}{27}\right)V_3^0 + \frac{(u-1)^3}{27}V_4^0$$



매개 변수 u의 위치에서
동일한 3차식을 얻을 수 있음
➡ de Boor Algorithm과
Cox de-Boor Algorithm은 동일한 것

Cox- de Boor algorithm

$$N_i^{k-1}(u) = \frac{u - u_{i-1}}{u_{i+k-2} - u_{i-1}} N_i^{k-2}(u) + \frac{u_{i+k-1} - u}{u_{i+k-1} - u_i} N_{i+1}^{k-2}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

$$\mathbf{r}_2(u) = \frac{(4-u)^3}{48}V_1^0 + (4-u)^2\left(\frac{u}{24} + \frac{u-1}{36}\right)V_2^0 +$$

$$(4-u)\left(\frac{u^2}{48} + \frac{u(u-1)}{36} + \frac{(u-1)^2}{27}\right)V_3^0 + \frac{(u-1)^3}{27}V_4^0$$

Bezier Curve and B-Spline Curve

항목		Bezier Curve	B-Spline Curve
Make Curve	Given	Bezier Control Point $\mathbf{b}_i$ Parameter t Bernstein Polynomial Func. $B_i^n(t)$	B-Spline Control Point $\mathbf{d}_i$ Parameter u B-Spline Basis Func. $N_i^n(u)$
	Find	Bezier Curve $\mathbf{r}(t)$ $\mathbf{r}(t)=\mathbf{b}_0B_0^n(t)+\mathbf{b}_1B_1^n(t)+.......+\mathbf{b}_nB_n^n(t).$	B-Spline Curve $\mathbf{r}(u)$ $\mathbf{r}(u)=\mathbf{d}_0N_0^3(u)+\mathbf{d}_1N_1^3(u)+\mathbf{d}_2N_2^3(u)+\cdots+\mathbf{d}_{D-1}N_{D-1}^3(u)$
-		Bernstein Polynomial Function $B_i^n(t)=\binom{n}{i}t^i(1-t)^{n-i},$ $\binom{n}{i}={}_nC_i=\begin{cases}\dfrac{n!}{i!(n-i)!} & \textbf{if } 0\le i\le n\\ \mathbf{0} & \textbf{else}\end{cases}$	B-Spline Basis Function (Cox-de boor Recursive Formula) $N_i^n(u)=\dfrac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}}N_i^{n-1}(u)+\dfrac{u_{i+n}-u}{u_{i+n}-u_i}N_{i+1}^{n-1}(u)$ $N_i^0(u)=\begin{cases}1 & \textbf{if } u_{i-1}\le u<u_i\\ 0 & \textbf{else}\end{cases}, \sum_{i=0}^{D-1}N_i^n(u)=1$
Constructive Approach		de Casteljau Algorithm $\mathbf{b}_i^k(t)=(1-t)\mathbf{b}_i^{k-1}+t\mathbf{b}_{i+1}^{k-1}$	de Boor Algorithm $\mathbf{d}_i^k(u)=\dfrac{u_{i+n-k}-u}{u_{i+n-k}-u_{i-1}}\mathbf{d}_{i-1}^{k-1}(u)+\dfrac{u-u_{i-1}}{u_{i+n-k}-u_{i-1}}\mathbf{d}_i^{k-1}(u)$
Inter-polation	Given	Points on Curve: $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$	Points on Curve: $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$
	Find	Bezier Control Point $\mathbf{b}_i$	B-Spline Control Point $\mathbf{d}_i$

Supplementary slide:
1) Determine the number of Bezier curve segments & knot values

Given:

- Points p_i on the curve
- Knots u_j of the given points on the curve
- Tangent vectors t_0, t_1 at both ends

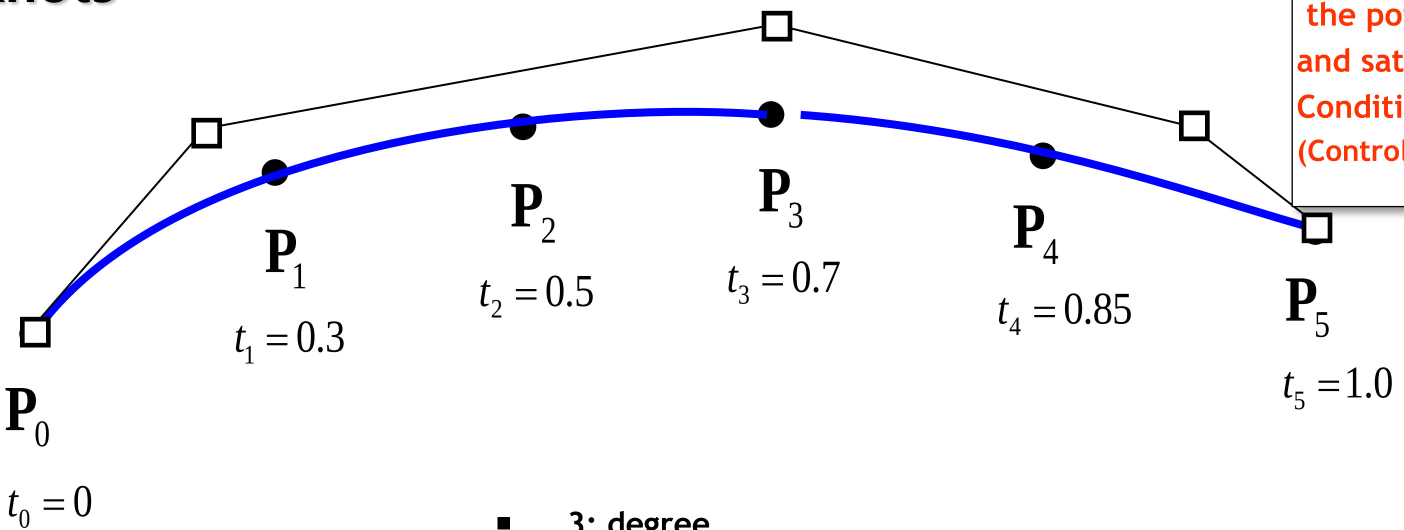
Find:

Cubic B-spline curve r(*u*) passing through the points p_i on the curve and satisfying C^2 continuity Condition:
(Control point of B-spline: d_i)

☑ **Given: fitting points $\mathbf{P}_i$ and corresponding parameter t_i**

where, $i = 0, 1, \ldots, m$ **and** $t_0 = 0, \; t_m = 1,$

☑ **First, determine number of Bezier curve segment and its knots**

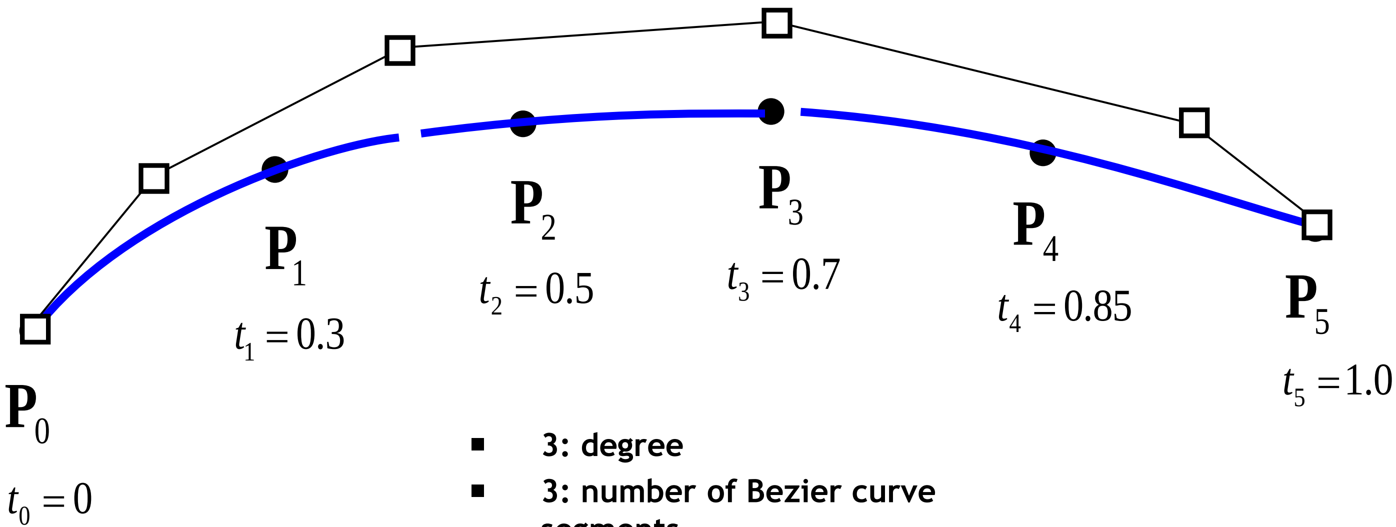


- 3: degree
- 2: number of Bezier curve segments
- number of control points = 4 + (2-1) = 5

☑ **Given: fitting points $\mathbf{P}_i$ and corresponding parameter t_i**

where, $i = 0, 1, \ldots, m$ **and** $t_0 = 0, \quad t_m = 1,$

☑ **First, determine number of Bezier curve segment and its knots**

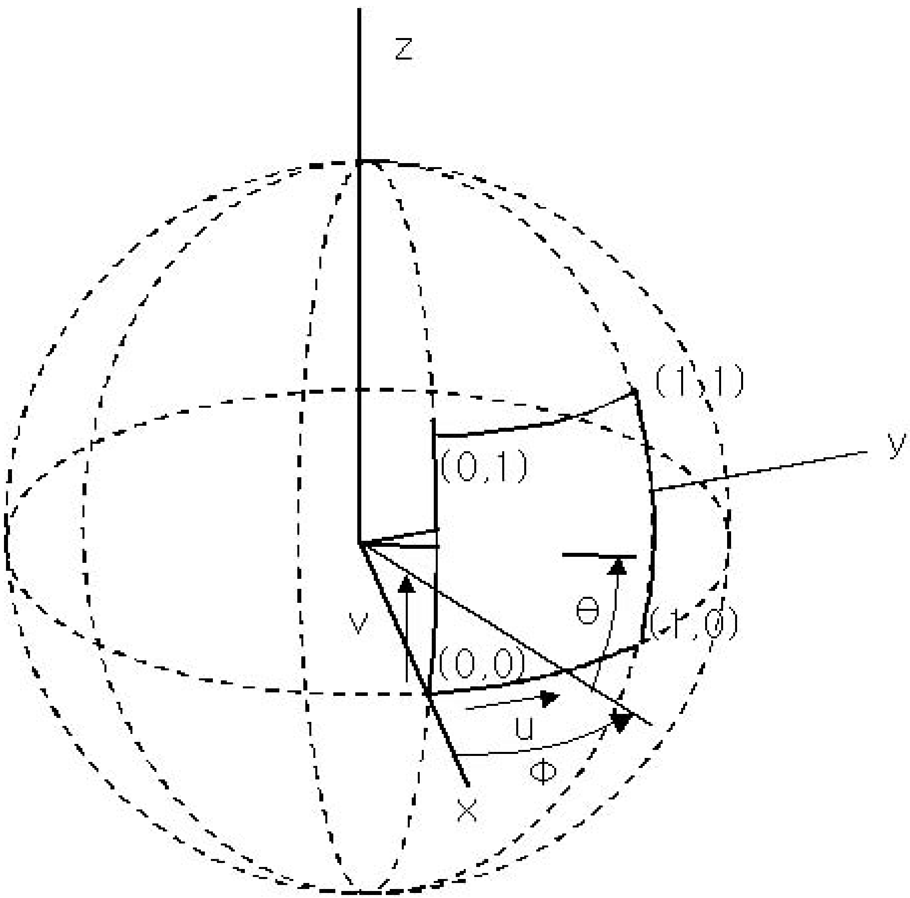


- 3: degree
- 3: number of Bezier curve segments
- number of control points
  = 4 + (3-1) = 6
- How do we determine Knots?
  (= start / end points of each
  cubic Bezier curve)

Chapter 4. Surfaces

4.1 Parametric Surfaces

4.1 Parametric Surfaces



	Surface
Explicit function	$z = \pm\sqrt{d^2 - x^2 - y^2}$
Implicit function	$x^2 + y^2 + z^2 = d^2$
Para-meter	$x = d\cos\phi\cos\theta$ $y = d\sin\phi\cos\theta$ $z = d\sin\theta$
	$\mathbf{r} = r(x(\phi,\theta), y(\phi,\theta), z(\phi,\theta))$

Sphere can be represented by three parameters (d, φ, θ)

4.2 Bezier Surfaces

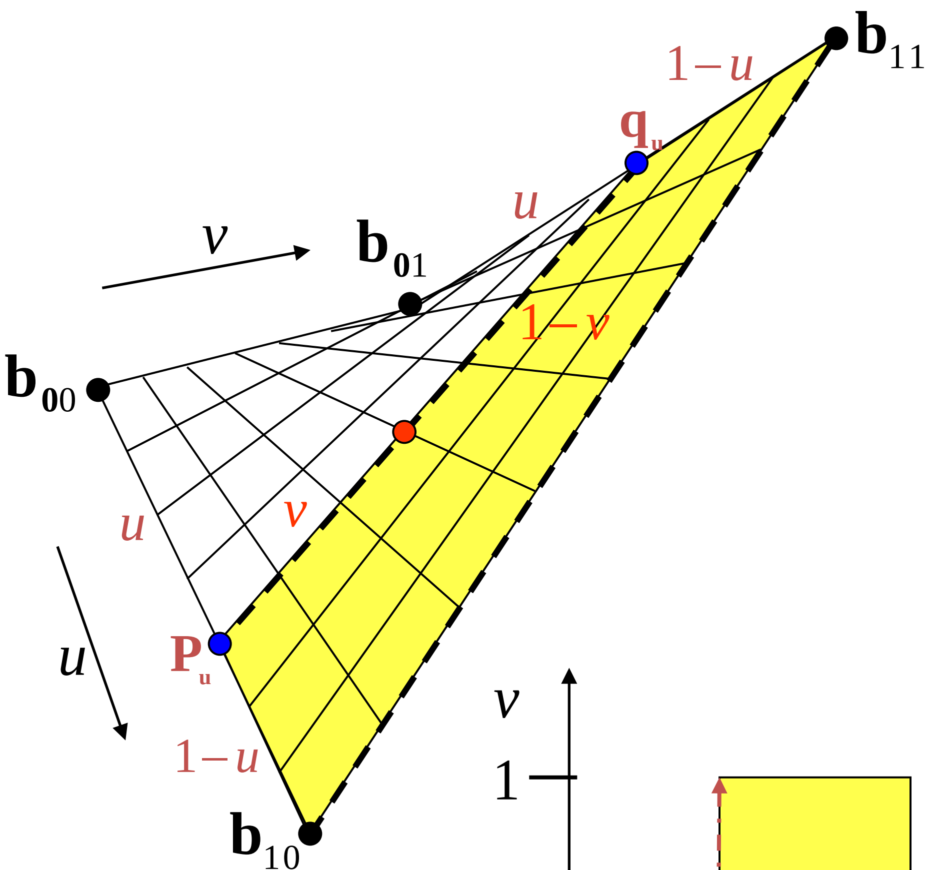
1) Generation of Bezier surfaces by de Casteljau algorithm
- Bilinear Bezier Surface Patch
- Biquadratic Bezier Surface Patch
- BiCubic Bezier Surface Patch

2) Generation of Bezier surfaces by tensor product approach

1) Bilinear Bezier Surface Patch

- Given: 2x2 Bezier control points
- Find: Points on the bilinear Bezier Surface Patch

Method : Applying 'de Casteljau algorithm' to the u- and v-directions



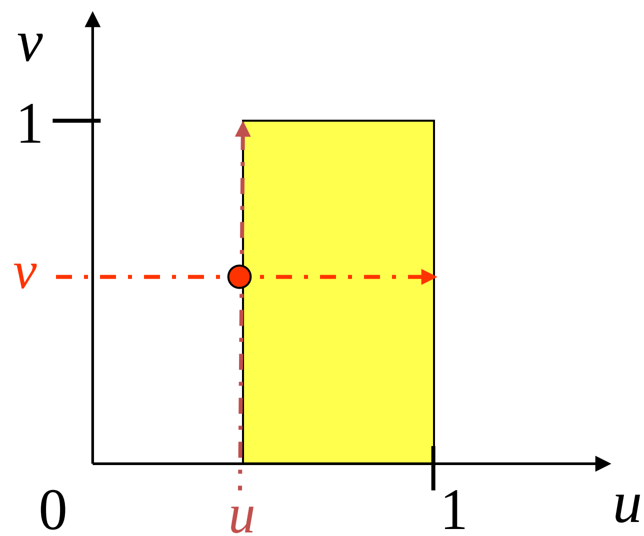
$$\mathbf{P}_u = (1-u)\mathbf{b}_{00} + u\,\mathbf{b}_{10}$$

$$\mathbf{q}_u = (1-u)\mathbf{b}_{01} + u\,\mathbf{b}_{11}$$

$$\mathbf{r}(u,v) = (1-v)\,\mathbf{P}_u + v\,\mathbf{q}_u$$

$$\mathbf{r}(u,v) = (1-v)(1-u)\mathbf{b}_{00} + (1-v)\,u\,\mathbf{b}_{10} + v\,(1-u)\mathbf{b}_{01} + v\,u\,\mathbf{b}_{11}$$

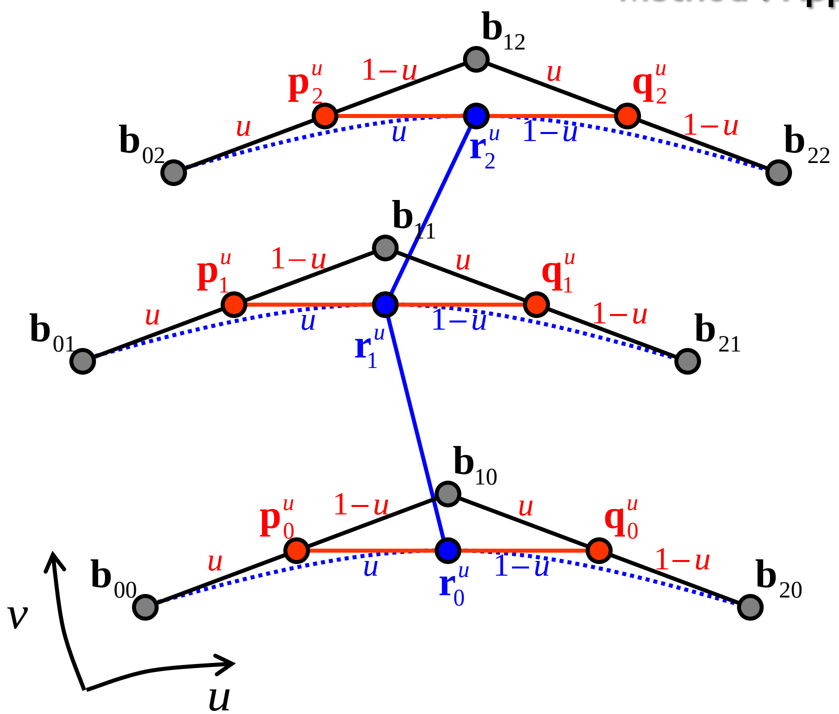
$$r(u,v) = \begin{bmatrix}(1-v) & v\end{bmatrix}\begin{bmatrix}\mathbf{b}_{00} & \mathbf{b}_{10} \\ \mathbf{b}_{01} & \mathbf{b}_{11}\end{bmatrix}\begin{bmatrix}(1-u) \\ u\end{bmatrix}$$



2) Biquadratic Bezier Surface Patch

- Given: 3x3 Bezier control points
- Find: Points on the biquadratic Bezier Surface Patch

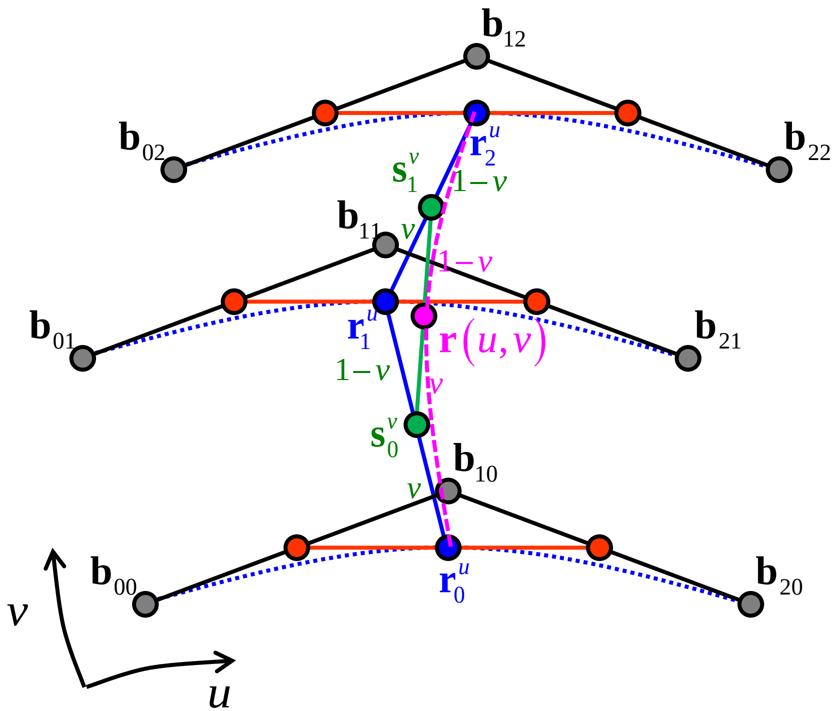
Method : Applying 'de Casteljau algorithm' to the u- and v-directions



$$\mathbf{p}_0^u = (1-u)\mathbf{b}_{00} + u\mathbf{b}_{10}$$
$$\mathbf{q}_0^u = (1-u)\mathbf{b}_{10} + u\mathbf{b}_{20}$$
$$\mathbf{p}_1^u = (1-u)\mathbf{b}_{01} + u\mathbf{b}_{11}$$
$$\mathbf{q}_1^u = (1-u)\mathbf{b}_{11} + u\mathbf{b}_{21}$$
$$\mathbf{p}_2^u = (1-u)\mathbf{b}_{02} + u\mathbf{b}_{12}$$
$$\mathbf{q}_2^u = (1-u)\mathbf{b}_{12} + u\mathbf{b}_{22}$$
$$\mathbf{r}_0^u = (1-u)\mathbf{p}_0^u + u\mathbf{q}_0^u$$
$$\mathbf{r}_1^u = (1-u)\mathbf{p}_1^u + u\mathbf{q}_1^u$$
$$\mathbf{r}_2^u = (1-u)\mathbf{p}_2^u + u\mathbf{q}_2^u$$

$$\begin{bmatrix} \mathbf{r}_0^u \\ \mathbf{r}_1^u \\ \mathbf{r}_2^u \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{10} & \mathbf{b}_{20} \\ \mathbf{b}_{01} & \mathbf{b}_{11} & \mathbf{b}_{21} \\ \mathbf{b}_{02} & \mathbf{b}_{12} & \mathbf{b}_{22} \end{bmatrix} \begin{bmatrix} (1-u)^2 \\ 2u(1-u) \\ u^2 \end{bmatrix}$$

$$\mathbf{r}_0^u = (1-u)^2\mathbf{b}_{00} + 2u(1-u)\mathbf{b}_{10} + u^2\mathbf{b}_{20}$$
$$\mathbf{r}_1^u = (1-u)^2\mathbf{b}_{01} + 2u(1-u)\mathbf{b}_{11} + u^2\mathbf{b}_{21}$$
$$\mathbf{r}_2^u = (1-u)^2\mathbf{b}_{02} + 2u(1-u)\mathbf{b}_{12} + u^2\mathbf{b}_{22}$$



$$\mathbf{s}_0^v = (1-v)\mathbf{r}_0^u + v\mathbf{r}_1^u$$

$$\mathbf{s}_1^v = (1-v)\mathbf{r}_1^u + v\mathbf{r}_2^u$$

$$\mathbf{r}(u,v) = (1-v)\mathbf{s}_0^v + v\mathbf{s}_1^v$$

$$\mathbf{r}(u,v) = (1-v)^2\mathbf{r}_0^u + 2v(1-v)\mathbf{r}_1^u + v^2\mathbf{r}_2^u$$

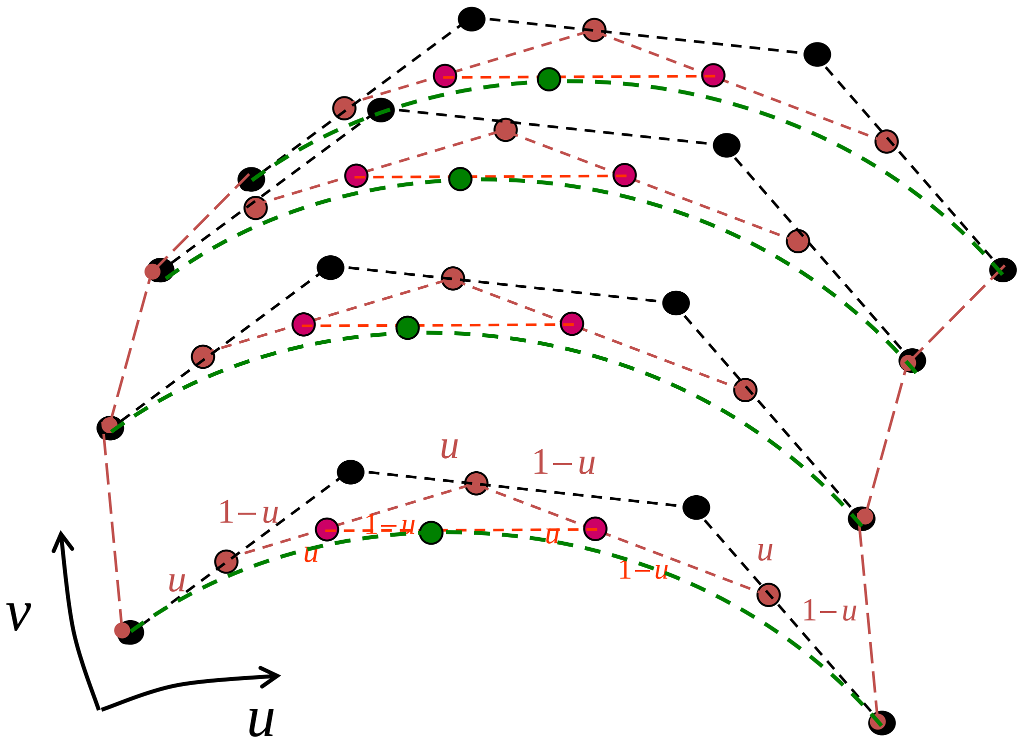
$$\mathbf{r}(u,v) = \begin{bmatrix}(1-v)^2 & 2v(1-v) & v^2\end{bmatrix}\begin{bmatrix}\mathbf{r}_0^u \\ \mathbf{r}_1^u \\ \mathbf{r}_2^u\end{bmatrix}$$

$$\begin{bmatrix}\mathbf{r}_0^u \\ \mathbf{r}_1^u \\ \mathbf{r}_2^u\end{bmatrix} = \begin{bmatrix}\mathbf{b}_{00} & \mathbf{b}_{10} & \mathbf{b}_{20} \\ \mathbf{b}_{01} & \mathbf{b}_{11} & \mathbf{b}_{21} \\ \mathbf{b}_{02} & \mathbf{b}_{12} & \mathbf{b}_{22}\end{bmatrix}\begin{bmatrix}(1-u)^2 \\ 2u(1-u) \\ u^2\end{bmatrix}$$

$$\mathbf{r}(u,v) = \begin{bmatrix}(1-v)^2 & 2v(1-v) & v^2\end{bmatrix}\begin{bmatrix}\mathbf{b}_{00} & \mathbf{b}_{10} & \mathbf{b}_{20} \\ \mathbf{b}_{01} & \mathbf{b}_{11} & \mathbf{b}_{21} \\ \mathbf{b}_{02} & \mathbf{b}_{12} & \mathbf{b}_{22}\end{bmatrix}\begin{bmatrix}(1-u)^2 \\ 2u(1-u) \\ u^2\end{bmatrix}$$

3) Bicubic Bezier Surface Patch

- Given: 4x4 Bezier control points
- Find: Points on the bicubic Bezier Surface Patch

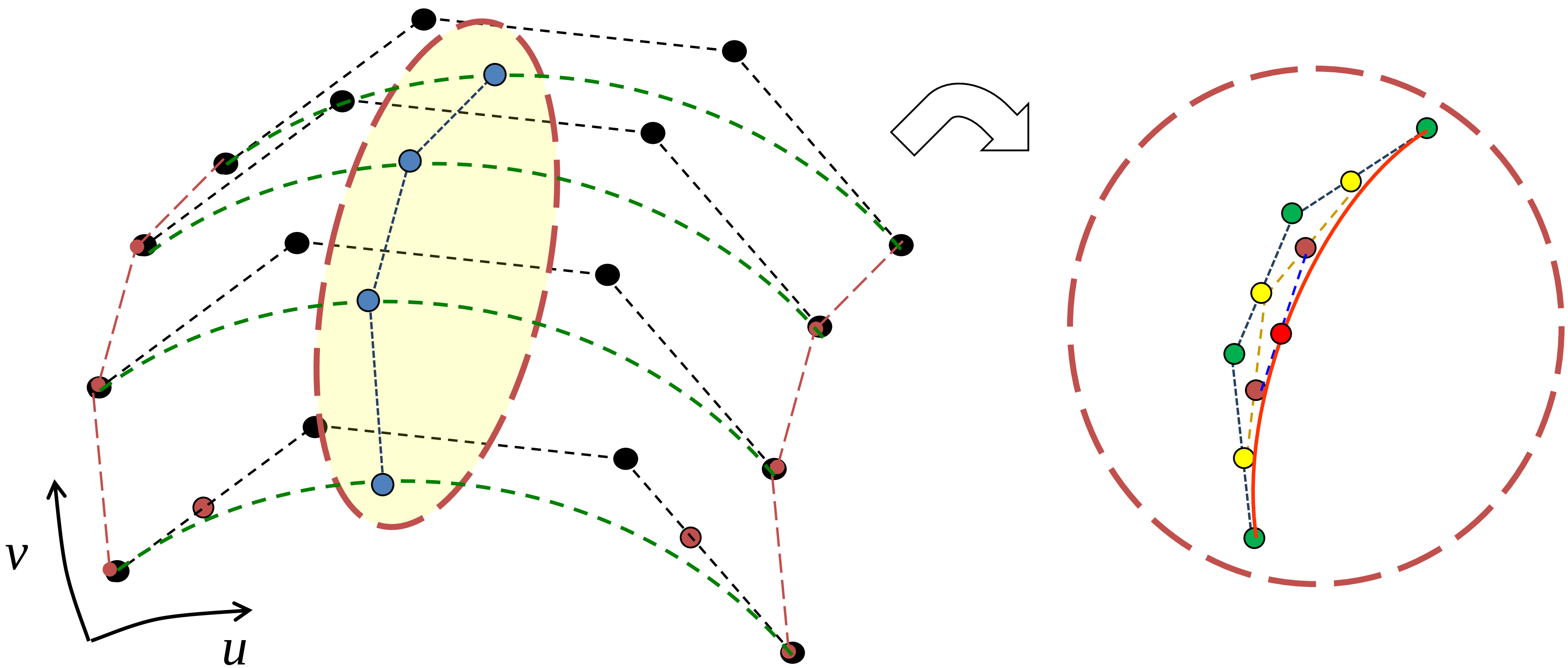


$$\mathbf{b}(u,v)=\begin{bmatrix} B_0^3(u) & B_1^3(u) & B_2^3(u) & B_3^3(u) \end{bmatrix}\begin{bmatrix} \mathbf{b}_{0,0} & \mathbf{b}_{1,0} & \mathbf{b}_{2,0} & \mathbf{b}_{3,0} \\ \mathbf{b}_{0,1} & \mathbf{b}_{1,1} & \mathbf{b}_{2,1} & \mathbf{b}_{3,1} \\ \mathbf{b}_{0,2} & \mathbf{b}_{1,2} & \mathbf{b}_{2,2} & \mathbf{b}_{3,2} \\ \mathbf{b}_{0,3} & \mathbf{b}_{1,3} & \mathbf{b}_{2,3} & \mathbf{b}_{3,3} \end{bmatrix}\begin{bmatrix} B_0^3(v) \\ B_1^3(v) \\ B_2^3(v) \\ B_3^3(v) \end{bmatrix}$$

3) Bicubic Bezier Surface Patch

- Given: 4x4 Bezier control points

- Find: Points on the bicubic Bezier Surface Patch

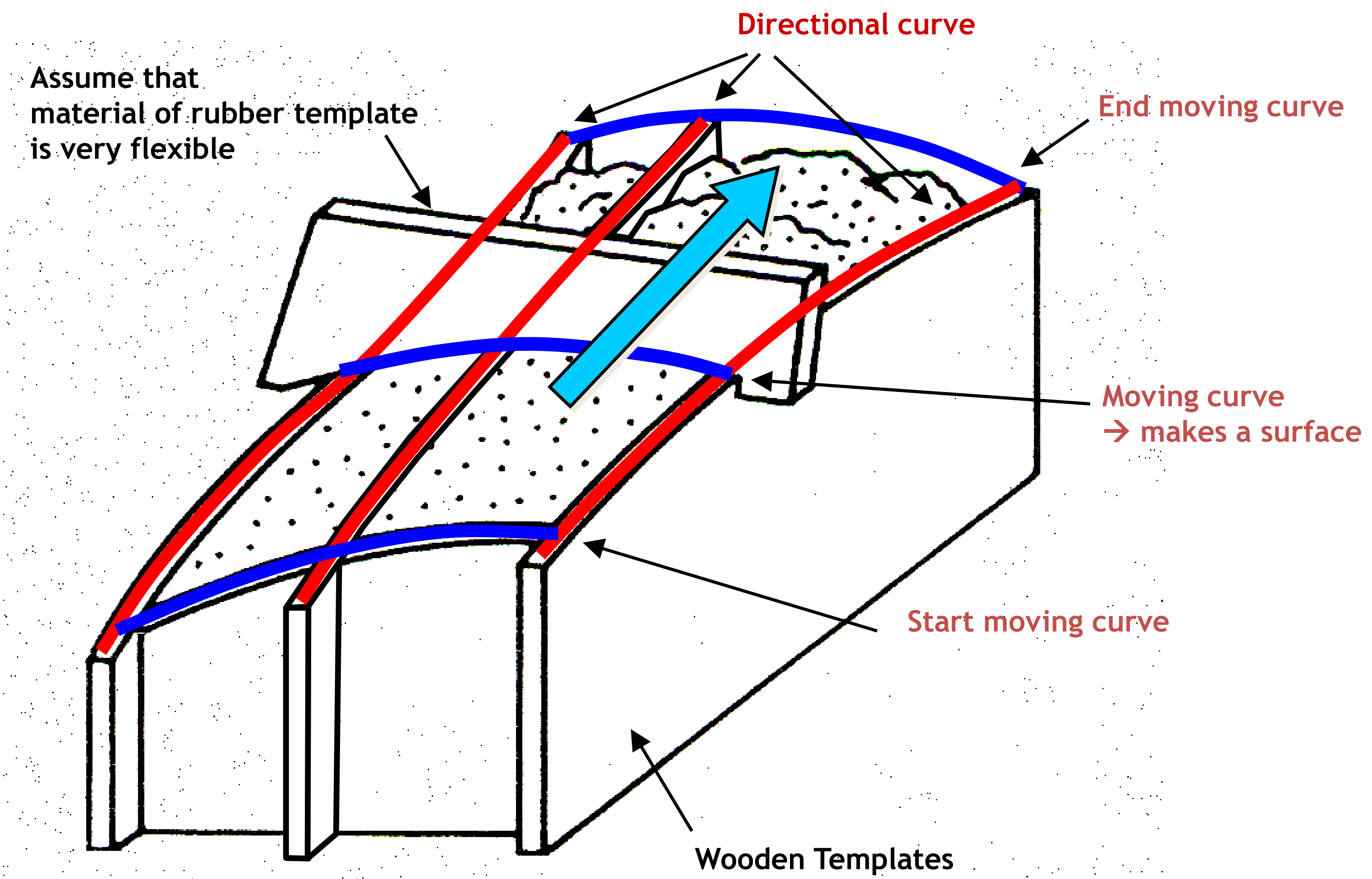


$$\mathbf{b}(u,v)=\begin{bmatrix} B_0^3(u) & B_1^3(u) & B_2^3(u) & B_3^3(u) \end{bmatrix}\begin{bmatrix} \mathbf{b}_{0,0} & \mathbf{b}_{1,0} & \mathbf{b}_{2,0} & \mathbf{b}_{3,0} \\ \mathbf{b}_{0,1} & \mathbf{b}_{1,1} & \mathbf{b}_{2,1} & \mathbf{b}_{3,1} \\ \mathbf{b}_{0,2} & \mathbf{b}_{1,2} & \mathbf{b}_{2,2} & \mathbf{b}_{3,2} \\ \mathbf{b}_{0,3} & \mathbf{b}_{1,3} & \mathbf{b}_{2,3} & \mathbf{b}_{3,3} \end{bmatrix}\begin{bmatrix} B_0^3(v) \\ B_1^3(v) \\ B_2^3(v) \\ B_3^3(v) \end{bmatrix}$$

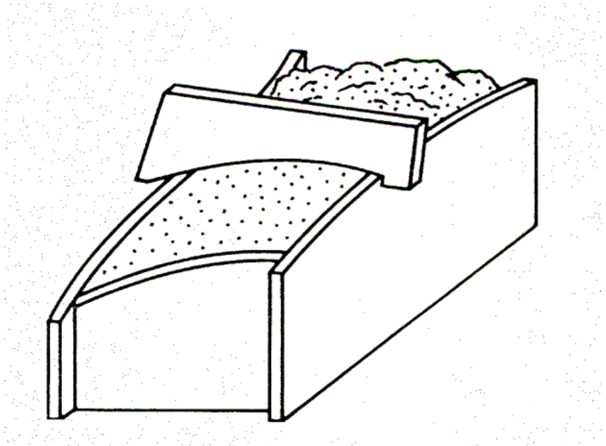
4.2 Bezier Surfaces

1) Generation of Bezier surfaces by de Casteljau algorithm
2) Generation of Bezier surfaces by tensor product approach
  - Tensor product approach
  - Tensor product biquadratic Bezier surface
  - Tensor product bicubic Bezier surface

1) Tensor product approach (1)

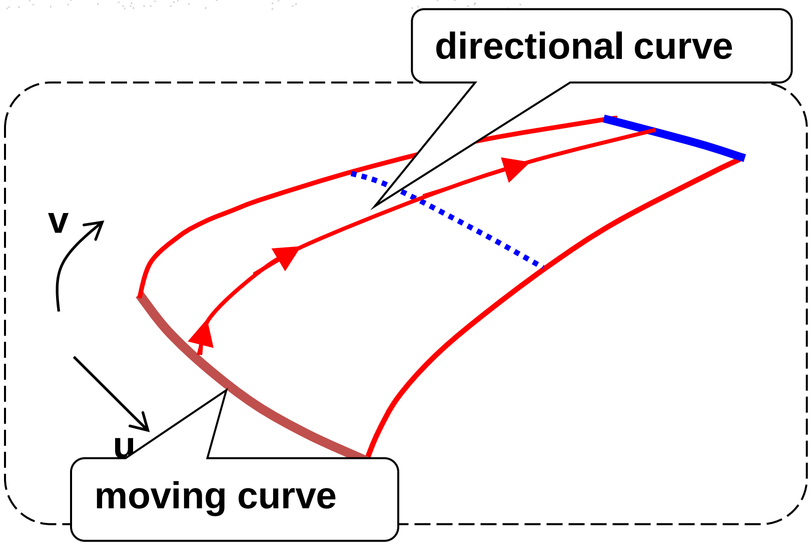


1) Tensor product approach* (2)

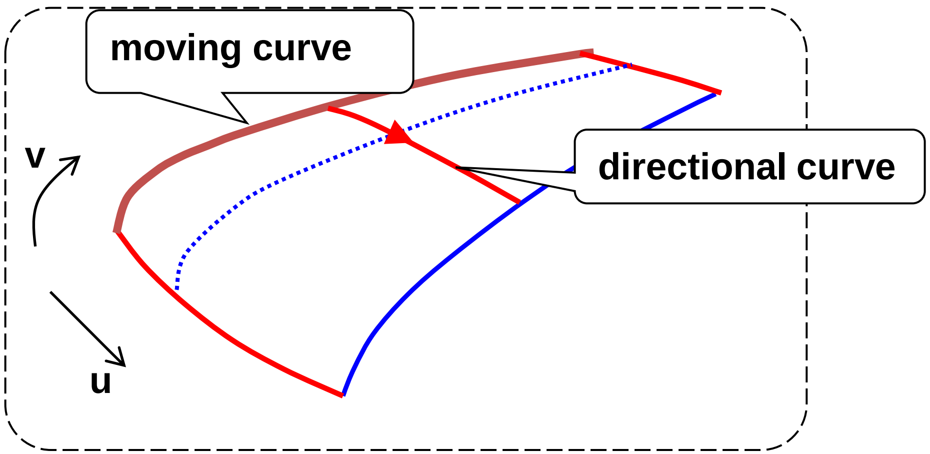


- -Moving curve is order n Bezier curve
  - Trajectories of the Bezier control points of the moving curve, i.e., directional curve is also order m Bezier curve

  The surface generated by sweeping the moving curve is called " Tensor product Bezier Surface Patch"



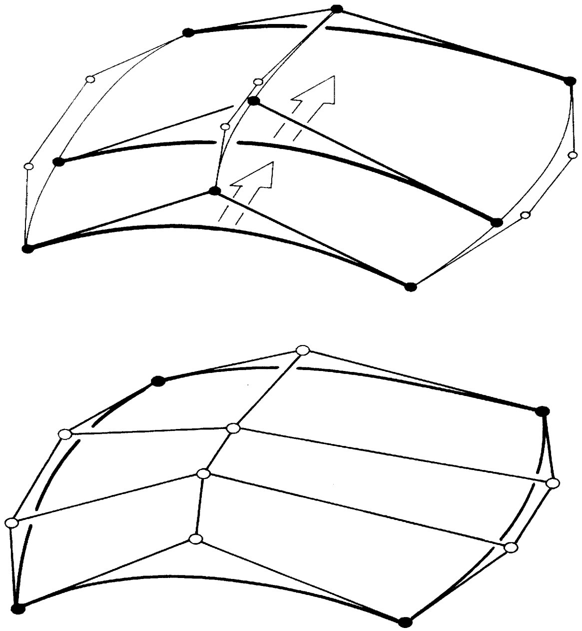
Curve r(*u*) is sweeping in the v-direction



Curve r(v) is sweeping in the u-direction

* Farin, CAGD, 5th Ed., 2002, Ch14.3, Tensor Product Approach

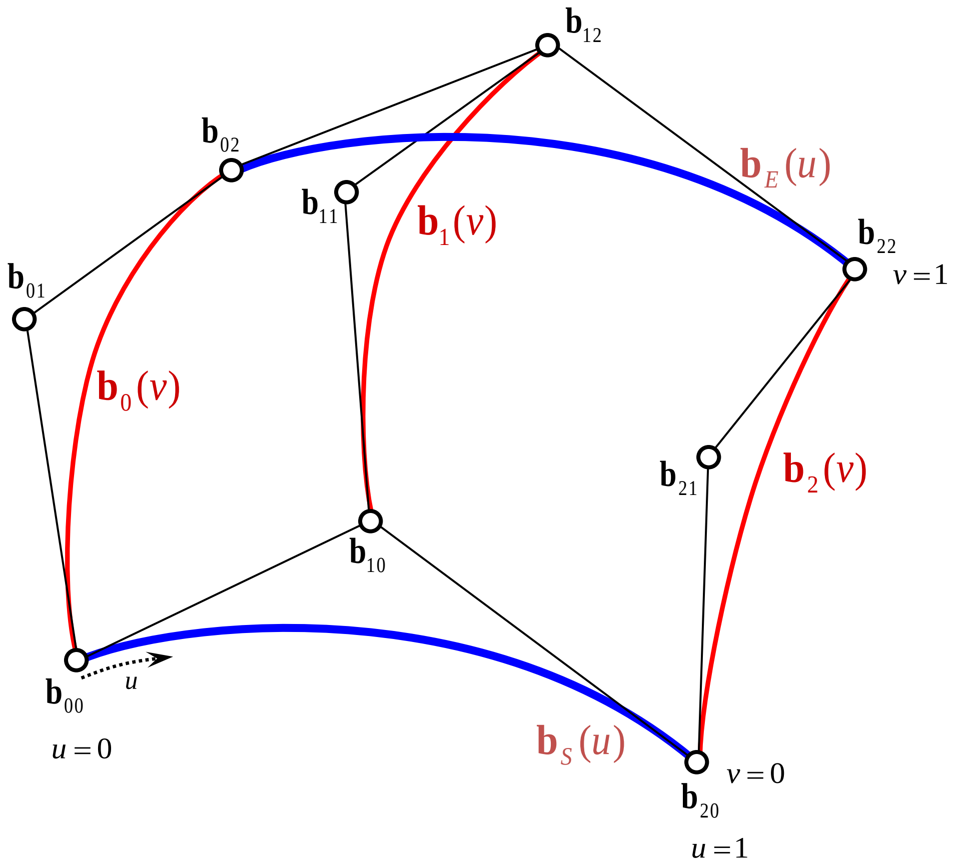
Tensor Bezier Surface



Move control points

2) Tensor product biquadratic Bezier surface (1)

- Given: Control Points of biquadratic Bezier Surface
- Find: Points on the biquadratic Bezier Surface



☑ **Given 3x3 Points $\mathbf{b}_{ij}$,**

☑ **Generate start/end moving curves and directional curves in quadratic Bezier form**

$$\mathbf{b}_E(u) = \mathbf{b}_{02}B_0^2(u) + \mathbf{b}_{12}B_1^2(u) + \mathbf{b}_{22}B_2^2(u)$$

$$\mathbf{b}_S(u) = \mathbf{b}_{00}B_0^2(u) + \mathbf{b}_{10}B_1^2(u) + \mathbf{b}_{20}B_2^2(u)$$

$$\mathbf{b}_0(v) = \mathbf{b}_{00}B_0^2(v) + \mathbf{b}_{01}B_1^2(v) + \mathbf{b}_{02}B_2^2(v)$$

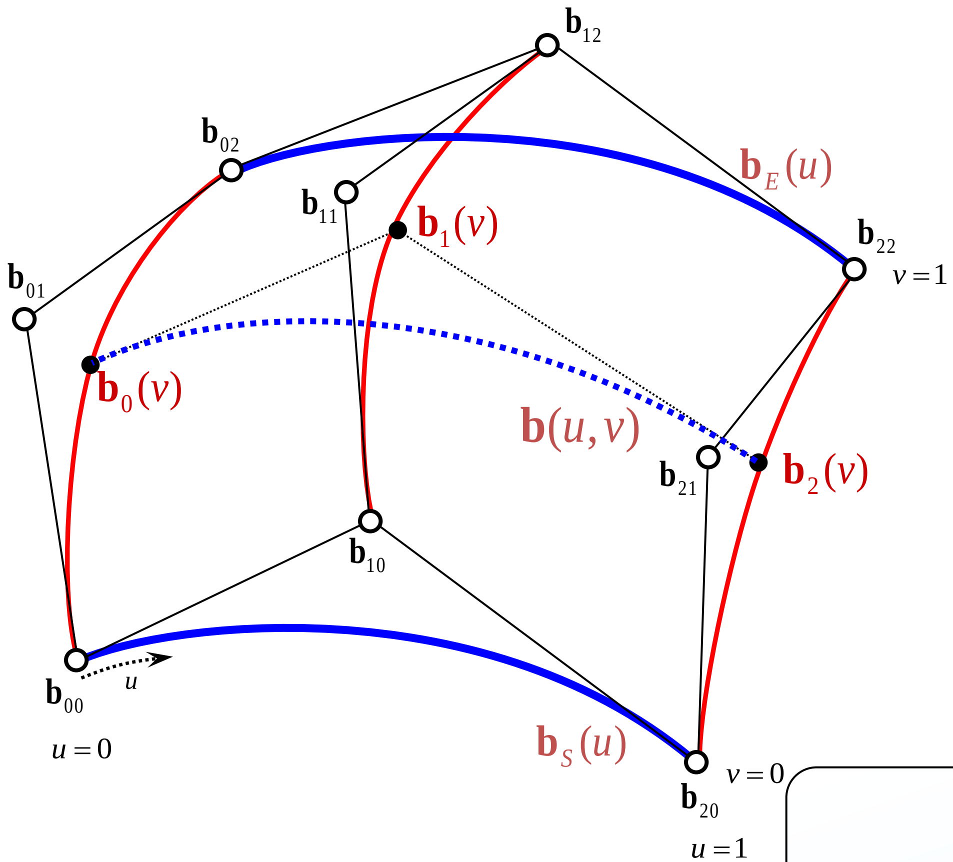
$$\mathbf{b}_1(v) = \mathbf{b}_{10}B_0^2(v) + \mathbf{b}_{11}B_1^2(v) + \mathbf{b}_{12}B_2^2(v)$$

$$\mathbf{b}_2(v) = \mathbf{b}_{20}B_0^2(v) + \mathbf{b}_{21}B_1^2(v) + \mathbf{b}_{22}B_2^2(v)$$

$$\begin{bmatrix} \mathbf{b}_0(v) \\ \mathbf{b}_1(v) \\ \mathbf{b}_2(v) \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{01} & \mathbf{b}_{02} \\ \mathbf{b}_{10} & \mathbf{b}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{20} & \mathbf{b}_{21} & \mathbf{b}_{22} \end{bmatrix} \begin{bmatrix} B_0^2(v) \\ B_1^2(v) \\ B_2^2(v) \end{bmatrix}$$

2) Tensor product biquadratic Bezier surface (2)

- Given: Control Points of biquadratic Bezier Surface
- Find: Points on the biquadratic Bezier Surface



☑ Given 3x3 Points $\mathbf{b}_{ij}$,

☑ Moving curve can be represented in the following form:

$$\mathbf{b}(u,v) = \mathbf{b}_0(v) B_0^2(u) + \mathbf{b}_1(v) B_1^2(u) + \mathbf{b}_2(v) B_2^2(u)$$

$$= \begin{bmatrix} B_0^2(u) & B_1^2(u) & B_2^2(u) \end{bmatrix} \begin{bmatrix} \mathbf{b}_0(v) \\ \mathbf{b}_1(v) \\ \mathbf{b}_2(v) \end{bmatrix}$$

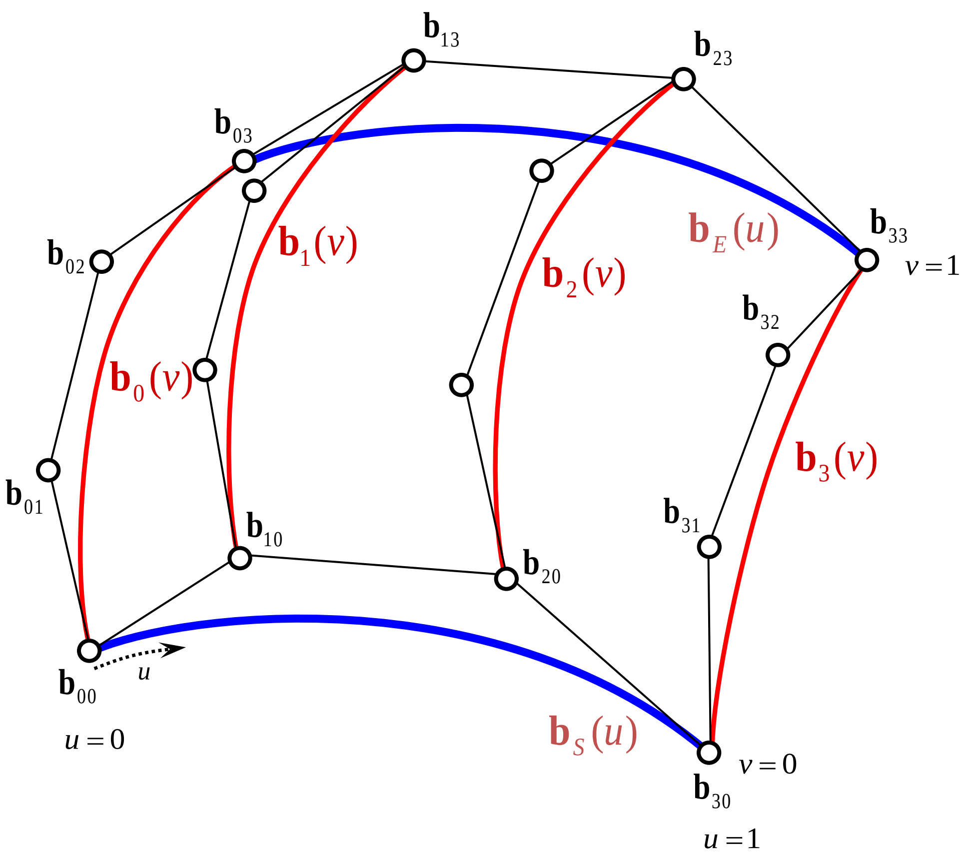
$$\begin{bmatrix} \mathbf{b}_0(v) \\ \mathbf{b}_1(v) \\ \mathbf{b}_2(v) \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{01} & \mathbf{b}_{02} \\ \mathbf{b}_{10} & \mathbf{b}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{20} & \mathbf{b}_{21} & \mathbf{b}_{22} \end{bmatrix} \begin{bmatrix} B_0^2(v) \\ B_1^2(v) \\ B_2^2(v) \end{bmatrix}$$

$$\mathbf{b}(u,v) = \begin{bmatrix} B_0^2(u) & B_1^2(u) & B_2^2(u) \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{01} & \mathbf{b}_{02} \\ \mathbf{b}_{10} & \mathbf{b}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{20} & \mathbf{b}_{21} & \mathbf{b}_{22} \end{bmatrix}}_{\text{Bezier surface control points}} \begin{bmatrix} B_0^2(v) \\ B_1^2(v) \\ B_2^2(v) \end{bmatrix}$$

$$= \sum_{j=0}^{2} \sum_{i=0}^{2} \mathbf{b}_{ij} B_i^2(u) B_j^2(v)$$

3) Tensor product bicubic Bezier surface (1)

- Given: Control Points of bicubic Bezier Surface
- Find: Points on the bicubic Bezier Surface



☑ Given 4x4 Points $\mathbf{b}_{ij}$,

☑ Generate start/end moving curves and directional curves in cubic Bezier form

$$\mathbf{b}_E(u)=\mathbf{b}_{03}B_0^3(u)+\mathbf{b}_{13}B_1^3(u)+\mathbf{b}_{23}B_2^3(u)+\mathbf{b}_{33}B_3^3(u)$$

$$\mathbf{b}_S(u)=\mathbf{b}_{00}B_0^3(u)+\mathbf{b}_{10}B_1^3(u)+\mathbf{b}_{20}B_2^3(u)+\mathbf{b}_{30}B_3^3(u)$$

$$\mathbf{b}_0(v)=\mathbf{b}_{00}B_0^3(v)+\mathbf{b}_{01}B_1^3(v)+\mathbf{b}_{02}B_2^3(v)+\mathbf{b}_{03}B_3^3(v)$$

$$\mathbf{b}_1(v)=\mathbf{b}_{10}B_0^3(v)+\mathbf{b}_{11}B_1^3(v)+\mathbf{b}_{12}B_2^3(v)+\mathbf{b}_{13}B_3^3(v)$$

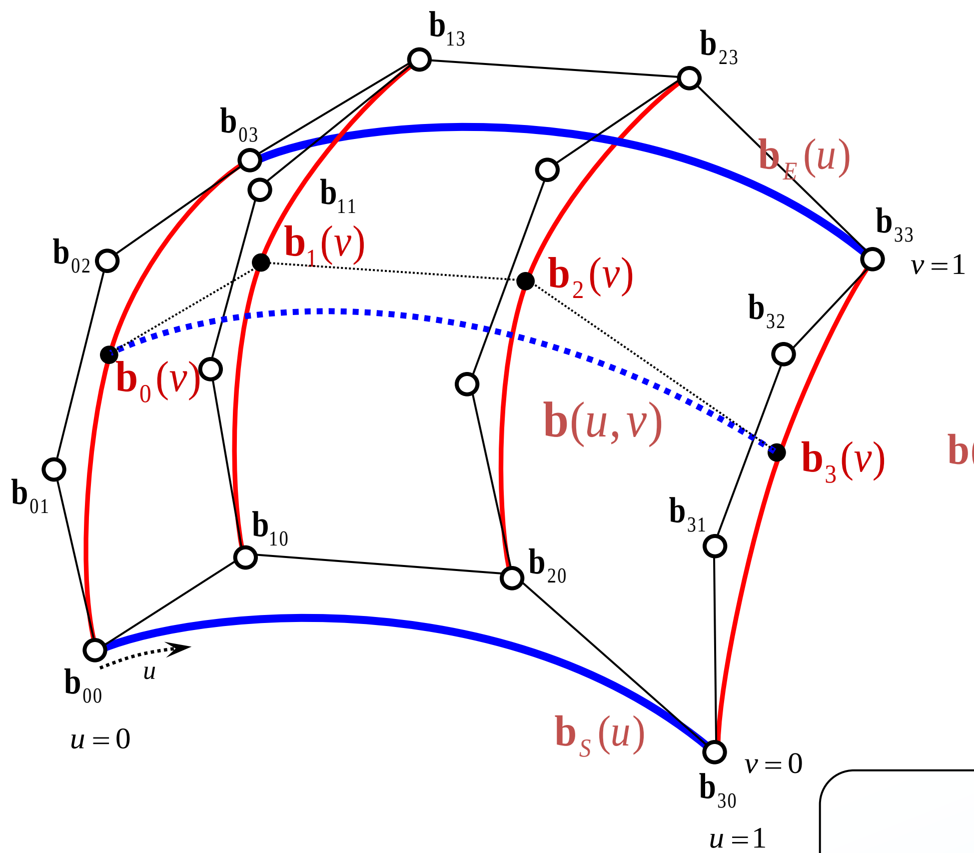
$$\mathbf{b}_2(v)=\mathbf{b}_{20}B_0^3(v)+\mathbf{b}_{21}B_1^3(v)+\mathbf{b}_{22}B_2^3(v)+\mathbf{b}_{23}B_3^3(v)$$

$$\mathbf{b}_3(v)=\mathbf{b}_{30}B_0^3(v)+\mathbf{b}_{31}B_1^3(v)+\mathbf{b}_{32}B_2^3(v)+\mathbf{b}_{33}B_3^3(v)$$

$$\begin{bmatrix}\mathbf{b}_0(v)\\ \mathbf{b}_1(v)\\ \mathbf{b}_2(v)\\ \mathbf{b}_3(v)\end{bmatrix}=\begin{bmatrix}\mathbf{b}_{00}&\mathbf{b}_{01}&\mathbf{b}_{02}&\mathbf{b}_{03}\\ \mathbf{b}_{10}&\mathbf{b}_{11}&\mathbf{b}_{12}&\mathbf{b}_{13}\\ \mathbf{b}_{20}&\mathbf{b}_{21}&\mathbf{b}_{22}&\mathbf{b}_{23}\\ \mathbf{b}_{30}&\mathbf{b}_{31}&\mathbf{b}_{32}&\mathbf{b}_{33}\end{bmatrix}\begin{bmatrix}B_0^3(v)\\ B_1^3(v)\\ B_2^3(v)\\ B_3^3(v)\end{bmatrix}$$

3) Tensor product bicubic Bezier surface (2)

- Given: Control Points of bicubic Bezier Surface
- Find: Points on the bicubic Bezier Surface



☑ **Given 4x4 Points $\mathbf{b}_{ij}$,**

☑ **Moving curve can be represented in the following form:**

$$\mathbf{b}(u,v) = \mathbf{b}_0(v)B_0^3(u) + \mathbf{b}_1(v)B_1^3(u) + \mathbf{b}_2(v)B_2^3(u) + \mathbf{b}_3(v)B_3^3(u)$$

$$= \begin{bmatrix} B_0^3(u) & B_1^3(u) & B_2^3(u) & B_3^3(u) \end{bmatrix} \begin{bmatrix} \mathbf{b}_0(v) \\ \mathbf{b}_1(v) \\ \mathbf{b}_2(v) \\ \mathbf{b}_3(v) \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{b}_0(v) \\ \mathbf{b}_1(v) \\ \mathbf{b}_2(v) \\ \mathbf{b}_3(v) \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{01} & \mathbf{b}_{02} & \mathbf{b}_{03} \\ \mathbf{b}_{10} & \mathbf{b}_{11} & \mathbf{b}_{12} & \mathbf{b}_{13} \\ \mathbf{b}_{20} & \mathbf{b}_{21} & \mathbf{b}_{22} & \mathbf{b}_{23} \\ \mathbf{b}_{30} & \mathbf{b}_{31} & \mathbf{b}_{32} & \mathbf{b}_{33} \end{bmatrix} \begin{bmatrix} B_0^3(v) \\ B_1^3(v) \\ B_2^3(v) \\ B_3^3(v) \end{bmatrix}$$

$$\mathbf{b}(u,v) = \begin{bmatrix} B_0^3(u) & B_1^3(u) & B_2^3(u) & B_3^3(u) \end{bmatrix} \begin{bmatrix} \mathbf{b}_{00} & \mathbf{b}_{01} & \mathbf{b}_{02} & \mathbf{b}_{03} \\ \mathbf{b}_{10} & \mathbf{b}_{11} & \mathbf{b}_{12} & \mathbf{b}_{13} \\ \mathbf{b}_{20} & \mathbf{b}_{21} & \mathbf{b}_{22} & \mathbf{b}_{23} \\ \mathbf{b}_{30} & \mathbf{b}_{31} & \mathbf{b}_{32} & \mathbf{b}_{33} \end{bmatrix} \begin{bmatrix} B_0^3(v) \\ B_1^3(v) \\ B_2^3(v) \\ B_3^3(v) \end{bmatrix}$$

$$= \sum_{j=0}^{3} \sum_{i=0}^{3} \mathbf{b}_{ij} B_i^3(u) B_j^3(v)$$

4.3 B-spline Surfaces

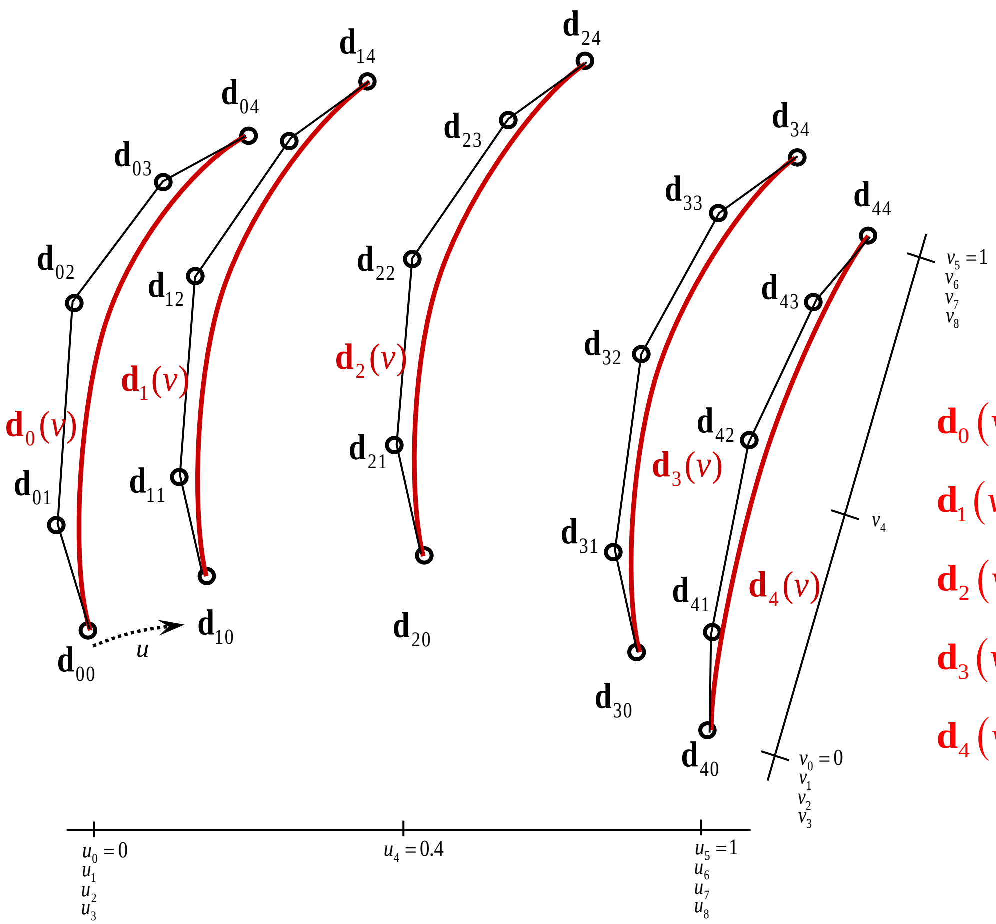
Generation of B-spline surfaces by tensor product approach

- Tensor product bicubic B-spline surface

- Programming Guide for Tensor product bicubic B-spline surface

1) Tensor product bicubic B-spline surface (1)

- Given: Control Points of bicubic B-spline surface
- Find: Points on the bicubic B-spline surface



☑ **Given 5x5 Control Points $\mathbf{d}_{ij}$, u-knots, v-knots, u-degree(=3), v-degree(=3),**

☑ **Generate start/end moving curves and directional curves in cubic B-spline form:**

$$\mathbf{d}_0(v)=\mathbf{d}_{00}N_0^3(v)+\mathbf{d}_{01}N_1^3(v)+\mathbf{d}_{02}N_2^3(v)+\mathbf{d}_{03}N_3^3(v)+\mathbf{d}_{04}N_4^3(v)$$

$$\mathbf{d}_1(v)=\mathbf{d}_{10}N_0^3(v)+\mathbf{d}_{11}N_1^3(v)+\mathbf{d}_{12}N_2^3(v)+\mathbf{d}_{13}N_3^3(v)+\mathbf{d}_{14}N_4^3(v)$$

$$\mathbf{d}_2(v)=\mathbf{d}_{20}N_0^3(v)+\mathbf{d}_{21}N_1^3(v)+\mathbf{d}_{22}N_2^3(v)+\mathbf{d}_{23}N_3^3(v)+\mathbf{d}_{24}N_4^3(v)$$

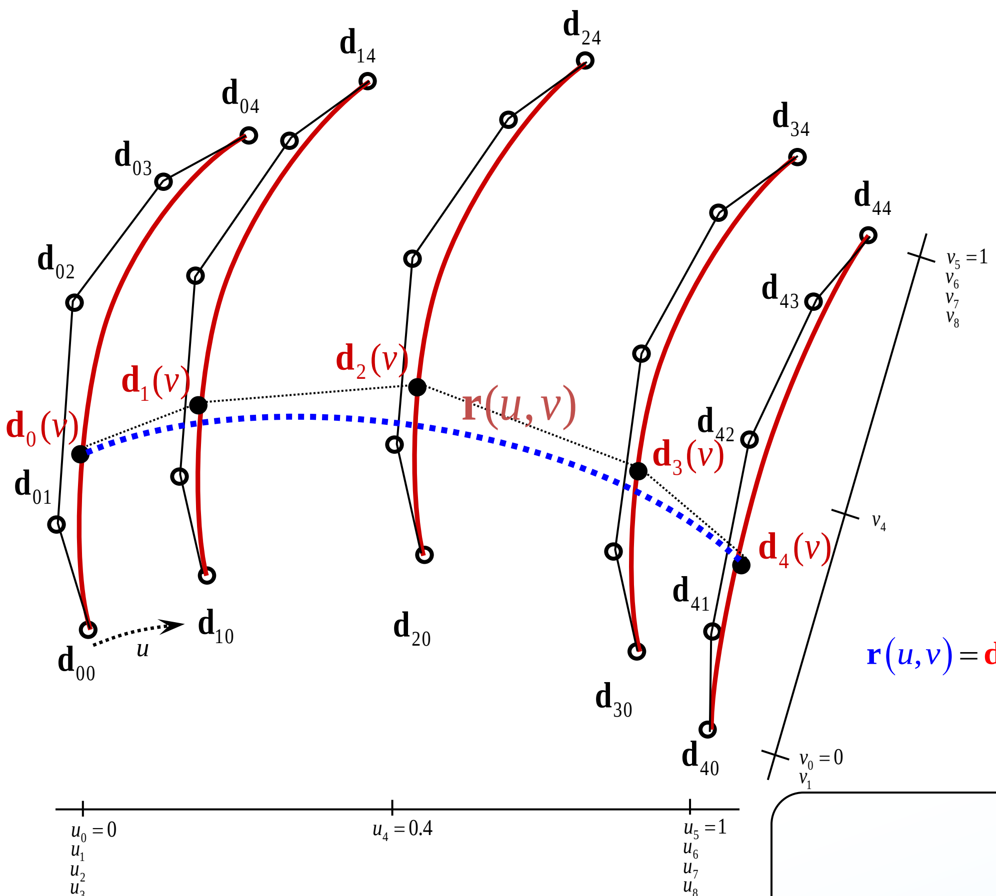
$$\mathbf{d}_3(v)=\mathbf{d}_{30}N_0^3(v)+\mathbf{d}_{31}N_1^3(v)+\mathbf{d}_{32}N_2^3(v)+\mathbf{d}_{33}N_3^3(v)+\mathbf{d}_{34}N_4^3(v)$$

$$\mathbf{d}_4(v)=\mathbf{d}_{40}N_0^3(v)+\mathbf{d}_{41}N_1^3(v)+\mathbf{d}_{42}N_2^3(v)+\mathbf{d}_{43}N_3^3(v)+\mathbf{d}_{44}N_4^3(v)$$

$$\begin{bmatrix}\mathbf{d}_0(v)\\ \mathbf{d}_1(v)\\ \mathbf{d}_2(v)\\ \mathbf{d}_3(v)\\ \mathbf{d}_4(v)\end{bmatrix}=\begin{bmatrix}\mathbf{d}_{00}&\mathbf{d}_{01}&\mathbf{d}_{02}&\mathbf{d}_{03}&\mathbf{d}_{04}\\ \mathbf{d}_{10}&\mathbf{d}_{11}&\mathbf{d}_{12}&\mathbf{d}_{13}&\mathbf{d}_{14}\\ \mathbf{d}_{20}&\mathbf{d}_{21}&\mathbf{d}_{22}&\mathbf{d}_{23}&\mathbf{d}_{24}\\ \mathbf{d}_{30}&\mathbf{d}_{31}&\mathbf{d}_{32}&\mathbf{d}_{33}&\mathbf{d}_{34}\\ \mathbf{d}_{40}&\mathbf{d}_{41}&\mathbf{d}_{42}&\mathbf{d}_{43}&\mathbf{d}_{44}\end{bmatrix}\begin{bmatrix}N_0^3(v)\\ N_1^3(v)\\ N_2^3(v)\\ N_3^3(v)\\ N_4^3(v)\end{bmatrix}$$

1) Tensor-product bicubic B-spline surface (1)

- Given: Control Points of bicubic B-spline surface
- Find: Points on the bicubic B-spline surface



☑ **Given 5x5 Control Points $\mathbf{d}_{ij}$, u-knots, v-knots, u-degree(=3), v-degree(=3),**

☑ **Moving curve can be represented in the following form:**

$$= \begin{bmatrix} N_0^3(u) & N_1^3(u) & N_2^3(u) & N_3^3(u) & N_4^3(u) \end{bmatrix} \begin{bmatrix} \mathbf{d}_0(v) \\ \mathbf{d}_1(v) \\ \mathbf{d}_2(v) \\ \mathbf{d}_3(v) \\ \mathbf{d}_4(v) \end{bmatrix}$$

$$\mathbf{r}(u,v) = \mathbf{d}_0(v)N_0^3(u) + \mathbf{d}_1(v)N_1^3(u) + \mathbf{d}_2(v)N_2^3(u) + \mathbf{d}_3(v)N_3^3(u) + \mathbf{d}_4(v)N_4^3(u)$$

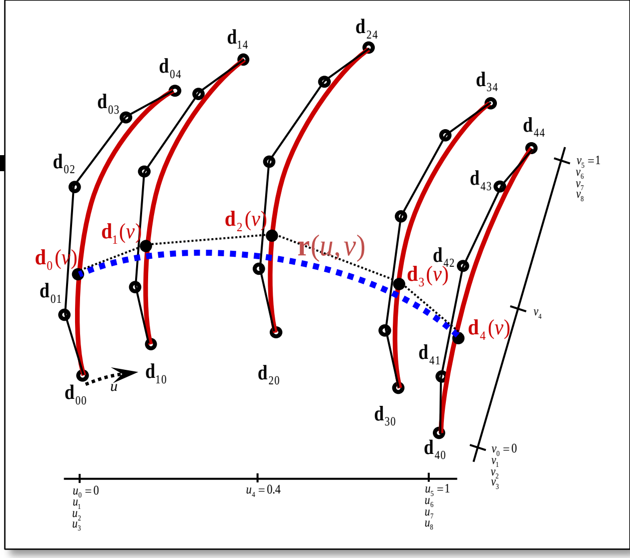
$$\begin{bmatrix} \mathbf{d}_0(v) \\ \mathbf{d}_1(v) \\ \mathbf{d}_2(v) \\ \mathbf{d}_3(v) \\ \mathbf{d}_4(v) \end{bmatrix} = \begin{bmatrix} \mathbf{d}_{00} & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} & \mathbf{d}_{04} \\ \mathbf{d}_{10} & \mathbf{d}_{11} & \mathbf{d}_{12} & \mathbf{d}_{13} & \mathbf{d}_{14} \\ \mathbf{d}_{20} & \mathbf{d}_{21} & \mathbf{d}_{22} & \mathbf{d}_{23} & \mathbf{d}_{24} \\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & \mathbf{d}_{33} & \mathbf{d}_{34} \\ \mathbf{d}_{40} & \mathbf{d}_{41} & \mathbf{d}_{42} & \mathbf{d}_{43} & \mathbf{d}_{44} \end{bmatrix} \begin{bmatrix} N_0^3(v) \\ N_1^3(v) \\ N_2^3(v) \\ N_3^3(v) \\ N_4^3(v) \end{bmatrix}$$

$$\mathbf{r}(u,v) = \begin{bmatrix} N_0^3(u) & N_1^3(u) & N_2^3(u) & N_3^3(u) & N_4^3(u) \end{bmatrix} \begin{bmatrix} \mathbf{d}_{00} & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} & \mathbf{d}_{04} \\ \mathbf{d}_{10} & \mathbf{d}_{11} & \mathbf{d}_{12} & \mathbf{d}_{13} & \mathbf{d}_{14} \\ \mathbf{d}_{20} & \mathbf{d}_{21} & \mathbf{d}_{22} & \mathbf{d}_{23} & \mathbf{d}_{24} \\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & \mathbf{d}_{33} & \mathbf{d}_{34} \\ \mathbf{d}_{40} & \mathbf{d}_{41} & \mathbf{d}_{42} & \mathbf{d}_{43} & \mathbf{d}_{44} \end{bmatrix} \begin{bmatrix} N_0^3(v) \\ N_1^3(v) \\ N_2^3(v) \\ N_3^3(v) \\ N_4^3(v) \end{bmatrix}$$

$$= \sum_{j=0}^{5} \sum_{i=0}^{5} \mathbf{d}_{ij} N_i^3(u) N_j^3(v)$$

2) Programming Guide for Tensor product bicubic B-spline surface(1)

- Member Variables of Class




```
class CBsplineSurface
{
public:
        // member variables
        int m_nDegree;

        double* m_pKnot_U;
        int m_nNumOfKnot_U;

        double* m_pKnot_V;
        int m_nNumOfKnot_V;

        Vector** m_pCP;
        int m_nNumOfCP_U;
        int m_nNumOfCP_V;

        // member functions
        ...
};
```

➔ Degree (n=3)

➔ Number and Value of Knot in *u* -direction (9)

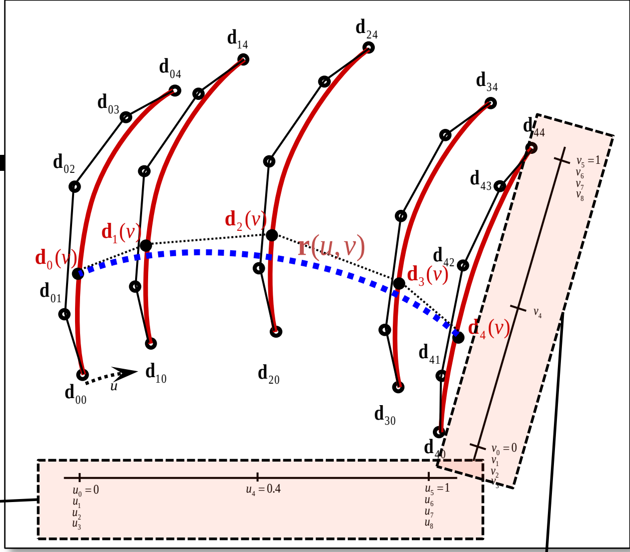
➔ Number and Value of Knot in *u* -direction (9)

➔ Number of Control Points in *u- and v-directions*

(*u* :5, *v* :5)

2) Programming Guide for Tensor product bicubic B-spline surface(2)

- Member Functions of Class



Knots in *u-direction*

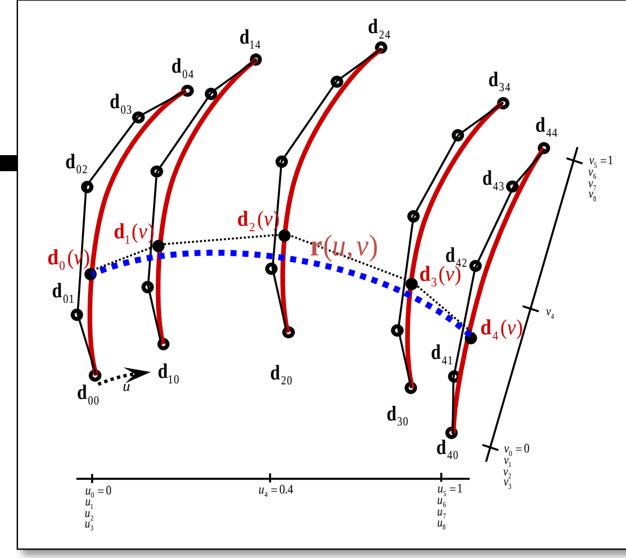
Knots in *u-direction*

```
class CBsplineSurface
{
public:
            // member variables
            …

            // member functions
            …

            void SetKnot(double* pKnot_U, int nNumOfKnot_U, double* pKnot_V, int
nNumOfKnot_V);

            double N(int n, int i, double u, int uv);

            Vector GetPoint(double u, double v);
};
```

➔ Define the Knots

2) Programming Guide for Tensor product bicubic B-spline surface(3)

- Member Functions of Class




```
class CBsplineSurface
{
public:
        // member variables
        ...

        // member functions
        ...

        void SetKnot(double* pKnot_U, int nNumOfKnot_U, double* pKnot_V, int
nNumOfKnot_V);

        double N(int n, int i, double u, int uv);

        Vector GetPoint(double u, double v);
};
```

Flag to distinguish u- and v-direction

➔ Calculate the B-spline Basis Function

B-spline Basis Function

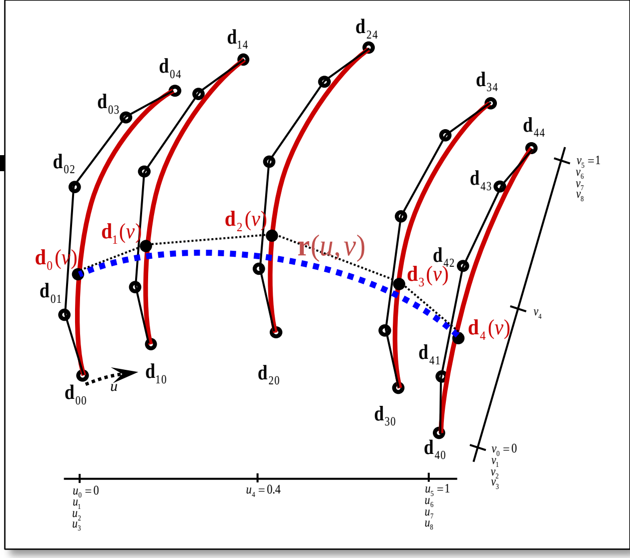
(Cox-de Boor Recurrence Formula)

$$N_i^n(u) = \frac{u-u_{i-1}}{u_{i+n-1}-u_{i-1}} N_i^{n-1}(u) + \frac{u_{i+n}-u}{u_{i+n}-u_i} N_{i+1}^{n-1}(u)$$

$$N_i^0(u) = \begin{cases} 1 & \text{if } u_{i-1} \le u < u_i \\ 0 & \text{else} \end{cases}$$

2) Programming Guide for Tensor product bicubic B-spline surface(4)

- Member Functions of Class




```
class CBsplineSurface
{
public:
            // member variables
            …

            // member functions
            …

            void SetKnot(double* pKnot_U, int nNumOfKnot_U, double* pKnot_V, int
nNumOfKnot_V);

            double N(int n, int i, double u, int uv);

            Vector GetPoint(double u, double v);
};
```

➔Calculate the points on the Surface for given Parameter ***u*, *v***

$$\mathbf{r}(u,v)=\begin{bmatrix}N_0^3(u) & N_1^3(u) & N_2^3(u) & N_3^3(u) & N_4^3(u)\end{bmatrix}\begin{bmatrix}\mathbf{d}_{00} & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} & \mathbf{d}_{04}\\ \mathbf{d}_{10} & \mathbf{d}_{11} & \mathbf{d}_{12} & \mathbf{d}_{13} & \mathbf{d}_{14}\\ \mathbf{d}_{20} & \mathbf{d}_{21} & \mathbf{d}_{22} & \mathbf{d}_{23} & \mathbf{d}_{24}\\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & \mathbf{d}_{33} & \mathbf{d}_{34}\\ \mathbf{d}_{40} & \mathbf{d}_{41} & \mathbf{d}_{42} & \mathbf{d}_{43} & \mathbf{d}_{44}\end{bmatrix}\begin{bmatrix}N_0^3(v)\\ N_1^3(v)\\ N_2^3(v)\\ N_3^3(v)\\ N_4^3(v)\end{bmatrix}$$

$$=\sum_{j=0}^{5}\sum_{i=0}^{5}\mathbf{d}_{ij}N_i^3(u)N_j^3(v)$$

2) Programming Guide for Tensor product bicubic B-spline surface(5)

- Member Function Example 'GetPoint'

```
Vector CBsplineSurface::GetPoint(double u, double v)
{
  // return value
  Vector r_u_v(0.0, 0.0, 0.0);

  // get curve
  for (int i=0; i<m_nNumOfCP_U; i++)
  {
    Vector r_v(0.0, 0.0, 0.0);

    for (int j=0; j<m_nNumOfCP_V; j++)
    {
      r_v = r_v + m_pCP[i][j] * N(m_nDegree, j, v, ID_V);
    }

    r_u_v = r_u_v + N(m_nDegree, i, u, ID_U) * r_v;
  }

  return r_u_v;
}
```

➔Calculate the points on the Surface for given Parameter *u*, *v*

$$\mathrm{d}_0(v) = \mathrm{d}_{00}N_0^3(v) + \mathrm{d}_{01}N_1^3(v) + \mathrm{d}_{02}N_2^3(v) + \mathrm{d}_{03}N_3^3(v) + \mathrm{d}_{04}N_4^3(v)$$

$$\mathbf{r}(u,v) = \begin{bmatrix} N_0^3(u) & N_1^3(u) & N_2^3(u) & N_3^3(u) & N_4^3(u) \end{bmatrix} \begin{bmatrix} \mathbf{d}_{00} & \mathbf{d}_{01} & \mathbf{d}_{02} & \mathbf{d}_{03} & \mathbf{d}_{04} \\ \mathbf{d}_{10} & \mathbf{d}_{11} & \mathbf{d}_{12} & \mathbf{d}_{13} & \mathbf{d}_{14} \\ \mathbf{d}_{20} & \mathbf{d}_{21} & \mathbf{d}_{22} & \mathbf{d}_{23} & \mathbf{d}_{24} \\ \mathbf{d}_{30} & \mathbf{d}_{31} & \mathbf{d}_{32} & \mathbf{d}_{33} & \mathbf{d}_{34} \\ \mathbf{d}_{40} & \mathbf{d}_{41} & \mathbf{d}_{42} & \mathbf{d}_{43} & \mathbf{d}_{44} \end{bmatrix} \begin{bmatrix} N_0^3(v) \\ N_1^3(v) \\ N_2^3(v) \\ N_3^3(v) \\ N_4^3(v) \end{bmatrix}$$

$$= \sum_{j=0}^{5}\sum_{i=0}^{5} \mathbf{d}_{ij} N_i^3(u) N_j^3(v)$$

4.4 B-spline Surface Interpolation

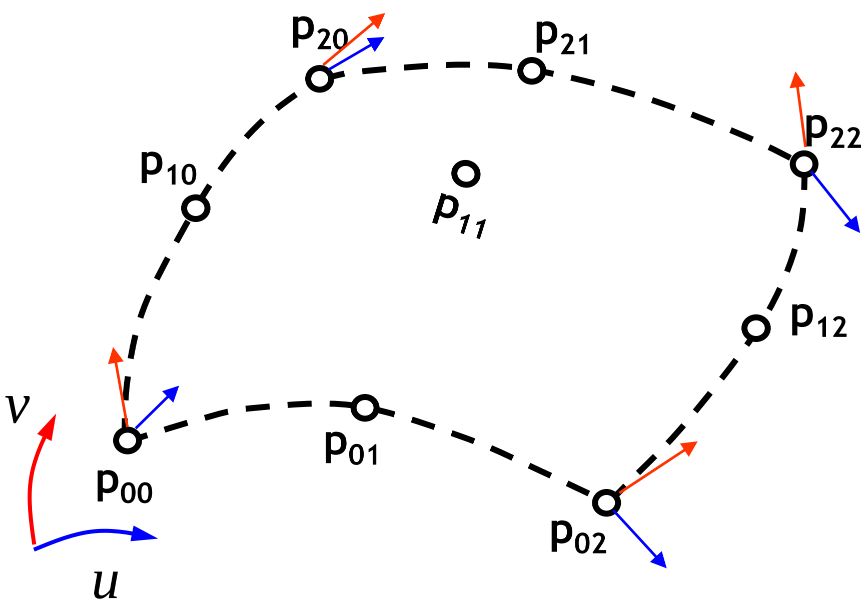
1) Bicubic B-spline surface interpolation

2) Determination of knot values

3) Sample code of bicubic B-spline Surface Interpolation

1) Bicubic B-spline Surface Interpolation (1)

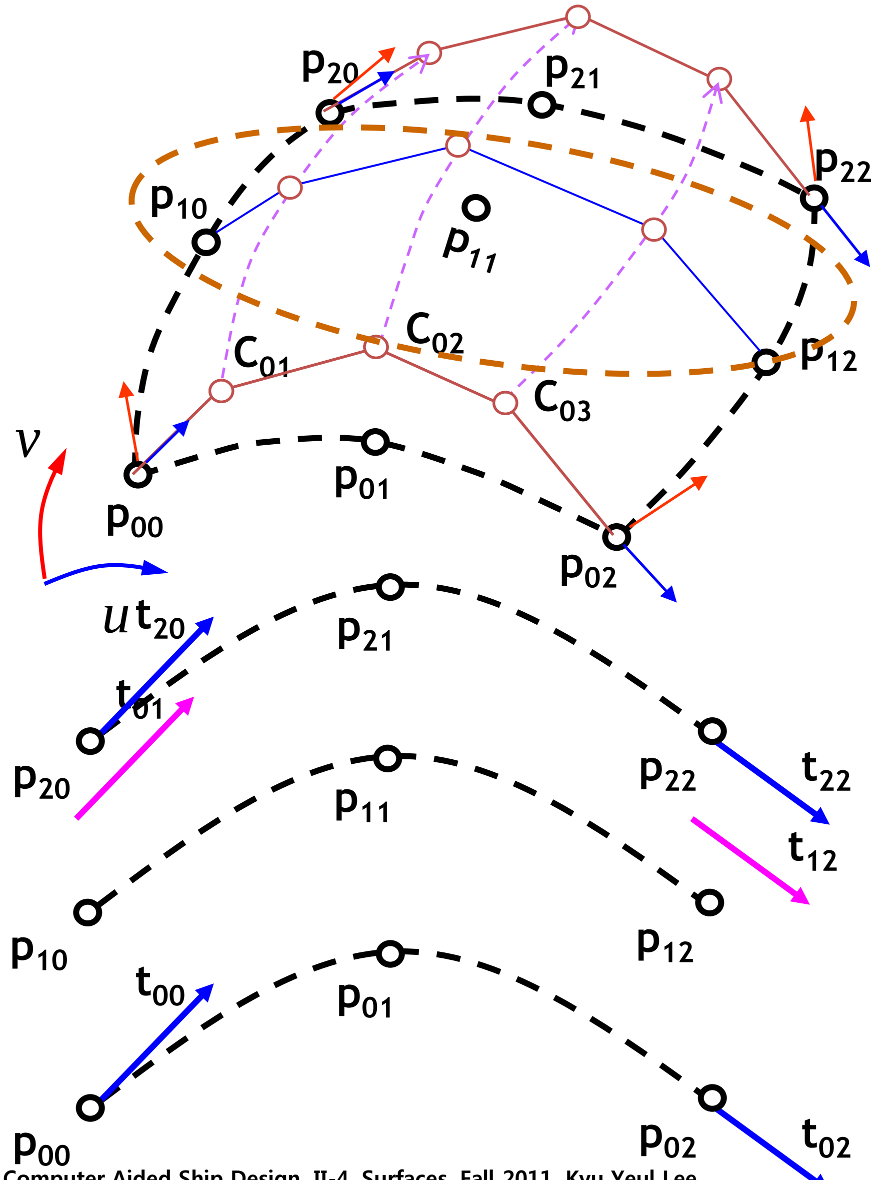
- Given: 9 points on the surface and Tangent vectors at four corners in u- and v- directions
- Find: Control Points of bicubic B-spline Surface



$$\mathbf{r}(u,v) = \begin{bmatrix} N_0^3(u) & N_1^3(u) & N_2^3(u) & N_3^3(u) & N_4^3(u) \end{bmatrix} \begin{bmatrix} \mathrm{d}_{00} & \mathrm{d}_{01} & \mathrm{d}_{02} & \mathrm{d}_{03} & \mathrm{d}_{04} \\ \mathrm{d}_{10} & \mathrm{d}_{11} & \mathrm{d}_{12} & \mathrm{d}_{13} & \mathrm{d}_{14} \\ \mathrm{d}_{20} & \mathrm{d}_{21} & \mathrm{d}_{22} & \mathrm{d}_{23} & \mathrm{d}_{24} \\ \mathrm{d}_{30} & \mathrm{d}_{31} & \mathrm{d}_{32} & \mathrm{d}_{33} & \mathrm{d}_{34} \\ \mathrm{d}_{40} & \mathrm{d}_{41} & \mathrm{d}_{42} & \mathrm{d}_{43} & \mathrm{d}_{44} \end{bmatrix} \begin{bmatrix} N_0^3(v) \\ N_1^3(v) \\ N_2^3(v) \\ N_3^3(v) \\ N_4^3(v) \end{bmatrix}$$

1) Bicubic B-spline Surface Interpolation (2)

- Given: 9 points on the surface and Tangent vectors at four corners in u- and v- directions
- Find: Control Points of bicubic B-spline Surface



Intermediate control points ($C_{i,j}$) are determined by the points on the surface ($P_{i,j}$) and tangent vectors at ends ($t_{i,j}$)

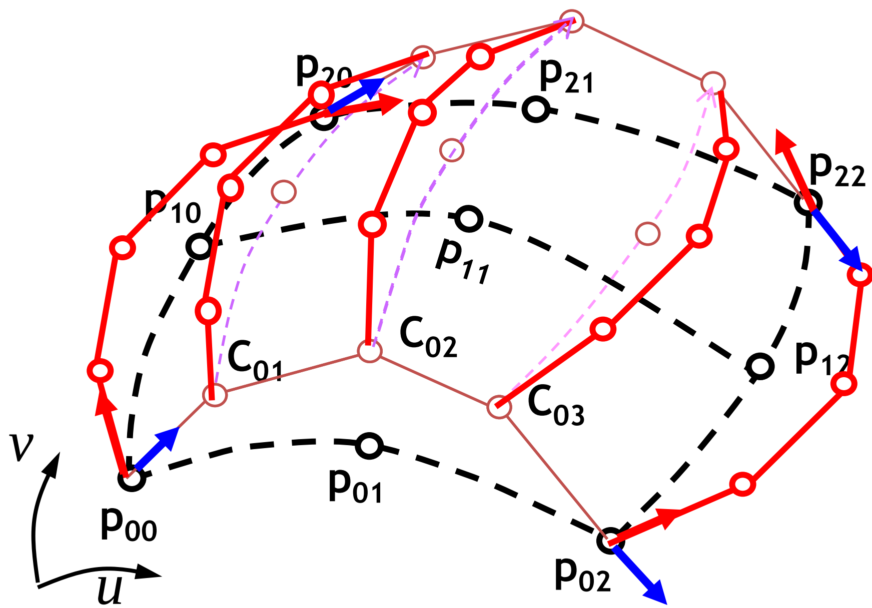
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -\frac{3}{\Delta_s} & \frac{3}{\Delta_s} & 0 & 0 & 0 \\ 0 & \alpha & \beta & \gamma & 0 \\ 0 & 0 & 0 & -\frac{3}{\Delta_E} & \frac{3}{\Delta_E} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{C}_{00} \\ \mathbf{C}_{01} \\ \mathbf{C}_{02} \\ \mathbf{C}_{03} \\ \mathbf{C}_{04} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{00} \\ \mathbf{t}_{00} \\ \mathbf{P}_{01} \\ \mathbf{t}_{02} \\ \mathbf{P}_{02} \end{bmatrix}$$

Tangent vectors at ends ($t_{i,j}$) are determined by using the Bessel end conditions

Bessel end condition:

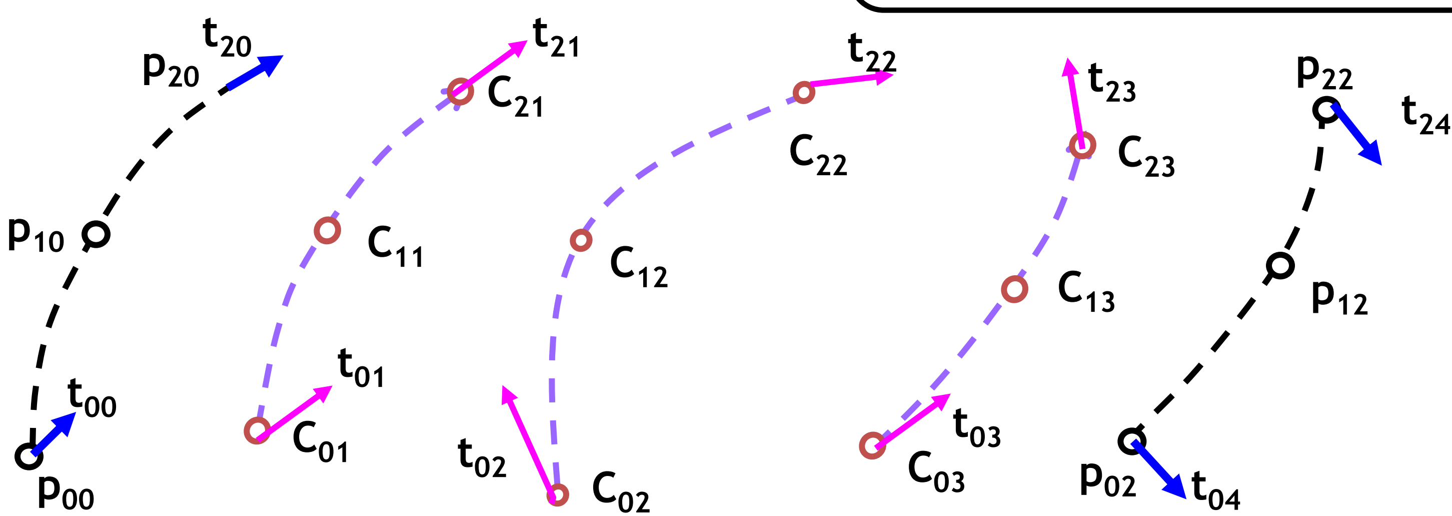
1) Bicubic B-spline Surface Interpolation (3)

- Given: 9 points on the surface and Tangent vectors at four corners in u- and v- directions
- Find: Control Points of bicubic B-spline Surface

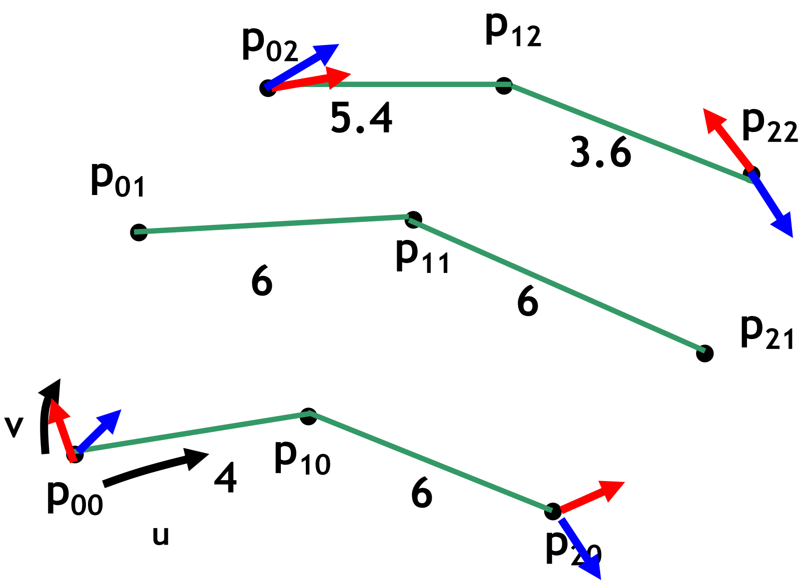


B-spline surface control points ($d_{i,j}$) are determined by the intermediate control points ($C_{i,j}$) and tangent vectors at ends ($t_{i,j}$)

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -\dfrac{3}{\Delta_s} & \dfrac{3}{\Delta_s} & 0 & 0 & 0 \\ 0 & \alpha & \beta & \gamma & 0 \\ 0 & 0 & 0 & -\dfrac{3}{\Delta_E} & \dfrac{3}{\Delta_E} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_{02} \\ d_{12} \\ d_{22} \\ d_{32} \\ d_{42} \end{bmatrix} = \begin{bmatrix} C_{02} \\ t_{02} \\ C_{12} \\ t_{22} \\ C_{22} \end{bmatrix}$$

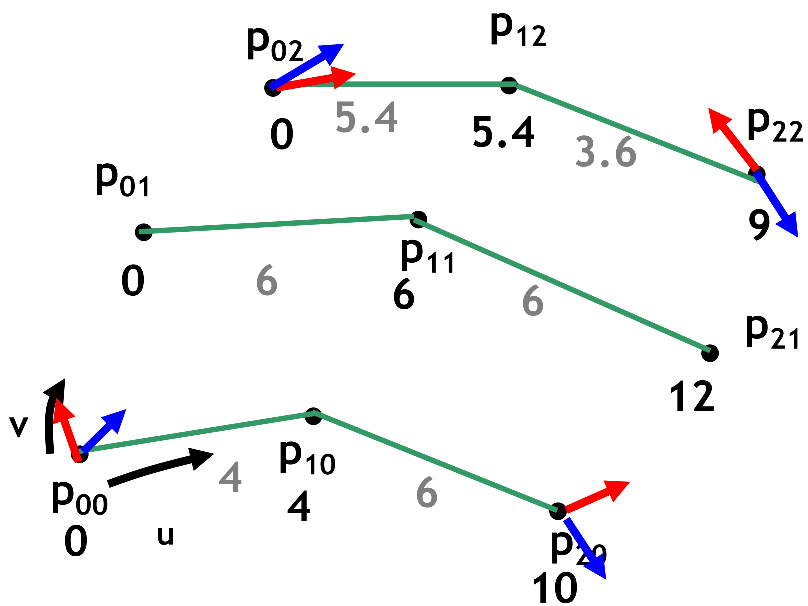


2) Determination of knot values (1)



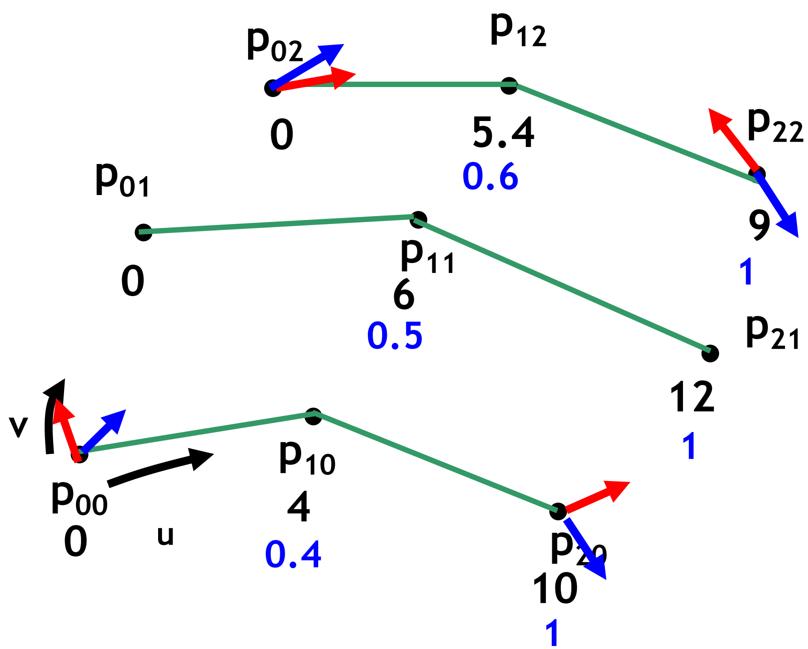
1. Determine the knots in u-direction
   - Calculate the distances between the points ($P_{i,j}$)

2) Determination of knot values (2)



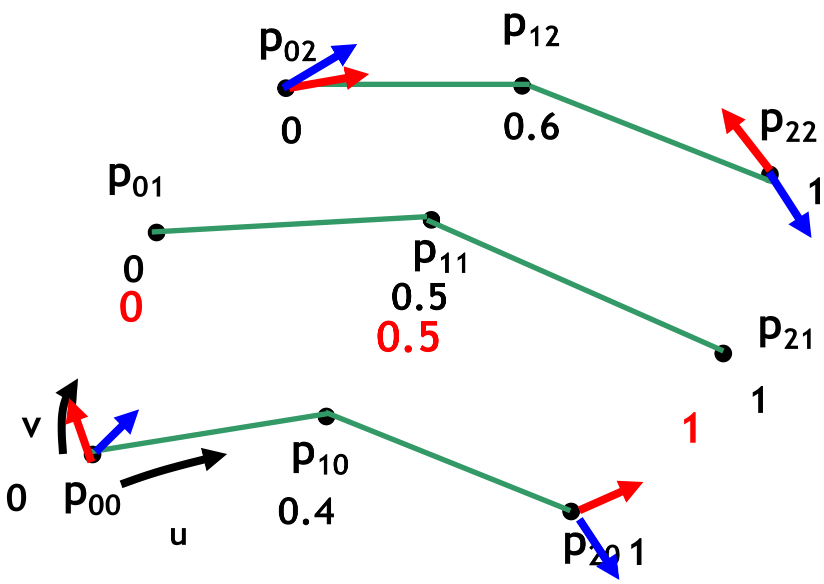
1. **Determine the knots in u-direction:**
   - Calculate the distances between the points ($P_{i,j}$)
   - Sum up the distances at each point. These accumulated distances are knots value in u-direction.

2) Determination of knot values (3)



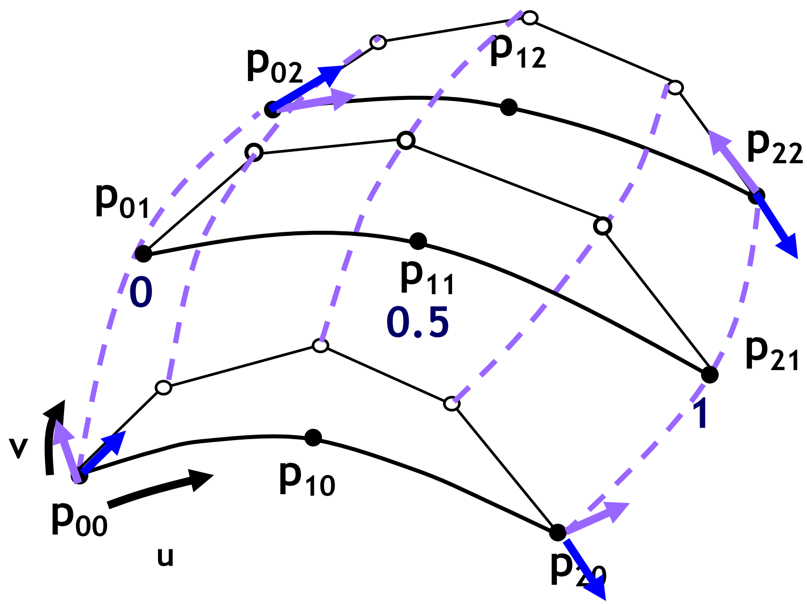
1. Determine the knots in u-direction:
   - Calculate the distances between the points ($P_{i,j}$)
   - Sum up the distances at each point. This accumulated distances are knots value in u-direction.
   - Normalize the knot values at each point by dividing with the knot value at end point.

2) Determination of knot values (4)



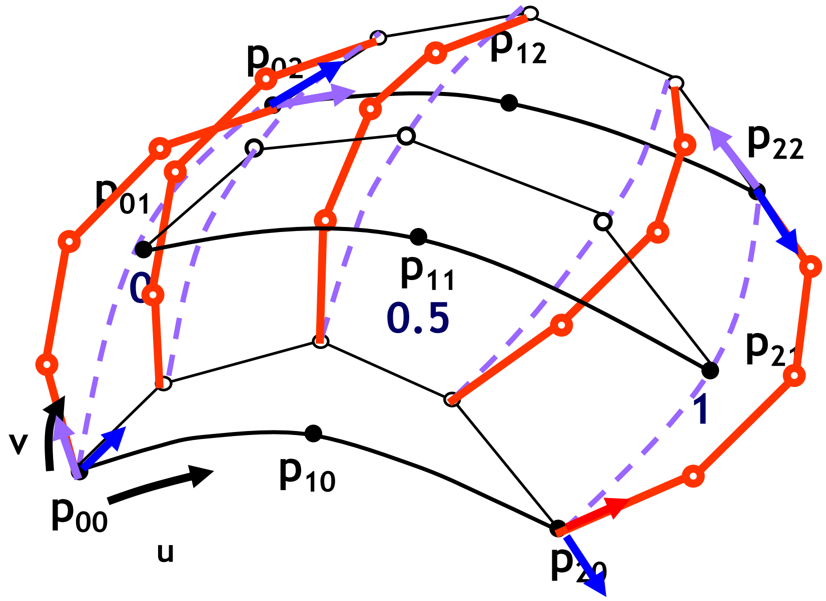
1. **Determine the knots in u-direction:**
   - Determine reference knot values in u-direction by calculating average knot values for each v-direction

2) Determination of knot values (5)



1. Determine the knots in u-direction

2) Determination of knot values (6)



3. In the same manner, determine the knots in v-direction

2) Determination of knot values (7)

Calculate chord length with given points on the curves

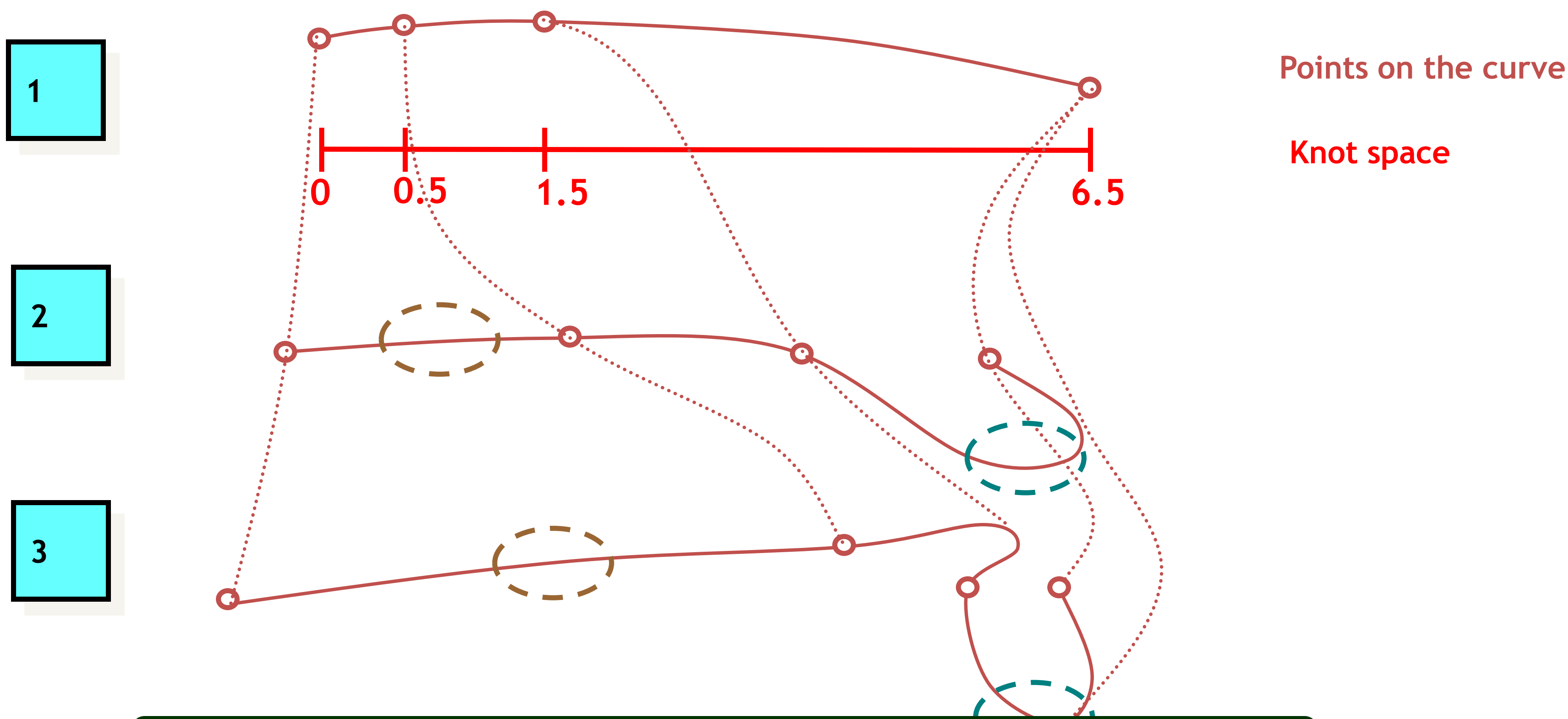
$$\frac{\Delta_i}{\Delta_{i+1}} = \sqrt{\frac{|p_{i-1} - p_{i-2}|}{|p_i - p_{i-1}|}}, \ (i = 2,3,....,n+2)$$

To generate a 'smooth' surface, positions of the points on the surface should be 'regular'

It means the space of the knots should be same.

2) Determination of knot values (8)
- Effect of the Knot space on the quality of the B-spline surfaces

Knot space can be regarded as the elapsed time to travel the points.



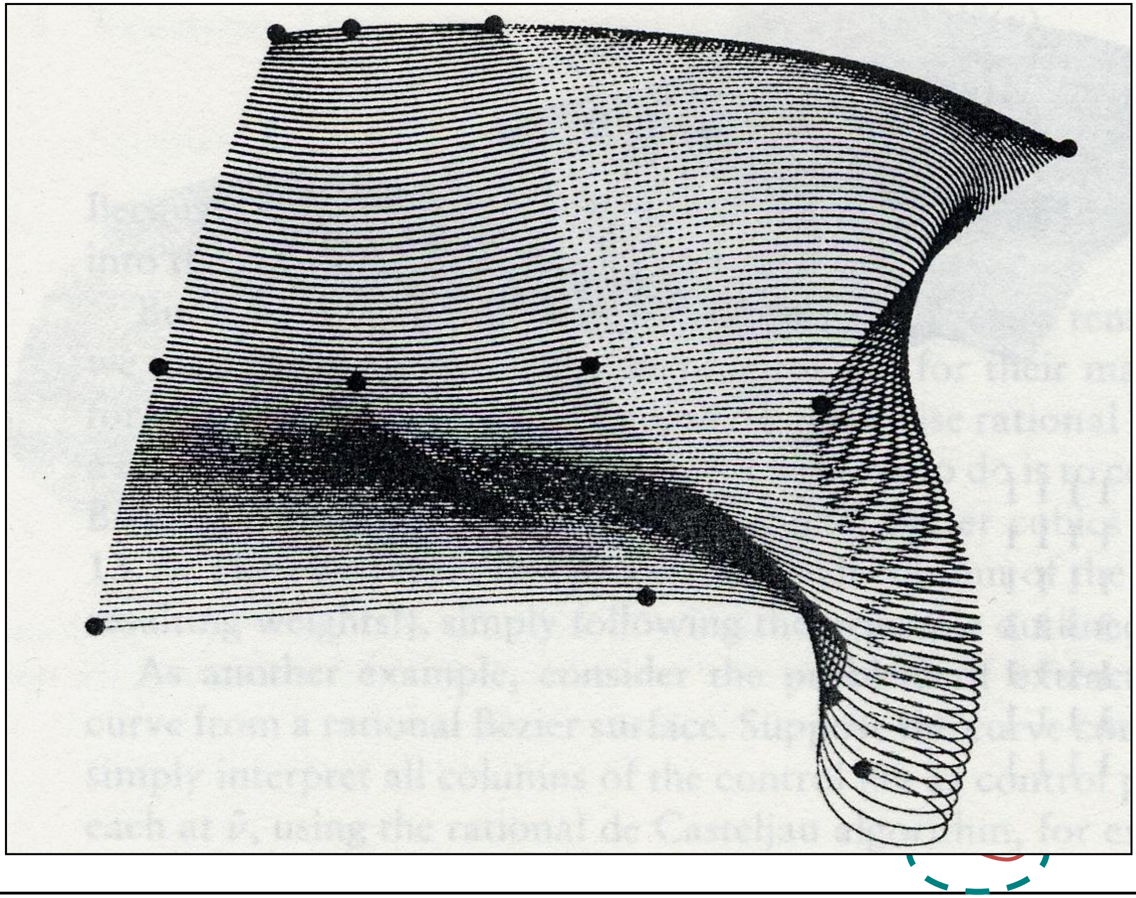
If the knot spaces are irregular , the surface will be strong distorted

2) Determination of knot values (9)
- Effect of the Knot space on the quality of the B-spline surfaces

1

2

3



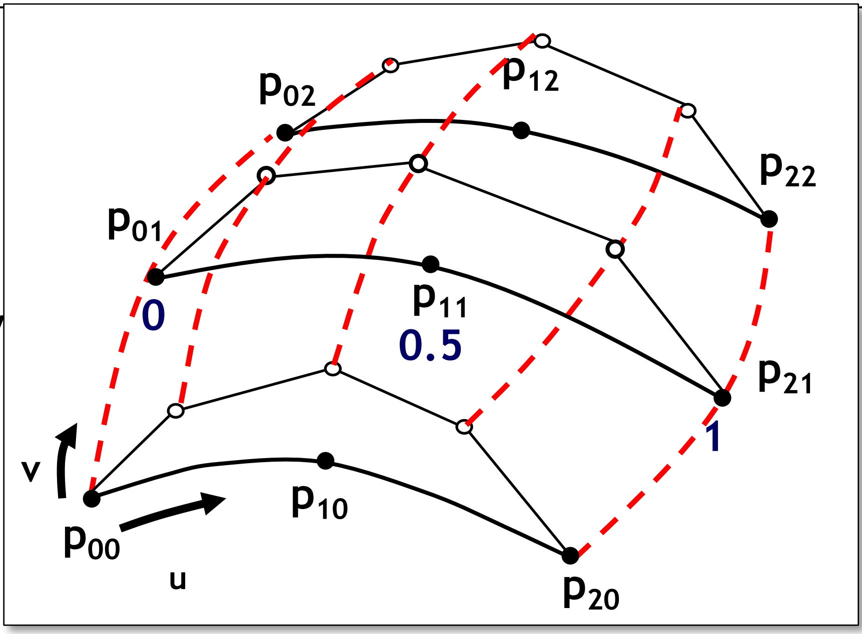
3) Sample code of bicubic B-spline Surface Interpolation (1)

```
void BicubicBsplineSurface::Interpolate(Vector **pFittingPoint, int nU, int nV) {
  // Generate u-Knot
    if(m_UKnot) delete[] m_UKnot;
    m_nUKnot = (m_nU - 2) + 2*(3+1);
    m_UKnot = new double [m_nUKnot];

    // Initial u-Knot
    double** tmpUKnots;
    tmpUKnots = new double*[nV];
    for(int j = 0; j < nV; j++){
        tmpUKnots[j] = new double[nU];
        for(int i = 0; i < nU; i++){
            tmpUKnot[j][i] = ...;   // chord length or centripetal
        }
    }
    // generate average u-Knot
   for(int i = 0; i < nU; i++){
        m_UKnot[i] = 0;
        for(int j=0; j<nV; j++) {        m_UKnot[i] += tmpUKnot[j][i];     }
        m_UKnot[i] = m_UKnot[i] / nV;
  }
```

➔ Calculate the knot values
➔ corresponding to the chord length

➔ Recalculate the knot values
➔ with average knot value

3) Sample code of bicubic B-spline Surface Interpolation (2)

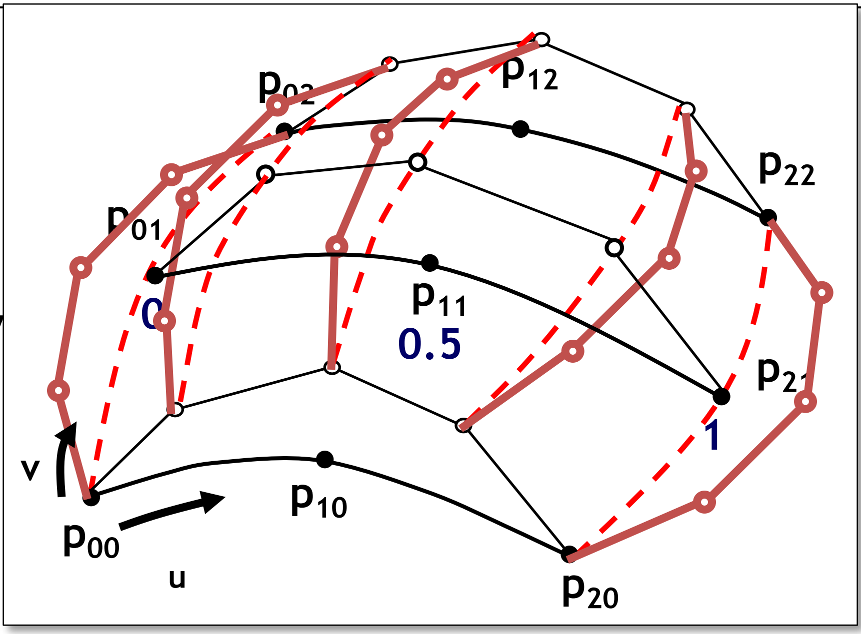



```
// Interpolate u-directional B-spline curve
 CubicBsplineCurve* u_curve = new CubicBsplineCurve[nV
 for(int j = 0; j < nV; j++){
     u_curve[j].SetKnot( m_UKnot );
     u_curve[j].Interpolate(  pFittingPoint[j], nU );
 }

 // Generate  Fitting Points in the  v-direction
 int nvFittingPoint = u_curve[0].m_nControlPoint;
 Vector** vFittingPoint = new Vector [ nvFittingPoint ];
 for(int j=0; j < nvFittingPoint; j++){
     vFittingPoint[j] = new Vector[ nV ];
     for( int i = 0; i < nV; i++){
          vFittingPoint[j][i] = u_curve[i].m_ControlPoint[j];
     }
 }
 ……………………….
```

3) Sample code of bicubic B-spline Surface Interpolation (3)

```
// Interpolate B-spline curves in the u-direction
 CubicBsplineCurve* u_curve = new CubicBsplineCurve[nV
 for(int j = 0; j < nV; j++){
     u_curve[j].SetKnot( m_UKnot );
     u_curve[j].Interpolate(  pFittingPoint[j], nU );
 }
```

```
// Generate v-directional Fitting Point
int nvFittingPoint = u_curve[0].m_nControlPoint;
Vector** vFittingPoint = new Vector [ nvFittingPoint ];
for(int j=0; j < nvFittingPoint; j++){
    vFittingPoint[j] = new Vector[ nV ];
    for( int i = 0; i < nV; i++){
        vFittingPoint[j][i] = u_curve[i].m_ControlPoint[j];
    }
}
……………………….
```



2011년 2학기 전산선박설계 강의자료
(Computer Aided Ship Design Lecture Note)

Part III. Finite Element Method

서울대학교 조선해양공학과
선박설계자동화연구실
이규열

CONTENTS

Naval Architecture & Ocean Engineering

Computer Aided Ship Design
-Part III Finite Element Method

November 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University

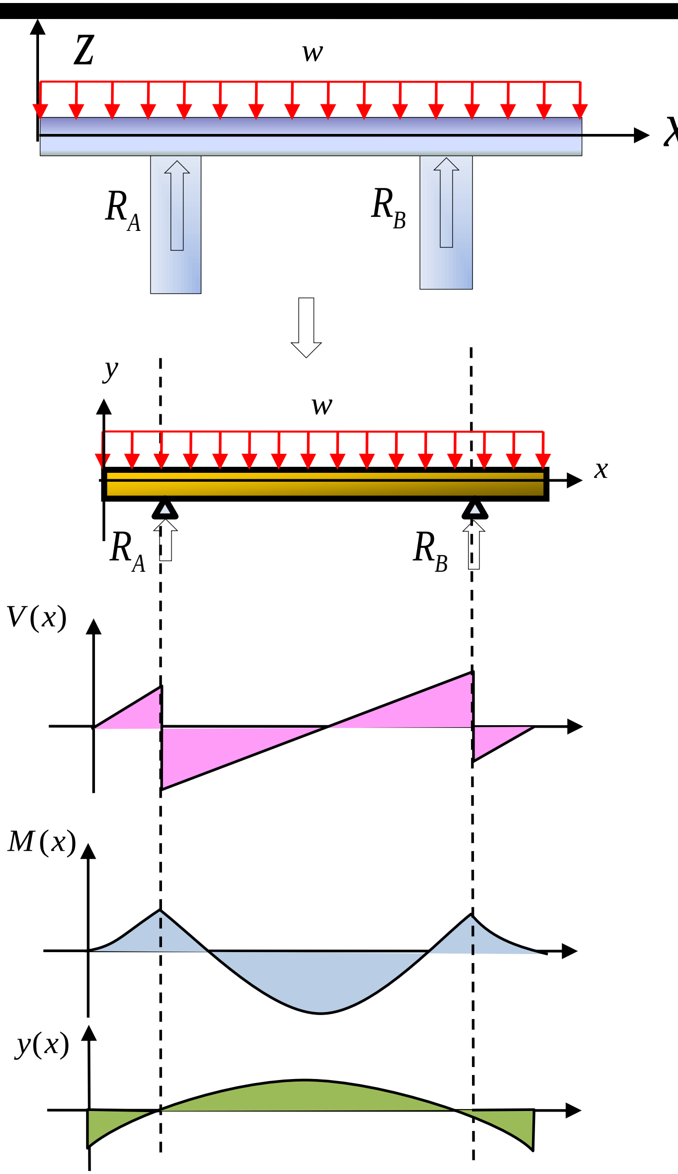
Seoul National Univ.

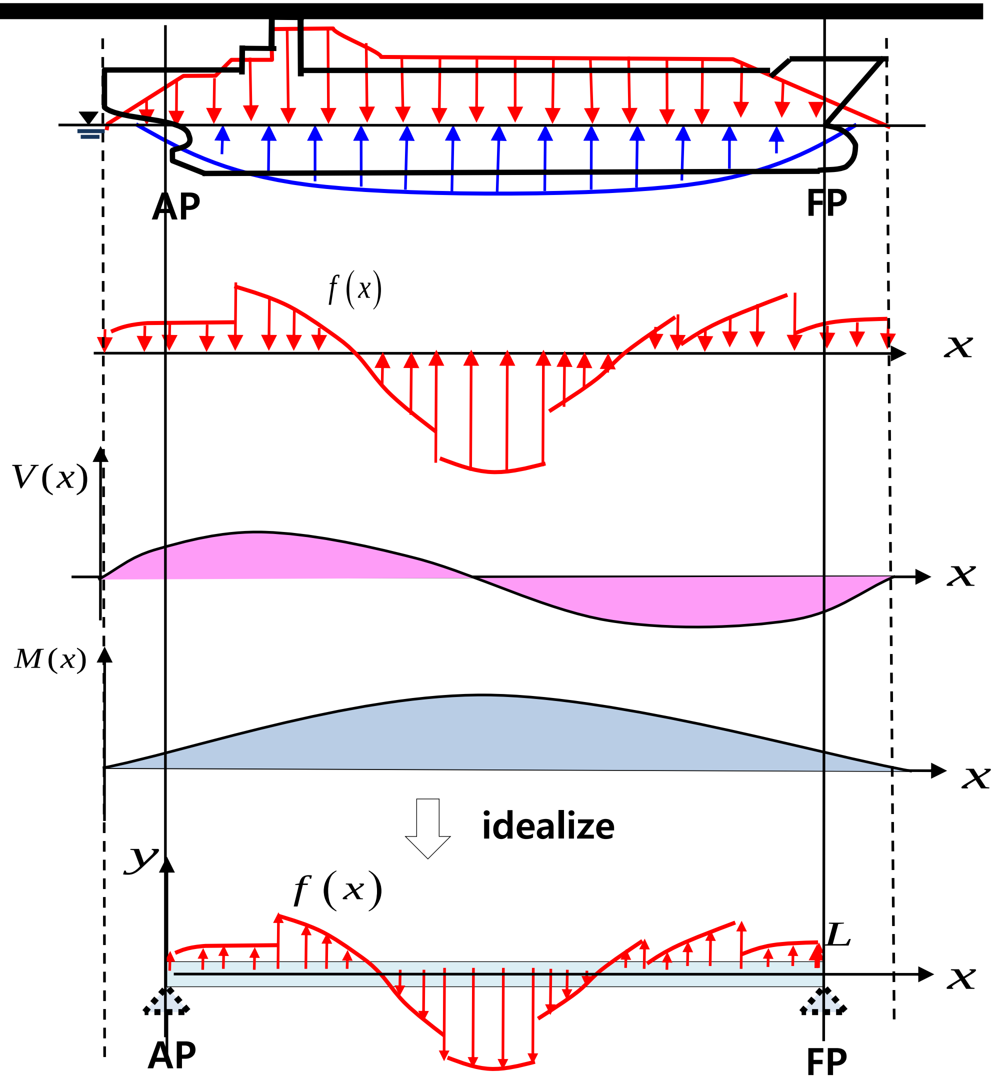
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

1. Beam Theory

1.1 Normal Stress and Strain, Shear Stress and Strain, and Torsion

Applying beam theory on a ship





Ship Structural Design

$$\boxed{Load: f(x)}$$

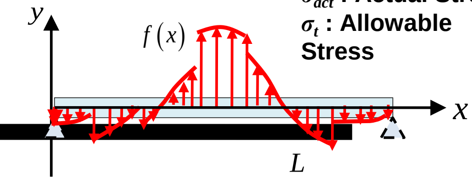
cause

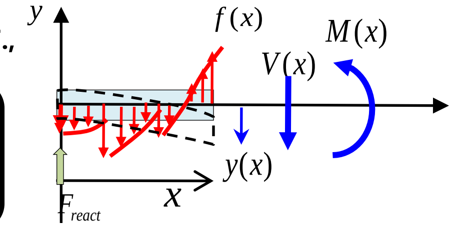
Shear Force: $V(x)$
Bending Moment: $M(x)$
Deflection: $y(x)$

'relations' of load, S.F., B.M., and deflection

$$\frac{dV(x)}{dx} = -f(x), \quad \frac{dM(x)}{dx} = V(x), \quad EI\frac{d^2y(x)}{dx^2} = M(x)$$

σ_{act} : Actual Stress
σ_l : Allowable Stress





● Ship Structural Design

what is designer's **major** interest?

● **Safety :**

Won't 'IT' fail under the load?

a ship	} global
a stiffener a plate	} local

a ship

what is our interest?

● **Safety :** *Won't it fail under the load?*

Stress should meet :

$$\boxed{\sigma_{act} \le \sigma_l}, \quad \text{where } \sigma_{act} = \frac{M}{I_{\bar{y}}/\bar{y}_i} = \frac{M}{Z}$$

● **Geometry :** *How much it would be bent under the load?*

Differential equations of the defection curve

$$\boxed{EI\frac{d^4y(x)}{dx^4} = -f(x)}$$

what kinds of load f cause hull girder moment?

$$\boxed{\sigma_{act.} \le \sigma_l}, \quad \sigma_{act.} = \frac{M_S + M_W}{Z_{mid}}$$

M_S = Still water bending moment
M_W = Vertical wave bending moment

$$f(x) = f_S(x) + f_W(x)$$
$$V(x) = V_S(x) + V_W(x)$$
$$M(x) = M_S(x) + M_W(x)$$

Hydrostatics

$f_S(x)$: load in still water = weight + buoyancy

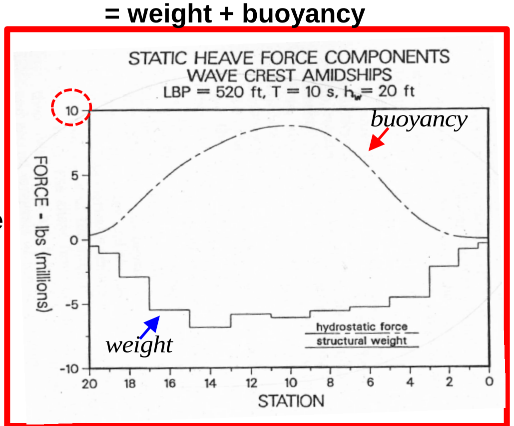
$f_S(x)$: load in still water

$$V_S(x) = \int_0^x f_S(x)dx$$

$V_S(x)$: still water shear force

$$M_S(x) = \int_0^x V_S(x)dx$$

$M_S(x)$: still water bending moment



Hydrodynamics

$f_W(x)$: load in wave = added mass + diffraction + damping + Froude-Krylov + mass inertia

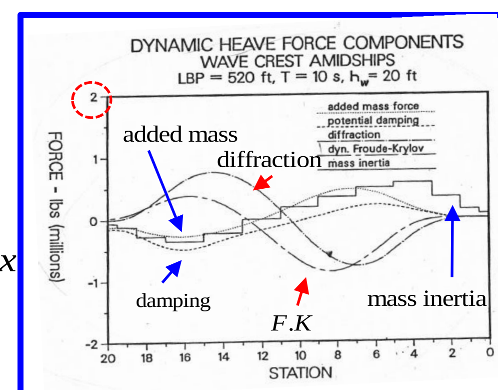
$f_W(x)$: load in wave

$$V_W(x) = \int_0^x f_W(x)dx$$

$V_W(x)$: wave shear force

$$M_W(x) = \int_0^x V_W(x)dx$$

$M_W(x)$: vertical wave bending moment



Ship Structural Design

- **Ship Structural Design**

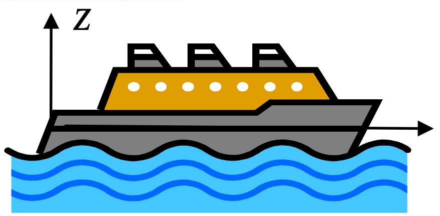
what is designer's major interest?

- **Safety :**

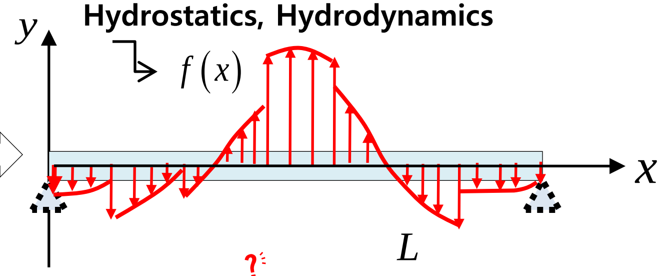
Won't 'IT' fail under the load?

a ship	global
a stiffener	local
a plate	local

a ship



Hydrostatics, Hydrodynamics



Actual stress on midship section should be less than allowable stress

$$\boxed{\sigma_{act.} \le \sigma_{allow}} \quad , \sigma_{act.} = \frac{M_{mid}}{Z_{mid}} = \frac{M_s + M_w}{I_{ship,N.A} / \bar{y}_i}$$

Allowable stress by Rule : (for example)

$$, \sigma_{allow} = 175 f_1 \ \ [N / mm^2]$$

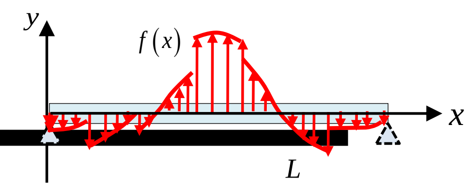
how we can meet the rule?

'Midship Design' is to <u>arrange</u> the structural members and <u>fix the thickness</u> of them to secure enough section modulus to the rule.

$\boxed{Load: f(x)}$

cause

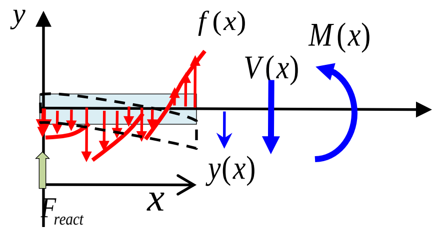
Shear Force: $V(x)$
Bending Moment: $M(x)$
Deflection : $y(x)$



'relations' of load, S.F., B.M., and deflection

$$\frac{dV(x)}{dx} = -f(x), \ \frac{dM(x)}{dx} = V(x)$$

$$, EI \frac{d^2 y(x)}{dx^2} = M(x)$$



what is our interest?

- **Safety :**

Won't it fail under the load?

Stress should meet :

$$\boxed{\sigma_{act} \le \sigma_{allow}} \quad , \text{where } \sigma_{act} = \frac{M}{I_{\bar{y}} / \bar{y}_i} = \frac{M}{Z}$$

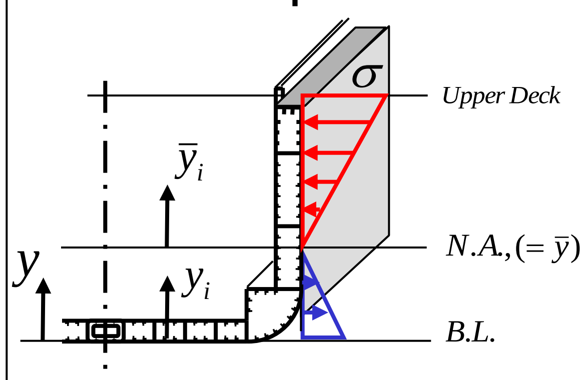
- **Geometry :**

How much it would be bent under the load?

Differential equations of the defection curve

$$\boxed{EI \frac{d^4 y(x)}{dx^4} = -f(x)}$$

<Midship section>



$, M_w$: vertical wave bending moment

$, M_s$: still water bending moment

$, I_{ship,N.A.}$: moment of inertia from N.A. of Midship section

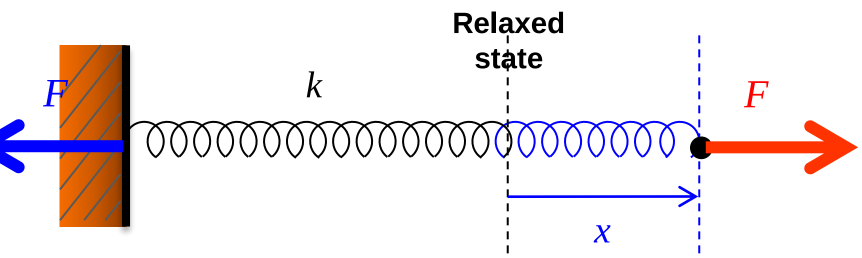
1. Normal Stress and Strain

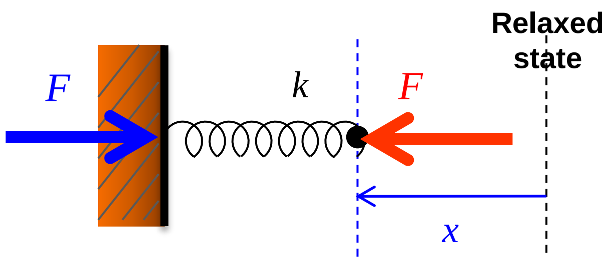
Restoring Force of a Spring; Hooke's Law

x : displacement of the spring length
k : spring constant
F : spring force (restoring force)

•Restoring force of a spring

x : positive → F : negative   x : negative → F : positive





✓Elastic : if a body suffers a deformation when a stretching force or compressing force is applied to the body and returns to its original shape when the force is removed, the body is said to be elastic.

✓Restoring force : the force with which a body resists deformation.

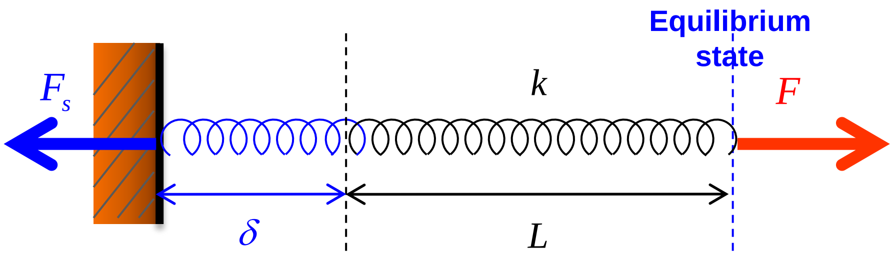
✓Hooke's Law : the magnitude of the restoring force is directly proportional to the deformation.

$$|F| \propto |x| \Rightarrow F = -kx$$: Hooke's Law

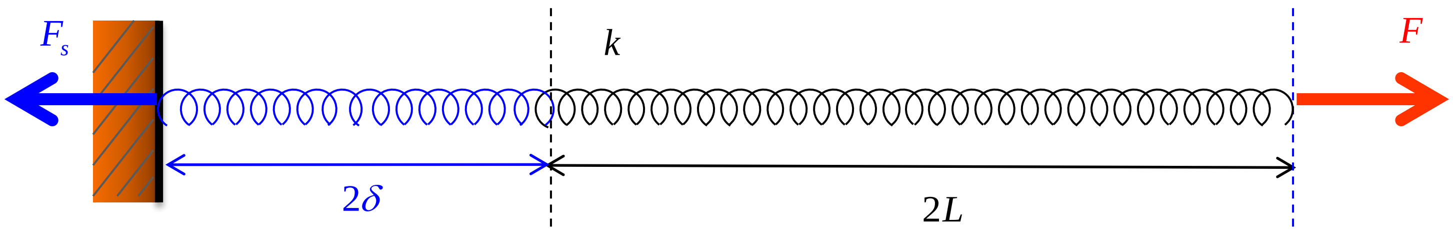
An example of Spring

L : spring length in a relaxed state
δ : displacement of the spring length
k : spring constant
F : external force
F_s : spring force (restoring force)

▪One spring : $-F_s = F$

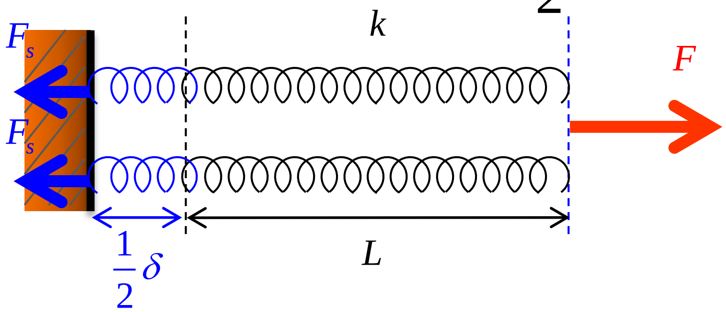


$\delta \propto F$



$\delta \propto L$

▪Two springs : $-F_s = \dfrac{1}{2}F$



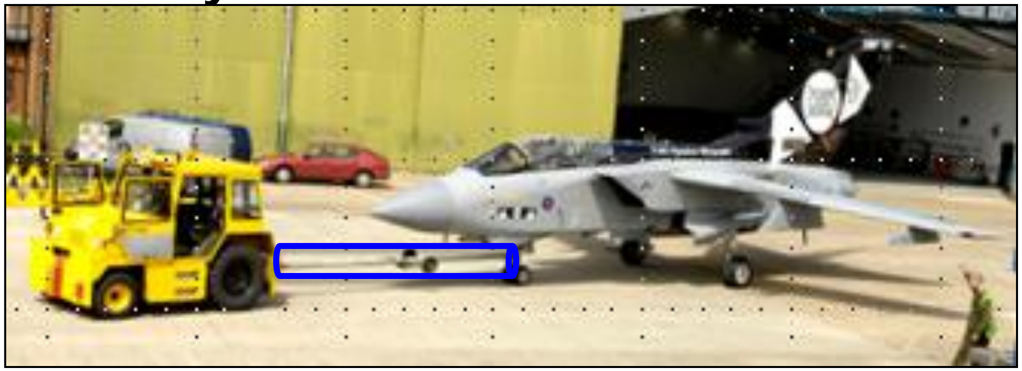
$$\delta \propto \frac{1}{\text{The number of springs}}$$

Hooke' Law

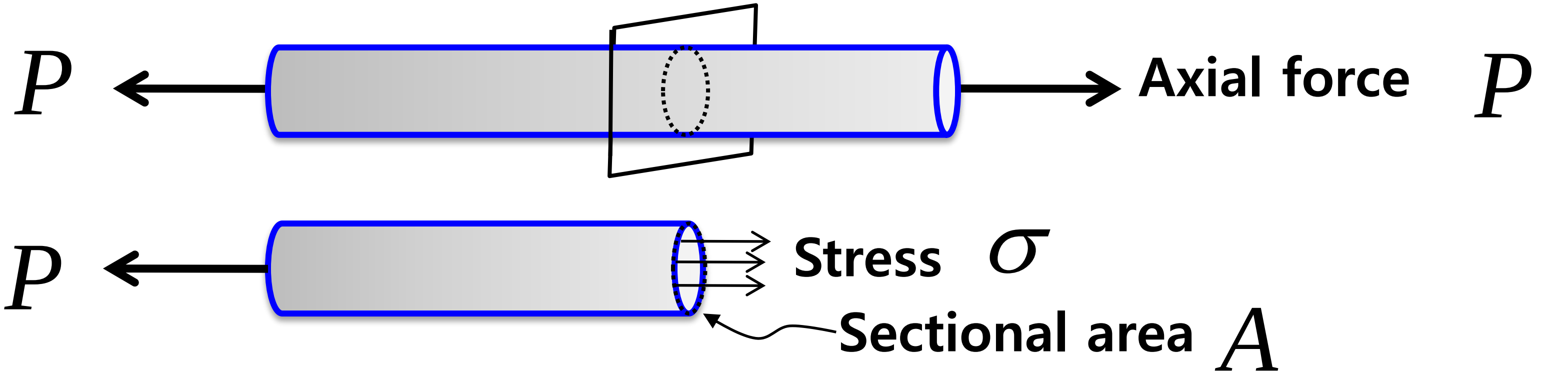
Normal Stress

$$\frac{P}{A} \propto \frac{\delta}{L}$$

An example of axially loaded member : aircraft tow bar



Free-body diagram of the bar



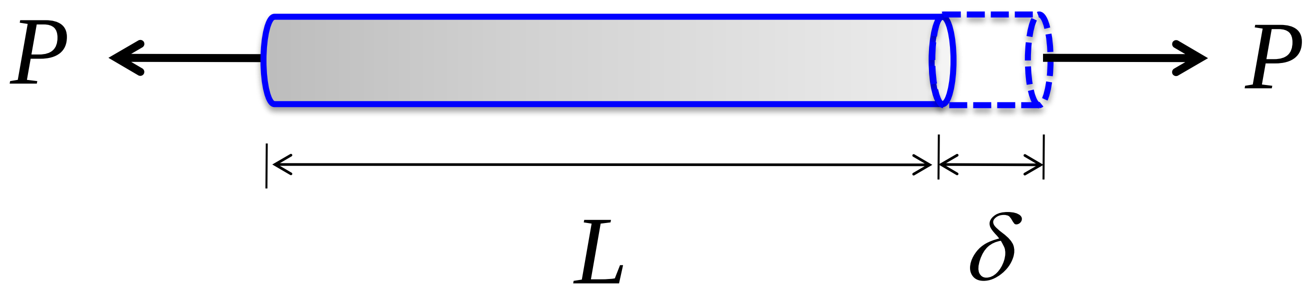
Normal Stress : force per unit area

$$\sigma = \frac{P}{A} \quad \text{, or} \quad P = \sigma \cdot A$$

Normal Strain

$$\frac{P}{A} \propto \frac{\delta}{L}$$

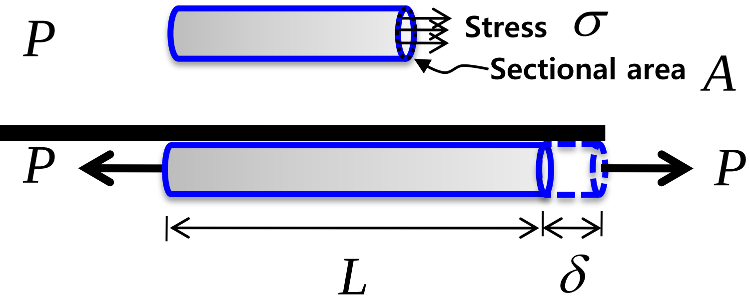
Elongation of the bar



Strain : elongation per unit length

$$\varepsilon = \frac{\delta}{L}$$

Hooke's Law



$$\delta \propto P$$
$$\delta \propto L$$
$$\delta \propto \frac{1}{A}$$

$$\Rightarrow \quad \delta \propto \frac{PL}{A}$$

$$\frac{P}{A} \propto \frac{\delta}{L} \quad \Rightarrow \quad \frac{P}{A} = E\frac{\delta}{L}$$

$$\sigma = \frac{P}{A}$$

$$\varepsilon = \frac{\delta}{L}$$

Relation between the normal stress and the strain

$$\sigma = E\varepsilon$$

"Hooke's law" (constitutive relation)

σ : normal stress
E : young's modulus
ε : strain

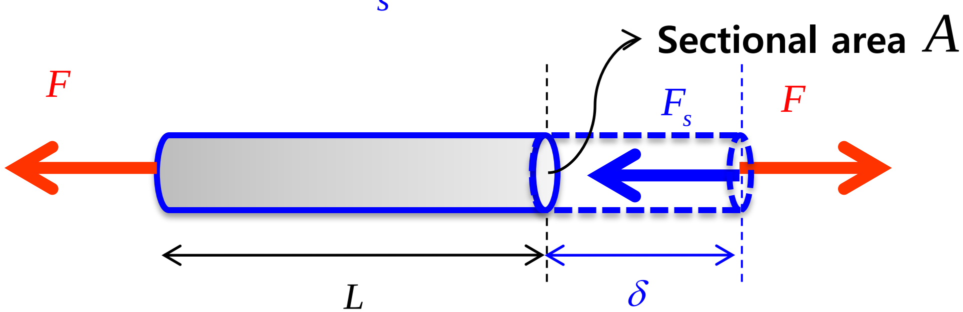
An example of bar

A : sectional area of the rod

$\delta \propto P$

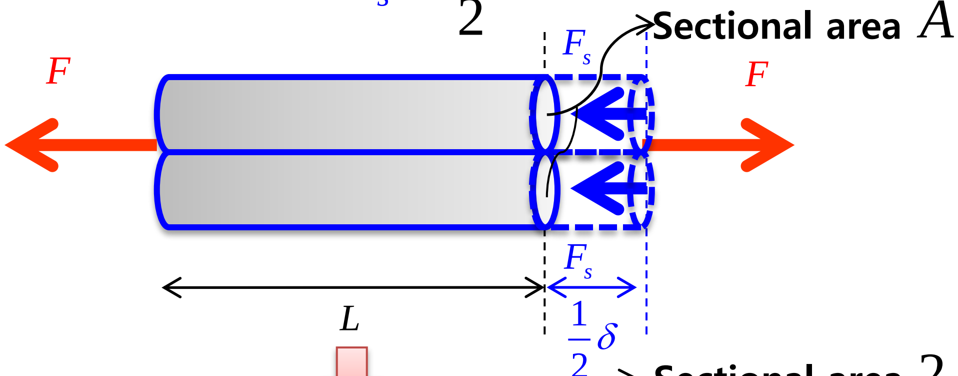
$\delta \propto L$

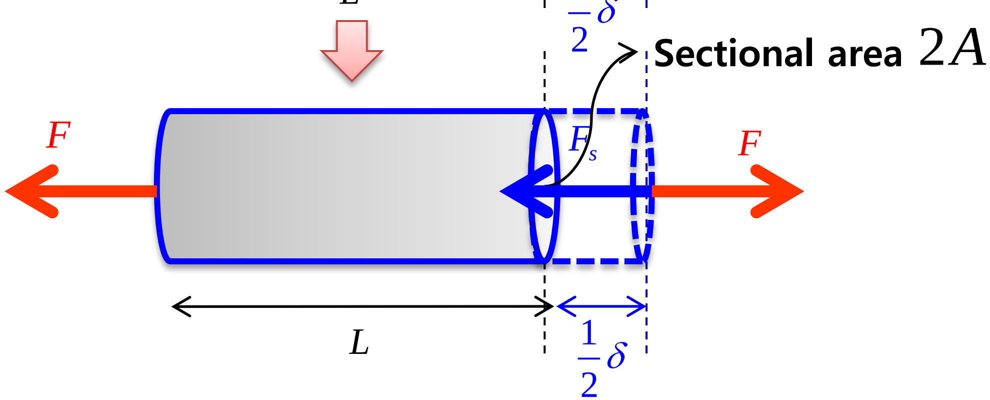
▪One rod : $-F_s = F$



$\Rightarrow \quad \delta \propto \dfrac{1}{A}$

▪Two rods : $-F_s = \dfrac{1}{2}F$





$$\delta \propto \frac{PL}{A}$$

$$\frac{P}{A} \propto \frac{\delta}{L} \quad \Rightarrow \quad \underbrace{\frac{P}{A}}_{\text{Stress}} = \overbrace{E}^{\text{Young's modulus}} \underbrace{\frac{\delta}{L}}_{\text{Strain}}$$

"Hooke's law": $\sigma = E\varepsilon$

(constitutive relation)

σ : normal stress

E : young's modulus

ε : strain

Relation between normal stress and strain

Relation between the normal stress and the strain

$$\sigma = E\varepsilon$$

"Hooke's law"
(constitutive relation)

Stress-Strain diagram for a typical structural steel in tension

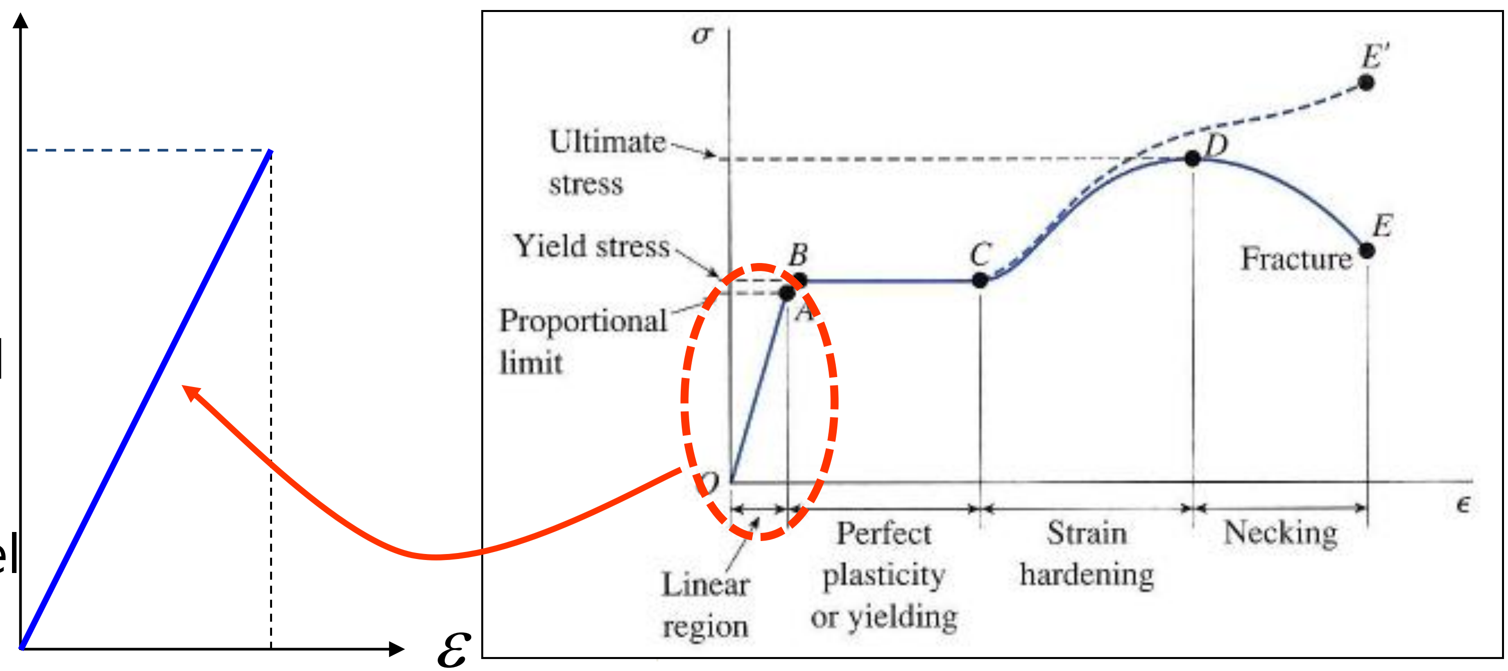
σ

The slope is called **the modulus of elasticity,** E **(**often called **Young's modulus)**

Ex.) mild steel

$200GPa$, or

$200 \times 10^9 N/m^2$



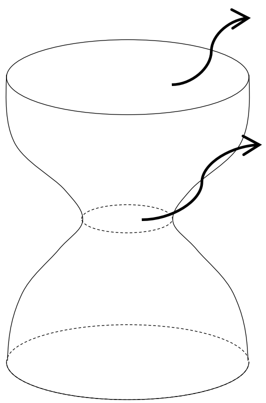
Relation between normal stress and strain

Relation between the normal stress and the strain

$$\sigma = E\varepsilon$$

“Hooke’s law” (constitutive relation)

Stress-Strain diagram for a typical structural steel in tension



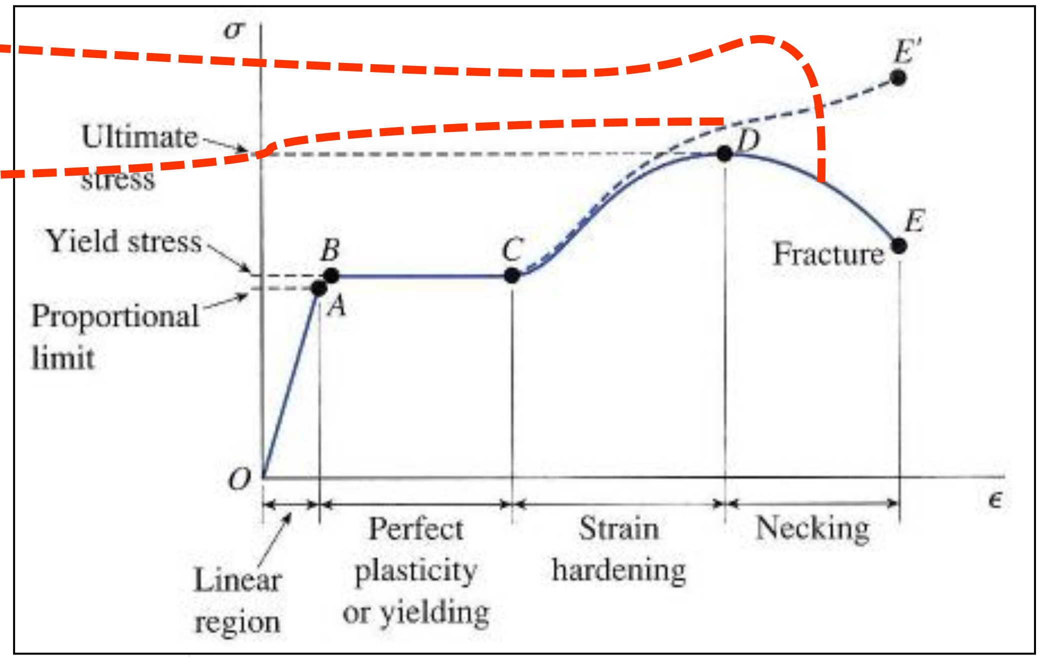
Sectional area A

Sectional area A'

$$\frac{F}{A'} > \frac{F}{A}$$

Increase $(C:E')$

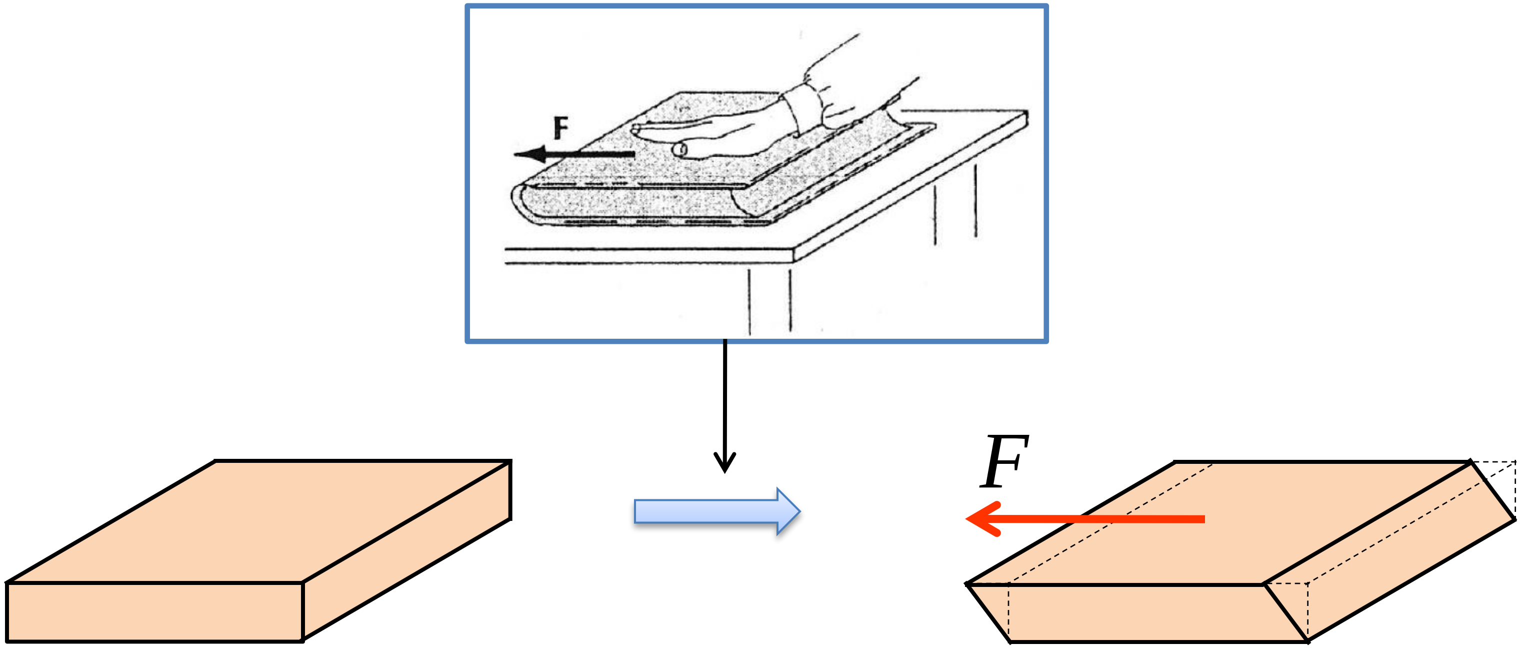
Decrease $(C:E)$



If the actual cross-sectional area at the narrow part of the neck is used to calculated the stress, **the true stress-strain curve** (the dashed line CE’) is obtained. The total load the bar can carry does indeed diminish after the ultimate stress is reached (the line DE), but this reduction is due to the decrease in area of the bar and not to a loss in strength of the material itself.

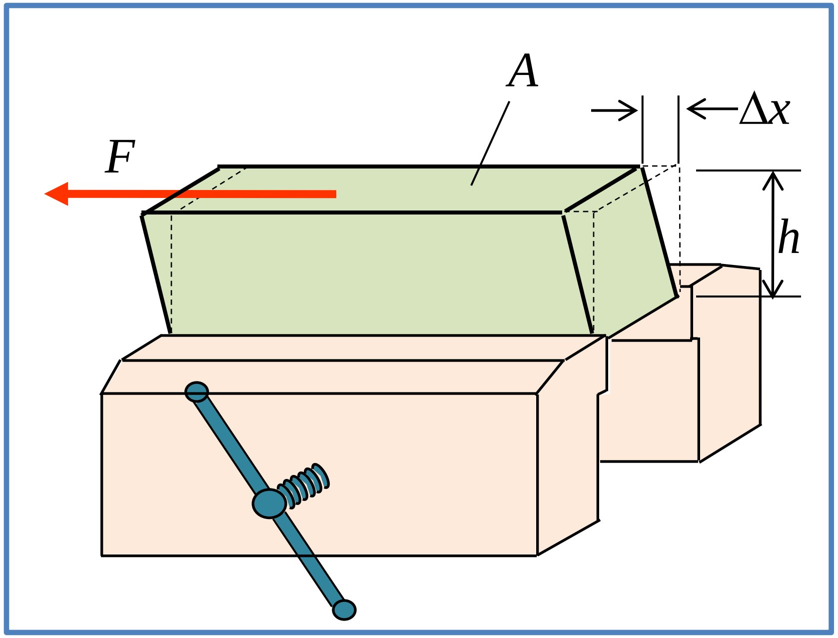
Shear Stress and Strain of the Block

Shear (1)



Shear : (noun) the deformation which changes the shape of the body from rectangular parallelepiped to a rhomboidal parallelepiped

Shear (2)



F : Axial force

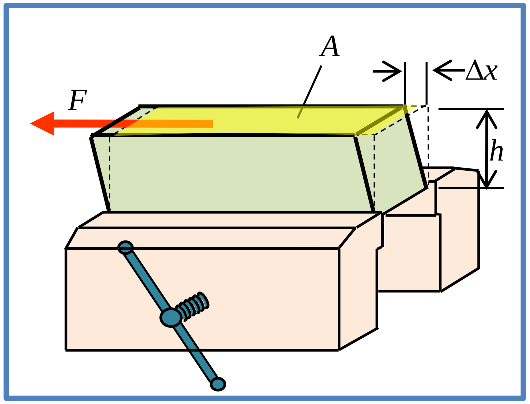
A : Sectional area

H : Height of the block

$\triangle x$: Displacement of the edge of the block

For example, if one side of the body is held fixed, and the, force(F) pushes tangentially along the other side, then the deformation is a shear.

Shear Stress in the block



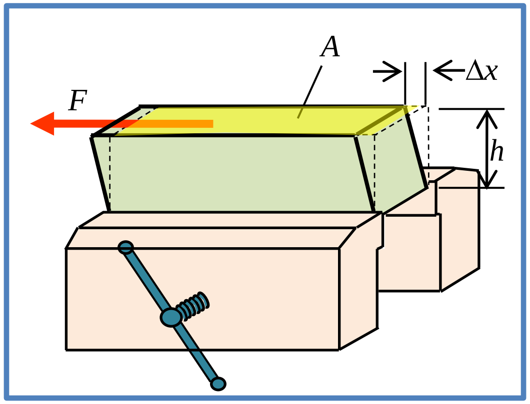
F : Axial force
A : Sectional area
H : Height of the block
$\triangle x$: Displacement of the edge of the block

A shear Stress(τ) : shear force(V, equal to force F) per unit area

<u>For example</u>, the shear stress is(τ) as follows ;

$$\tau = \frac{F}{A}$$

Shear Strain of the block



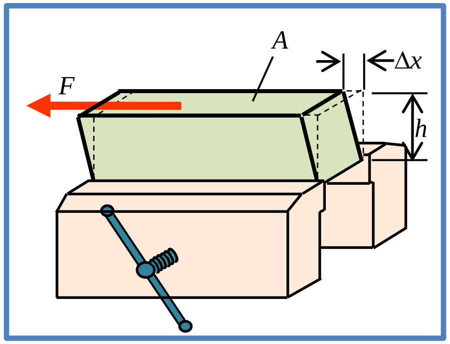
F : Axial force
A : Sectional area
H : Height of the block
$\triangle x$: Displacement of the edge of the block

Shear strain (γ) : A change in shape, or a measure of the distortion, of the element (measured in degrees or radians)

For example, the shear strain(γ) is as follows ;

$$\gamma = \frac{\Delta x}{h}$$

Relation between the shear stress and strain in the block



F : Axial force
A : Sectional area
H : Height of the block
$\triangle x$: displacement of the edge of the block

Shear strain $\gamma = \dfrac{\Delta x}{h}$

Relation between the shear strain(γ) and the force(F) $\qquad \gamma \propto F$

Relation between the shear strain(γ) and sectional area(A) $\qquad \gamma \propto \dfrac{1}{A}$

$$\gamma \propto \frac{F}{A}$$

Relation between the shear stress(τ) and strain (γ)

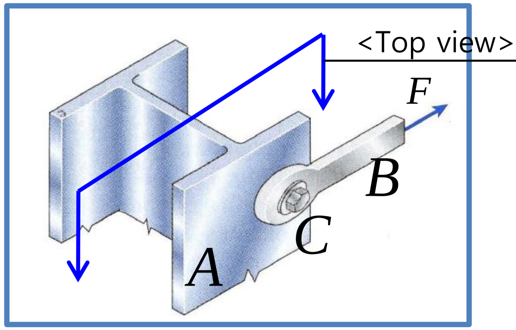
$$\boxed{\gamma \propto \frac{F}{A}} \Rightarrow \gamma \propto \tau \Rightarrow \boxed{G\gamma = \tau}$$

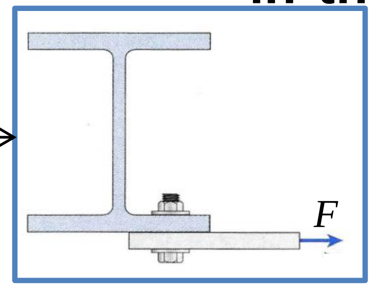
Shear modulus of elasticity

Shear Stress and Strain

Shear Stress

An example of axially loaded member : A bolt passing through holes in the bar and I-beam



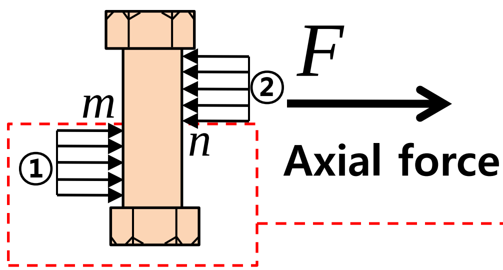


<Top view>

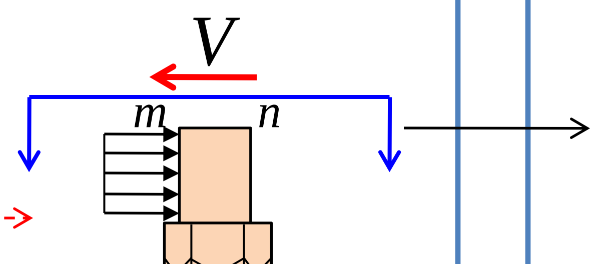
*shear it : (verb) cut through it.
In this case, the bar and I-beam tend to shear the bolts.

Free-body diagram of the bolt

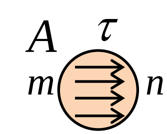
V : Shear force
A : Area



<Detail view of the bolt>



<Free-body diagram of the bolt>



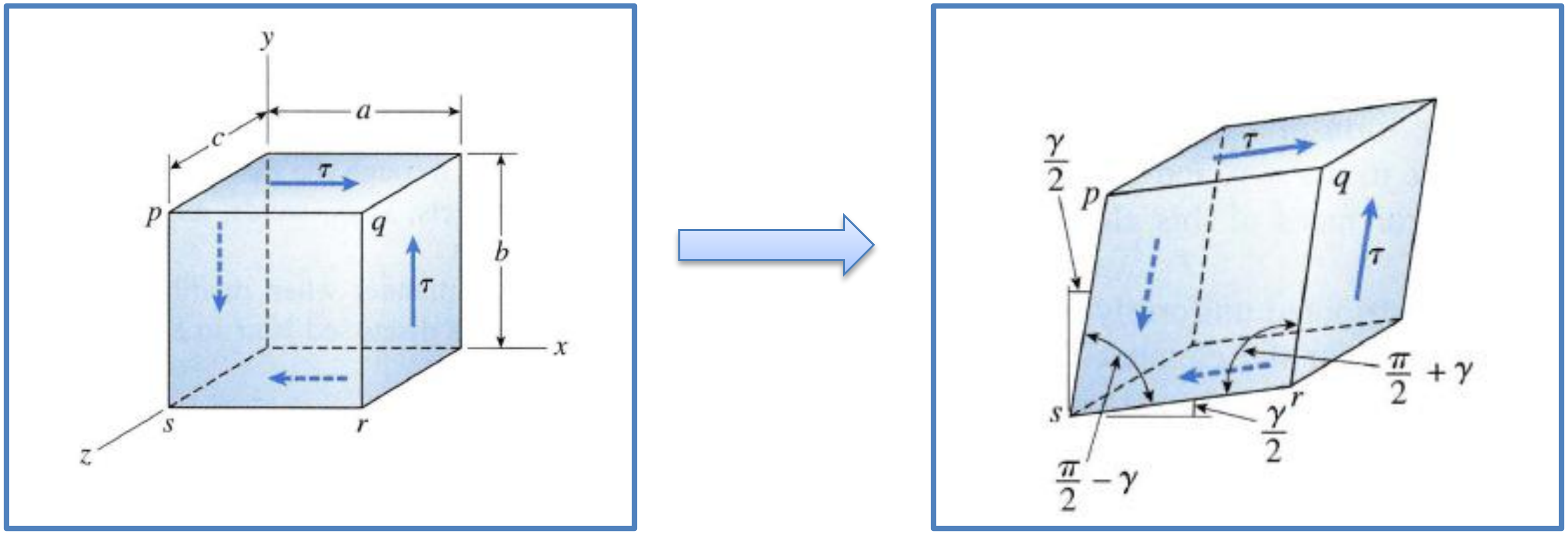
<Shear stress(τ) acting on the bolt>

A shear stress(τ) : shear force(V) per unit area

$$\tau = \frac{V}{A} \quad , \text{ or } \quad V = \tau \cdot A$$

Shear Strain

A change in the *shape* of the element



(a) Element of material

(b) Element of material subjected to shear stresses and strains

Shear strain(γ) : A change in shape, or a measure of the distortion, of the element

γ (measured in degrees or radians)

Relation between the shear stress and strain

Hooke's law in shear

$$\tau = G\gamma$$

τ : **Shear stress**

G : **Shear modulus of elasticity**

γ : **Shear strain**

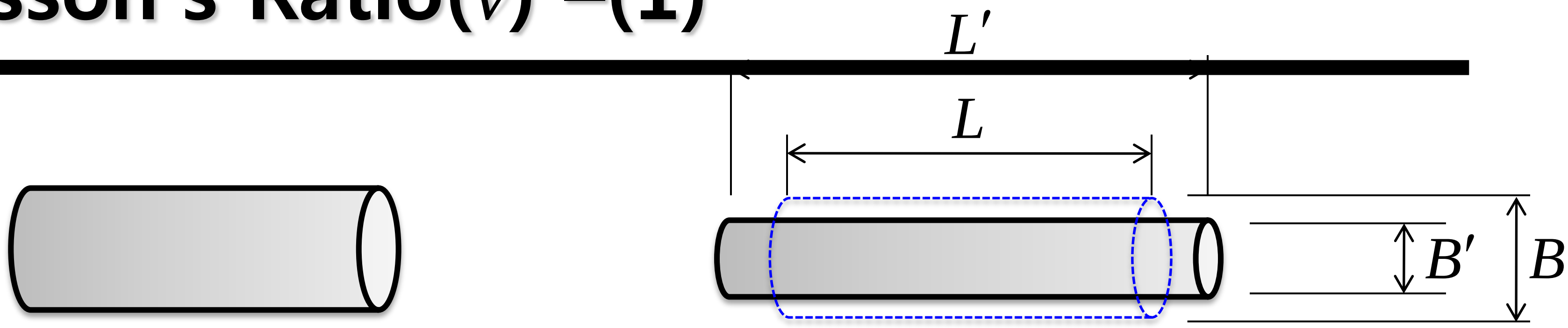
Cf.) $\sigma = E\varepsilon$

Ex.) Shear modulus of elasticity(G) of mild steel

$75GPa$, or

$75 \times 10^9 N / m^2$

Poisson's Ratio(ν) –(1)



(a) the bar before loading (b) the bar after loading P

When a prismatic bar is loaded in tension, the axial strain(ε) is accompanied by lateral contraction(ε').

$$\text{axial strain}(\varepsilon) = \frac{L' - L}{L}$$

$$\text{lateral strain}(\varepsilon') = \frac{B' - B}{B}$$

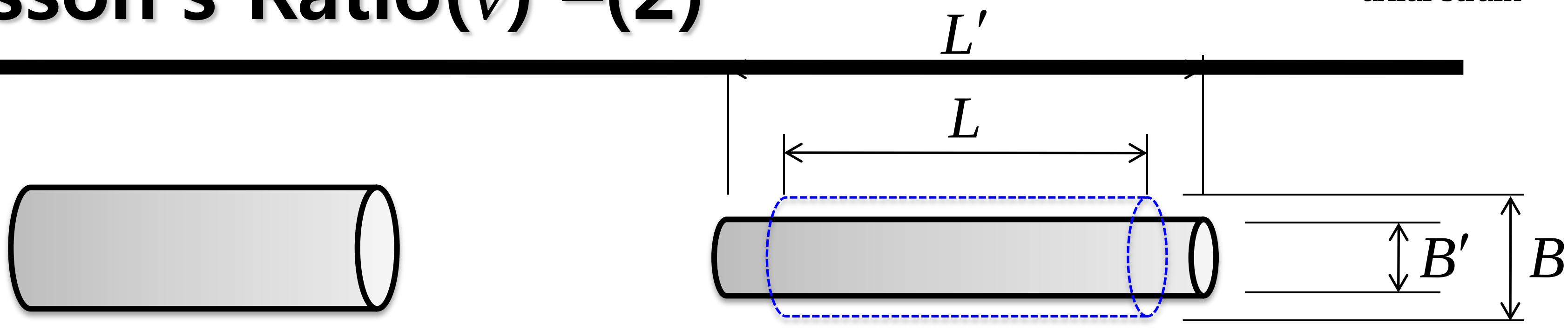
Poisson's ratio : The ratio of axial strain(ε) to lateral strain(ε')

$$\boxed{\nu = -\frac{\text{lateral strain}}{\text{axial strain}} = -\frac{\varepsilon'}{\varepsilon}} \Rightarrow \boxed{\varepsilon' = -\nu\varepsilon}$$

The minus sign is inserted in the equation to compensate for the fact that the lateral and axial strains normally have opposite signs. For instance, the axial strain in a bar in tension is positive and the lateral strain is negative (because the width of the bar decreases).

Poisson's Ratio(v) –(2)

$$v = -\frac{\text{lateral strain}}{\text{axial strain}} = -\frac{\varepsilon'}{\varepsilon}$$



(a) the bar before loading

(b) the bar after loading P

The range of the Poisson's ratio(v)

Most materials : 0.25 ~ 0.35

Theoretical limit : 0 ~ 0.5

Relation between the modulus of elasticity in tension(E) and shear(G)

Relation between the modulus of elasticity in tension(E) and shear(G)

$$G = \frac{E}{2(1+\nu)}$$

G : **Shear modulus of elasticity**

E : **Modulus of elasticity**

ν : **Poisson's ratio**

The range of the shear modulus of elasticity(G) which is relative to the modulus of elasticity(E)

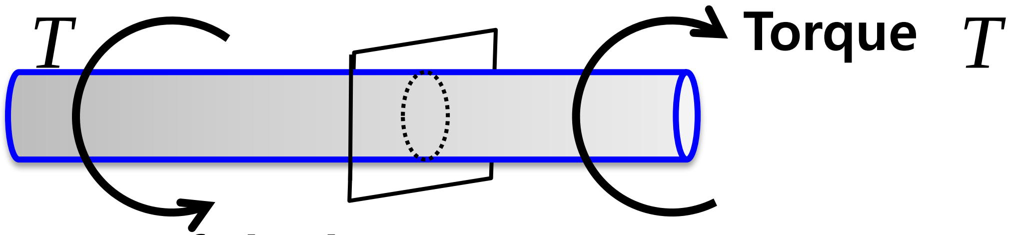
The range of the Poisson's ratio(ν) : 0 ~ 0.5

$$\left.\begin{aligned} \nu = 0 &\longrightarrow G = \frac{E}{2} \\ \nu = 0.5 &\longrightarrow G = \frac{E}{3} \end{aligned}\right\} \Rightarrow \frac{E}{3} \le G \le \frac{E}{2}$$

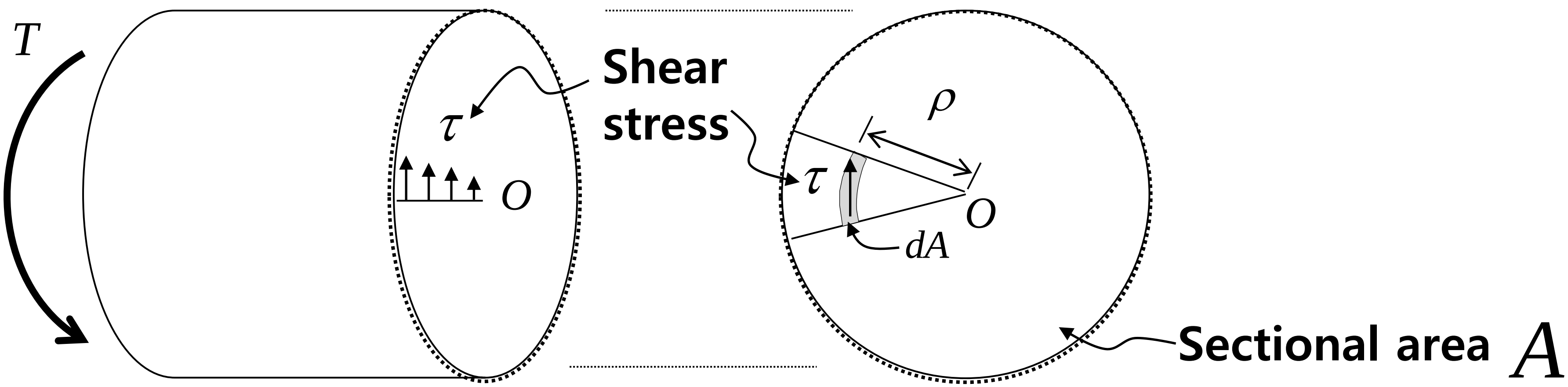
Shear Stress and Strain in Torsion

Shear Stress in torsion

Deformation of a circular bar in pure torsion



Free-body diagram of the bar



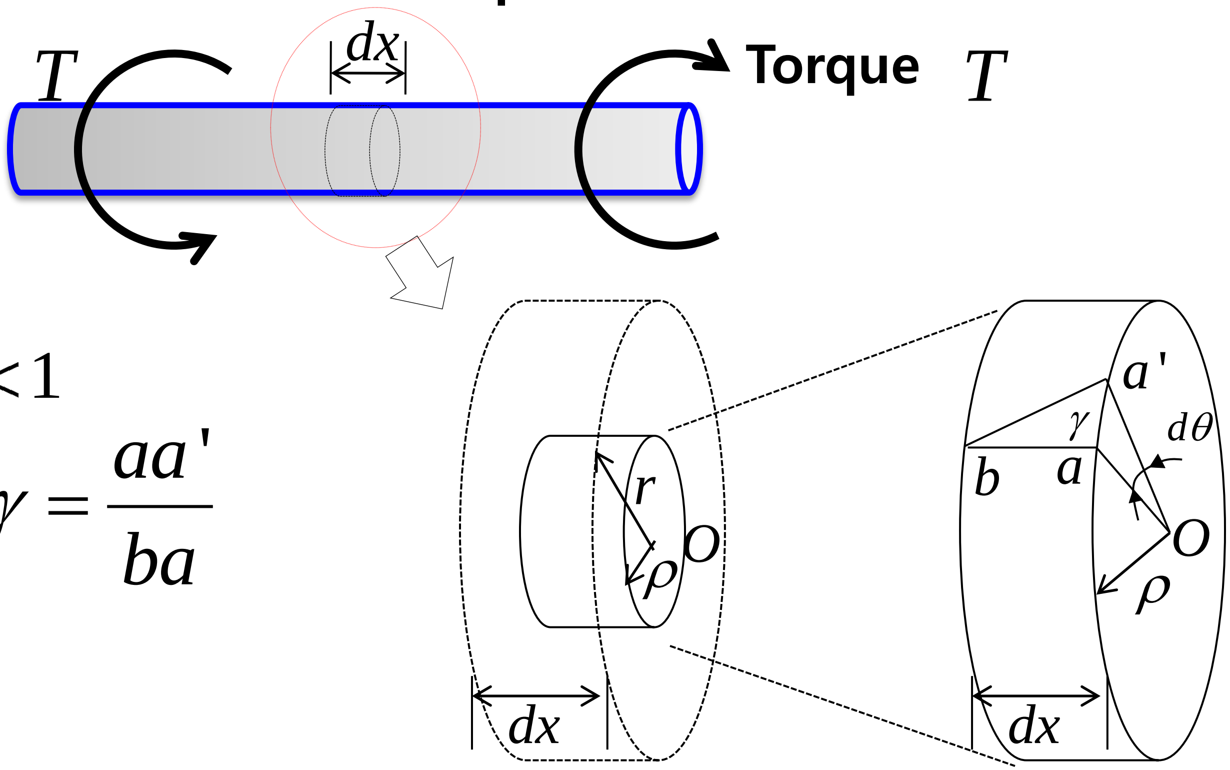
Shear force acting on the area dA : τdA

Resultant moment about a longitudinal axis through point O is equal to the torque :

$$T = \int_A \rho \tau dA$$

Shear Strain in torsion

Deformation of a circular bar in pure torsion



$$aa' = \rho\, d\theta$$

$$ba = dx$$

Assume, $\gamma \ll 1$

$$\gamma \approx \tan \gamma = \frac{aa'}{ba}$$

Shear strain

$$\gamma = \rho \frac{d\theta}{dx}$$

Relation between the torque and the angle of twist

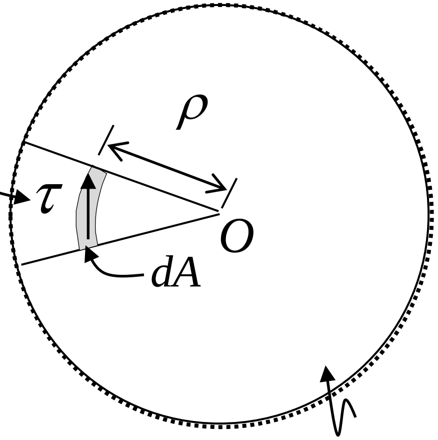
Shear force acting on area dA

$\tau\, dA$

Hooke's law in shear deformation

$$\tau = G\gamma$$

Shear strain $\gamma = \rho \dfrac{d\theta}{dx}$



Sectional area A

$$\tau\, dA = G\rho \frac{d\theta}{dx} dA$$

Resultant moment about a longitudinal axis through the point O is equal to the torque:

$$T = \int_A \rho\tau dA = \int_A G \frac{d\theta}{dx} \rho^2 dA = G \frac{d\theta}{dx} \int_A \rho^2 dA$$

Relation between the torque and the angle of twist

$$T = GJ \frac{d\theta}{dx}$$

Polar moment of inertia $J = \int_A \rho^2 dA$

Naval Architecture & Ocean Engineering

1. Beam Theory

1.2 Deflections of Beams

Derivation of Deflection Curve of Beam

Overview of Derivation Procedure

Geometry of Deformation

↓ ← Geometrical Linearization

Elongation

↓

Strain

↓ ← Hooke's Law

Stress

↓

Bending Moment ⇨ Relation between the deformation and the bending moment

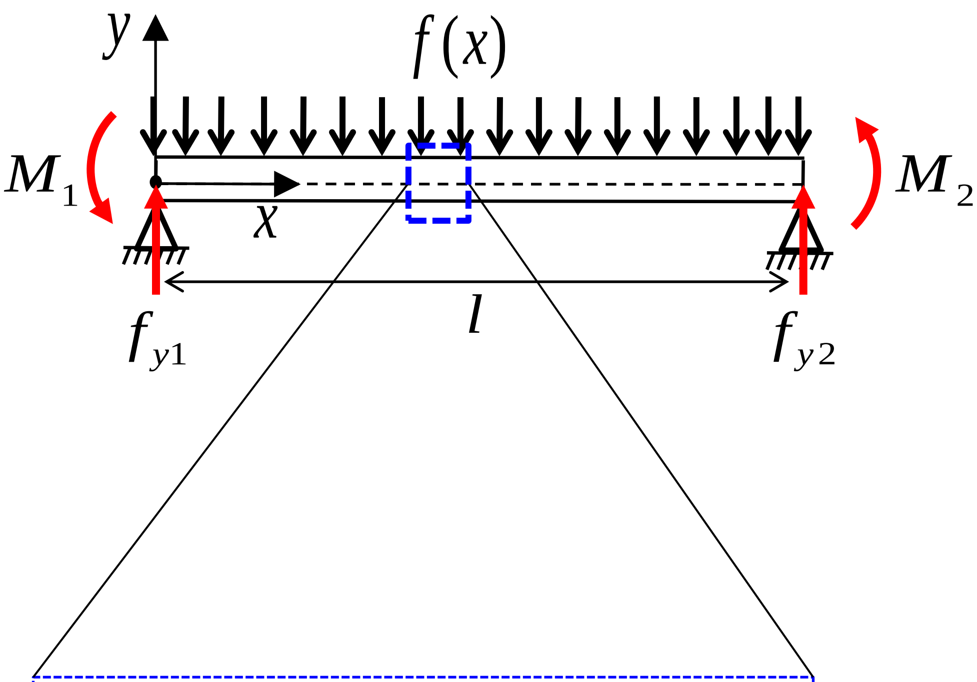
↓

Shear Force

↓

Distributed Load ⇨ Relation between the deformation and the distributed load

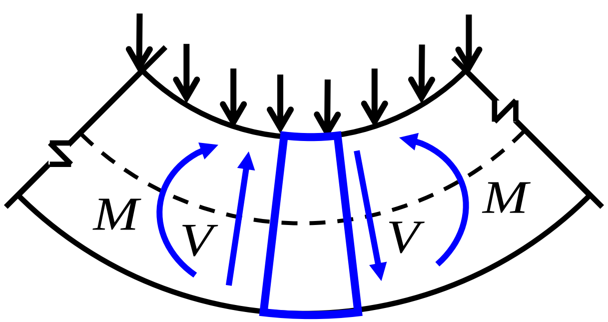
Geometry of Deformation



1) The concentrated forces f_{y1} and f_{y2} are exerted on the ends of the bar.

2) The moment M_1 and M_2 are exerted on the ends of the bar.

3) distributed force f(x) is applied to the element



This figure is expressed in very exaggerated way to recognize the deformation of a beam!

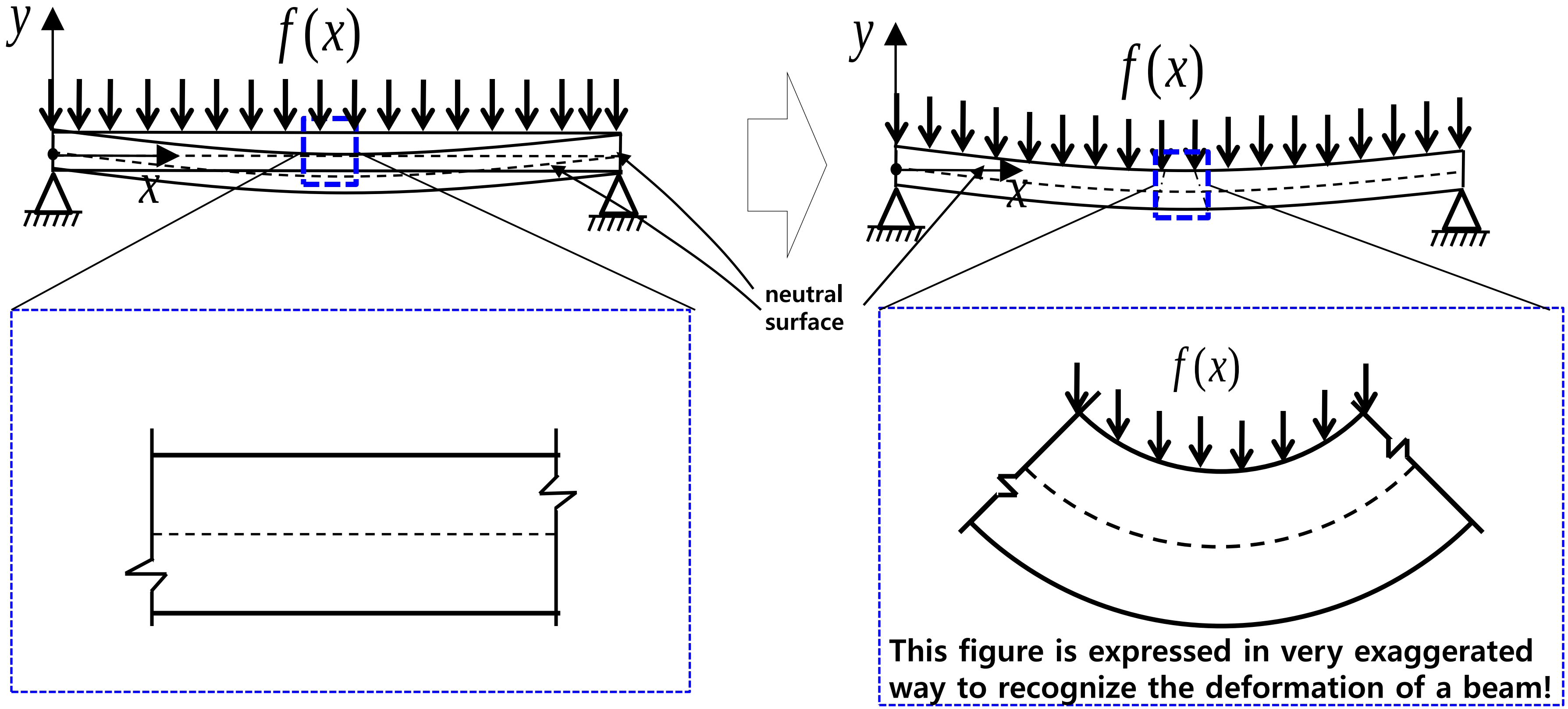
An infinitesimal element will be introduced for derivation deflection curve of a beam. Here, the **bending moment M** and the **shear force V** are **stress resultants**, and the positive directions of the stress resultants are shown in left figure.

Geometry of Deformation

Geometry of a beam

-before deformation

-after deformation



The x axis is a line along the neutral surface of the undeformed beam. Of course, when the beam deflects, the neutral surface moves with the beam, but the x axis remains fixed in position.

* neutral surface : Longitudinal lines on the lower part of the beam are elongated, whereas those on the upper part are shortened. Thus the lower part of the beam is in tension and the upper part is in compression. Somewhere between the top and bottom of the beam is a surface in which longitudinal lines do not change in length. This surface is called neutral surface

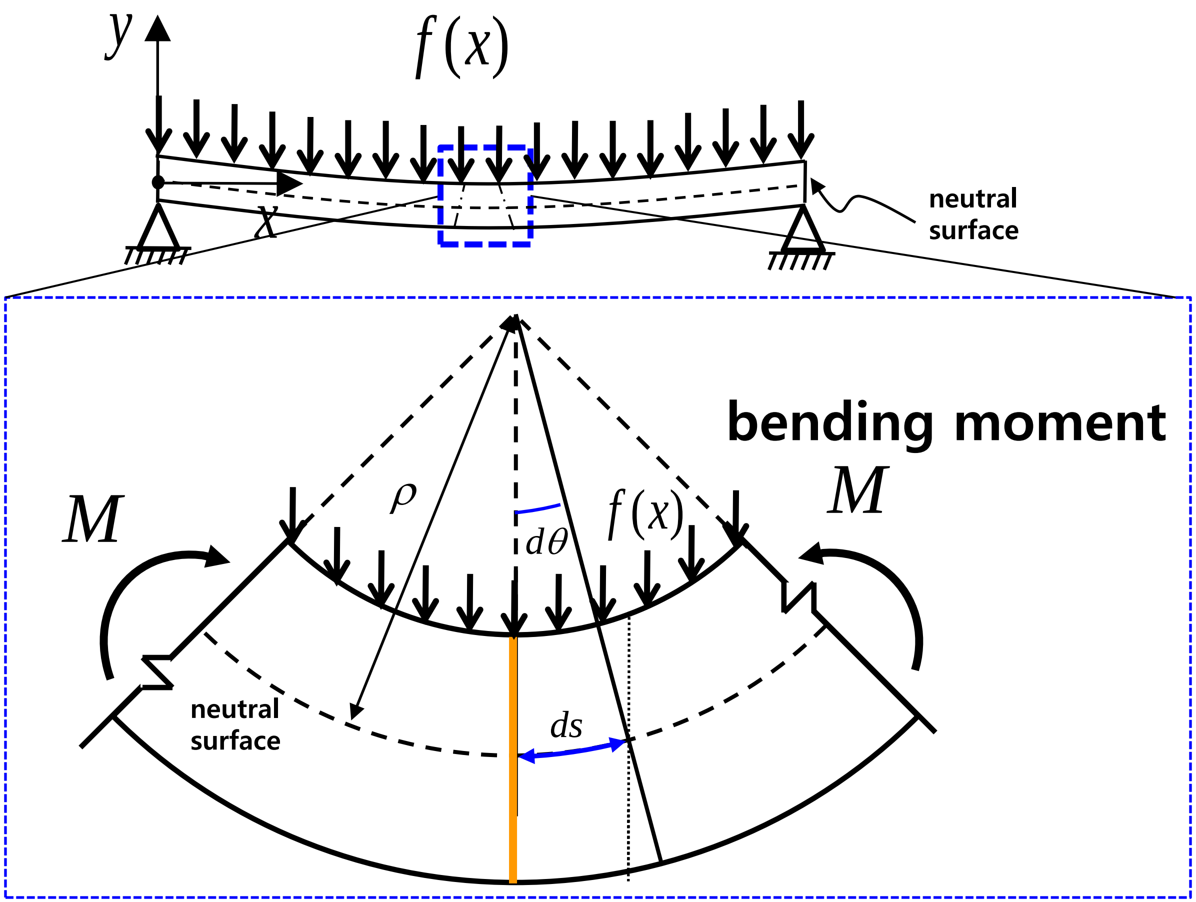
Geometry of Deformation

Geometry of a beam

-after deformation

$$\rho \cdot d\theta = ds$$

$$\kappa \equiv \frac{1}{\rho} = \frac{d\theta}{ds}$$



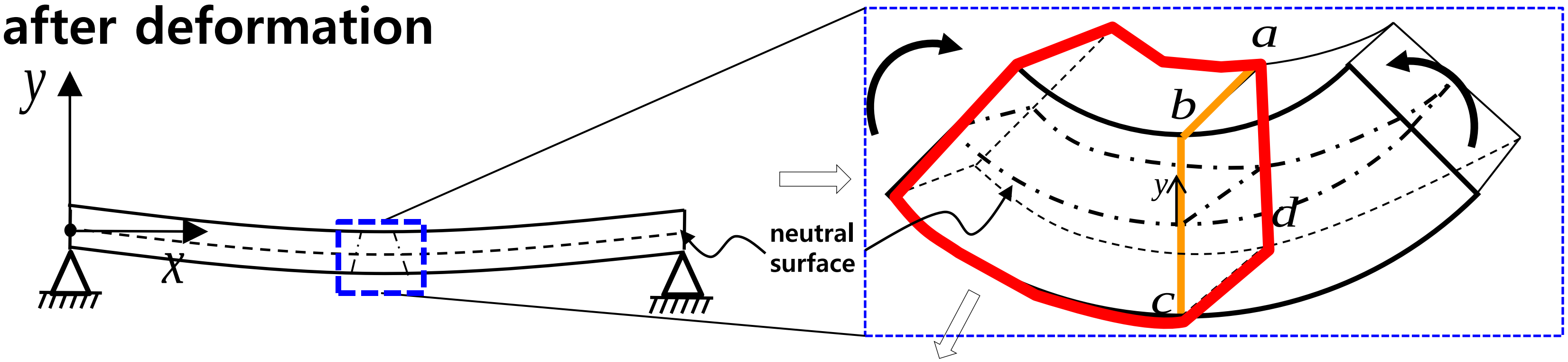
ρ: *radius of curvature*, κ : *curvature*

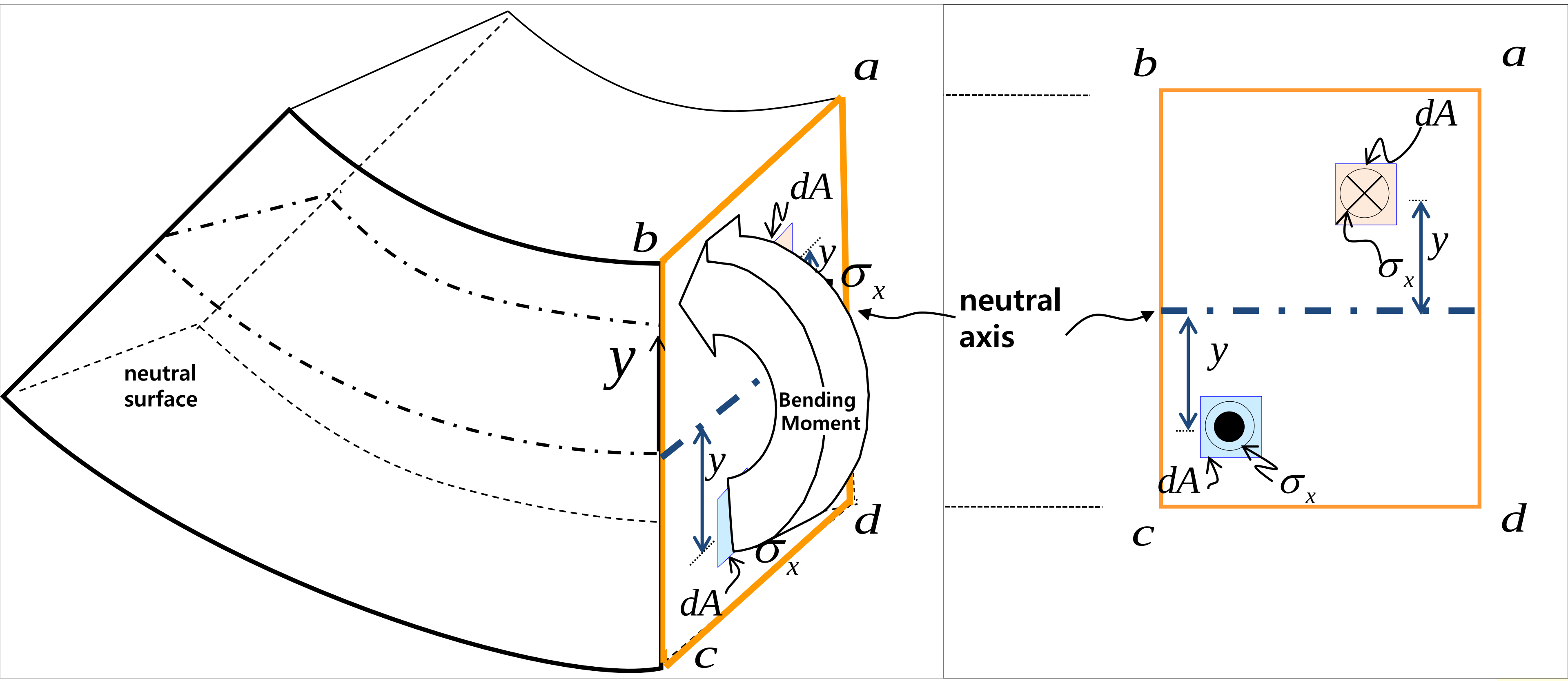
Geometry of Deformation

$$dM = y\ dF$$

$$dM = y\,\sigma_x\,dA$$

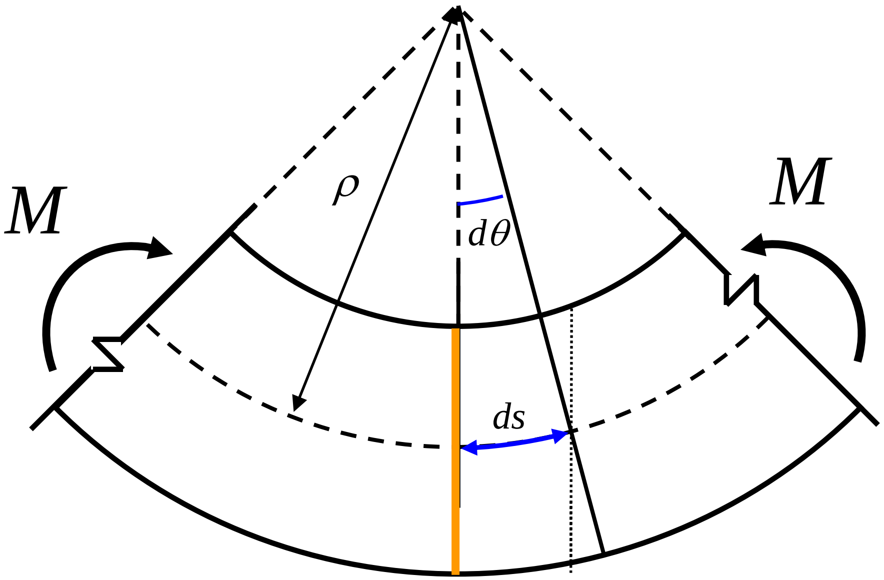
-after deformation



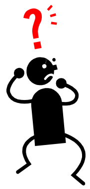


Geometry of Deformation

$$\kappa \equiv \frac{1}{\rho} = \frac{d\theta}{ds}$$



The differential equation for the deflection curve of a beam is supposed to be expressed based on the Cartesian coordinate system.



How can we express the geometry with dx **and** dy **instead of** ds **and** $d\theta$

Geometry of Deformation : Linearization

•IF WE ASSUME, $\theta << 1$

$$ds^2 = dx^2 + dy^2$$

$$\rightarrow ds = dx\boxed{\sqrt{1+\left(\frac{dy}{dx}\right)^2}}$$

$$let, z = \left(\frac{dy}{dx}\right)^2 \; then, \sqrt{1+\left(\frac{dy}{dx}\right)^2} = \sqrt{1+z}$$

$$f(z) = \boxed{\sqrt{1+z}} \xrightarrow[\text{Expansion}]{\text{•Taylor series}} f(z) = \boxed{1 + \frac{1}{2}z + \frac{1}{2}\left(-\frac{1}{4}\right)z^2 + \ldots}$$

•, since

$$\begin{cases} f(0) = 1 \\ f'(0) = \frac{1}{2}(1+z)^{-\frac{1}{2}}\Big|_{z=0} = \frac{1}{2} \\ f''(0) = -\frac{1}{4}(1+z)^{-\frac{3}{2}}\Big|_{z=0} = -\frac{1}{4} \end{cases}$$

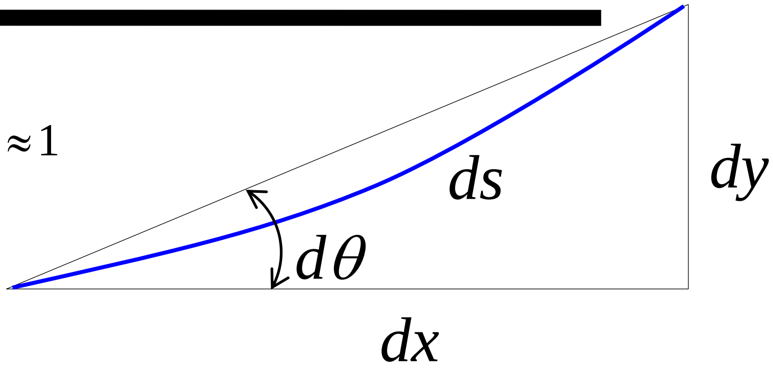
$$\rightarrow ds = dx\boxed{\left[1 + \frac{1}{2}\left(\frac{dy}{dx}\right)^2 + \frac{1}{2}\left(-\frac{1}{4}\right)\left(\frac{dy}{dx}\right)^4 + \cdots\right]}$$

$$if, \; \theta << 1$$

$$ds \approx dx$$

$$\therefore f(z) = \sqrt{1+z} \approx 1$$

$$\therefore ds \approx dx$$



Geometry of Deformation : Linearization

IF WE ASSUME, $\theta \ll 1$

$$\tan\theta = \theta + \frac{\theta^3}{3} + \frac{2\theta^5}{15} + \cdots$$

$$\theta \approx \tan(\theta)$$

$$\therefore \theta \approx \frac{dy}{dx}$$

Elongation

$$ds \approx dx$$

- at the neutral surface

No elongation

length of *AB* :

$$ds \approx dx$$

length of *A′B′* :

$$ds' = (\rho - y)d\theta$$

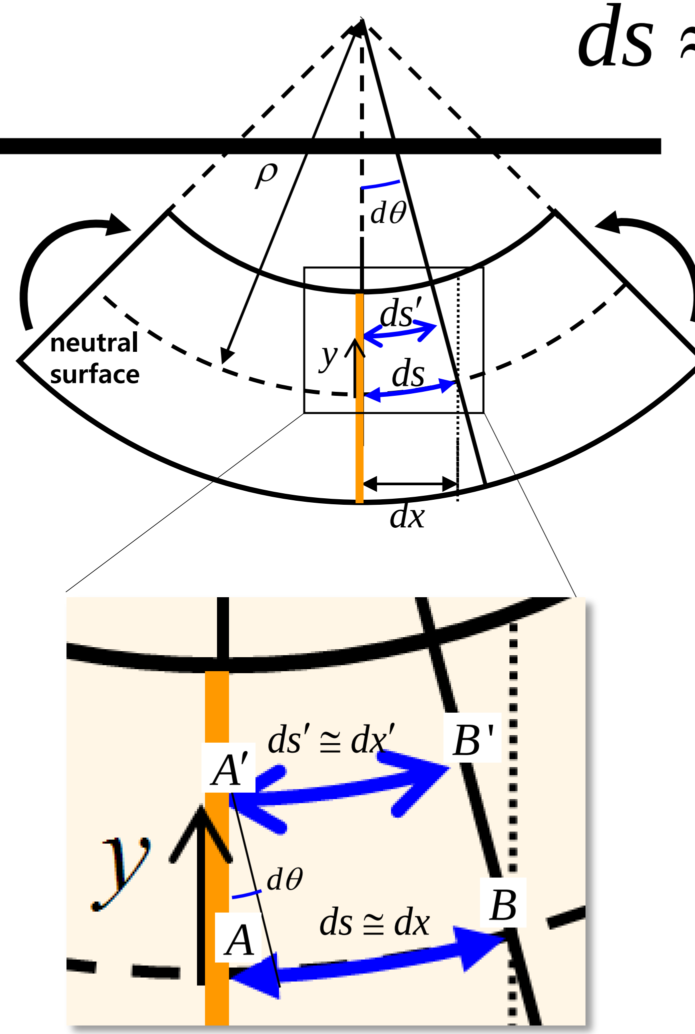
$$(ds' \approx dx')$$

$$dx' = (\rho - y)d\theta$$

$$dx' = \rho d\theta - y d\theta$$

$$dx' = dx - y d\theta$$

$$dx' = dx - y\frac{dx}{\rho}$$

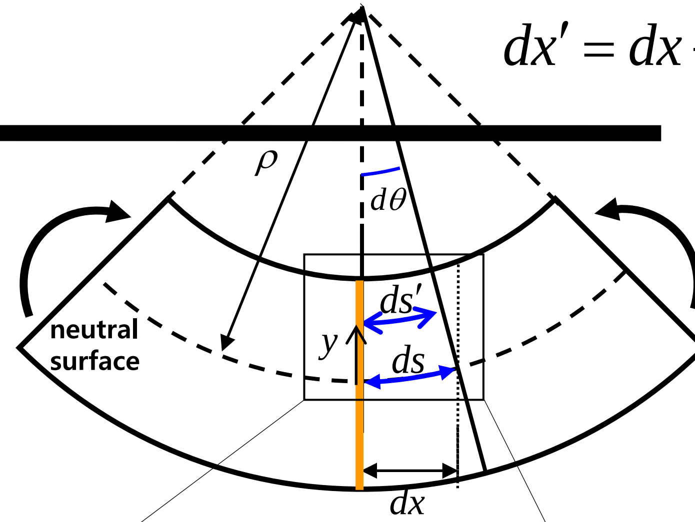


$$\rho d\theta = ds \approx dx$$

$$\therefore d\theta = \frac{dx}{\rho}$$

Strain

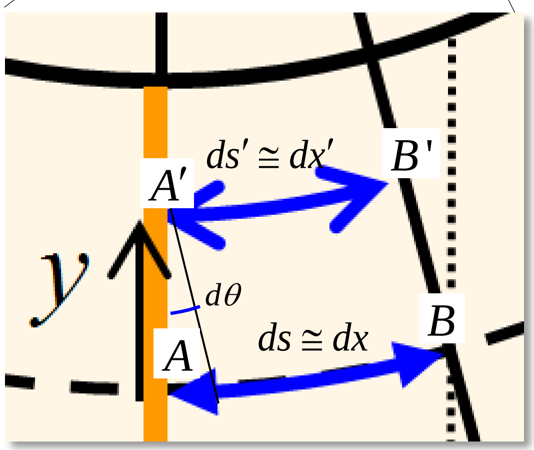
$$dx' = dx - y\frac{dx}{\rho}$$



$$dx' - dx = -y\frac{dx}{\rho}$$

$$\boxed{\frac{dx' - dx}{dx}} = -\frac{y}{\rho}$$

→ **Definition of strain**



$$\therefore \varepsilon_x = -\frac{y}{\rho}$$

Stress

$$\varepsilon_x = -\frac{y}{\rho}$$

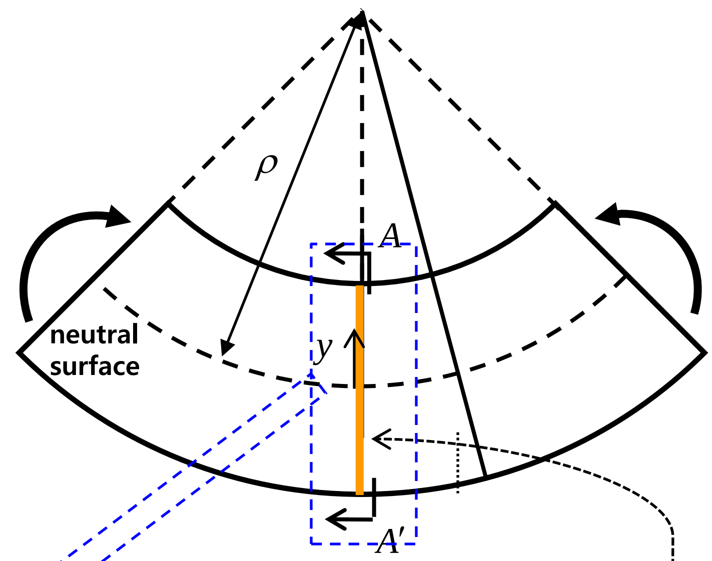
Hooke's Law

$$\sigma_x = E\varepsilon_x$$

$$\therefore \sigma_x = -E\frac{y}{\rho}$$

Bending Moment

$$\sigma_x = -E\frac{y}{\rho}$$

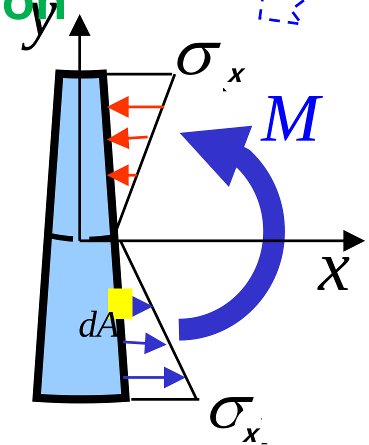


Bending moment about the neutral axis due to the normal stress acting on an infinitesimal area dA

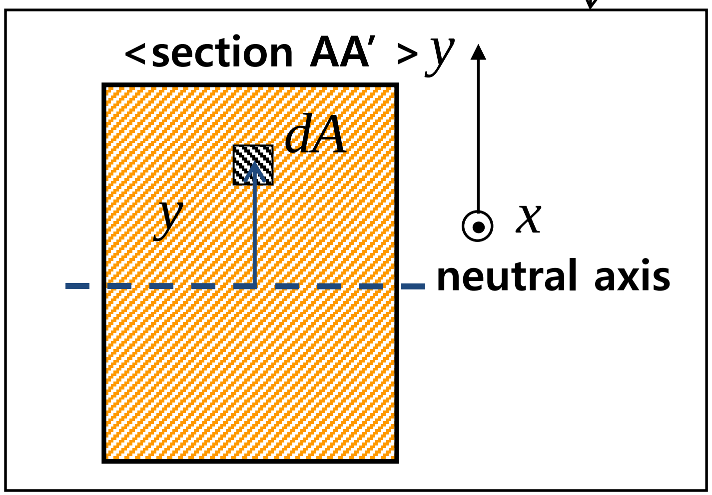
$$dM = y\sigma_x dA$$

Considering the sign convention, we need to add 'minus sign' for the bending moment. (Ref: To see this in detail, refer to the lecture on "sign convention")

$$\Rightarrow dM = \boxed{-}y\sigma_x dA$$



<Elevation view>



<Section view>

Bending Moment

$$dM = -\sigma_x y dA \quad , \sigma_x = -E\frac{y}{\rho}$$

The total Bending moment about the neutral axis due to the normal stress acting on the sectional area

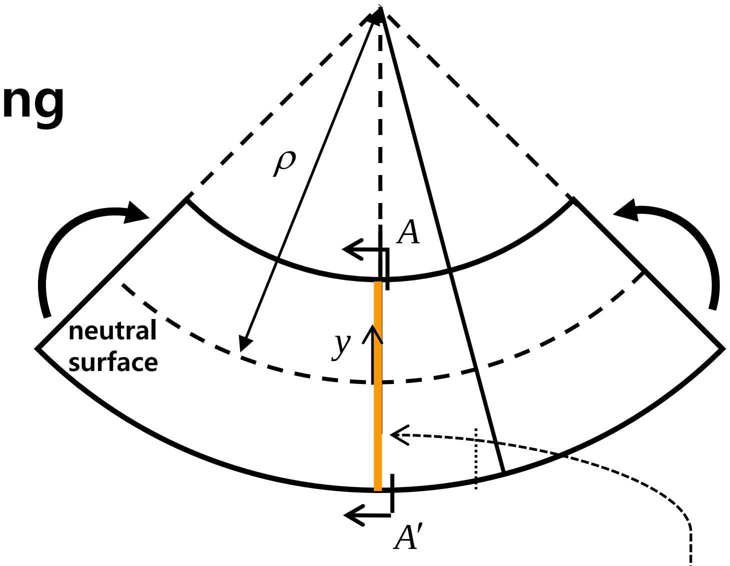
$$M = \int_A dM$$

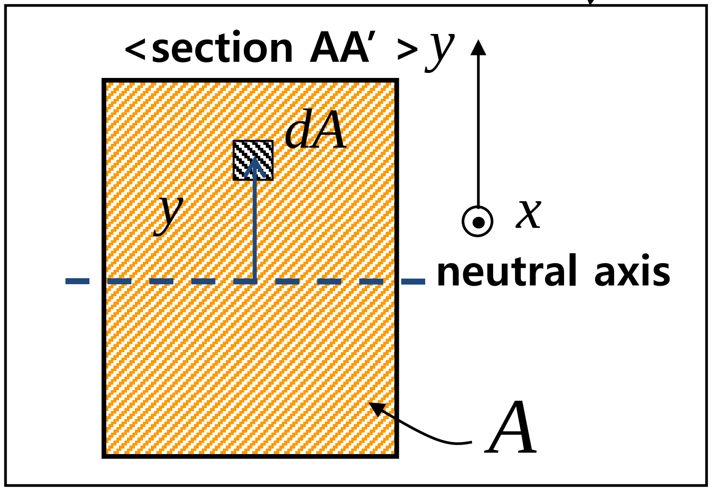
$$M = -\int_A y\sigma_x dA$$

$$= -\int_A y(-E\frac{y}{\rho})\, dA$$

$$= \frac{E}{\rho}\int_A y^2 dA$$

$$\therefore M = \frac{E}{\rho} I \qquad , I = \int_A y^2 dA$$



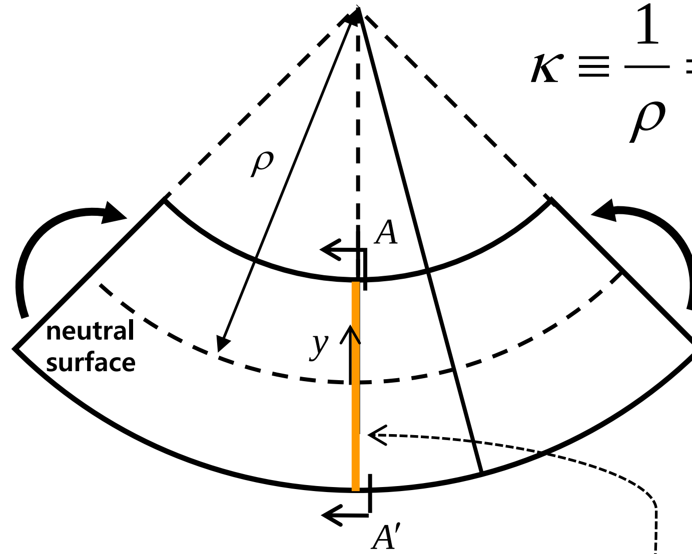


<Section view>

Relation between the deformation and the bending moment

$$M = \frac{E}{\rho} I \quad \Rightarrow \quad \frac{M}{EI} = \frac{1}{\rho} \quad \text{or} \quad \frac{M}{EI} = \kappa$$

$$\kappa \equiv \frac{1}{\rho} = \frac{d\theta}{ds}$$



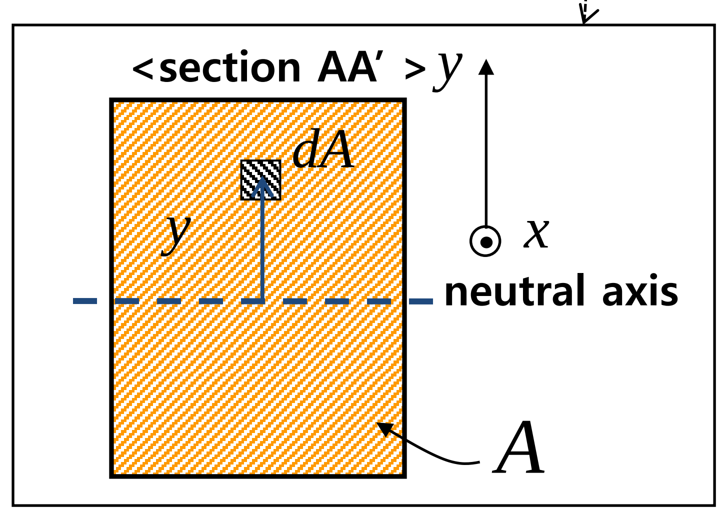
By performing linearization of the deformed geometry for small deflection

$$\kappa = \frac{d\theta}{ds}$$

$$\approx \frac{d}{ds}\tan(\theta) = \frac{d}{ds}\left(\frac{dy}{dx}\right)$$

$$\approx \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d^2 y}{dx^2}$$

$$\boxed{\therefore \frac{M}{EI} = \frac{d^2 y}{dx^2}}$$

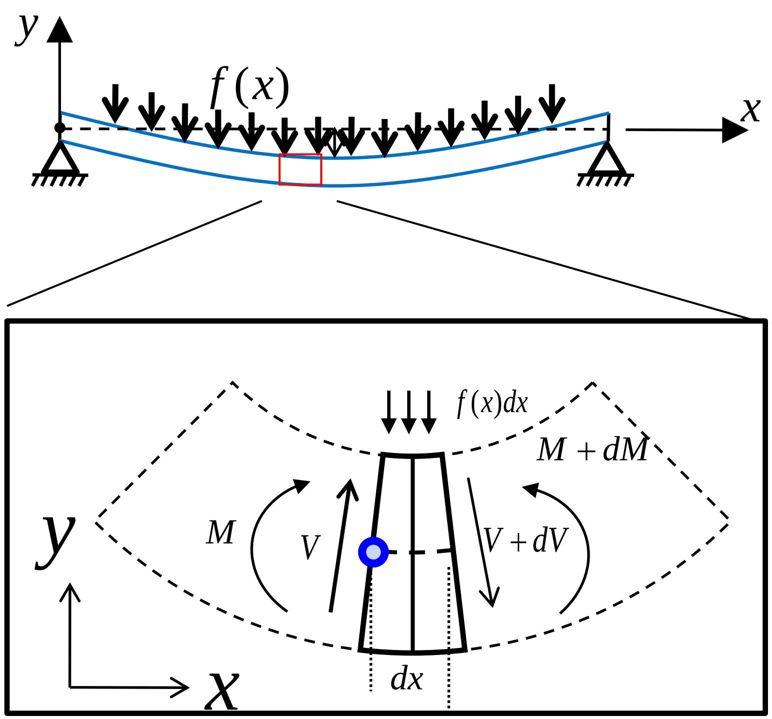


<Section view>

$$I = \int_A y^2 dA$$

Relation between bending moment and shear force

Let us consider the distributed load acting on a beam



From the moment equilibrium about z-axis through the point ○ , we obtain :

$$-M+(M+dM)-(V+dV)dx-\frac{1}{2}dx\cdot f(x)dx=0$$

$$dM-Vdx-dV\cdot dx-\frac{1}{2}(dx)^2\cdot f(x)=0$$

neglecting the high order terms

$$\boxed{\therefore \frac{dM}{dx}=V(x)}$$

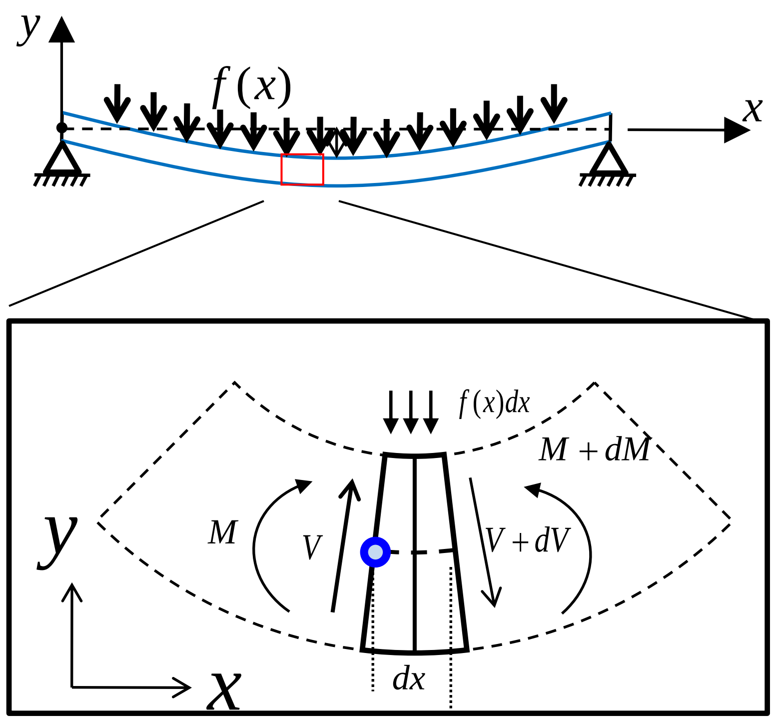
Ref: To see the direction of the shear forces and bending moments in the Fig. in detail, refer to the lecture on "sign convention"

Relation between shear force and distributed load

From the force equilibrium, we obtain

$$V - (V + dV) - f(x)dx = 0$$

$$\boxed{\therefore \frac{dV}{dx} = -f(x)}$$



Relation between the deformation and the distributed load

$$\boxed{\frac{d^2y}{dx^2} = \frac{M}{EI}} \rightarrow \boxed{EI\frac{d^2y}{dx^2} = M}$$

Differentiation with respect to *x*

$$\boxed{EI\frac{d^3y}{dx^3} = V}$$

Differentiation with respect to *x*

$$\boxed{EI\frac{d^4y}{dx^4} = -f(x)}$$

"Deflection Curve of a Beam"

$$\frac{M}{EI} = \frac{d^2y}{dx^2}$$

$$\frac{dM}{dx} = V(x)$$

$$\frac{dV}{dx} = -f(x)$$

What if the distributed load is not applied?

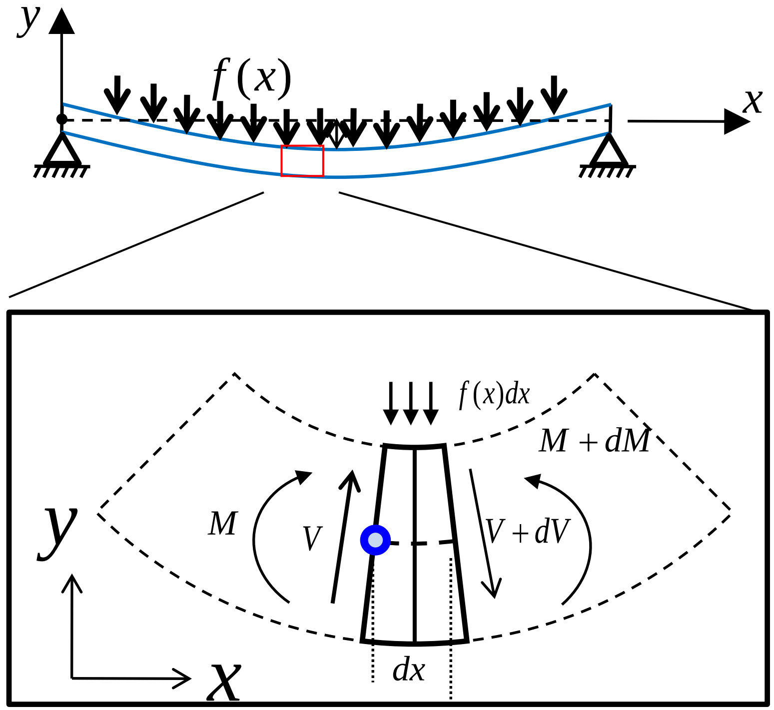
1) From the moment equilibrium about z-axis through the point ◦ , we obtain :

$$-M+(M+dM)-(V+dV)dx-\frac{1}{2}dx\cdot f(x)dx=0$$

$$dM-Vdx-dV\cdot dx-\frac{1}{2}(dx)^2\cdot f(x)=0$$

neglecting the high order terms, and $f(x)=0$

$$\frac{dM}{dx}=V(x) \quad \Rightarrow \quad \boxed{dM=V(x)dx}$$



2) From the force equilibrium, we obtain

$$V-(V+dV)-f(x)dx=0$$

Since $f(x)=0$

$$\boxed{dV=0}$$

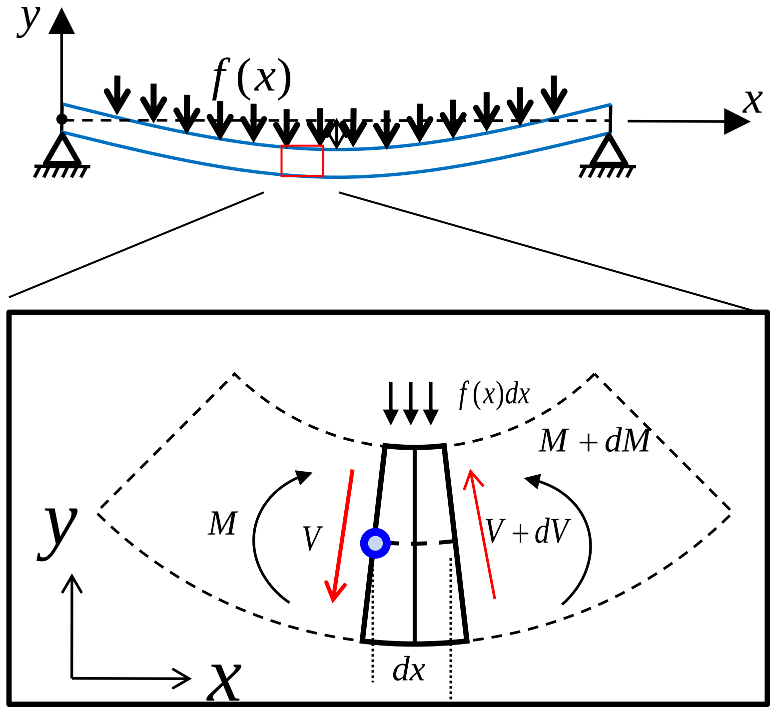
If the distributed load is not applied to the beam, the shear force V is constant, but the bending moment M is not.

Derivation of Deflection Curve of Beam **by applying opposite sign convention for shear force**

Relation between bending moment and shear force

…Continued from page 53

Let us consider the distributed load acting on a beam



From the moment equilibrium about z-axis through the point ○ , we obtain :

$$-M + (M + dM) + (V + dV)dx - \frac{1}{2}dx \cdot f(x)dx = 0$$

$$dM + V \cdot dx + dV \cdot dx - \frac{1}{2}(dx)^2 \cdot f(x) = 0$$

neglecting the high order terms

$$\boxed{\therefore \frac{dM}{dx} = -V(x)}$$

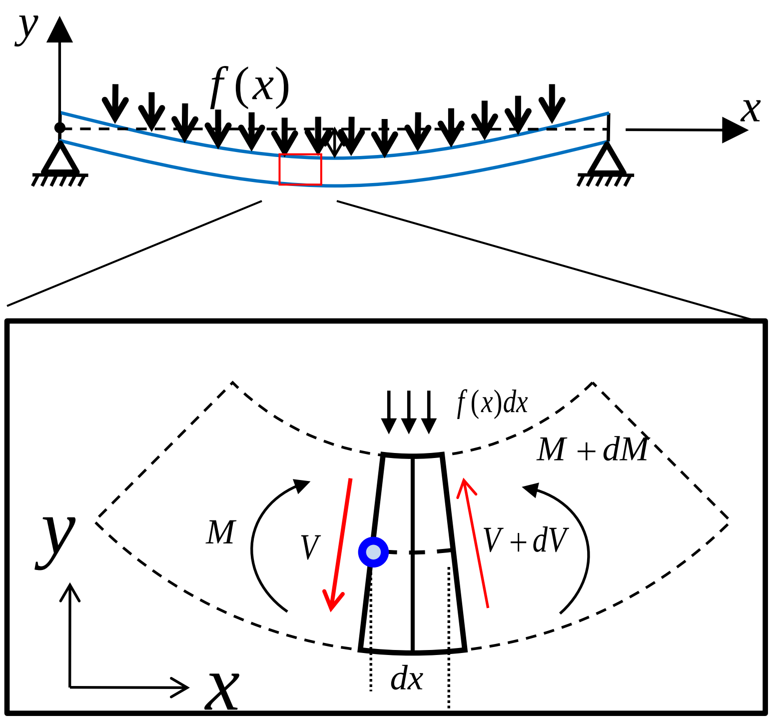
Ref: To see the direction of the shear forces and bending moments in the Fig. in detail, refer to the lecture on "sign convention"

Relation between shear force and distributed load

From the force equilibrium,
we obtain

$$-V + (V + dV) - f(x) \cdot dx = 0$$

$$\boxed{\therefore \frac{dV}{dx} = f(x)}$$



Relation between the deformation and the distributed load

$$\frac{d^2y}{dx^2} = \frac{M}{EI} \rightarrow EI\frac{d^2y}{dx^2} = M$$

Differentiation with respect to *x*

$$EI\frac{d^3y}{dx^3} = -V$$

Differentiation with respect to *x*

$$EI\frac{d^4y}{dx^4} = -f(x)$$

"Deflection Curve of a Beam"

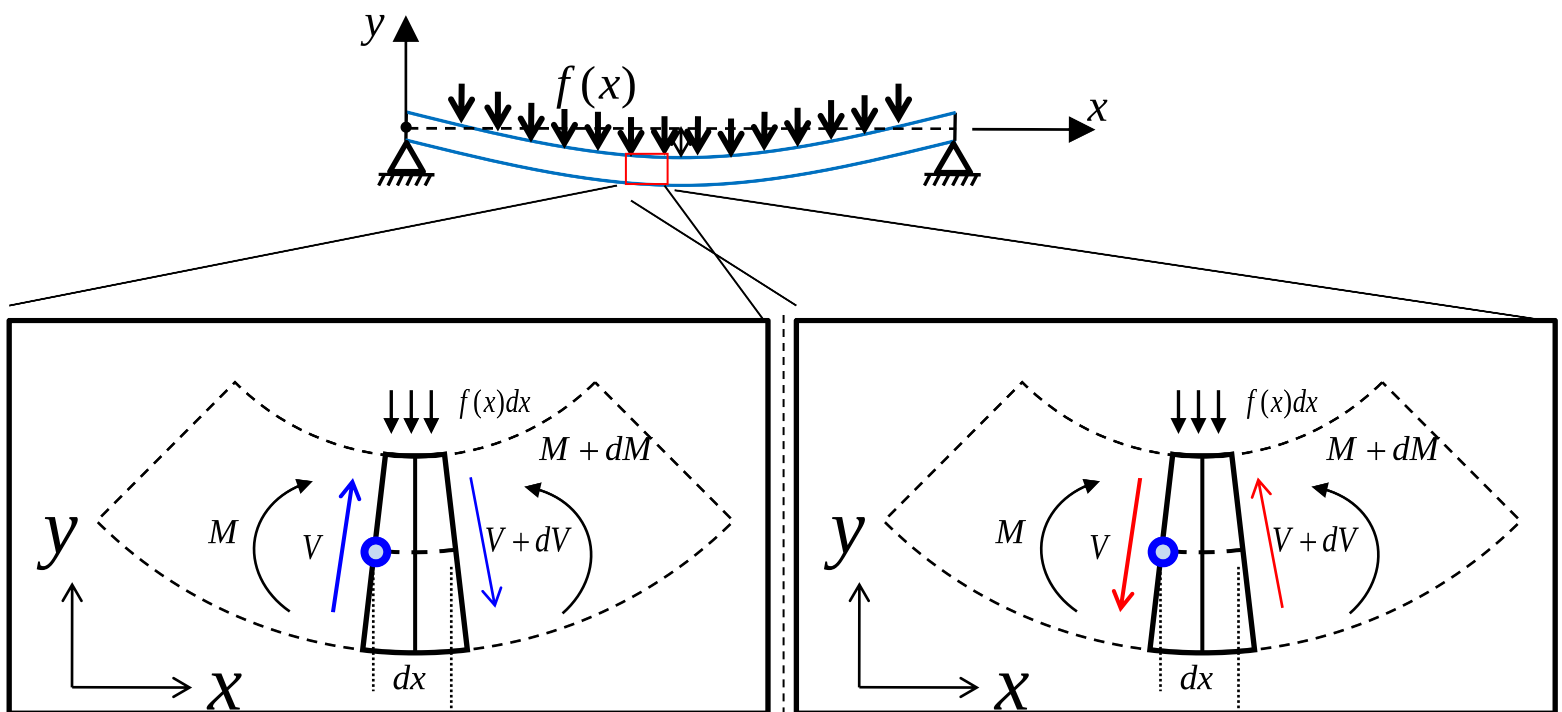
$$\frac{M}{EI} = \frac{d^2y}{dx^2}$$

$$\frac{dM}{dx} = -V(x)$$

$$\frac{dV}{dx} = f(x)$$

Comparison of Derivation of Deflection Curve of a Beam by applying different sign convention for shear force

Sign conventions for shear forces



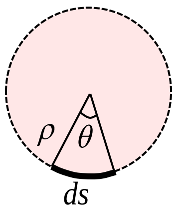
Bending moment	$EI\dfrac{d^2y}{dx^2}=M$	Same	$EI\dfrac{d^2y}{dx^2}=M$
Shear force	$EI\dfrac{d^3y}{dx^3}=V$	Not ~~same~~	$EI\dfrac{d^3y}{dx^3}=-V$
Deflection Curve of a Beam	$EI\dfrac{d^4y}{dx^4}=-f(x)$	Same	$EI\dfrac{d^4y}{dx^4}=-f(x)$

Derivation of Deflection Curve of Beam with Vector Notation

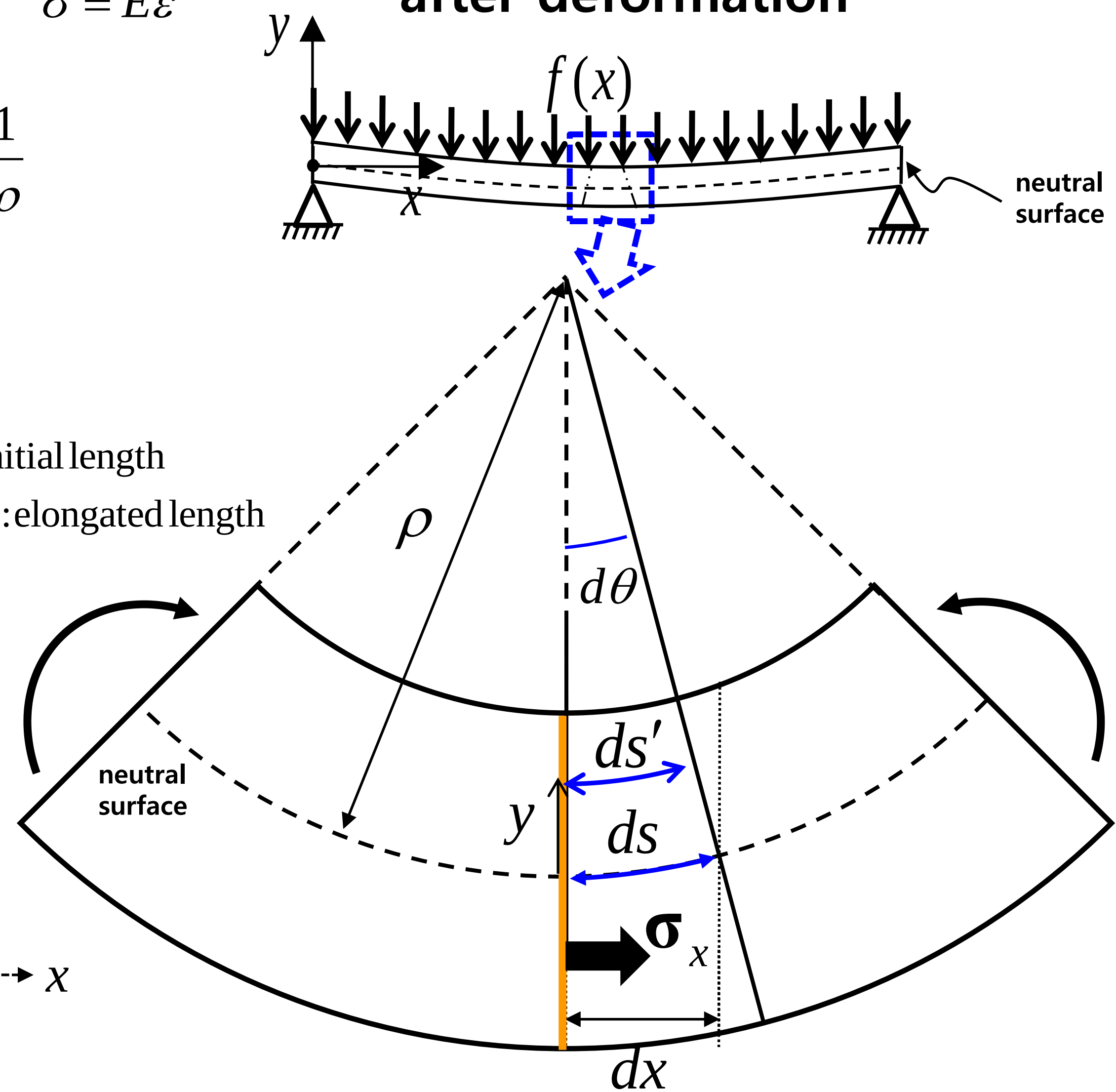
Deflection of Beam with Vector Notation

$\boldsymbol{\sigma}_x = \sigma \mathbf{i}$, $\boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}$, $\boldsymbol{\theta} = \theta \mathbf{k}$, $\mathbf{y} = y\mathbf{j}$ $\sigma = E\varepsilon$

-after deformation



$\rho \cdot d\theta = ds \Rightarrow \dfrac{d\theta}{ds} = \dfrac{1}{\rho}$



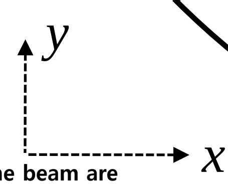
① **strain at** y **in x-direction :**

$$\boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}$$

$$\varepsilon = \frac{(\rho - y)\cdot d\theta - \rho d\theta}{ds} = -y\frac{d\theta}{ds} = -\frac{y}{\rho}$$

$, ds$: initial length

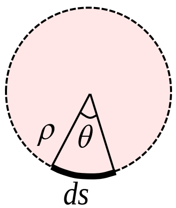
$, y \cdot d\theta$: elongated length



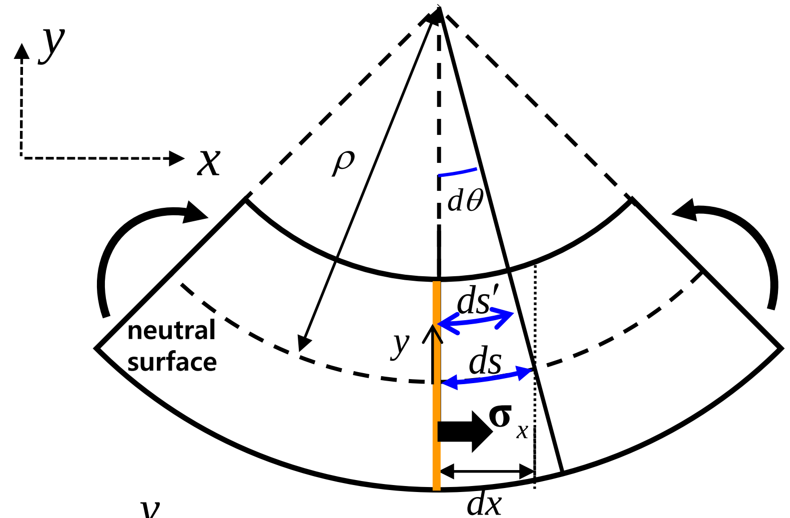
* neutral surface : Longitudinal lines on the lower part of the beam are elongated, whereas those on the upper part are shortened. Thus the lower part of the beam is in tension and the upper part is in compression. Somewhere between the top and bottom of the beam is a surface in which longitudinal lines do not change in length. This surface is called neutral surface

Deflection of Beam with Vector Notation

$$\boldsymbol{\sigma}_x = \sigma \mathbf{i} \;,\; \boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i} \;,\; \boldsymbol{\theta} = \theta \mathbf{k},\; \mathbf{y} = y\mathbf{j} \qquad \sigma = E\varepsilon$$



$$\rho \cdot d\theta = ds \Rightarrow \frac{d\theta}{ds} = \frac{1}{\rho}$$



② stress at y in x-direction : $\boldsymbol{\sigma}_x = \sigma \mathbf{i} = E \cdot \varepsilon \mathbf{i}$, where $\varepsilon = -\dfrac{y}{\rho}$

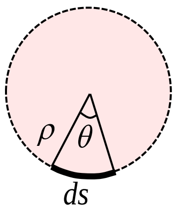
$$\therefore \boldsymbol{\sigma}_x = \sigma \mathbf{i} = -E\frac{y}{\rho}\mathbf{i}$$

③ force acting on dA in x-direction : $d\mathbf{F}_x = \boldsymbol{\sigma}_x dA = (\sigma \mathbf{i})dA = \sigma dA\mathbf{i}$ $\qquad \therefore d\mathbf{F} = -E\dfrac{y}{\rho}dA\mathbf{i}$

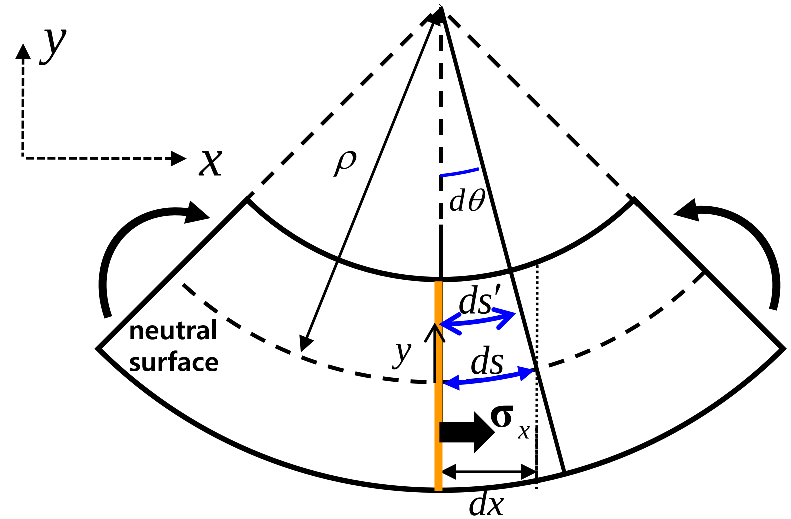
④ moment about z-axis : $d\mathbf{M} = \mathbf{y} \times d\mathbf{F} = (y\mathbf{j}) \times (-E\dfrac{y}{\rho}dA\mathbf{i}) = E\dfrac{y^2}{\rho}dA\mathbf{k}$

Deflection of Beam with Vector Notation

$\boldsymbol{\sigma}_x = \sigma \mathbf{i}$, $\boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}$, $\boldsymbol{\theta} = \theta \mathbf{k}$, $\mathbf{y} = y\mathbf{j}$ $\quad \sigma = E\varepsilon$



$$\rho \cdot d\theta = ds \Rightarrow \frac{d\theta}{ds} = \frac{1}{\rho}$$



④ **moment about z-axis :**

$$d\mathbf{M} = \mathbf{y} \times d\mathbf{F} = (y\mathbf{j}) \times (-E\frac{y}{\rho}dA\mathbf{i}) = E\frac{y^2}{\rho}dA\mathbf{k}$$

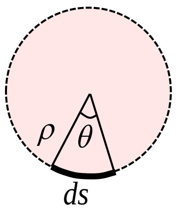
$$\therefore \mathbf{M} = \int_A d\mathbf{M} = \int_A E\frac{y^2}{\rho}dA\mathbf{k}$$

Define $I = \int_A y^2 dA$ then, $\mathbf{M} = \frac{EI}{\rho}\mathbf{k}$, $M = \frac{EI}{\rho}$ $\Rightarrow$ $\mathbf{M} = \frac{EI}{\rho}\mathbf{k}$

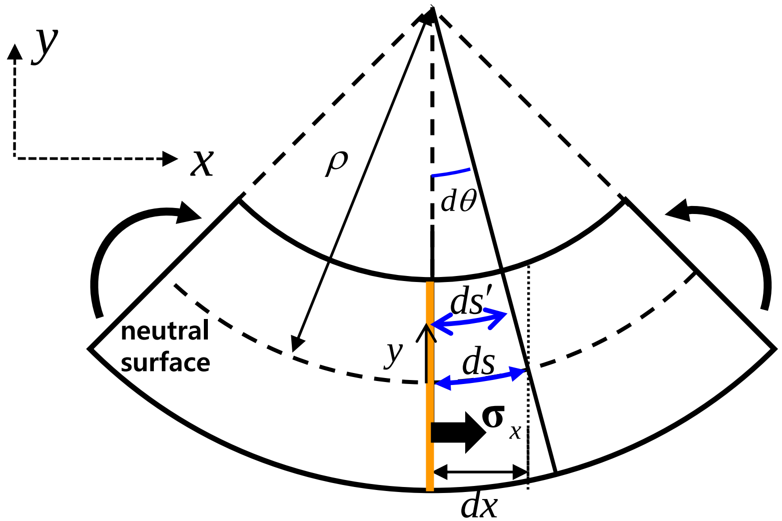
$$\mathbf{M} = EI\frac{d\theta}{ds}\mathbf{k}$$

Deflection of Beam with Vector Notation

$\boldsymbol{\sigma}_x = \sigma \mathbf{i}$, $\boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}$, $\boldsymbol{\theta} = \theta \mathbf{k}$, $\mathbf{y} = y\mathbf{j}$ $\quad \sigma = E\varepsilon$



$$\rho \cdot d\theta = ds \Rightarrow \frac{d\theta}{ds} = \frac{1}{\rho}$$



④ moment about z-axis :

$$\mathbf{M} = \frac{EI}{\rho}\mathbf{k}$$

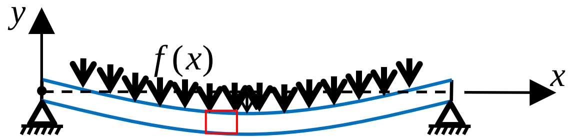
$$\mathbf{M} = EI\frac{d\theta}{ds}\mathbf{k}$$

⑤ assume $ds \approx dx,\ \theta \approx \tan(\theta) = \dfrac{dy}{dx}$

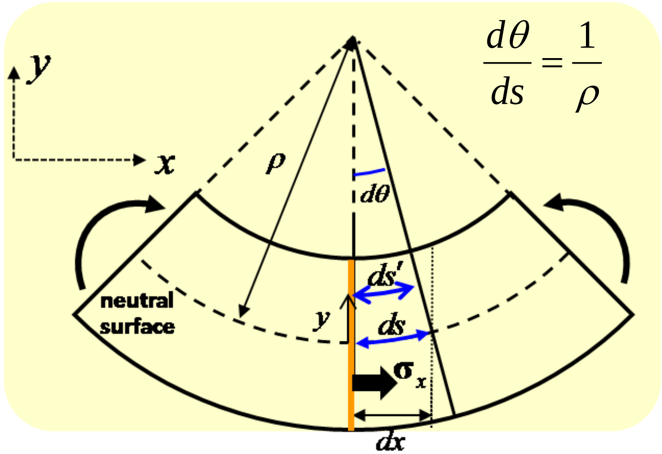
$$\frac{d\theta}{ds} = \frac{d}{ds}\left(\frac{dy}{dx}\right) = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d^2y}{dx^2} \Rightarrow \frac{d\theta}{ds} = \frac{d^2y}{dx^2}$$

$$\boxed{\mathbf{M} = EI\frac{d^2y}{dx^2}\mathbf{k} \qquad , M = EI\frac{d^2y}{dx^2}}$$

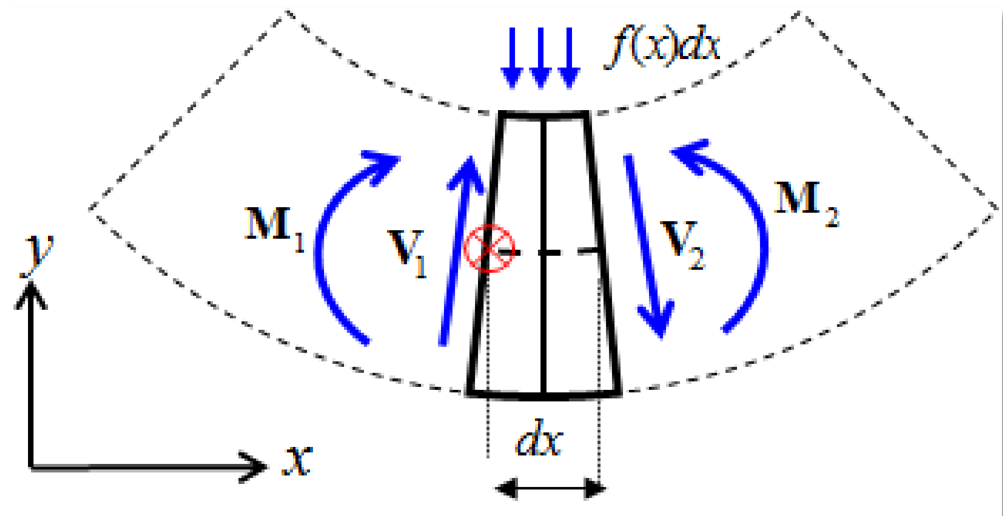
Deflection of Beam with Vector Notation



$$\boldsymbol{\sigma}_x = \sigma \mathbf{i}\ ,\ \boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}\ ,\ \boldsymbol{\theta} = \theta \mathbf{k},\ \mathbf{y} = y\mathbf{j} \qquad \sigma = E\varepsilon$$



⑥ relationships between loads, shear forces, and bending moments



$$\mathbf{V}_1 = V\mathbf{j},\ \ \mathbf{V}_2 = -\left(V + \frac{\partial V}{\partial x}dx\right)\mathbf{j}$$

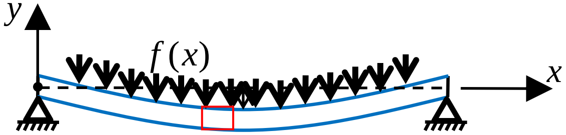
•force equilibrium $\sum \mathbf{F}_y = \mathbf{V}_1 + \mathbf{V}_2 + \mathbf{f}(x)\cdot dx = 0$

$$(V_1\mathbf{j}) + \left(-\left(V_1 + \frac{\partial V_1}{\partial x}dx\right)\mathbf{j}\right) + \left(-f(x)\cdot dx\mathbf{j}\right) = 0$$

$$\left(V_1 - V_1 - \frac{\partial V_1}{\partial x}dx - f(x)\cdot dx\right)\mathbf{j} = 0$$

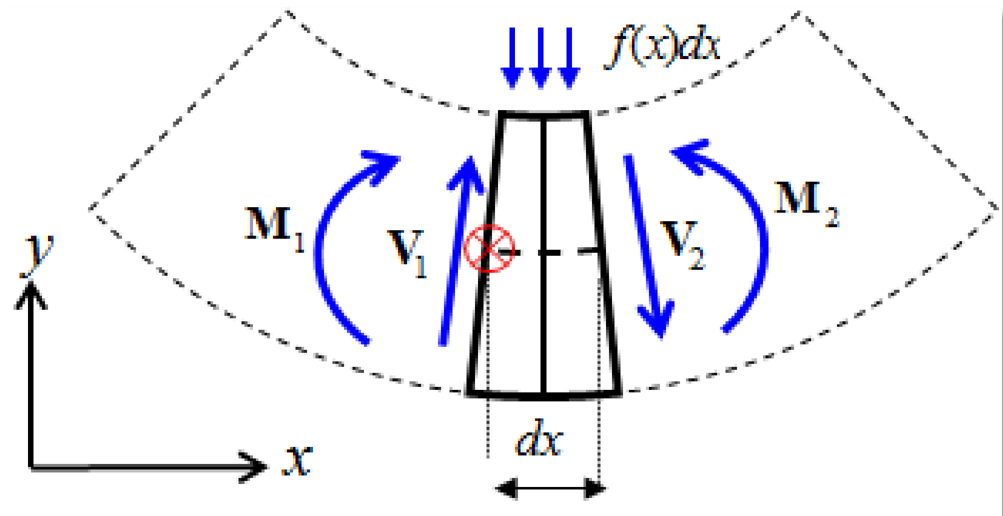
$$\boxed{\therefore \frac{dV}{dx} = -f(x)}$$

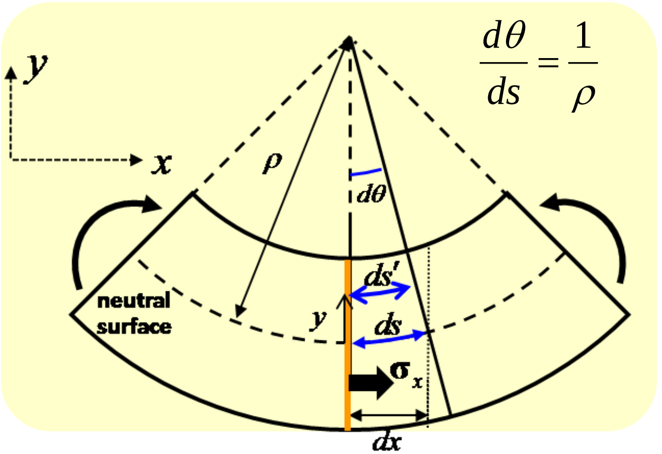
Deflection of Beam with Vector Notation



$$\boldsymbol{\sigma}_x = \sigma\mathbf{i}\ ,\ \boldsymbol{\varepsilon}_x = \varepsilon\mathbf{i}\ ,\ \boldsymbol{\theta} = \theta\mathbf{k},\ \mathbf{y} = y\mathbf{j} \qquad \sigma = E\varepsilon$$

⑥ **relationships between loads, shear forces, and bending moments**





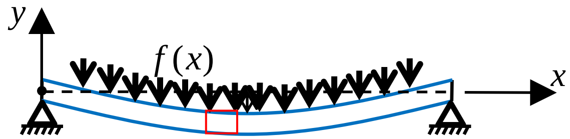
$$\mathbf{V}_1 = V\mathbf{j},\ \ \mathbf{V}_2 = -\left(V + \frac{\partial V}{\partial x}dx\right)\mathbf{j},\ \ \mathbf{M}_1 = -M\mathbf{k},\ \ \mathbf{M}_2 = \left(M + \frac{\partial M}{\partial x}dx\right)\mathbf{k}$$

•moment equilibrium $\sum \mathbf{M}_z = \mathbf{M}_1 + \mathbf{M}_2 + d\mathbf{x}\times\mathbf{V}_2 + \frac{1}{2}d\mathbf{x}\times\left(\mathbf{f}(x)\cdot dx\right) = 0$

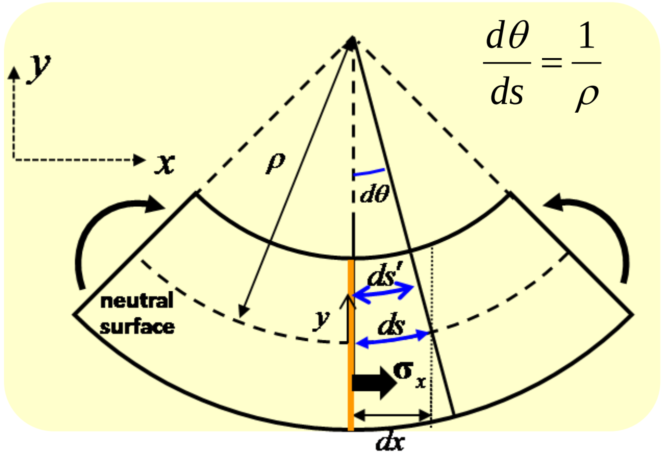
$$-M\mathbf{k} + \left(M + \frac{\partial M}{\partial x}dx\right)\mathbf{k} + (dx\,\mathbf{i})\times\left(-\left(V + \frac{\partial V}{\partial x}dx\right)\mathbf{j}\right) + \left(\frac{1}{2}dx\,\mathbf{i}\right)\times\left(-f(x)\cdot dx\mathbf{j}\right) = 0$$

$$\boxed{\therefore\ \frac{dM}{dx} = V(x)}$$

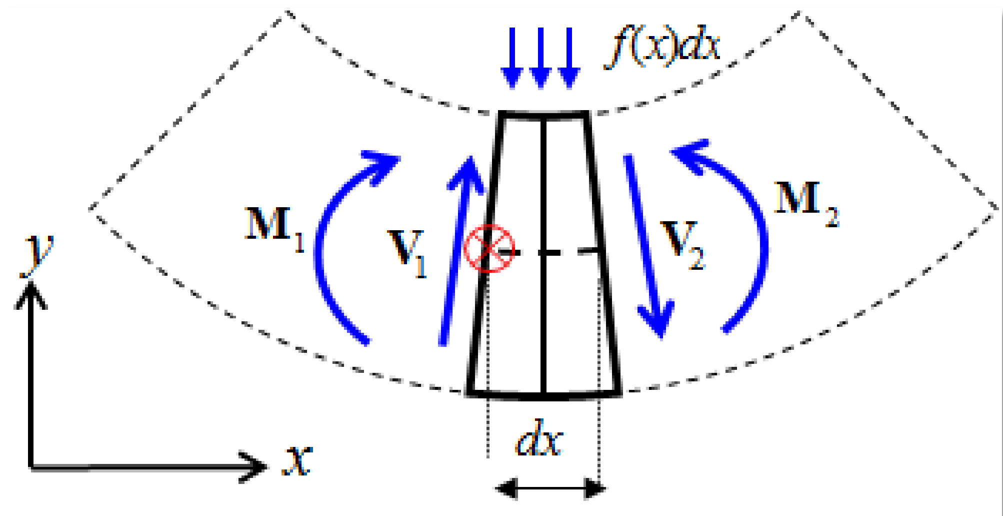
Deflection of Beam with Vector Notation



$$\boldsymbol{\sigma}_x = \sigma\mathbf{i}\ ,\ \boldsymbol{\varepsilon}_x = \varepsilon\mathbf{i}\ ,\ \boldsymbol{\theta} = \theta\mathbf{k},\ \mathbf{y} = y\mathbf{j} \qquad \sigma = E\varepsilon$$



⑥ relationships between loads, shear forces, and bending moments



•force equilibrium $\frac{dV}{dx} = -f(x)$

•moment equilibrium $\frac{dM}{dx} = V(x)$

$$\mathbf{V}_1 = V\mathbf{j},\ \ \mathbf{V}_2 = -\left(V + \frac{\partial V}{\partial x}dx\right)\mathbf{j},\ \ \mathbf{M}_1 = -M\mathbf{k},\ \ \mathbf{M}_2 = \left(M + \frac{\partial M}{\partial x}dx\right)\mathbf{k}$$

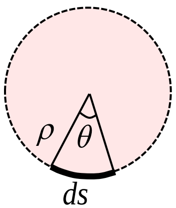
$$\frac{d^2y}{dx^2} = \frac{M}{EI} \rightarrow \frac{d^3y}{dx^3} = \frac{1}{EI}\cdot\frac{dM}{dx} = \frac{1}{EI}\cdot V(x) \rightarrow \frac{d^4y}{dx^4} = \frac{1}{EI}\cdot\frac{dV}{dx} = -\frac{1}{EI}\cdot f(x)$$

$$\boxed{\therefore EI\frac{d^4y}{dx^4} = -f(x)}$$

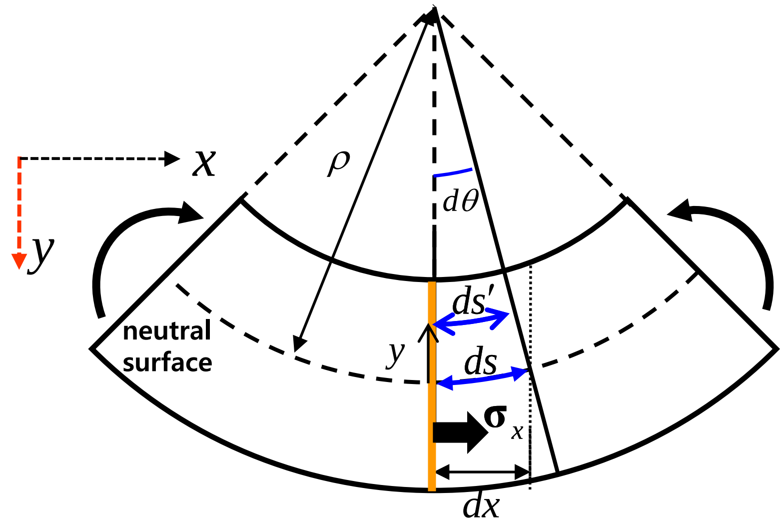
Deflection of Beam with Vector Notation

what happen
if we take the direction of
y axis reversed?

$\boldsymbol{\sigma}_x = \sigma\mathbf{i}$, $\boldsymbol{\varepsilon}_x = \varepsilon\mathbf{i}$, $\boldsymbol{\theta} = \theta\mathbf{k}$, $\mathbf{y} = y\mathbf{j}$ $\quad \sigma = E\varepsilon$



$\rho \cdot d\theta = ds \Rightarrow \dfrac{d\theta}{ds} = \dfrac{1}{\rho}$



① **strain at y in x-direction :** $\varepsilon = \dfrac{(\rho + y)\cdot d\theta - \rho \cdot d\theta}{ds} = y\dfrac{d\theta}{ds}$

$\boldsymbol{\varepsilon}_x = \varepsilon\mathbf{i}$ $\quad$ $, d\theta$: initial length, $\; y\cdot d\theta$: elongated length

② **stress at y in x-direction :** $\boldsymbol{\sigma}_x = \sigma\mathbf{i} = E\dfrac{y}{\rho}\mathbf{i}$

③ **force acting on dA in x-direction :** $d\mathbf{F} = E\dfrac{y}{\rho}dA\mathbf{i}$

④ **moment about z-axis :** $d\mathbf{M} = \mathbf{y} \times d\mathbf{F} = (y\mathbf{j}) \times (E\dfrac{y}{\rho}dA\mathbf{i}) = -E\dfrac{y^2}{\rho}dA\mathbf{k} \;\therefore \mathbf{M} = \int_A d\mathbf{M} = -\int_A E\dfrac{y^2}{\rho}dA\mathbf{k}$

Define $I = \int_A y^2 dA$ then, $\mathbf{M} = -\dfrac{EI}{\rho}\mathbf{k}$, $M = -\dfrac{EI}{\rho}$

$\Rightarrow \mathbf{M} = -\dfrac{EI}{\rho}\mathbf{k}$

$\Downarrow$

$\mathbf{M} = -EI\dfrac{d\theta}{ds}\mathbf{k}$

$\Downarrow$

⑤ **assume** $ds \approx dx$, $\theta \approx \tan(\theta) = \dfrac{dy}{dx}$

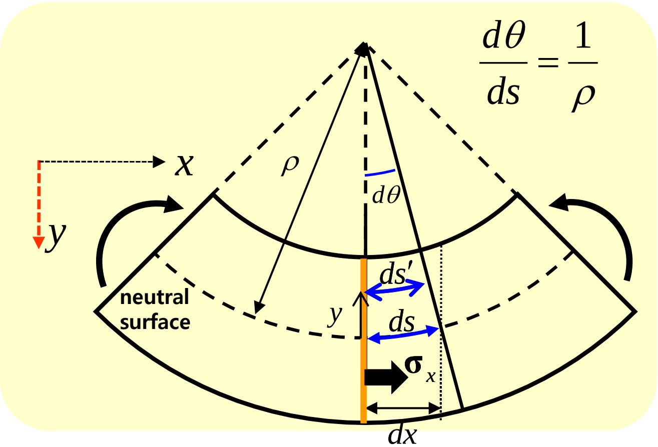
$\dfrac{d\theta}{ds} = \dfrac{d}{ds}\left(\dfrac{dy}{dx}\right) = \dfrac{d}{dx}\left(\dfrac{dy}{dx}\right) = \dfrac{d^2y}{dx^2} \Rightarrow \dfrac{d\theta}{ds} = \dfrac{d^2y}{dx^2}$

$$\boxed{\mathbf{M} = -EI\frac{d^2y}{dx^2}\mathbf{k} \quad , M = -EI\frac{d^2y}{dx^2}}$$

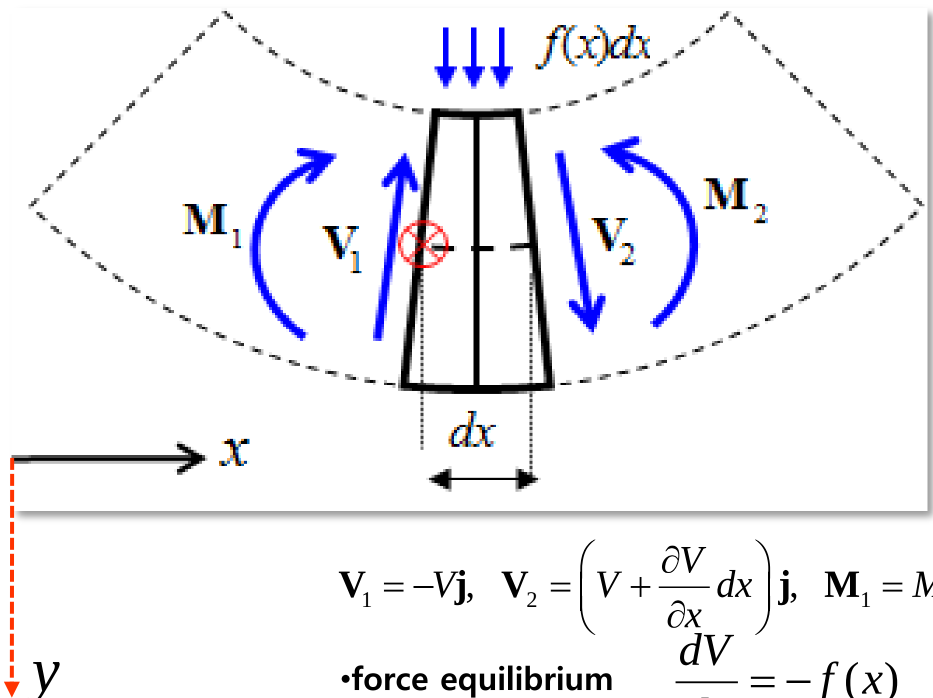
Deflection of Beam with Vector Notation

what happen
if we take the direction of
y axis reversed?

$$\boldsymbol{\sigma}_x = \sigma \mathbf{i}\ ,\ \boldsymbol{\varepsilon}_x = \varepsilon \mathbf{i}\ ,\ \boldsymbol{\theta} = \theta \mathbf{k},\ \mathbf{y} = y\mathbf{j} \qquad \sigma = E\varepsilon$$



⑥ relationships between loads, shear forces, and bending moments



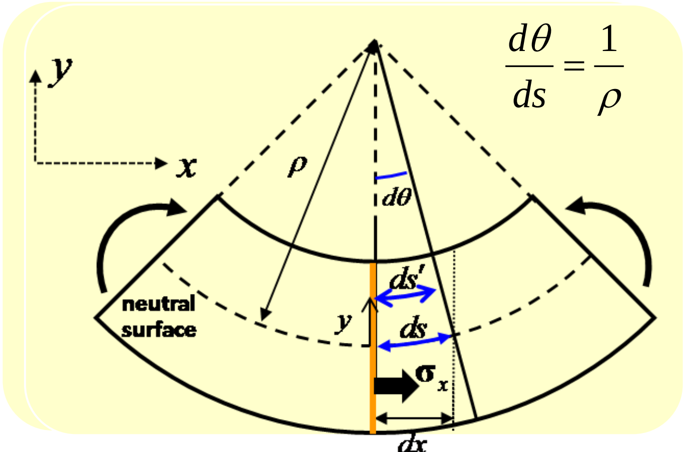
$$\mathbf{V}_1 = -V\mathbf{j},\ \ \mathbf{V}_2 = \left(V + \frac{\partial V}{\partial x}dx\right)\mathbf{j},\ \ \mathbf{M}_1 = M\mathbf{k},\ \ \mathbf{M}_2 = -\left(M + \frac{\partial M}{\partial x}dx\right)\mathbf{k}$$

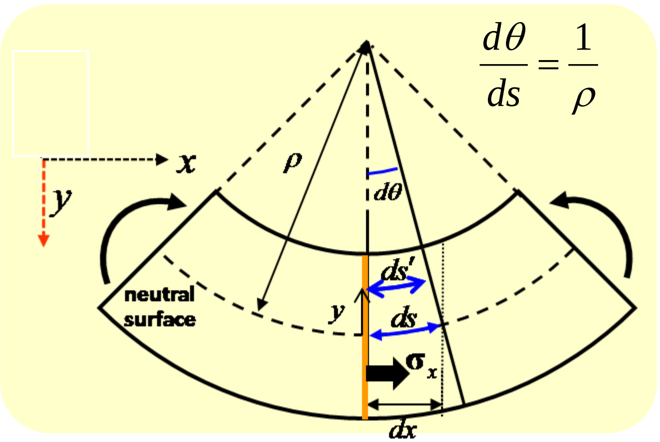
•force equilibrium $\dfrac{dV}{dx} = -f(x)$ •moment equilibrium $\dfrac{dM}{dx} = V(x)$

$$\frac{d^2y}{dx^2} = -\frac{M}{EI} \rightarrow \frac{d^3y}{dx^3} = -\frac{1}{EI}\cdot\frac{dM}{dx} = -\frac{1}{EI}\cdot V(x) \rightarrow \frac{d^4y}{dx^4} = -\frac{1}{EI}\cdot\frac{dV}{dx} = \frac{1}{EI}\cdot f(x)$$

$$\boxed{\therefore EI\frac{d^4y}{dx^4} = f(x)}$$

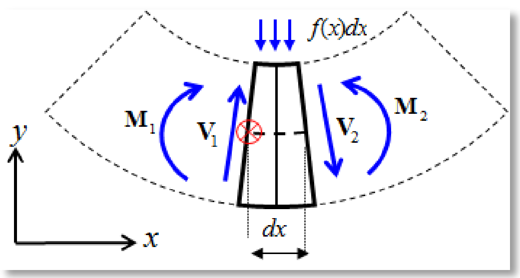
Deflection of Beam with Vector Notation



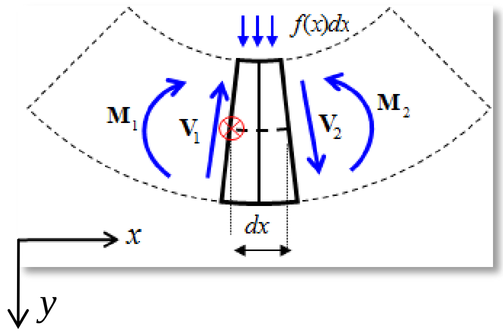


$$M = EI\frac{d^2y}{dx^2}$$

$$M = -EI\frac{d^2y}{dx^2}$$



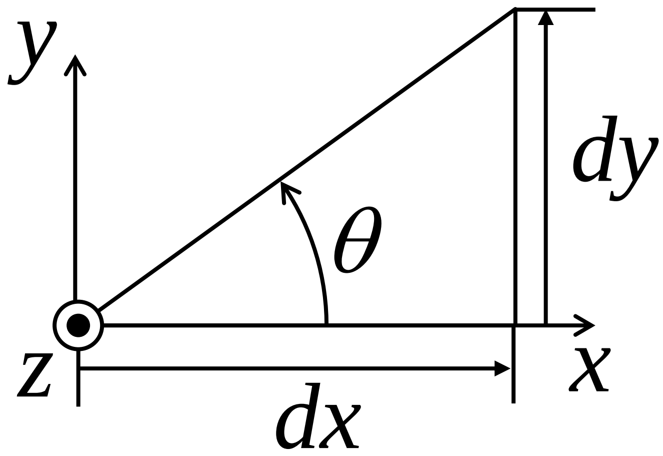
$$\frac{dV}{dx} = -f(x) \ , \frac{dM}{dx} = V(x)$$



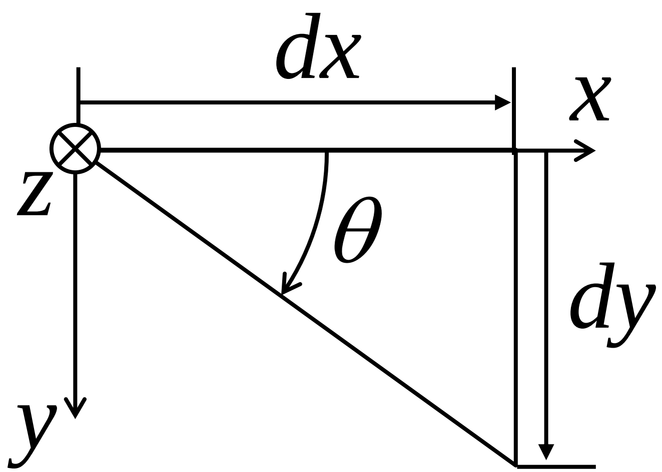
$$\boxed{\therefore EI\frac{d^4y}{dx^4} = -f(x)}$$

$$\boxed{\therefore EI\frac{d^4y}{dx^4} = f(x)}$$

Reference) Linearization of tanθ in the different coordinate system



$$\theta \approx \tan\theta = \frac{dy}{dx}$$



$$\theta \approx \tan\theta = \frac{dy}{dx}$$

The positive direction is opposite not only for the y-coordinate, but also for the angle θ. Therefore, the results of the linearization of tanθ is same in the different coordinate system

1. Beam Theory

1.3 Sign Convention

Sign Convections

References : Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006	
Sign Convention for Normal Stress	Sec. 1.2 p4
Deformation Sign Convention and Static Sign Convention	Sec. 4.3, p270~p271
Curvature Sign Convention	Sec. 5.3, p303
Differential Equation of the Deflection Curve	Sec. 9.2, p594~p599

Sign Conventions*

Static sign conventions

When writing equations of equilibrium
we use static sign conventions,
in which forces are positive or negative
according to their directions along the coordinates axes.
They depend upon the coordinates system

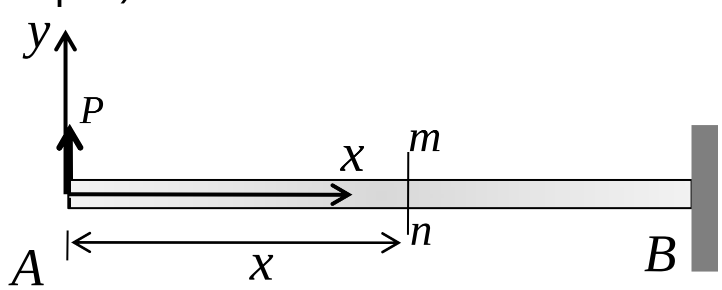
*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, pp270-271

Static Sign Conventions for Equation of Equilibrium*

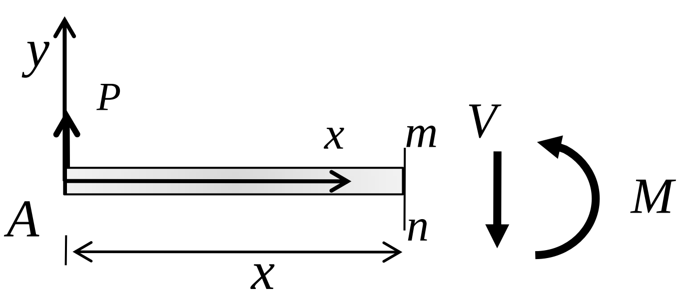
Static Sign Convention

-When writing equations of equilibrium, forces are positive or negative according to their direction along the coordinate axes

Example)



Free-body diagram of the left part



Force equilibrium

$$+p-V=0$$

Moment equilibrium about *z*-axis through the point *A*

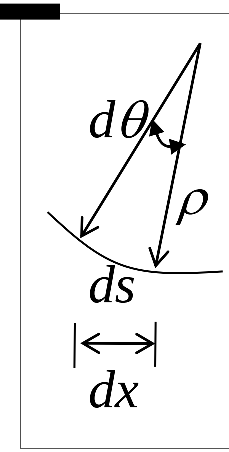
$$+M-x\cdot V=0$$

*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, Sec.4.3 p271

Sign Conventions for Curvature

$$\kappa \equiv \frac{1}{\rho} = \frac{d\theta}{ds} \cong \frac{d\theta}{dx}$$

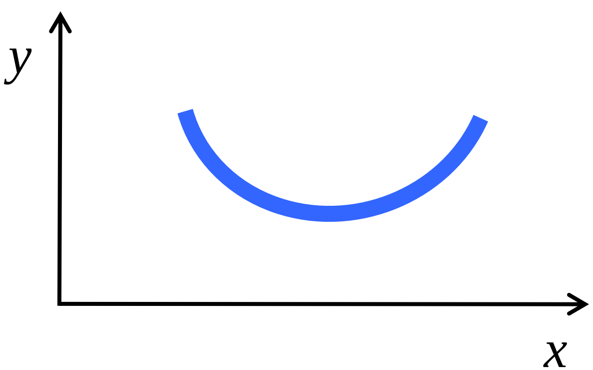
Static Sign Convention



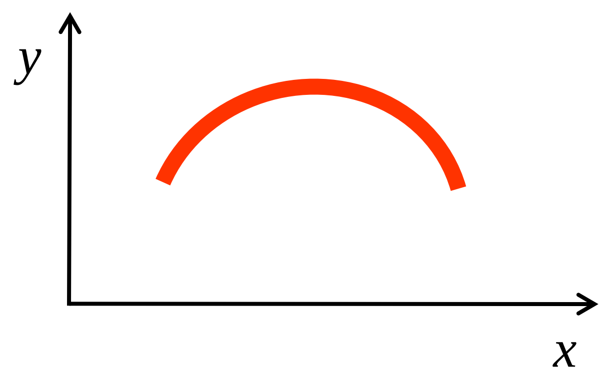
-The sign convention for curvature depends upon the orientation of <u>the coordinate axes*</u>

-Curvature is positive when the angle of rotation increase as moving along the beam in the positive *x*-direction

Positive Curvature



Negative Curvature



*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, Sec.5.3 p303

Deformation Sign Conventions*

Deformation sign conventions

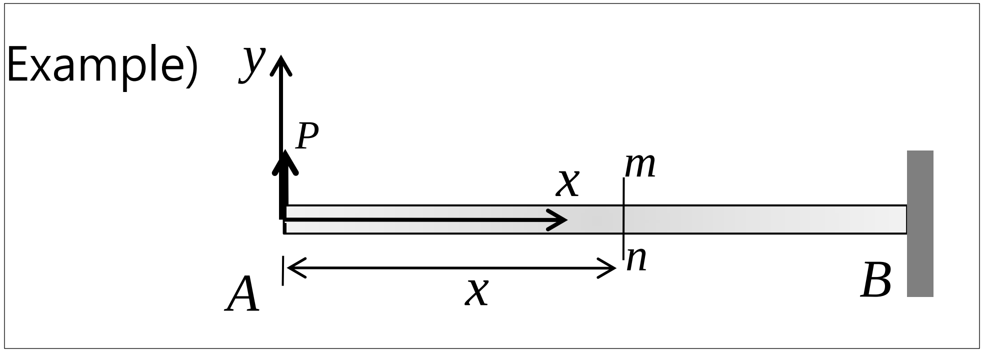
Sign conventions for stress resultants are called deformation sign conventions.

The algebraic sign of a stress resultant is determined by how it deforms the material on which it acts rather than by its direction in space

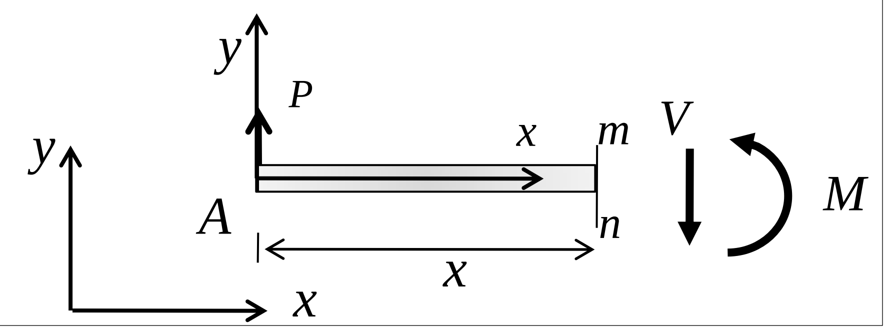
They are independent of the coordinates system

*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, pp270-271

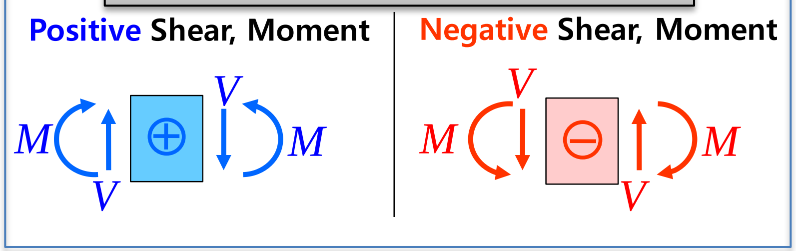
Static sign conventions and Deformation sign conventions



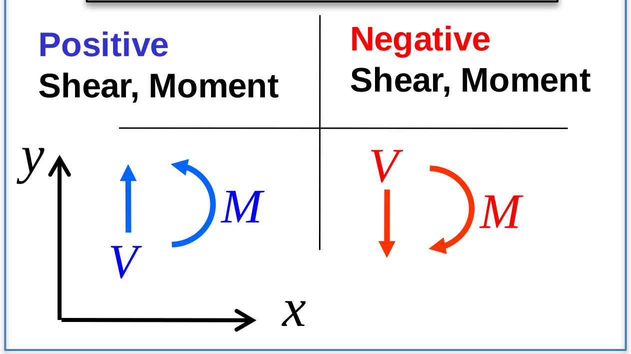
Free-body diagram of the left part



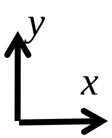
Deformation Sign Conventions



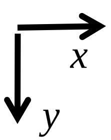
Static Sign Conventions



The shear force *V*, which is a positive shear force according to the deformation sign convention, is given a negative sign in the equation of equilibrium because it acts downward to the *y*-axis.

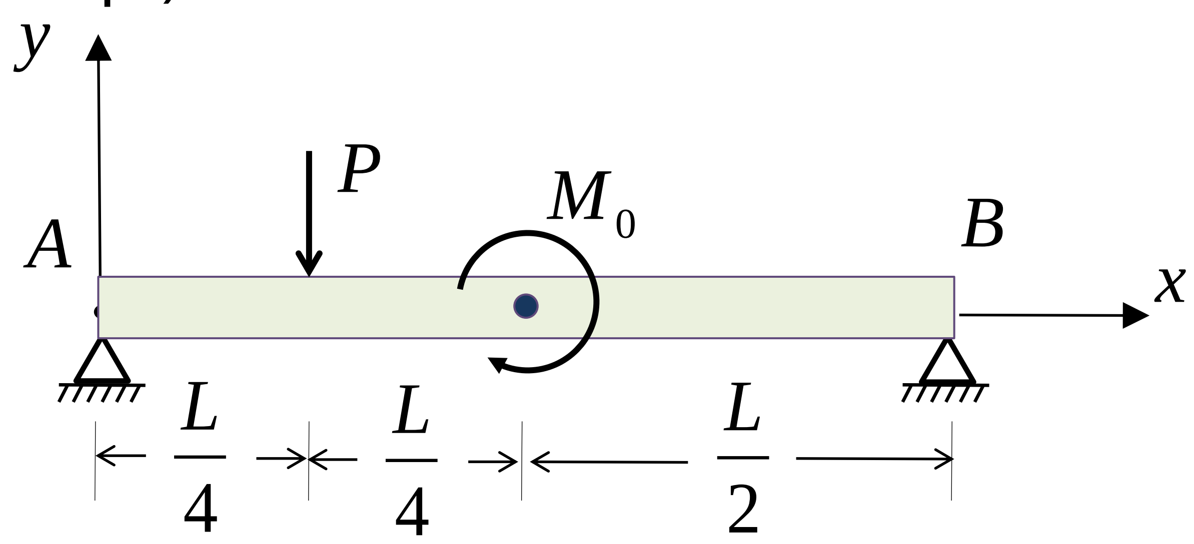


The shear force *V*, is given a positive sign in the equation of equilibrium if the *y*-axis is opposite



Example of Static Sign Convention for Reaction forces

Example)

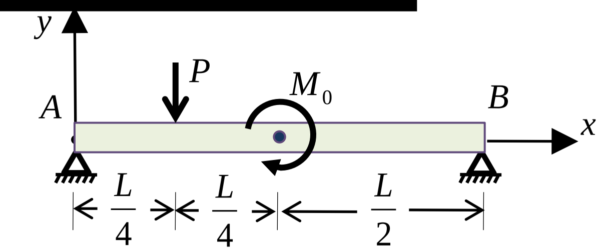


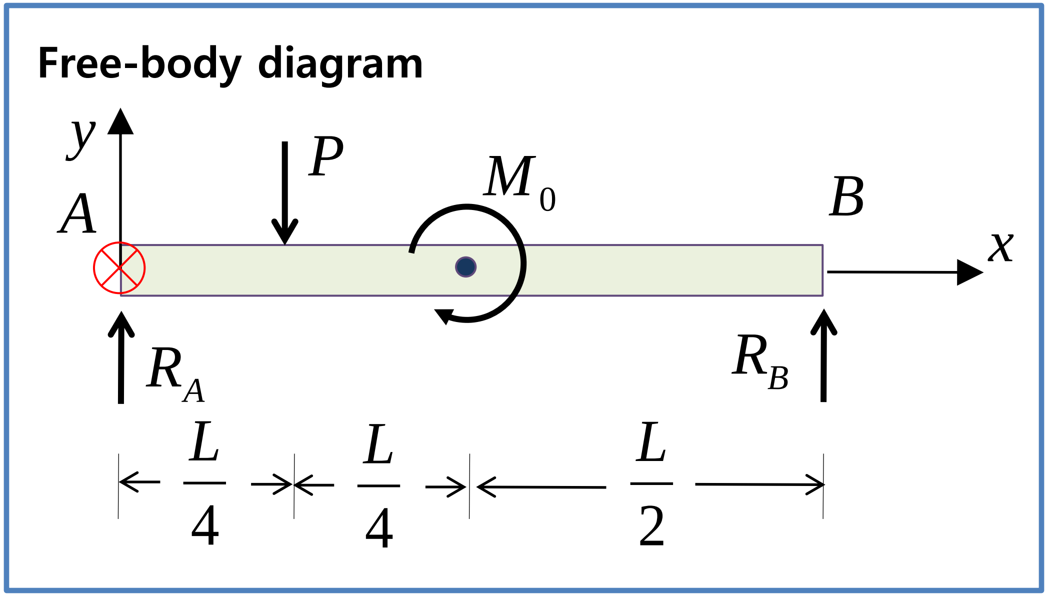
Given : force P and M_0 moment

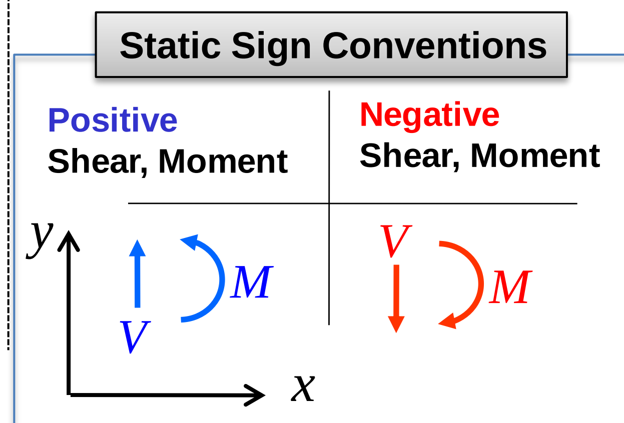
Find : reaction forces at point A and B

Example of Static Sign Convention for Reaction forces

Given : force P and M_0 moment
Find : the reaction forces at point A and B

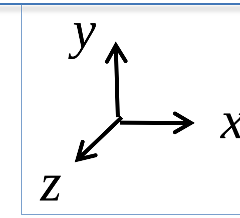






Moment Equilibrium about *z*-axis through the point *A*

$$M_{z\ at\,A} = -P\cdot\frac{L}{4} - M_0 + R_B\cdot L = 0 \quad\Rightarrow\quad \therefore R_B = \frac{P}{4} + \frac{M_0}{L}$$



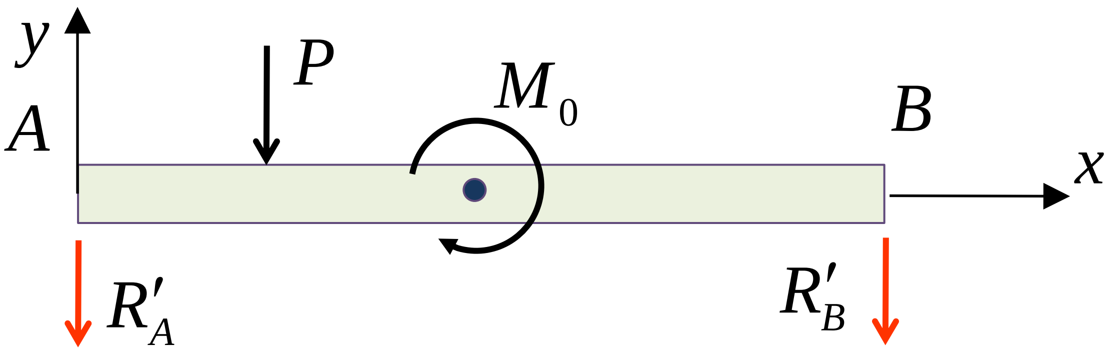
Moment Equilibrium about *z*-axis through the point *B*

$$M_{z\ at\,B} = +P\cdot\frac{3L}{4} - M_0 - R_A\cdot L = 0 \quad\Rightarrow\quad \therefore R_A = \frac{3P}{4} - \frac{M_0}{L}$$

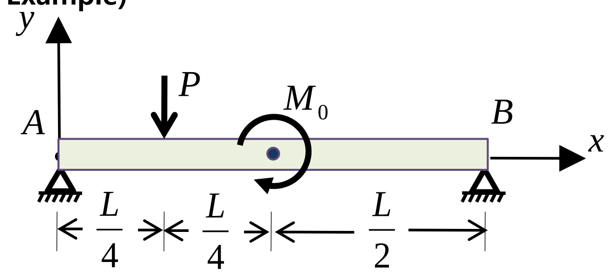
Example of Static Sign Convention for opposite directing reaction forces

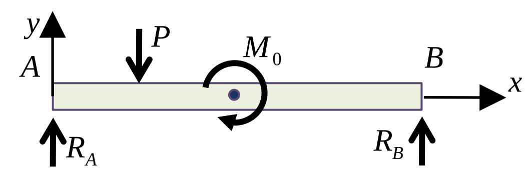
What will happen if the directions of the reaction forces are assumed to be opposite?

Free-body diagram



Example)

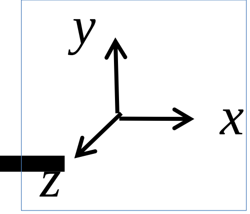




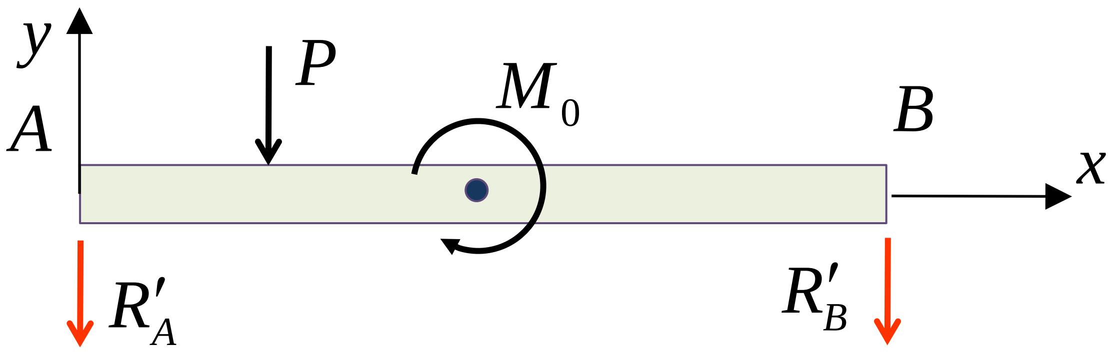
$$R_B = \frac{P}{4} + \frac{M_0}{L}$$

$$R_A = \frac{3P}{4} - \frac{M_0}{L}$$

Example of Static Sign Convention for opposite directing reaction forces



Free-body diagram



Moment Equilibrium about z-axis through the point *A*

$$M_{z\ at\ A} = -P\cdot\frac{L}{4} - M_0 \boxed{-} R'_B\cdot L = 0 \qquad \therefore R'_B = \boxed{-}\left(\frac{P}{4}+\frac{M_0}{L}\right)$$

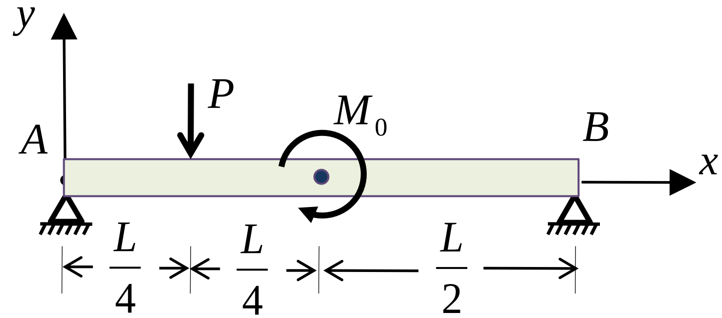
Moment Equilibrium about z-axis through the point *B*

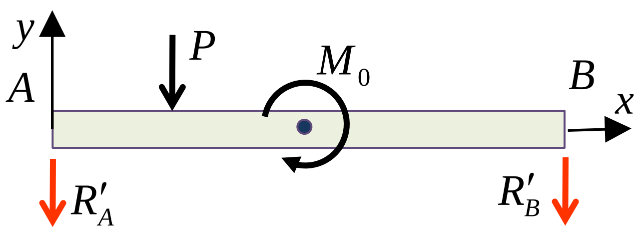
$$M_{z\ at\ B} = +P\cdot\frac{3L}{4} - M_0 \boxed{+} R'_A\cdot L = 0 \qquad \therefore R'_A = \boxed{-}\left(\frac{3P}{4}-\frac{M_0}{L}\right)$$

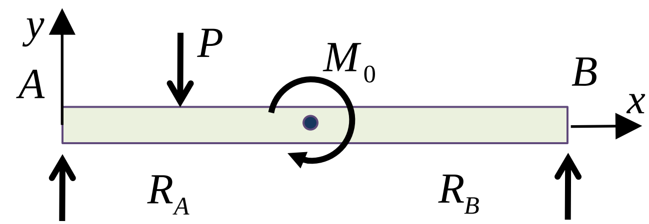
Example of Static Sign Convention for opposite directing reaction forces

What will happen if the direction are assumed to be opposite?

Example)







$$R'_B = \boxed{-}\left(\frac{P}{4} + \frac{M_0}{L}\right)$$

$$R'_A = \boxed{-}\left(\frac{3P}{4} - \frac{M_0}{L}\right)$$

$$R_B = \frac{P}{4} + \frac{M_0}{L}$$

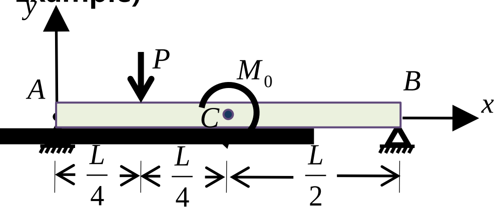
$$R_A = \frac{3P}{4} - \frac{M_0}{L}$$

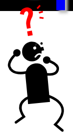
$$R'_B = -R_B$$

$$R'_A = -R_A$$

Example of Static Sign Convention for opposite directing reaction forces

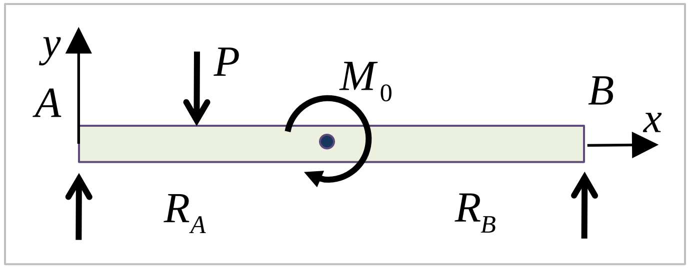
Example)





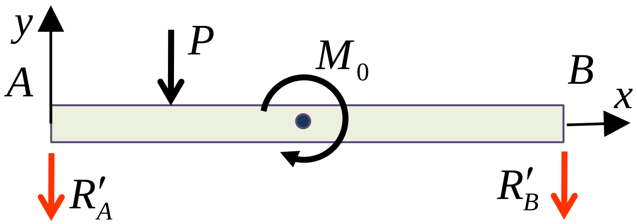
What will happen if the direction are assumed to be opposite?

① **for instance,** $R_A = 7[N],\ R_B = 3[N]$



$$R'_A = -R_A$$
$$R'_B = -R_B$$

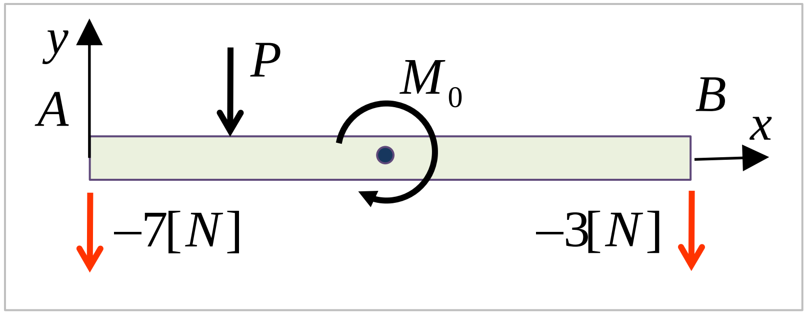
②



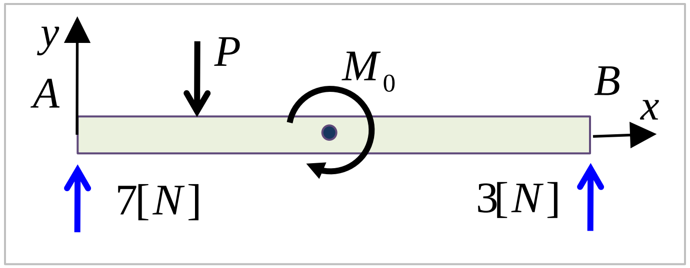
$$R'_A = -7[N],\quad R'_B = -3[N]$$

the same!!

③ **it means**

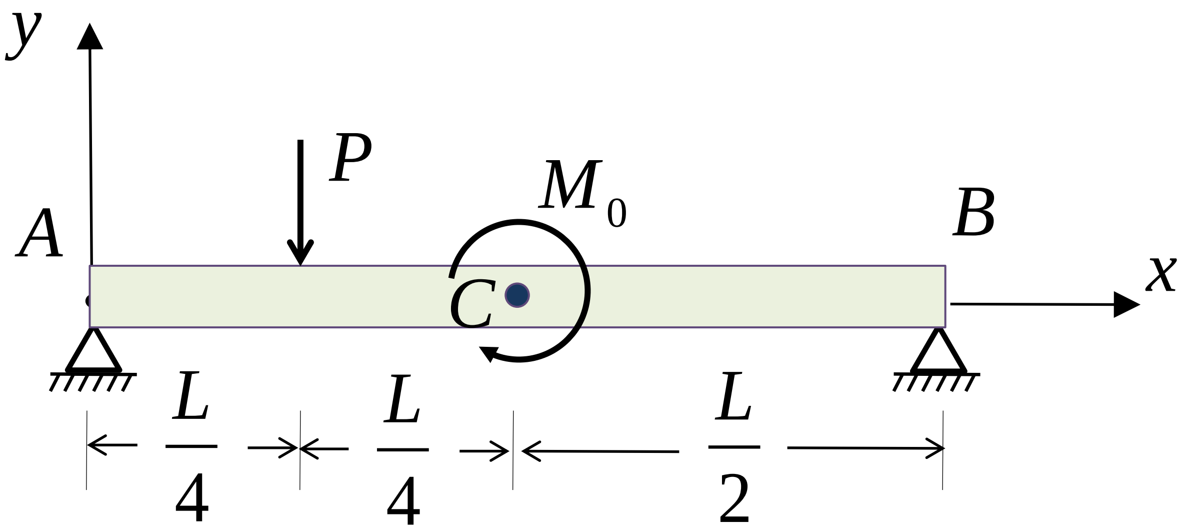


④ **it means**



Example of Static Sign Convention for Shear force and bending moment

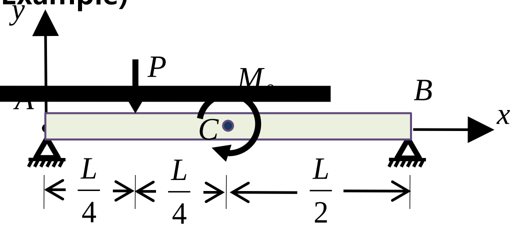
Example)



Given : force P and M_0 moment

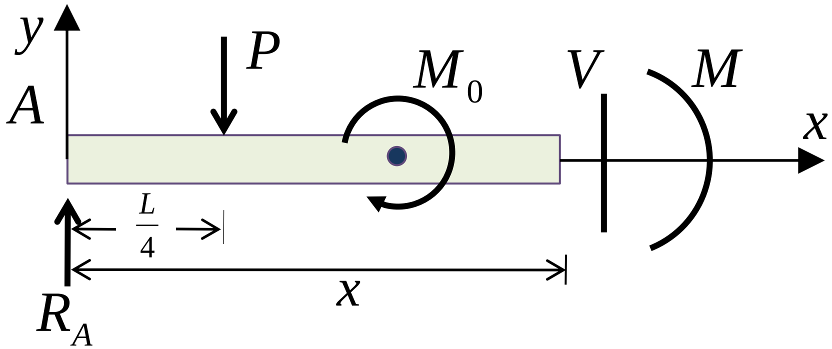
Find : shear force and bending moment at a point between C and B

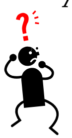
Example of Static Sign Convention for Shear force and bending moment



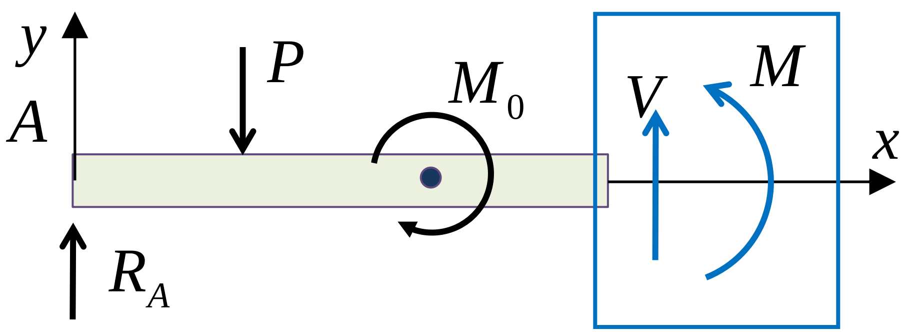
Given : force P and M_0 moment
Find : shear force and bending moment at a point between C and B

Free-body diagram



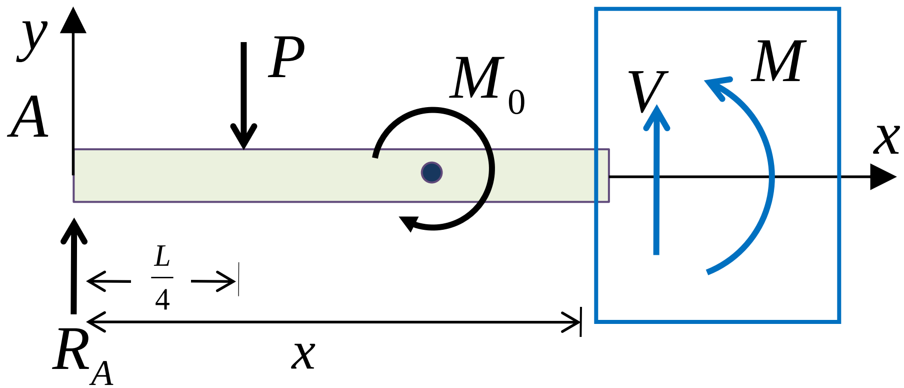

How can we assume the directions of the shear force and the bending moment?

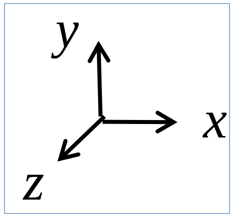
Let us assume that the material would deform in the direction of the figure below (case1)



Example of Static Sign Convention for Shear force and bending moment

$$R_A = \frac{3P}{4} - \frac{M_0}{L}$$





Force Equilibrium

$$F_y = + R_A - P + V = 0 \qquad \therefore V = P - R_A$$

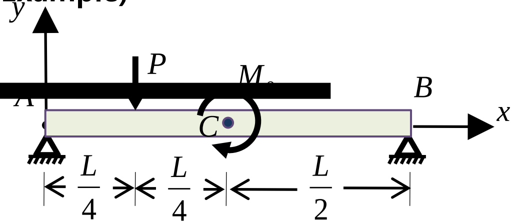
Moment Equilibrium about *z*-axis through the point *A*

$$M_{z\ at\,A} = - P \cdot \frac{L}{4} - M_0 + V \cdot x + M = 0$$

$$\therefore M = \frac{PL}{4} + M_0 - V \cdot x$$

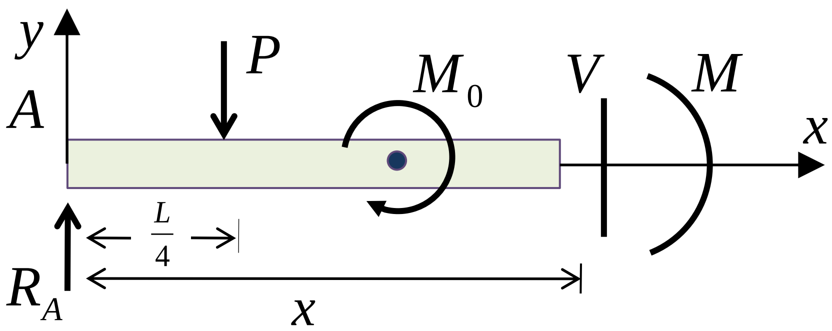
Example of Static Sign Convention for Opposite directing shear forces and bending moments

Example)



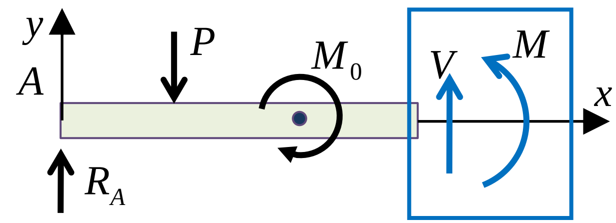
Given : force P and M_0 moment
Find : shear force and bending moment at a point between C and B

Free-body diagram



What will happen if we assume that the material would deform in the opposite direction of case1 ?

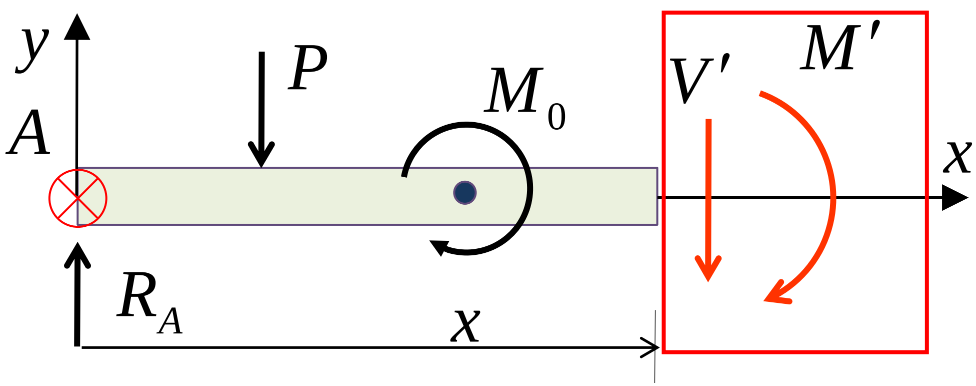
case1



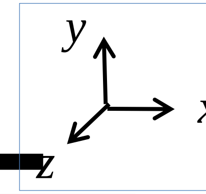
$$V = P - R_A$$

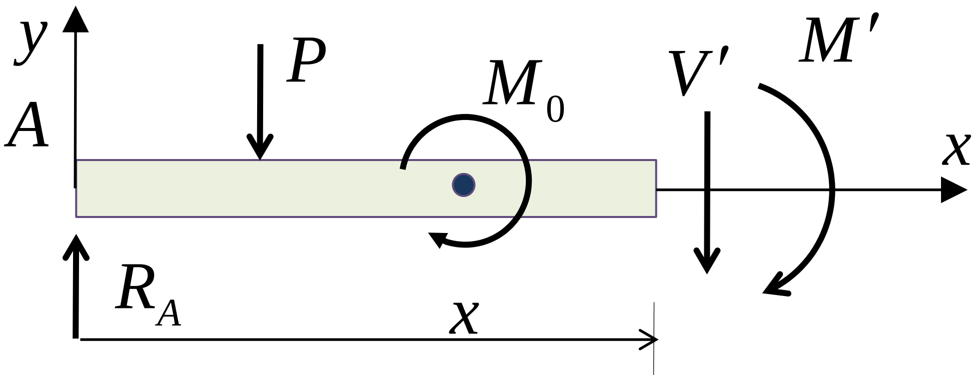
$$M = \frac{PL}{4} + M_0 - V \cdot x$$

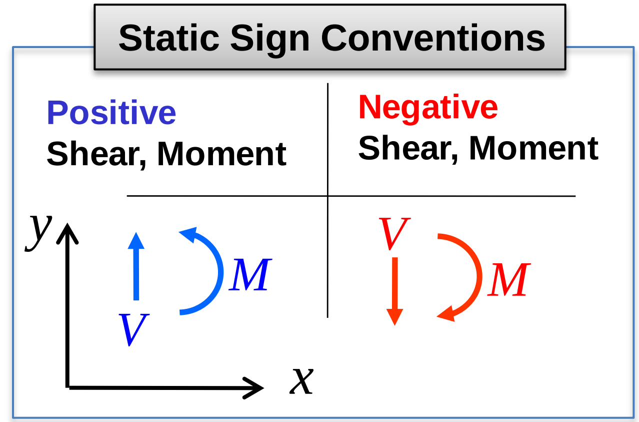
case2



Static Sign Convention for opposite directing shear forces and bending moments







Force Equilibrium

$$F_y = + R_A - P \boxed{-} V' = 0$$

$$\therefore V' = R_A - P$$

Moment Equilibrium about *z*-axis through the point A

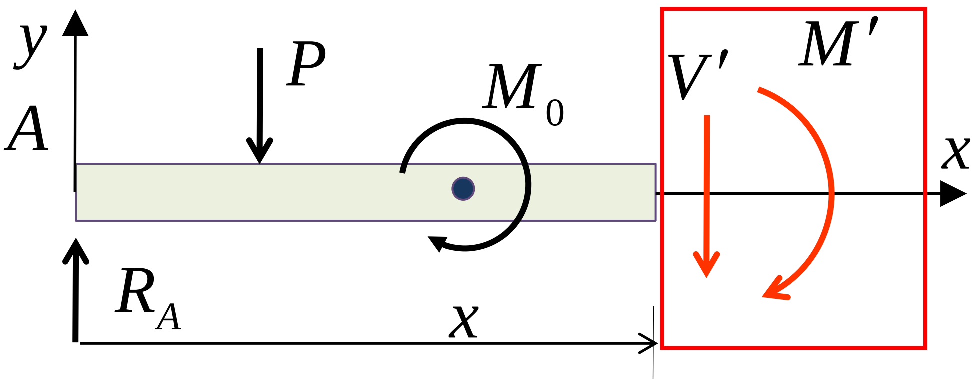
$$M_{z\ at\ A} = - P\cdot\frac{L}{4} - M_0 \boxed{-} V'\cdot x \boxed{-} M' = 0$$

$$\therefore M' = -\frac{PL}{4} - M_0 - V'\cdot x$$

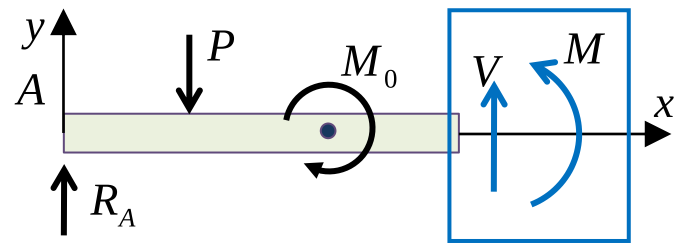
Static Sign Convention for opposite directing shear forces and bending moments

What will happen if we assume that the material would deform in the opposite direction of case1 ?

case2



case1



$$V' = R_A - P = -(P - R_A)$$

$$M' = -\frac{PL}{4} - M_0 - V' \cdot x$$

$$M' = -\frac{PL}{4} - M_0 + (P - R_A) \cdot x$$

$$= -\left(\frac{PL}{4} + M_0 + (R_A - P) \cdot x\right)$$

$$V = P - R_A$$

$$M = \frac{PL}{4} + M_0 - V \cdot x$$

$$M = \frac{PL}{4} + M_0 + (R_A - P) \cdot x$$

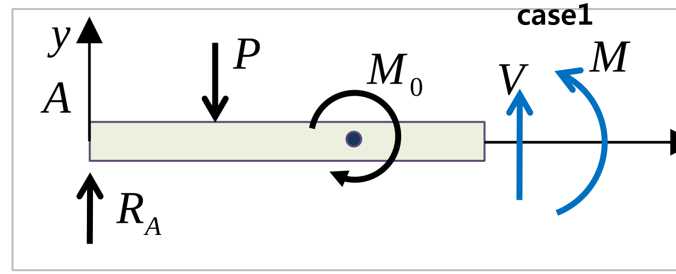
$$V' = -V$$

$$M' = -M$$

Static Sign Convention for opposite directing shear forces and bending moments

What will happen if we assume that the material would deform in the opposite direction of case1 ?

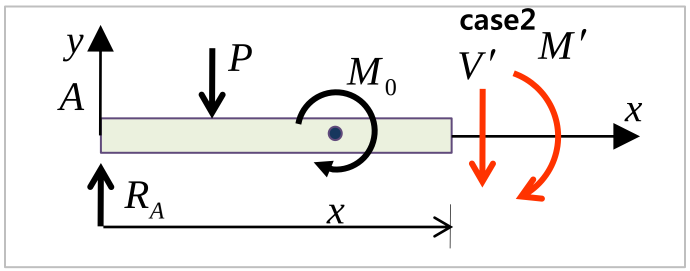
① for instance, $V = 3[N], M = 25.5[Nm]$



$$V' = -V$$

$$M' = -M$$

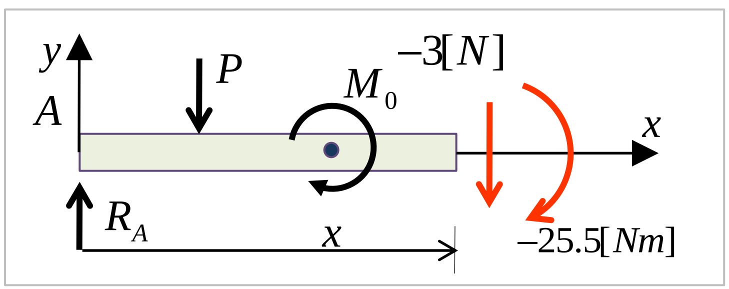
②



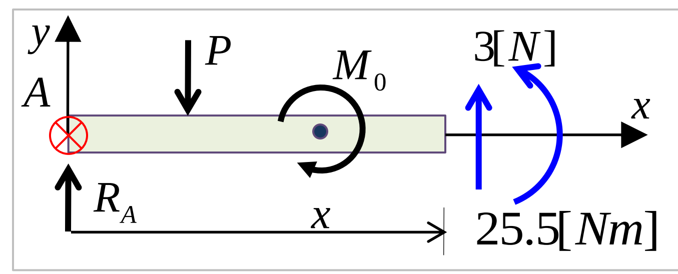
$$V' = -3[N],\ M' = -25.5[Nm]$$

the same!!

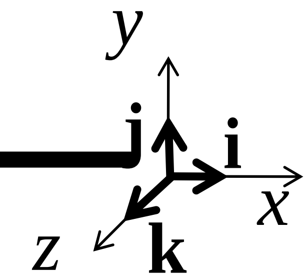
③ it means



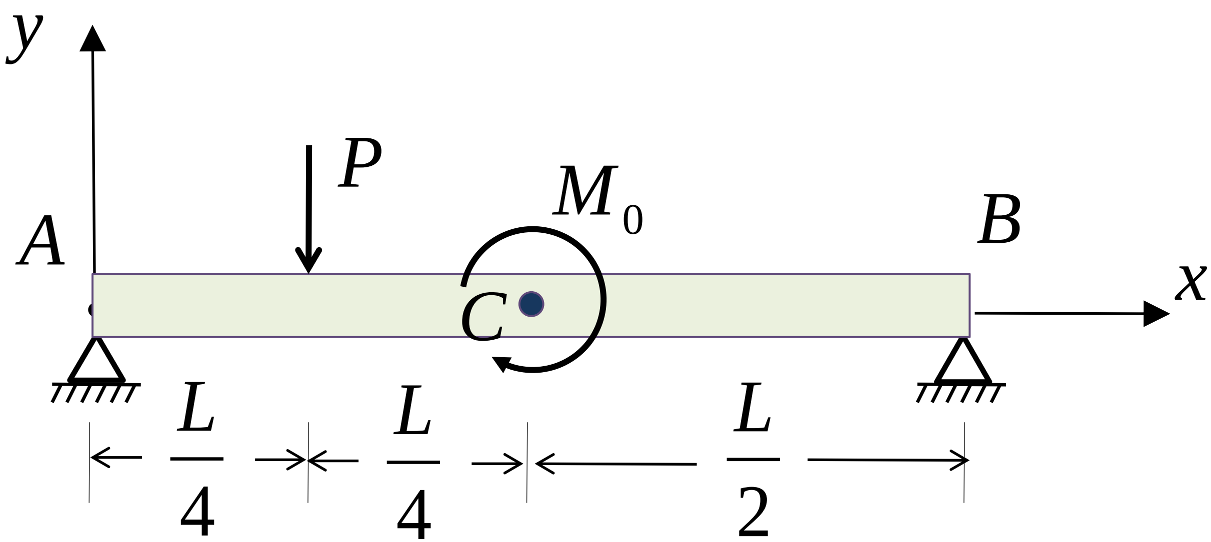
④ it means



Example of Static Sign Convention with Vector notation



Example)

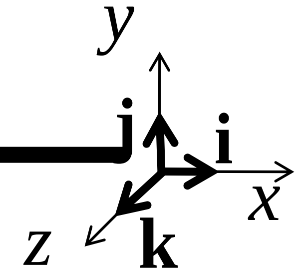


Let us use the vector notation for the example

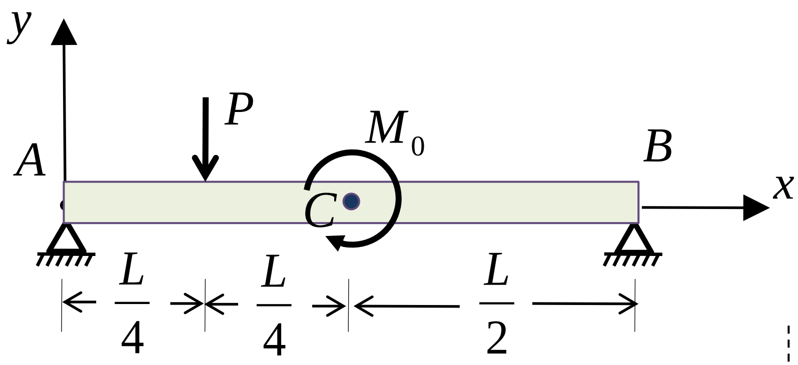
Given : force P and M_0 moment

Find : 1) reaction forces at point A and B
2) shear force and bending moment at a point between C and B

Find reaction forces at point A and B with Vector notation



Example)

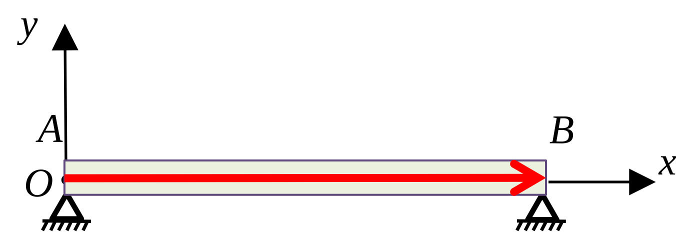


Given values in the vector notation

$$\mathbf{P} = -10\mathbf{j},\ P = -10$$

$$\mathbf{M_0} = -15\mathbf{k},\ M_0 = -15$$

$$\mathbf{L} = 30\mathbf{i}, L = 30$$

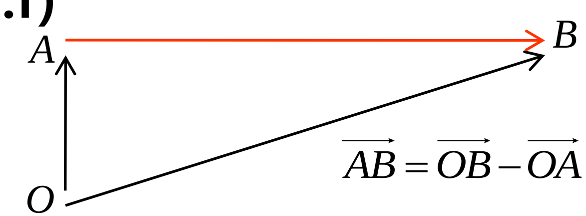


$$\mathbf{L} = \overrightarrow{AB}$$

$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = 30\mathbf{i} - 0\mathbf{i} = 30\mathbf{i}$$

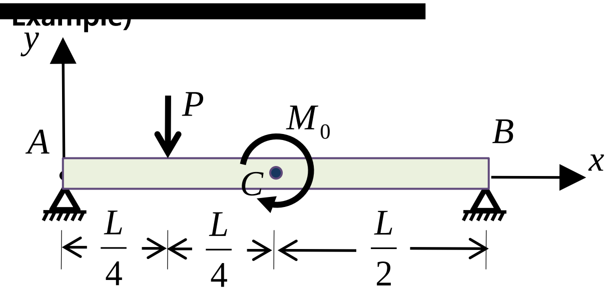
$\overrightarrow{OB}$, $\overrightarrow{OA}$: position vector

c.f)



$$\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$$

Find reaction forces at point A and B with Vector notation



Moment Equilibrium about *z*-axis through the point *A*

$$\mathbf{M}_{z\ at\,A} = \frac{\mathbf{L}}{4}\times\mathbf{P}+\mathbf{M_0}+\mathbf{L}\times\mathbf{R}_B$$

$$= \frac{L}{4}\mathbf{i}\times P\mathbf{j}+M_{\mathbf{0}}\mathbf{k}+L\mathbf{i}\times R_B\mathbf{j}$$

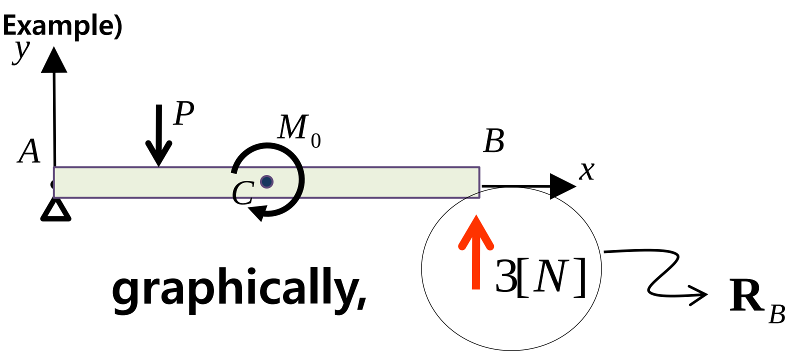
$$\therefore \frac{L}{4}\cdot P\mathbf{k}+M_{\mathbf{0}}\mathbf{k}+L\cdot R_B\mathbf{k}=0 \;\Rightarrow\; \left(\frac{L}{4}\cdot P+M_{\mathbf{0}}+L\cdot R_B\right)\mathbf{k}=0$$

for instance, $\left(\frac{30}{4}\cdot(-10)+(-15)+30\cdot R_B\right)\mathbf{k}=0 \;\Rightarrow\; 30\cdot R_B=\frac{30}{4}\cdot 10+15$

$$R_B=\frac{10}{4}+\frac{15}{30}=\frac{5}{2}+\frac{1}{2}=3[N]$$

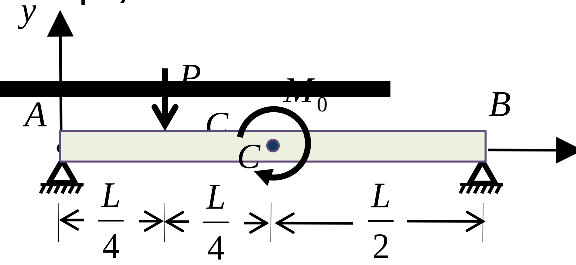
it means, $\mathbf{R}_B=R_B\mathbf{j}=3\mathbf{j}$

graphically,



Find reaction forces at point A and B with Vector notation

Example)



Moment Equilibrium about *z*-axis through the point *B*

$$\mathbf{M}_{z\ \ at\,B} = \frac{3\mathbf{L}}{4} \times \mathbf{P} + \mathbf{M_0} + \mathbf{L} \times \mathbf{R}_A$$

$$= \frac{3L}{4}\mathbf{i} \times P\mathbf{j} + M_0\mathbf{k} + L\mathbf{i} \times R_A\mathbf{j}$$

$$\therefore \frac{3L}{4} \cdot P\mathbf{k} + M_0\mathbf{k} + L \cdot R_A\mathbf{k} = 0 \Rightarrow \left(\frac{3L}{4} \cdot P + M_0 + L \cdot R_A\right)\mathbf{k} = 0$$

for instance,

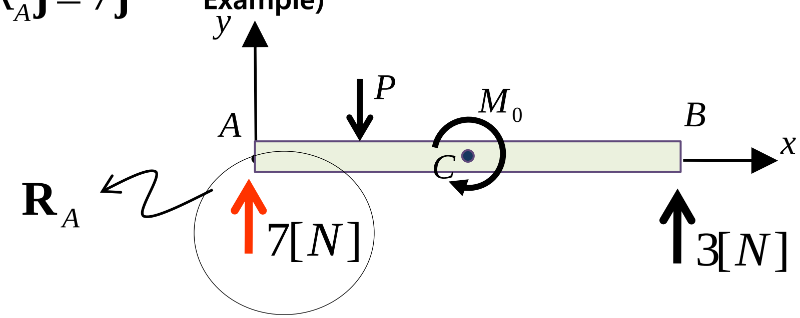
$$\left(\frac{3 \cdot (-30)}{4} \cdot (-10) + (-15) + (-30) \cdot R_A\right)\mathbf{k} = 0$$

$$\Rightarrow 30 \cdot R_A = \frac{3 \cdot (-30)}{4} \cdot (-10) + (-15) \Rightarrow \therefore R_A = 7[N]$$

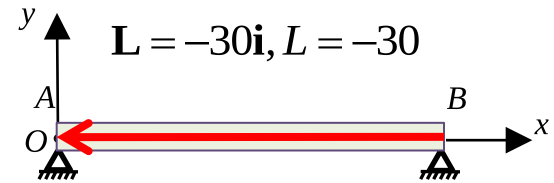
it means, $\mathbf{R}_A = R_A\mathbf{j} = 7\mathbf{j}$

graphically,

Example)



Caution,

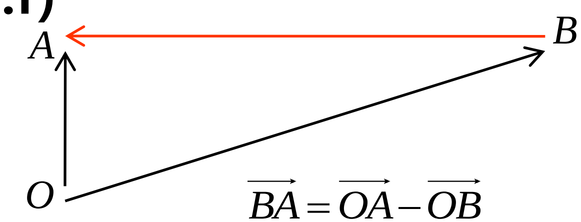


$\mathbf{L} = \overrightarrow{BA}$

$\overrightarrow{BA} = \overrightarrow{OA} - \overrightarrow{OB} = 0\mathbf{i} - 30\mathbf{i} = -30\mathbf{i}$

$\overrightarrow{OA}, \overrightarrow{OB}$: position vector

c.f)



$\frac{3}{4}\mathbf{L} = \frac{3}{4}\overrightarrow{BA}$

$\overrightarrow{BA} = \overrightarrow{OA} - \overrightarrow{OB} = 0\mathbf{i} - 30\mathbf{i} = -30\mathbf{i}$

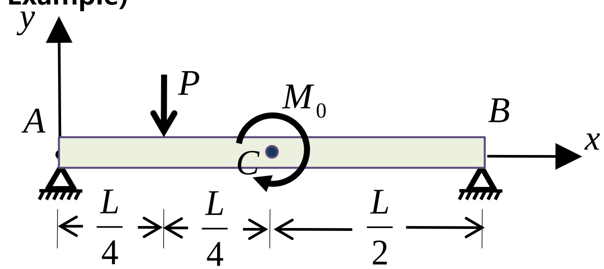
$\frac{3}{4}\overrightarrow{BA} = \frac{3 \cdot (-30)}{4}$

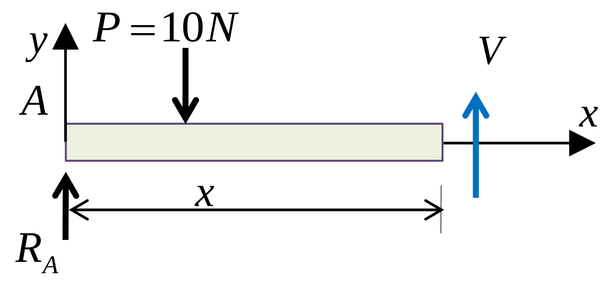
Find shear force at x with Vector notation

$\mathbf{P} = -10\mathbf{j},\ P = -10$

$\mathbf{R}_A = R_A\mathbf{j} = 7\mathbf{j}$

Example)





Force Equilibrium

$$\mathbf{F}_y = \mathbf{R}_A + \mathbf{P} + \mathbf{V} = R_A\mathbf{j} + P\mathbf{j} + V\mathbf{j} = 0$$

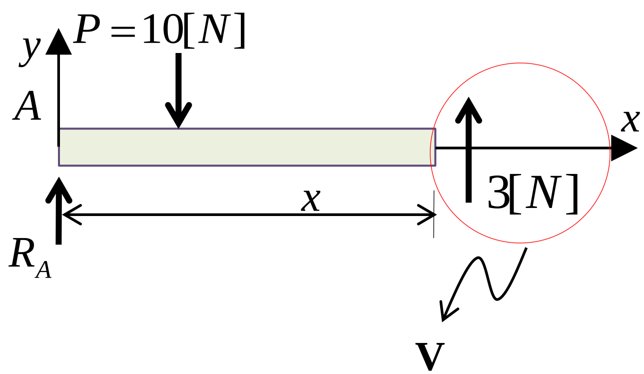
$$(R_A + P + V)\mathbf{j} = 0$$

for instance, $(7 + (-10) + V)\mathbf{j} = 0$

$$V = 3[N]$$

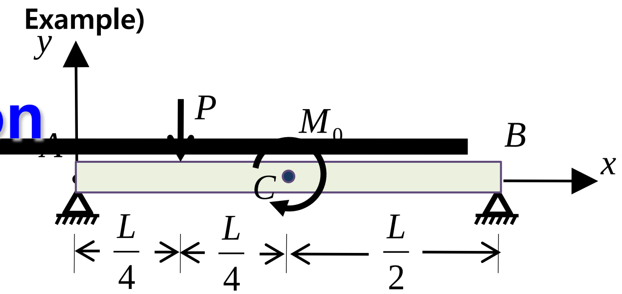
it means, $\mathbf{V} = V\mathbf{j} = 3\mathbf{j}$

graphically,

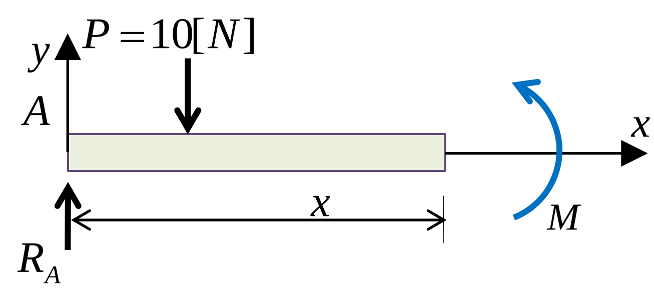


Find **bending moment** about *z*-axis through the point *x* with **Vector notation**

Example)



$\mathbf{R}_A = R_A\mathbf{j} = 7\mathbf{j}$



Moment Equilibrium about z-axis through the point A

$$\mathbf{M}_{z\ at\ A} = \frac{\mathbf{L}}{4}\times\mathbf{P}+\mathbf{M}_0+\mathbf{x}\times\mathbf{V}+\mathbf{M}$$

$$= \frac{L}{4}\mathbf{i}\times P\mathbf{j}+M_0\mathbf{k}+x\mathbf{i}\times V\mathbf{j}+\mathbf{M}$$

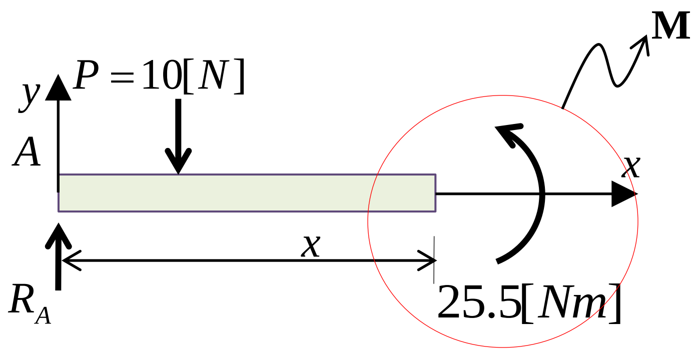
$$\therefore \frac{L}{4}\cdot P\mathbf{k}+M_0\mathbf{k}+x\cdot V\mathbf{k}+M\mathbf{k}=0$$

$$\Rightarrow \left(\frac{L}{4}\cdot P+M_0+x\cdot V+M\right)\mathbf{k}=0$$

for instance, at $x=\frac{3}{4}L, R_x = 3, \left(\frac{30}{4}\cdot(-10)+(-15)+\frac{3}{4}(30)\cdot(3)+M\right)\mathbf{k}=0 \Rightarrow M = 25.5[Nm]$

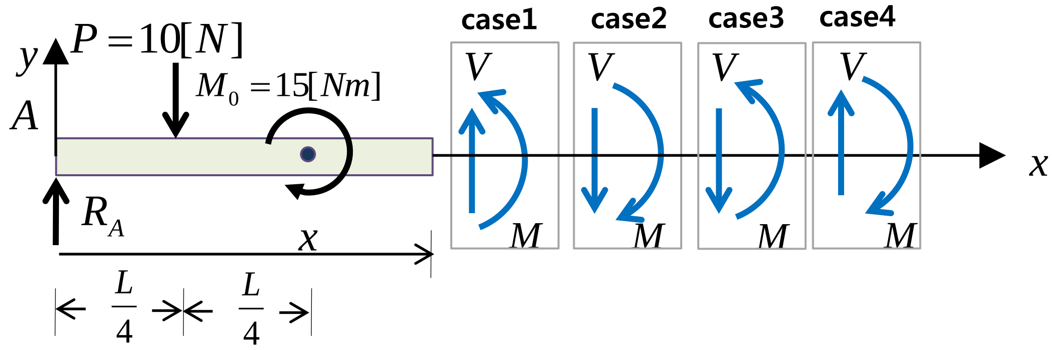
it means, $\mathbf{M} = M\mathbf{k} = 25.5\mathbf{k}$

graphically,



Static Sign Convention for Shear force and bending moment

There were no differences between the case1 and the case2 for the solution of the problem.

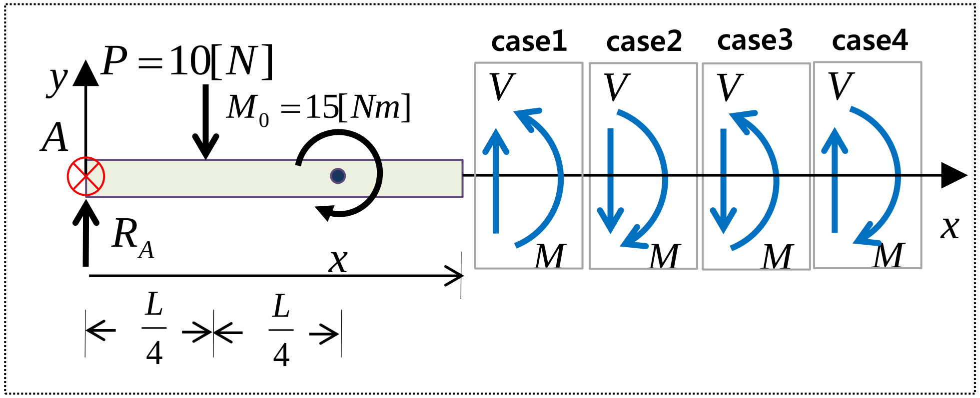


Do you think we will have (the) same solutions if we assume the directions of the shear forces and the bending moments as the case3 and case4?

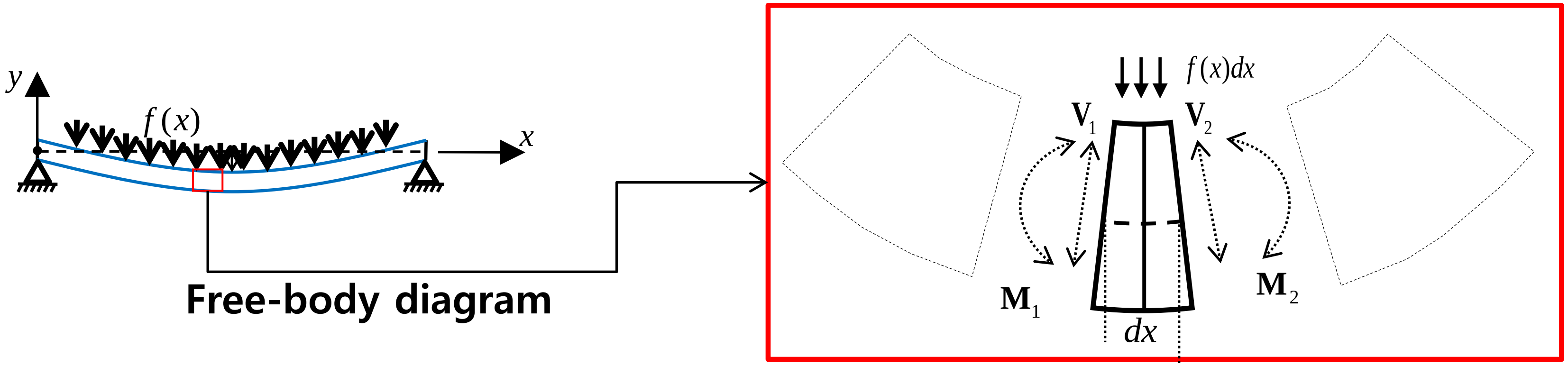
Exercise: Do it yourself and explain why.

Static sign conventions and Deformation sign conventions

We have (the) same solutions for all the cases. It means we can assume arbitrary directions for the shear forces and the bending moments to solve this problem.



Can we assume any arbitrary directions for the shear forces and the bending moments for the problem below?



What is the difference between the two problems?

Deformation Sign Conventions*

Deformation sign conventions

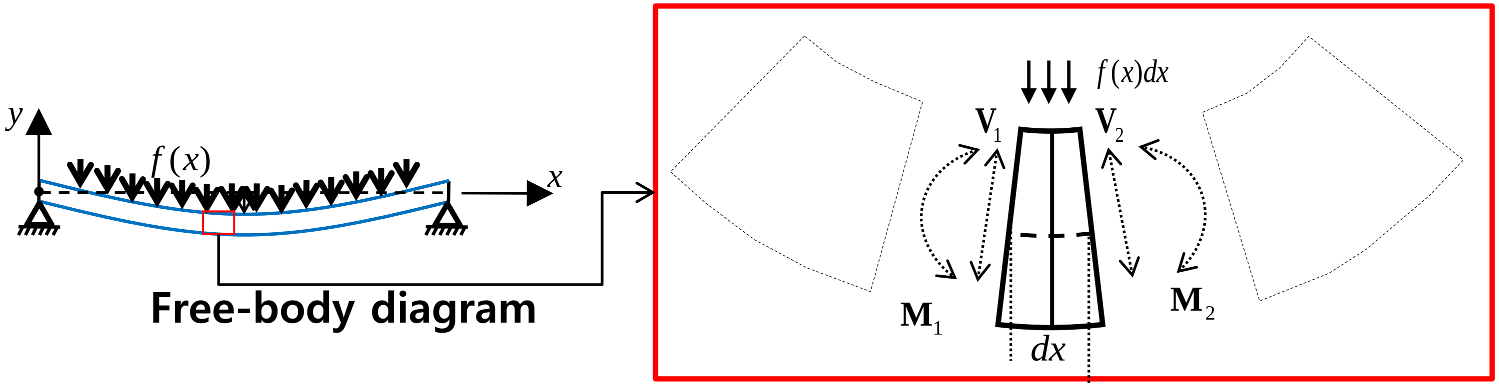
The algebraic sign of a stress resultant
is determined by <u>how it deforms</u> the material
on which it acts rather than
by its direction in space.

They are independent of the coordinates system

*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, pp270-271

Directions of shear forces and bending moments for the free body diagram of a beam element

Can we assume any arbitrary directions for the shear forces and the bending moments for the free body diagram of a beam element as below?



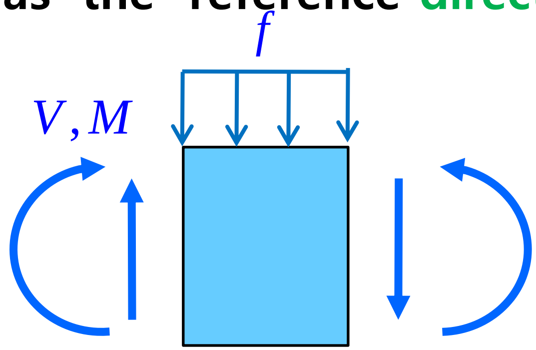
the directions of two unknowns should be defined → 'a reference' is required

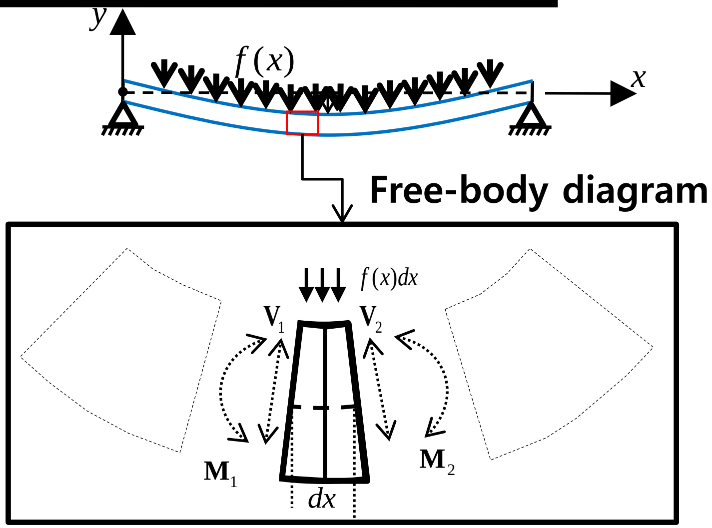
'a reference' is required in which the directions of shear forces and bending moments can be explained the bending in accordance with the physical phenomenon (or to make the equation 'solvable')

Directions of shear forces and bending moments for the free body diagram of a beam element

For this bending,

Let us introduce the following directions as 'the reference directions'



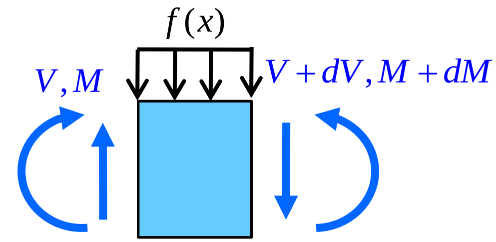


The deformation sign convention is adopted for 'the reference.'

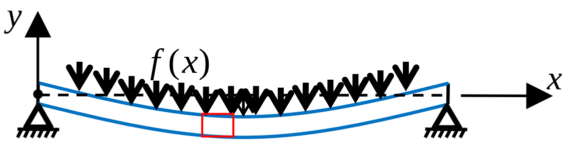
Why these directions of the shear forces and bending moments are able to explain the bending in accordance with the physical phenomenon? (in other words, are the directions reasonable for describing the bending?)

Deformation sign conventions* for distributed load

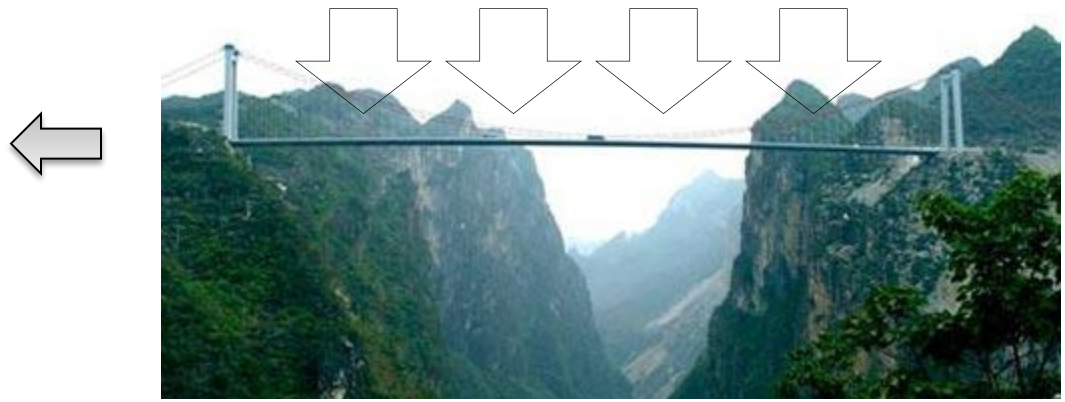
Are the directions reasonable?



1) Distributed Load $f(x)$



Gravitational force



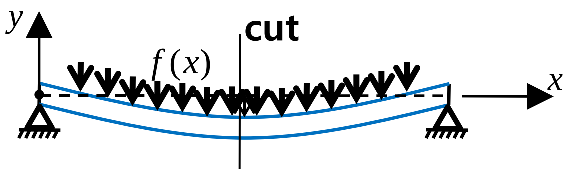
Since, all the structures on earth are subject to the downward gravitational force, it is very natural to consider the direction of the distributed load as vertically downward.

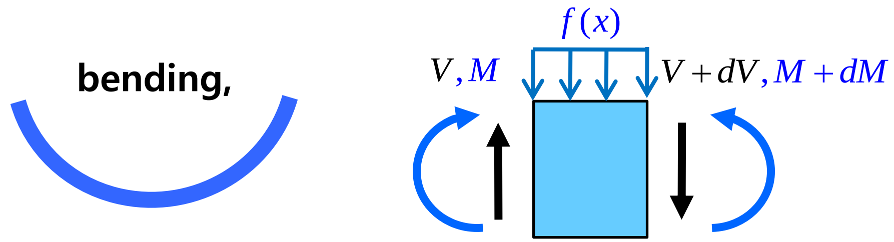
* Sign conventions for stress resultants are called deformation sign conventions
because they are based upon how the material is deformed.
The sign of deformation convention depends upon how it deforms the material, not upon its direction in space.
By contrast, when writing equations of equilibrium we use static sign conventions,
in which forces are positive or negative according to their directions along the coordinates axes.

Deformation sign conventions for shear forces

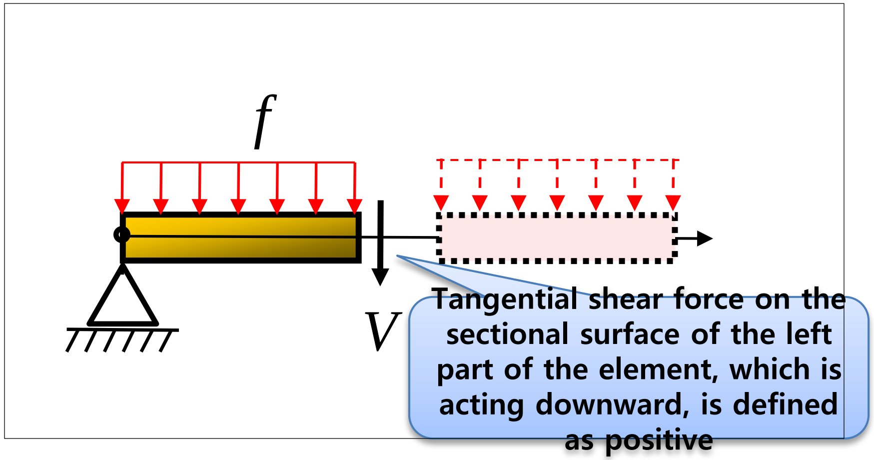
Are the directions reasonable?

2) Shear force V : direction

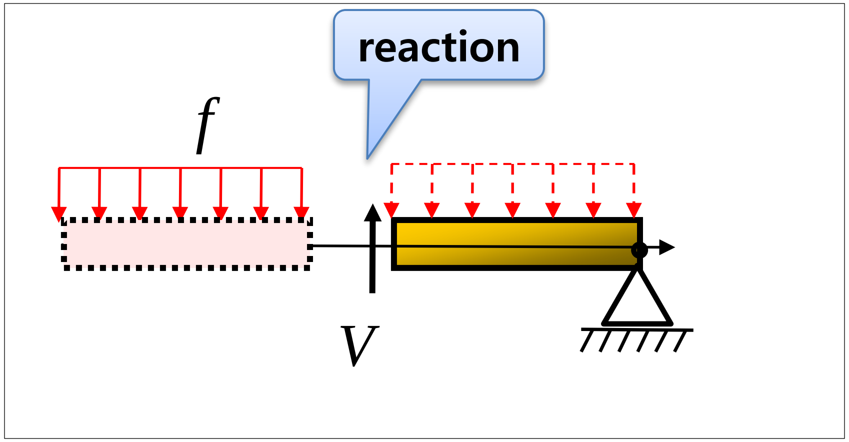




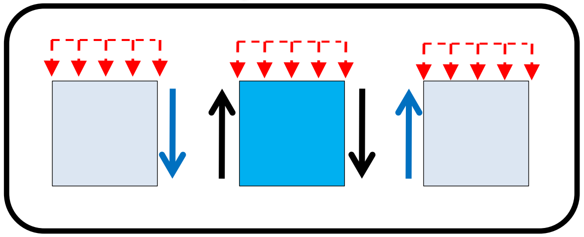
Free-body diagram of the left part



Free-body diagram of the right part

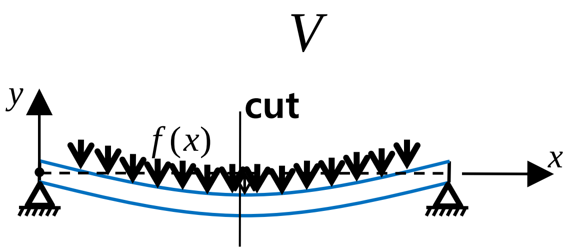


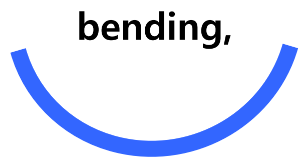
in the same manner,

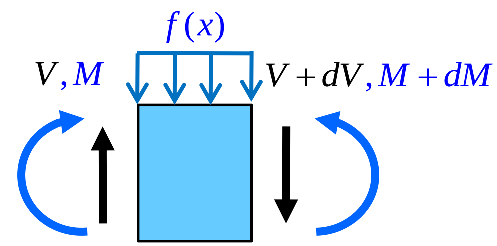


If we consider the free-body diagram in accordance with the Newton's third law, the directions of the shear forces are reasonable.

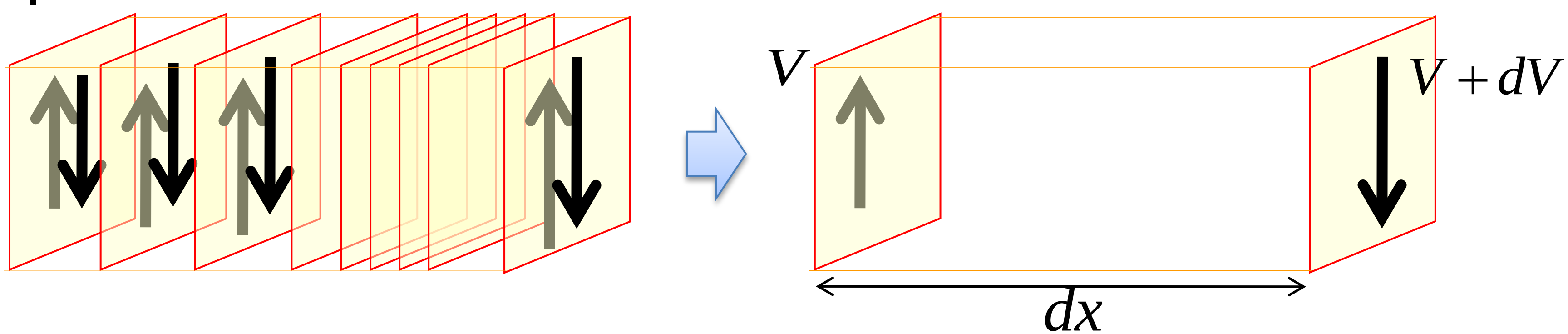
Magnitudes of shear forces







While the magnitude of the shear forces are same for the surface at a point in accordance with the Newton's third law, it is, however, reasonable to assume the shear forces are different at different positions.



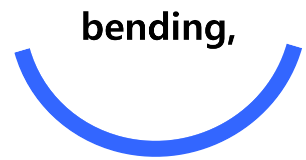
It shows that the length of element, dx, converges to zero.

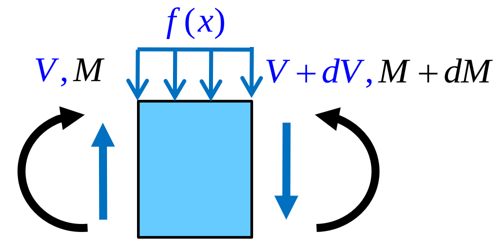
It shows that the internal shear forces are cancelled out.

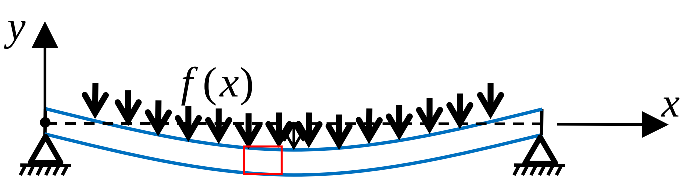
Deformation sign conventions for bending moments

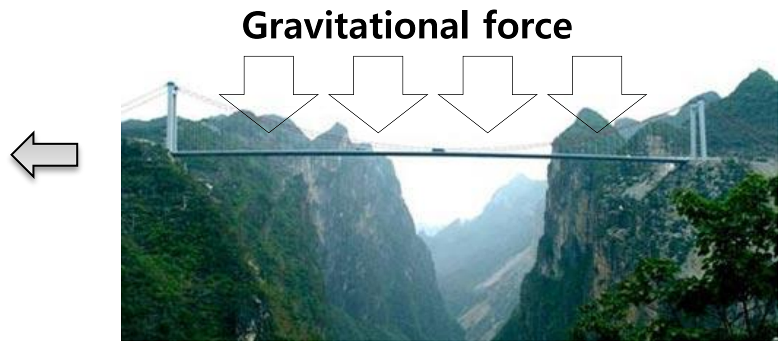
Are the direction reasonable?

3) Bending moment M : direction









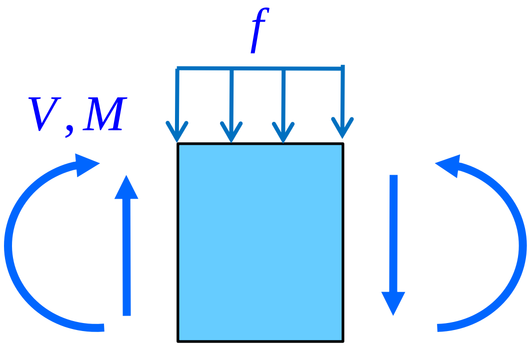
If we consider the deflected geometry of a structure subjected to the gravitational force, the direction of bending moments are reasonable

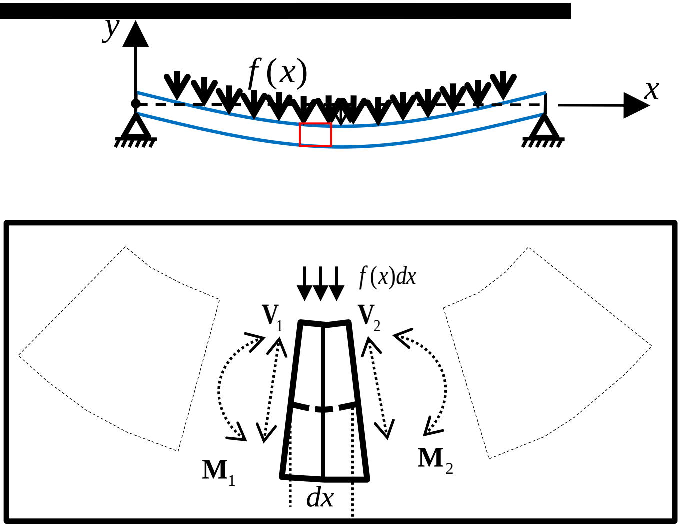
Bending moment M : magnitude

By analogy with the magnitudes of the shear forces, it is reasonable to assume the bending moments are different at different positions.

Deformation sign conventions for bending moments

For this bending,





- The directions are reasonable for describing the bending due to the distributed load.
- We define these directions as 'positive' for the 'positive' bending

Deformation Sign Conventions

Deformation Sign Conventions for Normal Stress*

When a sign convention for normal stresses is required, it is customary to define tensile stress as positive and compressive stresses as negative

positive

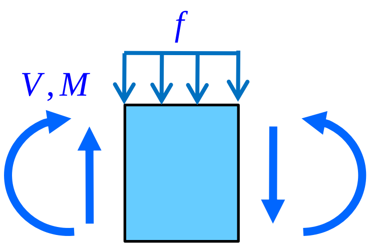
negative

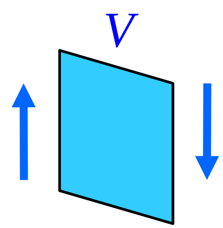
*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, Sec.1.2 p4

Summary : Deformation Sign Conventions for Shear Forces, Bending Moments and Distributed Loads*

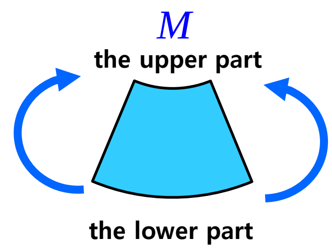
-The algebraic sign of a stress resultant is determined by how it deforms the material on which it acts rather than by its direction in space

Positive
-Shear,
-Bending Moment,
-Intensity of distributed load

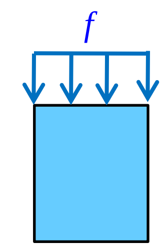




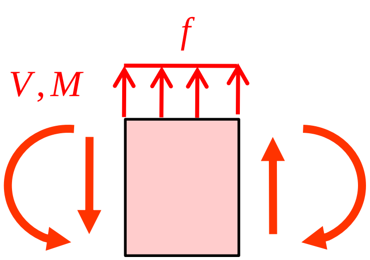
a positive shear force acts clockwise against the material

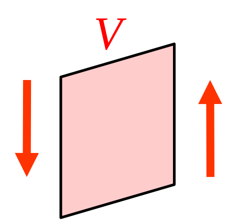


a positive bending compresses the upper part of the beam

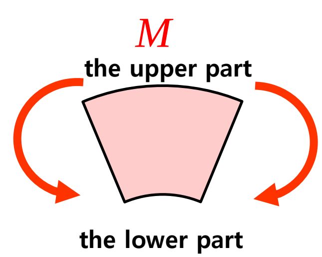


Negative
-Shear,
-Bending Moment,
-Intensity of distributed load

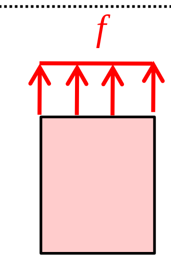




a negative shear force acts counterclockwise against the material



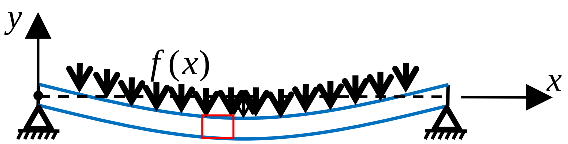
a negative bending compresses the lower part of the beam

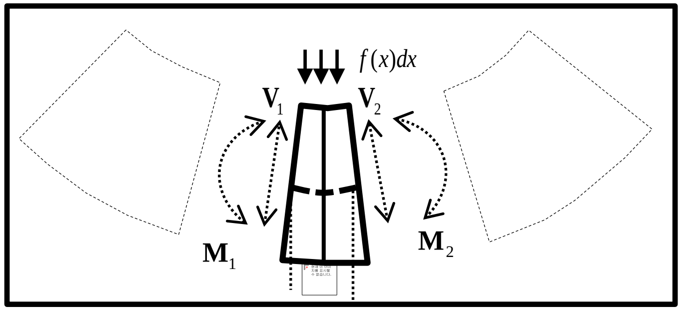


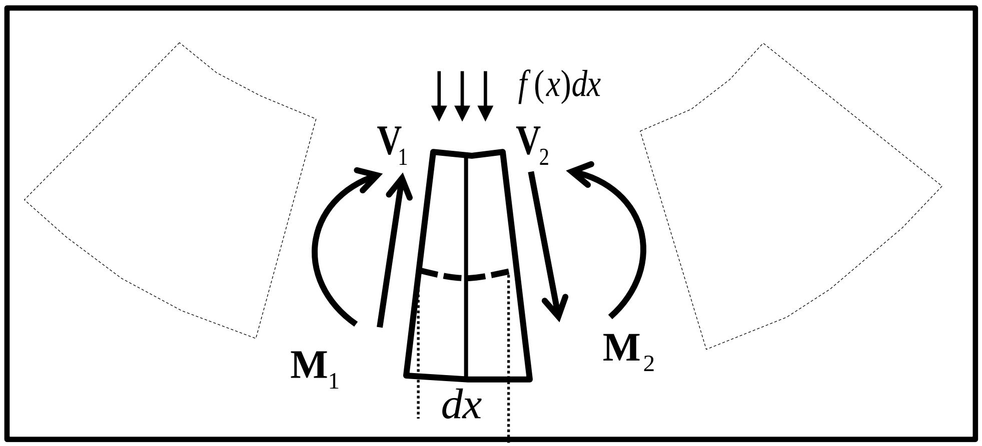
*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, Sec.4.3 p271
*Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006, Sec.9.2 p598 (figure9-4)

Deformation sign conventions for distributed load, shear forces, and bending moments

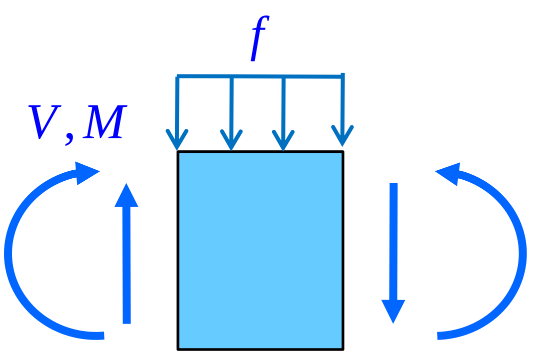
In accordance with the deformation sign conventions, we assume directions of the distributed load, the shear forces, and the bending moments are positive to obtain the relations of them.







'Positive'



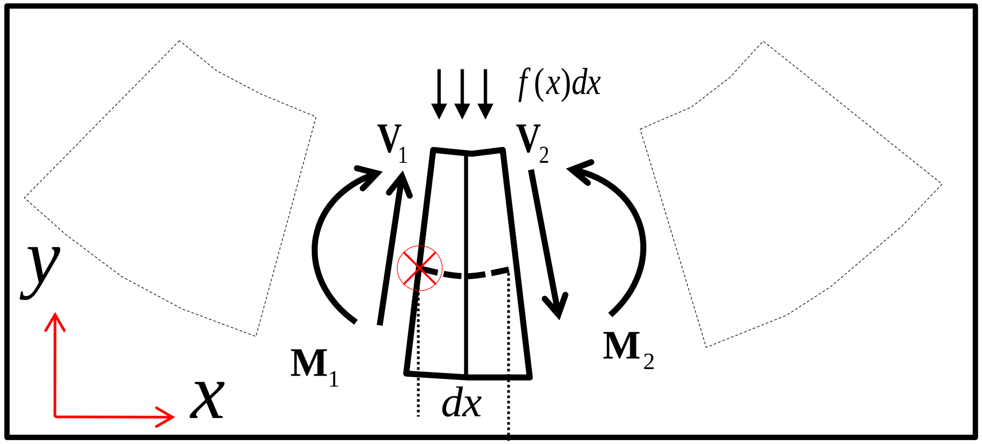
Relations of Distributed Load, Shear Forces, and Bending Moments

We obtain the relations of the distributed load, shear forces, and the bending moments by using the equations of equilibrium.

Force equilibrium

$$V-(V+dV)-f(x)dx=0$$

$$\boxed{\therefore \frac{dV}{dx}=-f(x)}$$



Moment equilibrium about *z*-axis through the point ⊗

$$-M+(M+dM)-(V+dV)dx-\frac{1}{2}dx\cdot f(x)dx=0$$

$$dM-Vdx-\cancel{dV\cdot dx}-\frac{1}{2}\cancel{(dx)^2}\cdot f(x)=0$$

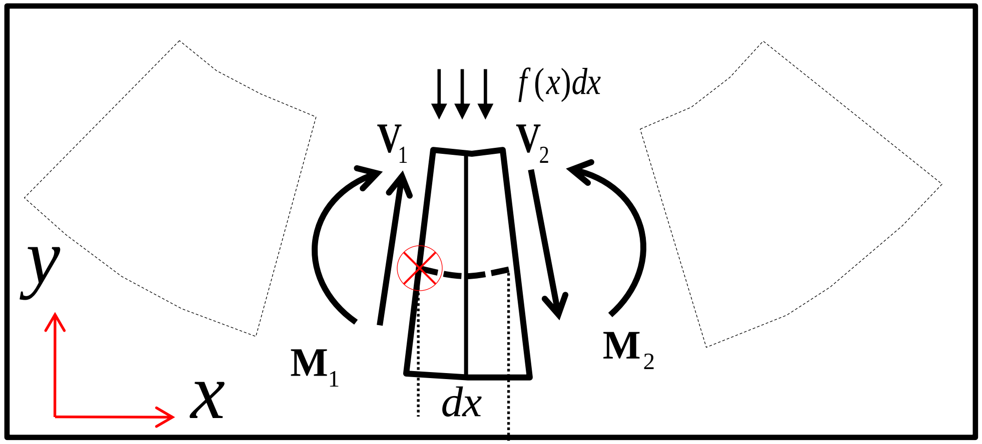
neglecting the second order or high terms

$$\boxed{\therefore \frac{dM}{dx}=V(x)}$$

Relations of Distributed Force, Shear Forces, and Bending Moments with Vector notation

$\mathbf{f}(x)$: given in vector

e.g., $\mathbf{f}(x) = -f(x)\mathbf{j}$



$\mathbf{V}_1, \mathbf{V}_2, \mathbf{M}_1, \mathbf{M}_2$: *unknown*

$$\mathbf{V}_1 = V_1\mathbf{j}, \quad \mathbf{V}_2 = V_2\mathbf{j},$$

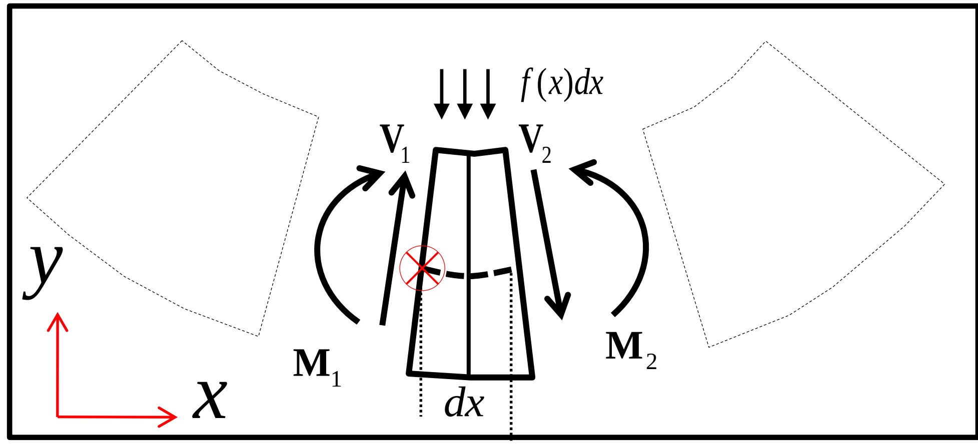
$$\mathbf{M}_1 = M_1\mathbf{k}, \quad \mathbf{M}_2 = M_2\mathbf{k}$$

consider $V_1 = V,\ M_1 = -M$ at $\otimes$

then, $V_2 = -(V + dV), M_2 = (M + dM)$

$$\mathbf{V}_1 = V\mathbf{j}, \quad \mathbf{V}_2 = -(V + dV)\mathbf{j}, \quad \mathbf{M}_1 = -M\mathbf{k}, \quad \mathbf{M}_2 = (M + dM)\mathbf{k}$$

Relations of Distributed Force, Shear Forces, and Bending Moments with Vector notation



$$\mathbf{f}(x) = -f(x)dx\mathbf{j}$$

$$\mathbf{V}_1 = V\mathbf{j},\quad \mathbf{V}_2 = -(V + dV)\mathbf{j},$$

$$\mathbf{M}_1 = -M\mathbf{k},\quad \mathbf{M}_2 = (M + dM)\mathbf{k}$$

Force equilibrium

$$\mathbf{F}_y = \mathbf{V}_1 + \mathbf{V}_2 + \mathbf{f}(x) = 0$$

$$(V\mathbf{j}) + \left(-(V + dV)\mathbf{j}\right) + \left(-f(x)dx\mathbf{j}\right) = 0$$

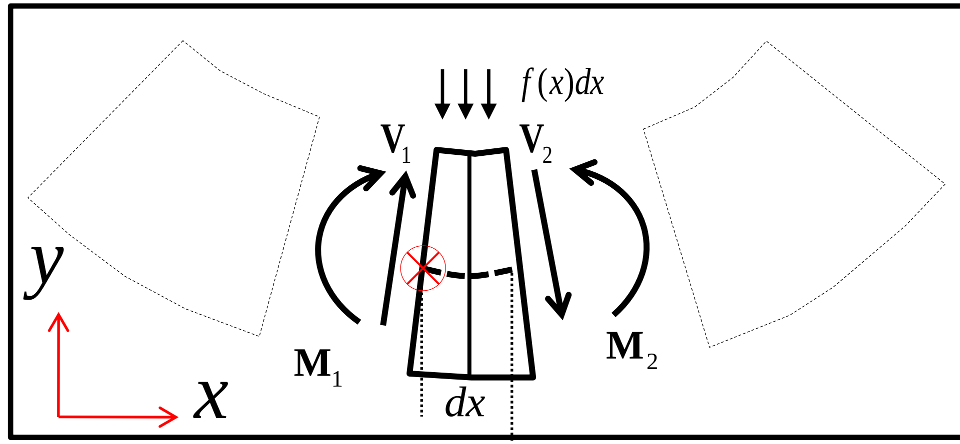
$$\left(V - V - dV - f(x)dx\right)\mathbf{j} = 0$$

$$\boxed{\therefore \frac{dV}{dx} = -f(x)}$$

Relations of Distributed Force, Shear Forces, and Bending Moments with Vector notation

Moment equilibrium about *z*-axis through the point $\otimes$

$$\mathbf{M}_z = \mathbf{M}_1 + \mathbf{M}_2 + d\mathbf{x} \times \mathbf{V}_2 + \frac{1}{2} d\mathbf{x} \times \left(\mathbf{f}(x)\right) =$$



$$-M\mathbf{k} + \left(M + dM\right)\mathbf{k} + \left(dx\,\mathbf{i}\right) \times \left(-\left(V + dV\right)\mathbf{j}\right) + \left(\frac{1}{2} dx\,\mathbf{i}\right) \times \left(-f(x)dx\mathbf{j}\right) = 0$$

$$\mathbf{f}(x) = -f(x)dx\mathbf{j}$$

$$\mathbf{V}_1 = V\mathbf{j},\ \ \mathbf{V}_2 = -\left(V + dV\right)\mathbf{j},$$

$$\mathbf{M}_1 = -M\mathbf{k},\ \ \mathbf{M}_2 = \left(M + dM\right)\mathbf{k}$$

$$\left(-M + M + dM - Vdx - \cancel{dV \cdot dx} - \cancel{\frac{1}{2} f(x)(dx)^2}\right)\mathbf{k} = 0$$

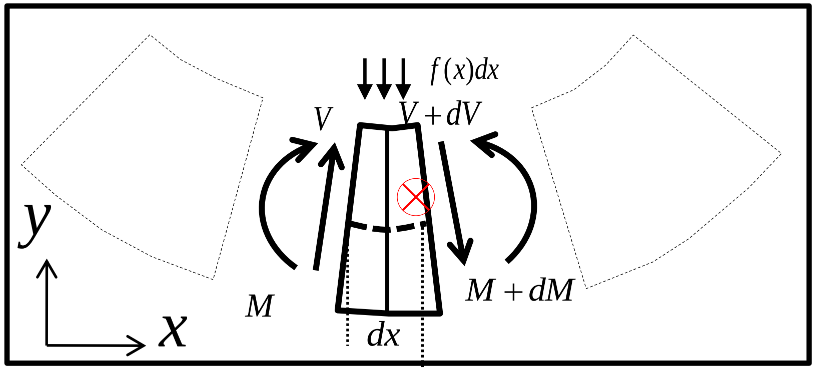
neglecting the second order or higher terms

$$\left(dM - Vdx\right)\mathbf{k} = 0$$

$$\boxed{\therefore \frac{dM}{dx} = V(x)}$$

Sign Conventions and Differential Equation of Deflection Curve of Beam

Recall, $M = EI\dfrac{d^2y}{dx^2}$



This equation is derived with the positive shear forces and the positive bending moments in deformation sign conventions

$$\frac{d^2y}{dx^2} = \frac{M}{EI}$$

$$\frac{d^3y}{dx^3} = \frac{1}{EI}\cdot\frac{dM}{dx} = \frac{1}{EI}\cdot V(x)$$

$$\frac{dV}{dx} = -f(x)$$

$$\frac{dM}{dx} = V(x)$$

$$\frac{d^4y}{dx^4} = \frac{1}{EI}\cdot\frac{dV}{dx} = -\frac{1}{EI}\cdot f(x)$$

Differential equation of deflection curve of a beam

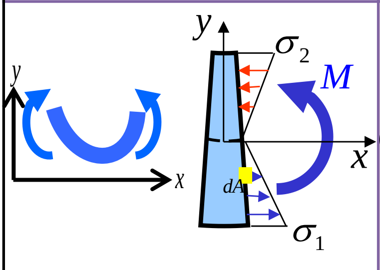
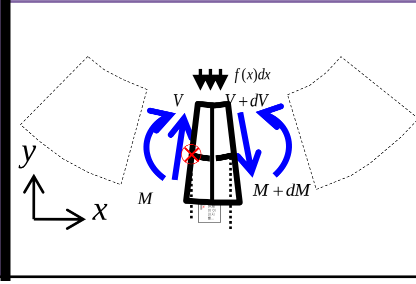
$$EI\frac{d^4y}{dx^4} = -f(x)$$

the equation is derived with the positives directions of the distributed load, the shear forces, and the bending moments.

Sign Conventions and Differential Equation of Deflection Curve of Beam

$$\varepsilon \Rightarrow \sigma = E \cdot \varepsilon \Rightarrow y\sigma dA \rightarrow dM$$

Distributed load, shear force, bending moment, curvature, bending : positive

deformation (ref : Gere)	B.M.	K =1/ρ	y	ε	σ	check	dM= yσdA	dM	$M = \int_A dM$	relation btw V, M, f(x)	
	⊕	+	+	$-\kappa y$ or $-\dfrac{y}{\rho}$	-	comp.	⊖	$-\sigma y dA$ modified	$M = -\int_A y\sigma dA$ $= -\int_A y(-\frac{E}{\rho}y)dA$ $M = \frac{E}{\rho}\int_A y^2 dA$ $\dfrac{M}{EI} = \dfrac{d^2y}{dx^2}$		$V - f(x)dx - (V + dV) = 0$ $\Rightarrow \dfrac{dV}{dx} = -f(x)$ $-M + f(x)dx \cdot \frac{1}{2}dx + (M + dM) - (V + dV)dx = 0$ $\Rightarrow \dfrac{dM}{dx} = V(x)$ $EI\dfrac{d^4y}{dx^4} = -f(x)$
			-		+	tension					

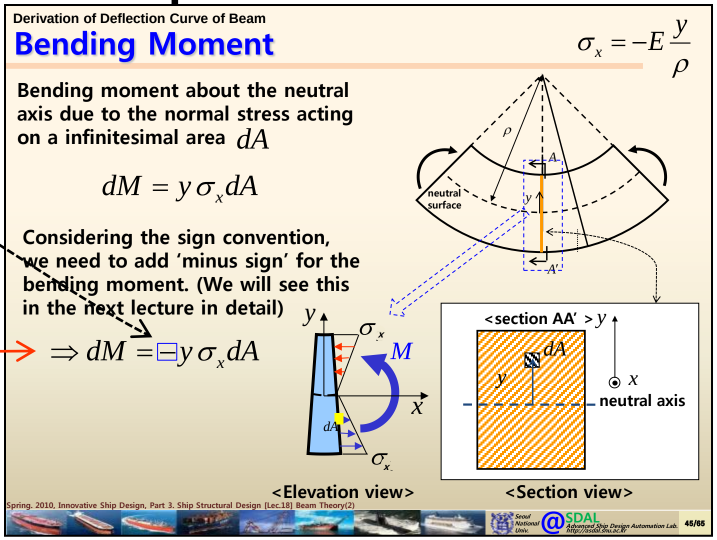
not match!

Recall, p.47

$$d\mathbf{M} = \mathbf{y} \times d\mathbf{F}$$
$$= (y\mathbf{j}) \times (\sigma_x \, dA\mathbf{i})$$
$$= \boxed{-} y\,\sigma_x dA\mathbf{k}$$
$$= -y\left(-E\frac{y}{\rho}\right)dA\mathbf{k} = E\frac{y^2}{\rho}dA\mathbf{k}$$

modified

Recall, p.47


Derivation of Deflection Curve of Beam

Bending Moment

$$\sigma_x = -E\frac{y}{\rho}$$

Bending moment about the neutral axis due to the normal stress acting on a infinitesimal area dA

$$dM = y\,\sigma_x dA$$

Considering the sign convention, we need to add 'minus sign' for the bending moment. (We will see this in the next lecture in detail)

$$\Rightarrow dM = \boxed{-} y\,\sigma_x dA$$

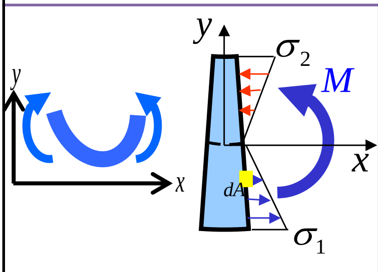
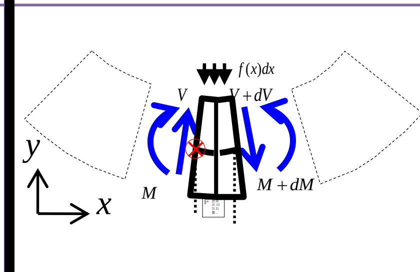
Sign Conventions and Differential Equation of Deflection Curve of Beam - Comparison

convention

Derived

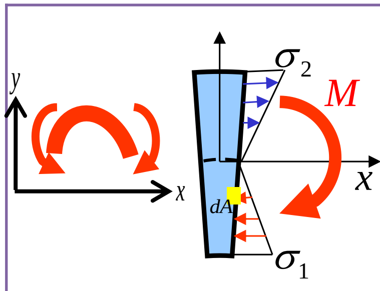
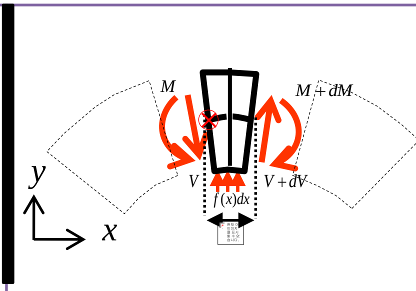
$$\varepsilon \Rightarrow \sigma = E\cdot\varepsilon \Rightarrow y\sigma dA \rightarrow dM$$

Distributed load, shear force, bending moment, curvature, bending : positive

deformation (ref : Gere)	B.M.	K =1/ρ	y	ε	σ	check	dM= yσdA	dM	$M=\int_A dM$	relation btw V, M, f(x)	
	+	+	+	$-\kappa y$ or $-\dfrac{y}{\rho}$	-	comp.	-	$-\sigma y dA$ modified	$M=-\int_A y\sigma dA$ $=-\int_A y(-\frac{E}{\rho}y)dA$ $M=\frac{E}{\rho}\int_A y^2 dA$ $\frac{M}{EI}=\frac{d^2y}{dx^2}$		$V-f(x)dx-(V+dV)=0$ $\Rightarrow \frac{dV}{dx}=-f(x)$ $-M+f(x)dx\cdot\frac{1}{2}dx+(M+dM)-(V+dV)dx=$ $\Rightarrow \frac{dM}{dx}=V(x)$ $EI\frac{d^4y}{dx^4}=-f(x)$
			-		+	tension					

not match!

Distributed load, shear force, bending moment, curvature, bending : negative

deformation	B.M.	K =1/ρ	y	ε	σ	check	dM= yσdA	dM	$M=\int_A dM$	relation btw V, M, f(x)	
	-	-	+	$-\kappa y$ or $-\dfrac{y}{\rho}$	+	tension	+	$-\sigma y dA$ modified	$M=-\int_A y\sigma dA$ $=-\int_A y(-\frac{E}{\rho}y)dA$ $M=\frac{E}{\rho}\int_A y^2 dA$ $\frac{M}{EI}=\frac{d^2y}{dx^2}$		$-V+f(x)dx+(V+dV)=0$ $\Rightarrow \frac{dV}{dx}=-f(x)$ $M+f(x)dx\cdot\frac{1}{2}dx-(M+dM)+(V+dV)dx=$ $\Rightarrow \frac{dM}{dx}=V(x)$ $EI\frac{d^4y}{dx^4}=-f(x)$
			-		-	comp.					

not match!

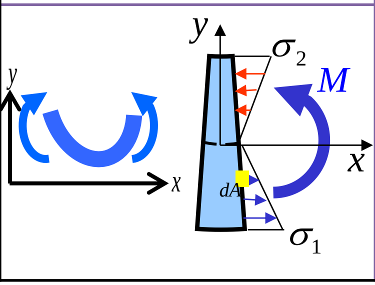
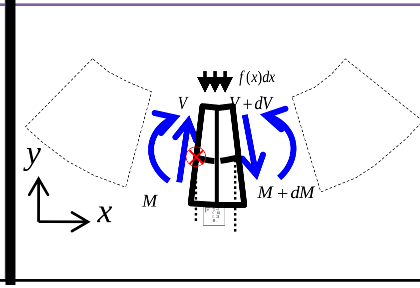
Comparison : Gere[1)] and 임상전[2)]

convention

Derived

$$\varepsilon \Rightarrow \sigma = E \cdot \varepsilon \Rightarrow y\sigma dA \rightarrow dM$$

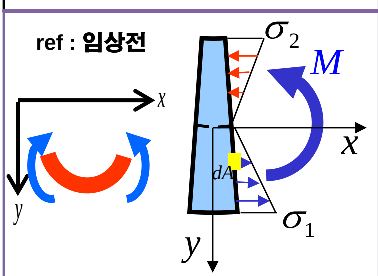
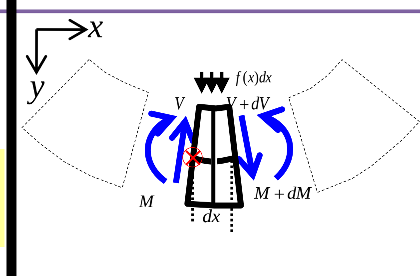
1) Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006

deformation (ref : Gere)	B.M.	K =1/ρ	y	ε	σ	check	dM= yσdA	dM	$M = \int_A dM$	relation btw V, M, f(x)	
	⊕	+	+	$-\kappa y$ or $-\dfrac{y}{\rho}$	-	comp.	⊖	$-\sigma ydA$	$M = -\int_A y\sigma dA$ $= -\int_A y(-\frac{E}{\rho}y)dA$ $M = \frac{E}{\rho}\int_A y^2 dA$ $\dfrac{M}{EI} = \dfrac{d^2y}{dx^2}$		$V - f(x)dx - (V+dV) = 0$ $\Rightarrow \dfrac{dV}{dx} = -f(x)$ $-M + f(x)dx \cdot \frac{1}{2}dx + (M+dM) - (V+dV)dx = 0$ $\Rightarrow \dfrac{dM}{dx} = V(x)$ $EI\dfrac{d^4y}{dx^4} = -f(x)$
			-		+	tension					

modified

not match!

2) 임상전 편저, 재료역학, 2002년 ,문운당 (Timoshenko S., Young D.H., Elements of strength of materials, 5th edition, Van Nostrand, 1968

	B.M.	K =1/ρ	y	ε	σ	check	dM= yσdA	dM	$M = \int_A dM$	relation btw V, M, f(x)	
ref : 임상전 	⊕	≠ -	-	κy or $\dfrac{y}{\rho}$	-	comp.	⊕	σydA	$\dfrac{d^2y}{dx^2} = \pm\dfrac{M}{EI}$ ⇩ $\dfrac{d^2y}{dx^2} = -\dfrac{M}{EI}$		$-V + f(x)dx + (V+dV) = 0$ $\Rightarrow \dfrac{dV}{dx} = -f(x)$ $M - (M+dM) + (V+dV)dx - f(x)dx \cdot \frac{1}{2}dx = 0$ $\Rightarrow \dfrac{dM}{dx} = V(x)$ $EI\dfrac{d^4y}{dx^4} = -f(x)$
			+		+	tension					

match

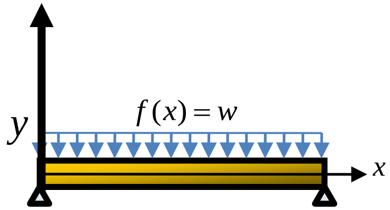
modify considering the curvature

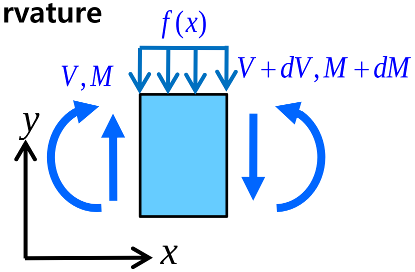
All sign conventions used in ‘임상전 ‘are same as those of ‘Gere’ except the opposite direction of y-axis.

Comparison of Solutions : Gere[1)] and 임상전[2)]

'Gere'

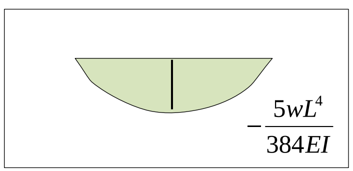
Positive Bending Moment
Positive Curvature



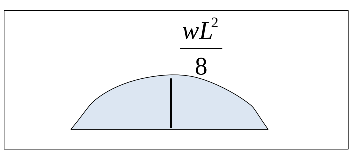


$$EI\frac{d^4y(x)}{dx^4} = -w$$

$$y(x) = -\frac{wx}{24EI}(L^3 - 2Lx^2 + x^3)$$

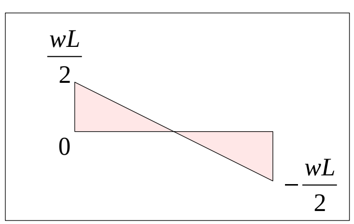


$$M(x) = \frac{wLx}{2} - \frac{wx^2}{2}$$



match

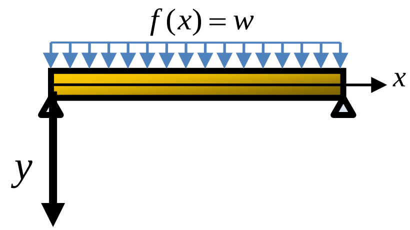
$$V(x) = \frac{wL}{2} - wx$$

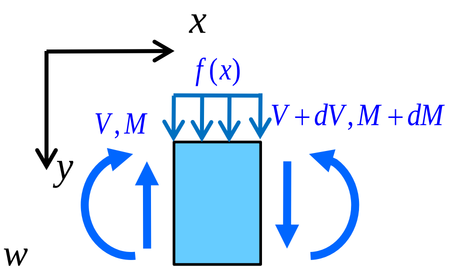


match

'임상전'

Positive Bending Moment
Negative Curvature





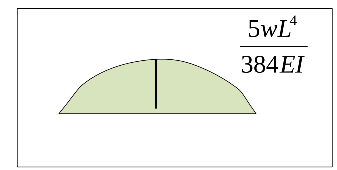
$$EI\frac{d^4y(x)}{dx^4} = w$$

$$y'(x) = \frac{w}{24EI}(L^3 - 6Lx^2 + 4x^3)$$

$$y''(x) = \frac{w}{24EI}(-12Lx + 12x^2)$$

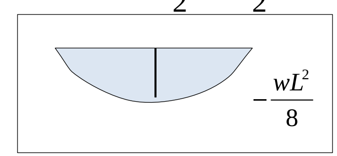
$$= \frac{w}{2EI}(-Lx + x^2)$$

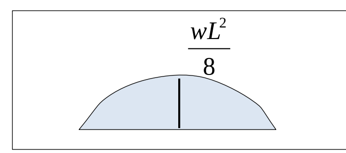
$$y(x) = \frac{wx}{24EI}(L^3 - 2Lx^2 + x^3)$$



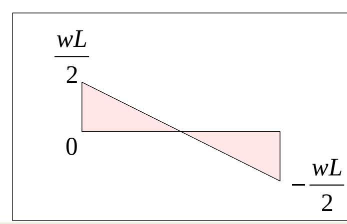
Since 'Negative Curvature' is used for the 'Positive Bending Moment'

$$M(x) = -\frac{wLx}{2} + \frac{wx^2}{2}$$





correction



Both of the solutions can explain the physical phenomenon correctly as long as they are interpreted by the used sign conventions.

1) Gere J.M., Mechanics of Materials, 6th edition, Thomson, 2006

2) 임상전 편저, 재료역학, 2002년 ,문운당 (Timoshenko S., Young D.H., Elements of strength of materials, 5th edition, Van Nostrand, 1968

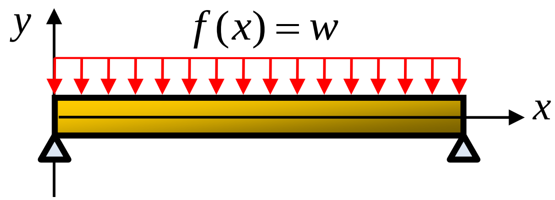
•Com|

1. Beam Theory

1.4 Examples of Deflection Curve of Beam

Example of Deflection Curve of Beam – Simply supported beam

$$\boxed{EI\frac{d^4y}{dx^4}=-f(x)}$$



distributed load $f(x)=w$

Boundary Condition

①No displacement at x=0 → $y(0)=0$

②No displacement at x=L → $y(L)=0$

③No bending moment x=0 → $EIy''(0)=0$

④No bending moment x=L → $EIy''(L)=0$

$$EI\frac{d^4y(x)}{dx^4}=-w$$

Integrate four times

$$y'''(x)=-\frac{w}{EI}x+c_1$$

$$y''(x)=-\frac{w}{2EI}x^2+c_1x+c_2$$

$$y'(x)=-\frac{w}{6EI}x^3+\frac{1}{2}c_1x^2+c_2x+c_3$$

$$y(x)=-\frac{w}{24EI}x^4+\frac{1}{6}c_1x^3+\frac{1}{2}c_2x^2+c_3x+c_4$$

by the boundary condition *(E≠0, I≠0)*

$$③\ \left.\begin{aligned}EIy''(0)&=c_2\\0&=c_2\end{aligned}\right\}\Rightarrow\ \therefore c_2=0$$

$$①\ \left.\begin{aligned}y(0)&=c_4\\0&=c_4\end{aligned}\right\}\Rightarrow\ \therefore c_4=0$$

$$④\ \left.\begin{aligned}EIy''(L)&=-\frac{wL^2}{2EI}+c_1L\\0&=-\frac{wL^2}{2EI}+c_1L\end{aligned}\right\}\Rightarrow\ \therefore c_1=\frac{wL}{2EI}$$

$$②\ \left.\begin{aligned}y(L)&=-\frac{L^4w}{24EI}+\frac{wL^4}{12EI}+c_3L\\0&=-\frac{L^4w}{24EI}+\frac{wL^4}{12EI}+c_3L\end{aligned}\right\}\Rightarrow\ \therefore c_3=-\frac{wL^3}{24EI}$$

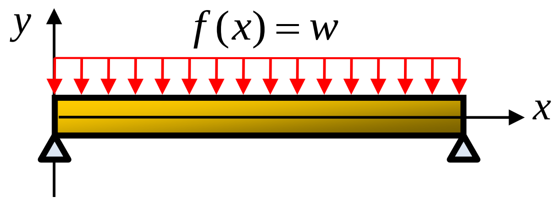
$$y(x)=-\frac{w}{24EI}x^4+\frac{wL}{12EI}x^3-\frac{wL^3}{24EI}x=-\frac{wx}{24EI}(L^3-2Lx^2+x^3)$$

$$M(x)=EIy''=\frac{wLx}{2}-\frac{wx^2}{2}$$

$$V(x)=EIy'''=\frac{wL}{2}-wx$$

Example of Deflection Curve of Beam
– Simply supported beam

$$\boxed{EI\frac{d^4y}{dx^4} = -f(x)}$$



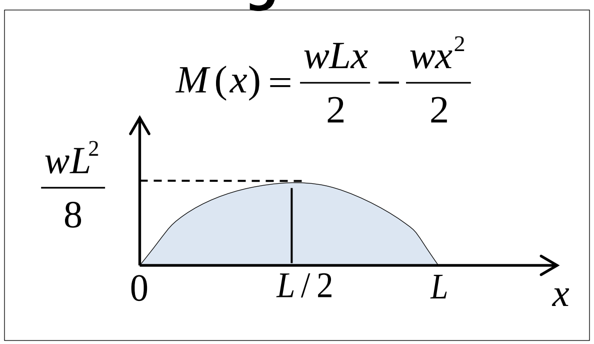
distributed load $f(x) = w$

$$y(x) = -\frac{wx}{24EI}(L^3 - 2Lx^2 + x^3)$$

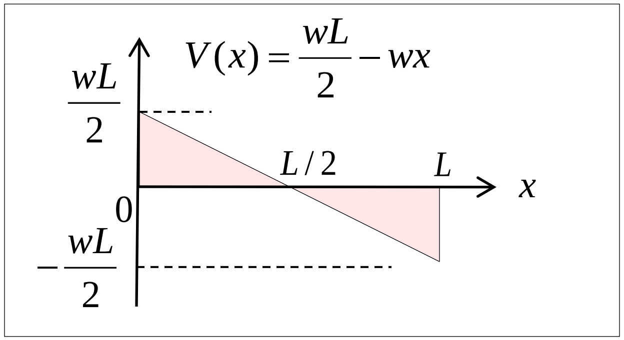
$$M(x) = \frac{wLx}{2} - \frac{wx^2}{2}$$

$$V(x) = \frac{wL}{2} - wx$$

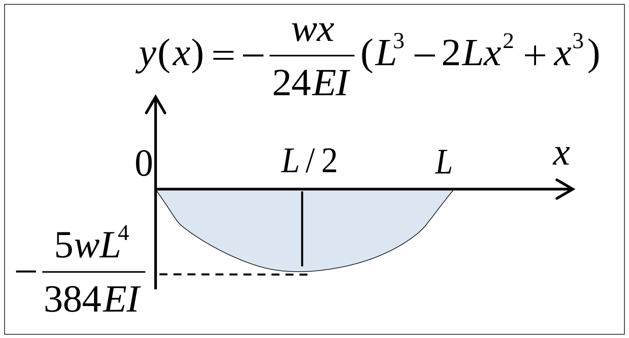
Bending moment



Shear force

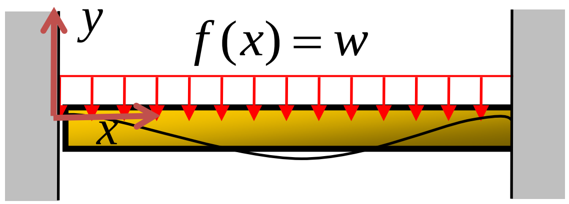


Deflection



Example of Deflection Curve of Beam – fixed-end beam

$$EI\frac{d^4y}{dx^4}=-f(x)$$



distributed load $f(x)=w$

Boundary Condition

①No displacement at x=0 → $y(0)=0$

②No displacement at x=L → $y(L)=0$

③No slope at x=0 → $y'(0)=0$

④No slope at x=0 → $y'(L)=0$

$$EI\frac{d^4y(x)}{dx^4}=-w$$

Integrate four times

$$y'''(x)=-\frac{w}{EI}x+c_1$$

$$y''(x)=-\frac{w}{2EI}x^2+c_1x+c_2$$

$$y'(x)=-\frac{w}{6EI}x^3+\frac{1}{2}c_1x^2+c_2x+c_3$$

$$y(x)=-\frac{w}{24EI}x^4+\frac{1}{6}c_1x^3+\frac{1}{2}c_2x^2+c_3x+c_4$$

by the boundary condition *(E≠0, I≠0)*

$$① \left.\begin{aligned} y(0)&=c_4\\ 0&=c_4\end{aligned}\right\} \Rightarrow \therefore c_4=0$$

$$③ \left.\begin{aligned} y'(0)&=c_3\\ 0&=c_3\end{aligned}\right\} \Rightarrow \therefore c_3=0$$

$$② \; y(L)=-\frac{1}{24EI}wL^4+\frac{1}{6}c_1L^3+\frac{1}{2}c_2L^2$$

$$0=-\frac{1}{24EI}wL^4+\frac{1}{6}c_1L^3+\frac{1}{2}c_2L^2$$

$$④ \; y'(L)=-\frac{1}{6EI}L^3+\frac{1}{2}c_1L^2+c_2L$$

$$0=-\frac{1}{6EI}L^3+\frac{1}{2}c_1L^2+c_2L$$

$$\rightarrow \quad c_2=-\frac{wL^2}{12EI},\quad c_1=\frac{wL}{2EI}$$

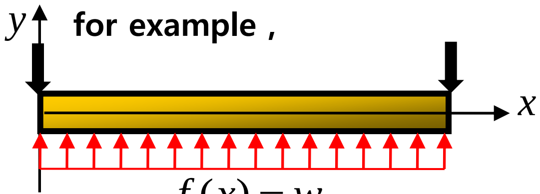
$$y(x)=-\frac{w}{24EI}x^4+\frac{wL}{12EI}x^3-\frac{wL^2}{24EI}x^2=-\frac{wL^2x^2}{24EI}\left(\frac{x}{L}-1\right)^2$$

$$M(x)=EIy''=-\frac{wL^2}{12}\left(\frac{6x^2}{L^2}-\frac{6x}{L}+1\right)$$

$$V(x)=EIy'''=-\frac{wL}{2}\left(\frac{2x}{L}-1\right)$$

Example of Deflection Curve of Beam by Negative Deformation Sign Conventions

for example ,



$$f(x) = w$$

$$EI\frac{d^4y(x)}{dx^4} = -w \qquad y(0)=0 \qquad y(L)=0 \qquad y''(0)=0 \qquad y''(L)=0$$

After integrate four times, $y(x) = c_1 + c_2x + c_3x^2 + c_4x^3 - \frac{w}{24EI}x^4$

$$y''(x) = 2c_3 + 6c_4x + \frac{w}{2EI}x^2$$

$y(0)=0$, $y(0)=c_1$ ⇨ $\therefore c_1 = 0$

$y''(0)=0$, $y''(0)=2c_3$ ⇨ $\therefore c_3 = 0$

$y''(L)=0$, $y''(x) = 6c_4L - \frac{w}{2EI}L^2$ ⇨ $\therefore c_4 = \frac{w}{12EI}L$

$y(L)=0$, $y(L) = c_2L + c_4L^3 - \frac{w}{24EI}L^4$ ⇨ $\therefore c_2L + c_4L^3 - \frac{w}{24EI}L^4 = 0$

$$c_2 = -\frac{w}{24EI}L^3$$

$$\begin{aligned} c_2 &= -c_4L^2 + \frac{w_0}{24EI}L^3 \\ &= -(\frac{w}{12EI}L)L^2 + \frac{w}{24EI}L^3 \\ &= -\frac{w}{12EI}L^3 + \frac{w}{24EI}L^3 \\ &= -\frac{w}{24EI}L^3 \end{aligned}$$

$$\begin{aligned} y(x) &= -\frac{w}{24EI}L^3x + \frac{w}{12EI}Lx^3 - \frac{w}{24EI}x^4 \\ &= -\frac{wx}{24EI}\left(L^3 - 2Lx^2 + x^3\right) \end{aligned}$$

⇨

$$y(x) = -\frac{wx}{24EI}(L^3 - 2Lx^2 + x^3)$$

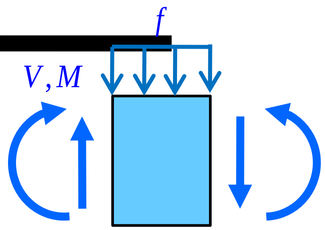
$$V(x) = \frac{wL}{2} - wx$$

$$M(x) = \frac{wLx}{2} - \frac{wx^2}{2}$$

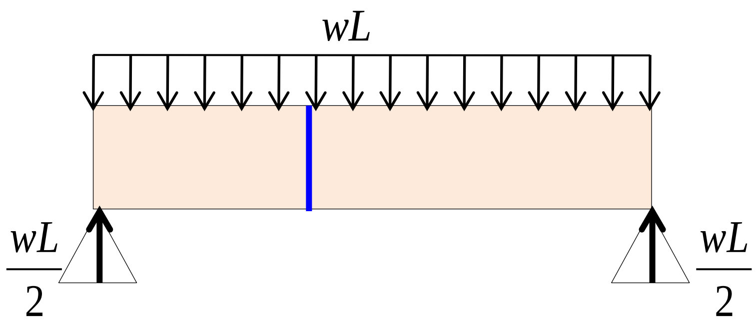
1. Beam Theory

1.5 Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces and Bending Moments

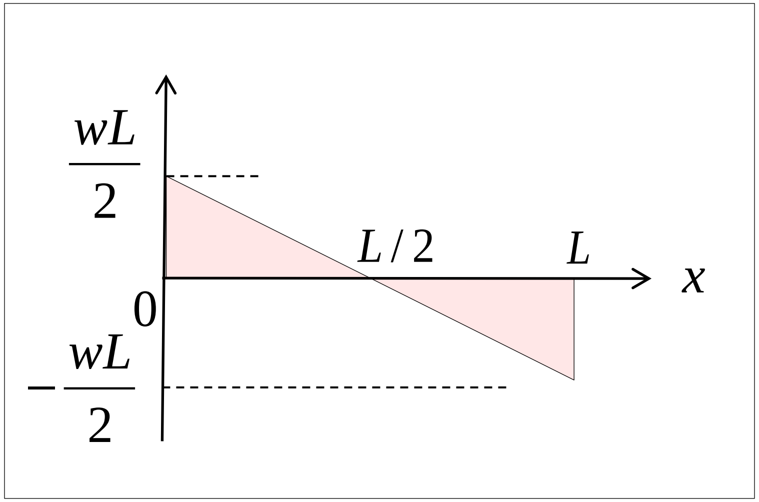
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces



The values of the function V(x) at x are the values of the shear force acting on the cross section of the beam

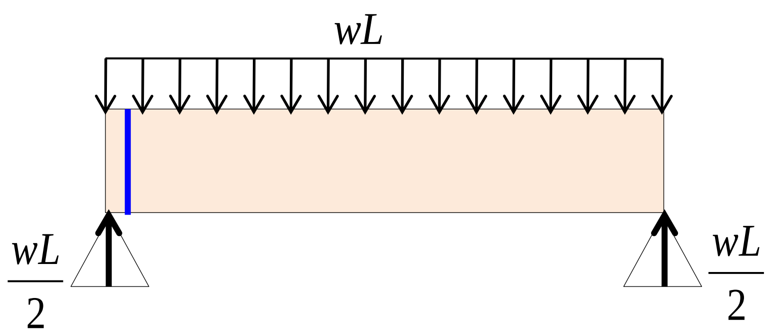


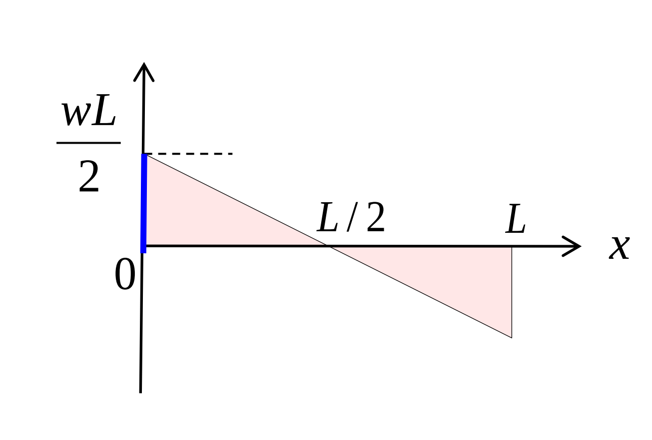
Value : $V(x) = \frac{wL}{2} - wx$



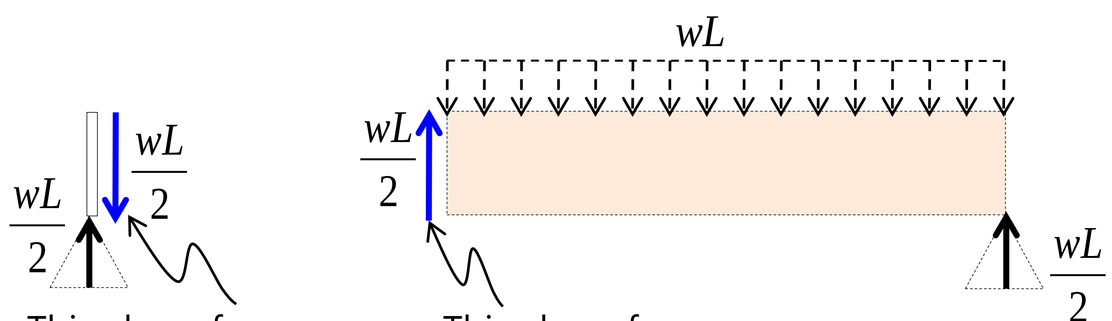
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces

$at\ \ x=0$



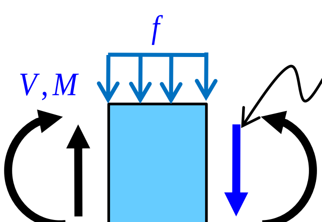


The free-body diagram



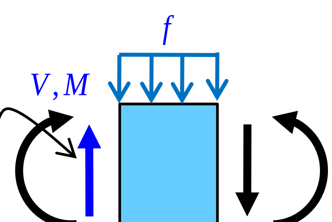
This shear force acts downward.

This shear force acts upward.



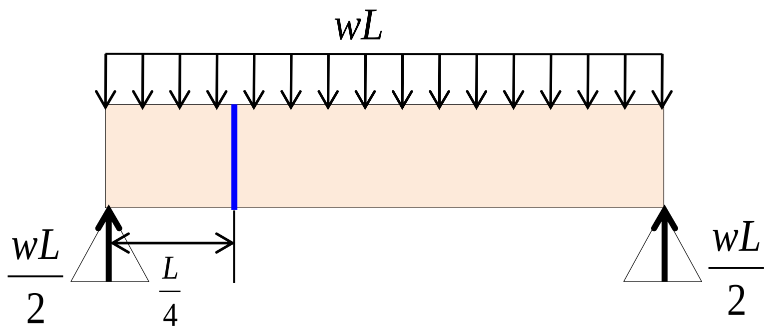
And the shear force acting downward is positive for right side of the element

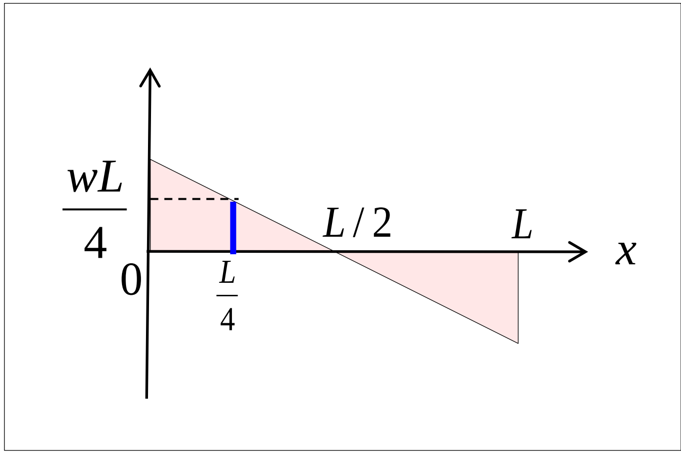
And the shear force acting upward is positive for the left side of the element



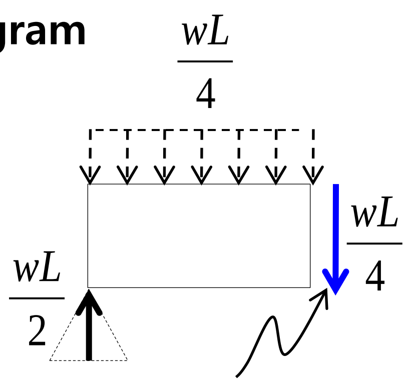
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces

$$at\ x = \frac{L}{4}$$

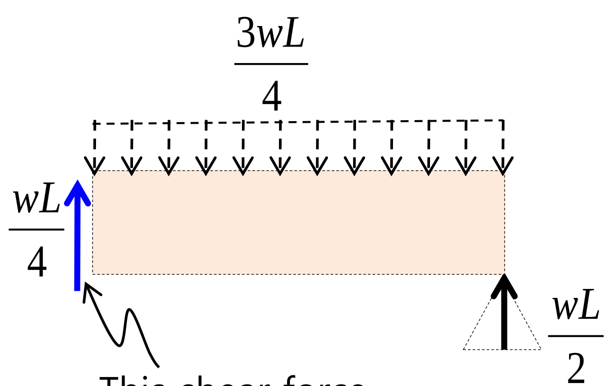




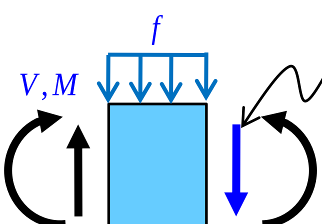
The free-body diagram



This shear force acts downward.

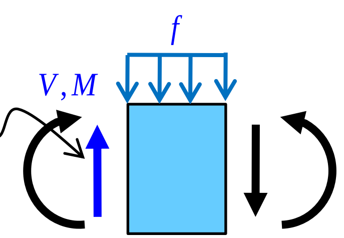


This shear force acts upward.



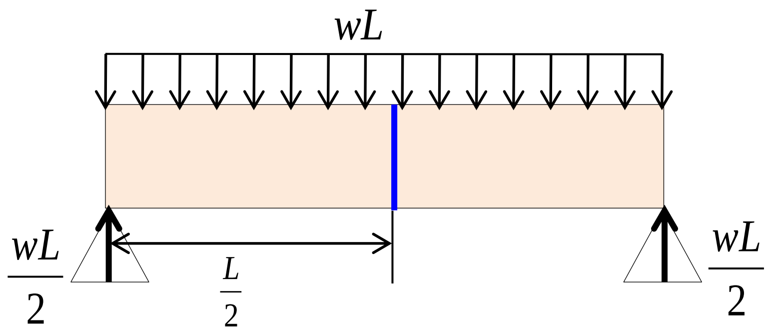
And the shear force acting downward is positive for right side of the element

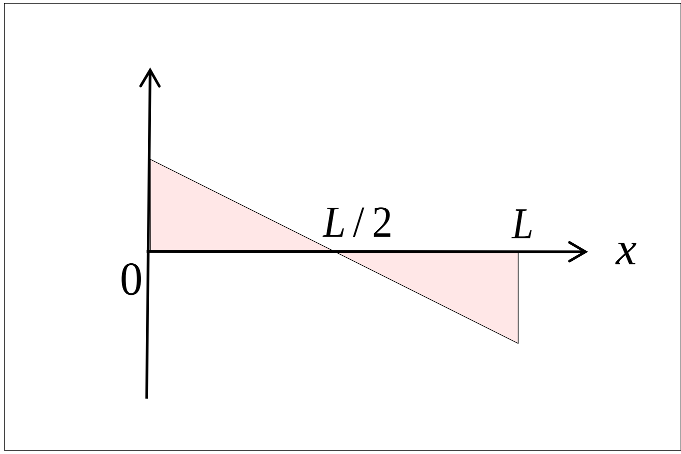
And the shear force acting upward is positive for the left side of the element



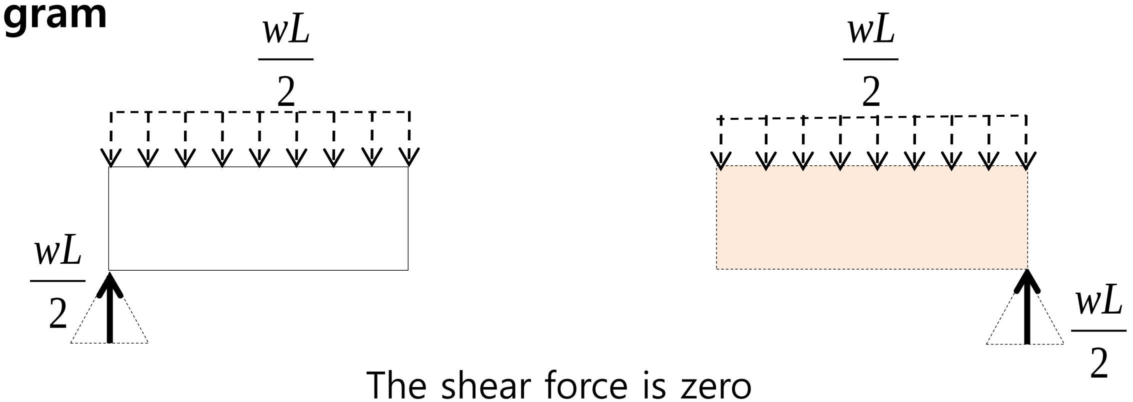
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces

$$at\ \ x = \frac{L}{2}$$





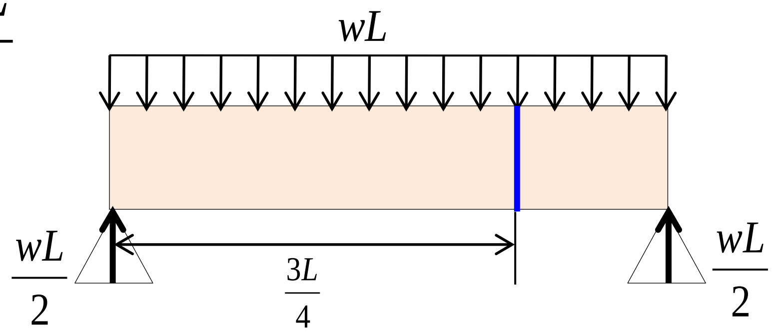
The free-body diagram

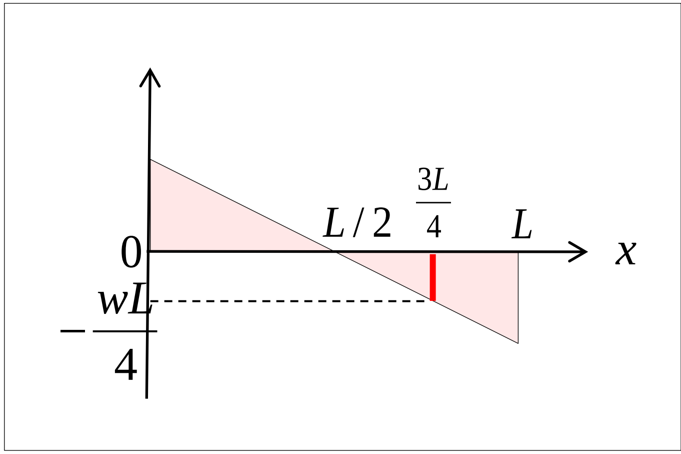


The shear force is zero

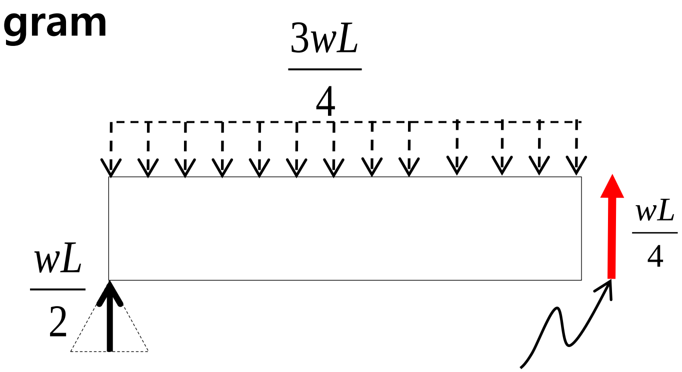
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces

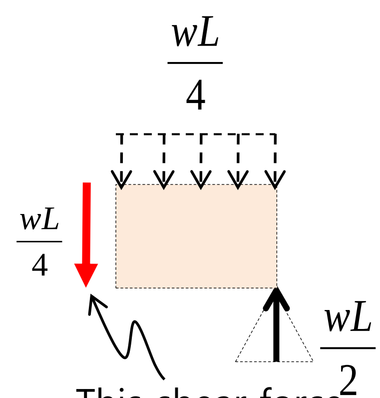
$$at\ \ x = \frac{3L}{4}$$





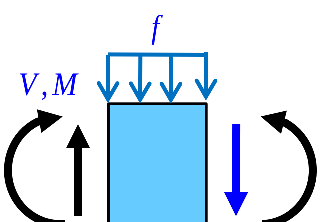
The free-body diagram





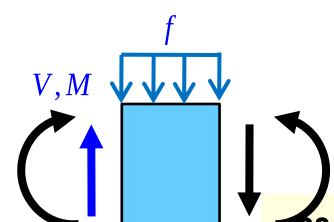
This shear force acts upward.

This shear force acts downward.



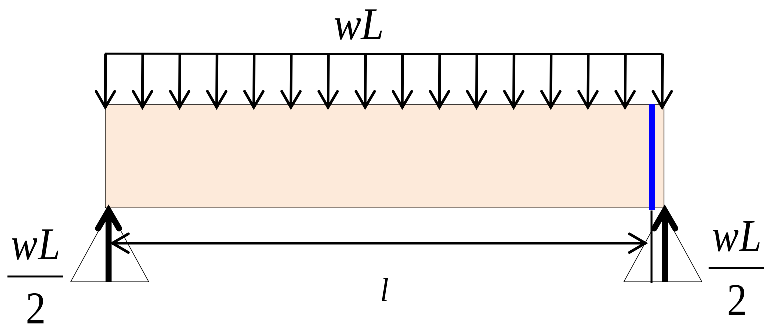
And the shear force acting upward is negative for right side of the element

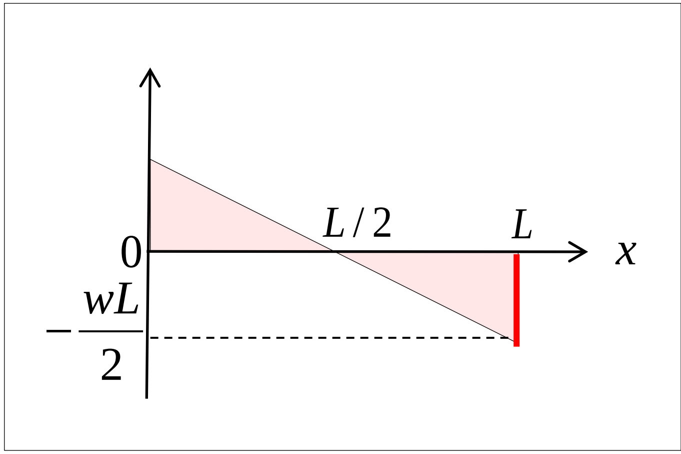
And the shear force acting downward is negative for the left side of the element



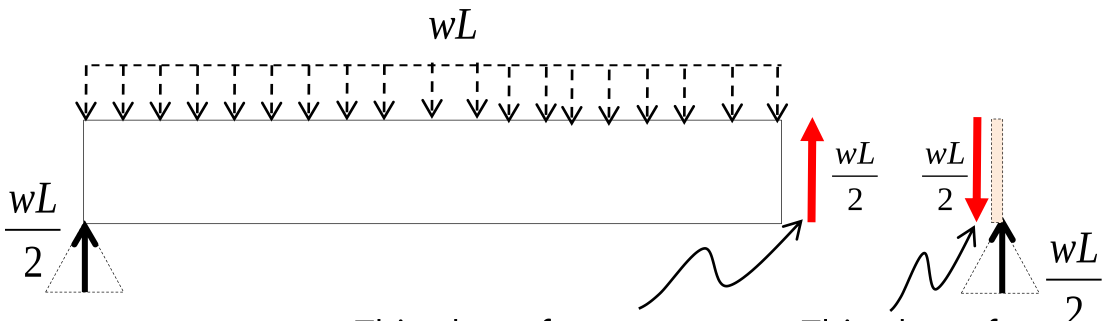
Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Shear Forces

$$at\ \ x = L$$



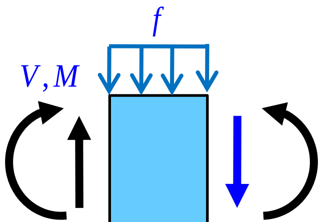


The free-body diagram



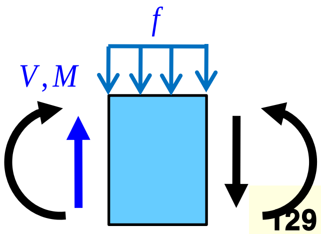
This shear force acts upward.

This shear force acts downward.

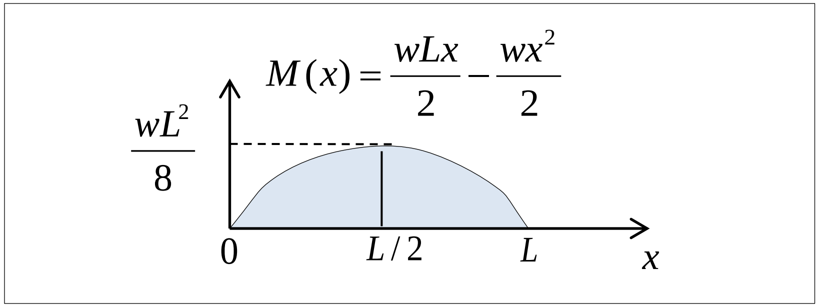


And the shear force acting upward is negative for right side of the element

And the shear force acting downward is negative for the left side of the element



Sign Conventions and Differential Equation of Deflection Curve of Beam – Interpretation of Bending Moments



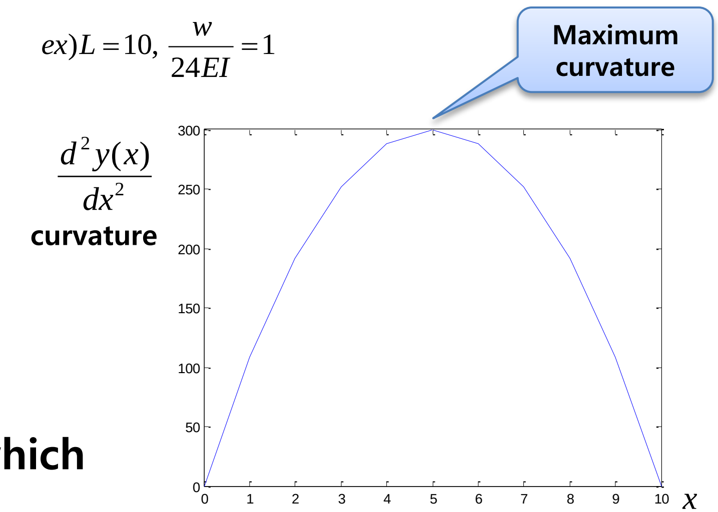
Why is the bending moment maximum at $x = \frac{L}{2}$ **?**

$$y(x) = -\frac{wx}{24EI}(L^3 - 2Lx^2 + x^3)$$

$$\frac{d^2y(x)}{dy^2} = -\frac{wx}{24EI}(12x - 6L)$$

since, $\frac{M(x)}{EI} = \frac{d^2y(x)}{dx^2}$

The bending moment is maximum at the point at which the curvature is maximum.



Sign Conventions for Stress Analysis

Fall 2011

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University

Equations of Motions

Newton's second law

$$m\frac{d\mathbf{V}}{dt}=\mathbf{F}$$

Classification of the resultant force

$$\mathbf{F}=\mathbf{F}_{Body}+\mathbf{F}_{Surface}$$

Derivation of Equations of Motions

$$\boxed{m}\boxed{\frac{d\mathbf{V}}{dt}}=\boxed{\mathbf{F}_{Body}}+\boxed{\mathbf{F}_{Surface}}$$

→surface force

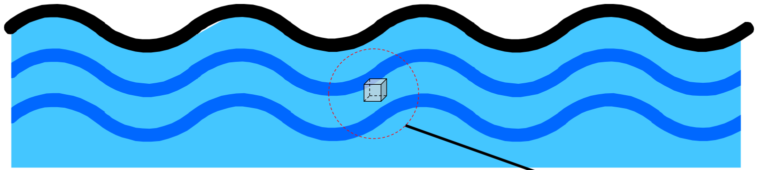
→body force

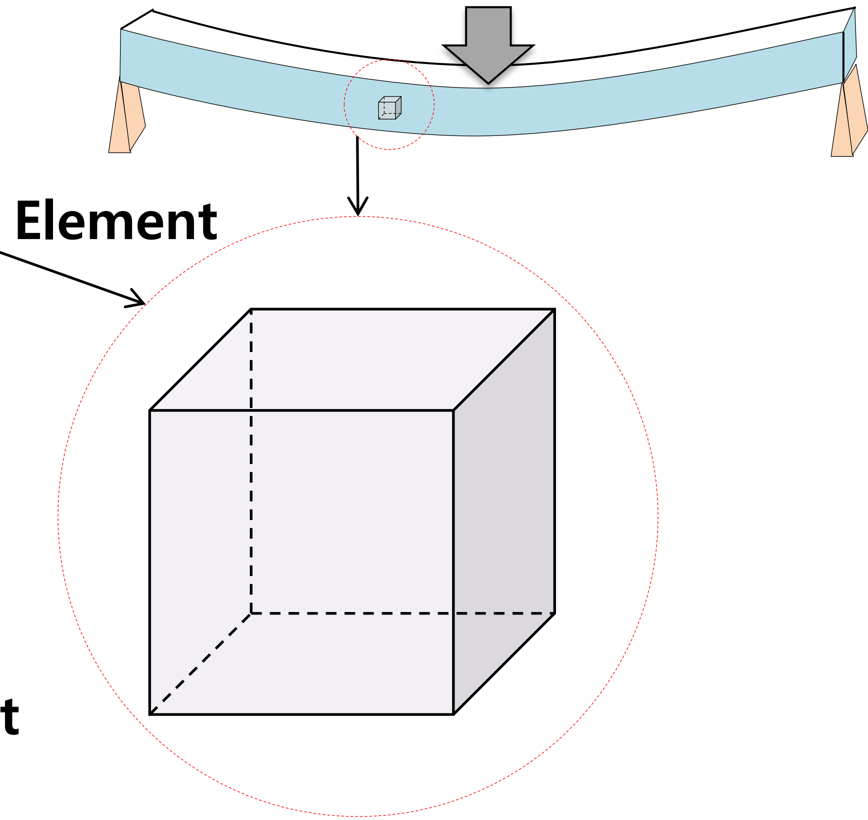
→acceleration

→mass

Sign convention for stress analysis

Element of Material





Derivation of Equations of Motions

$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \mathbf{F}_{Surface}$$

surface force of the element

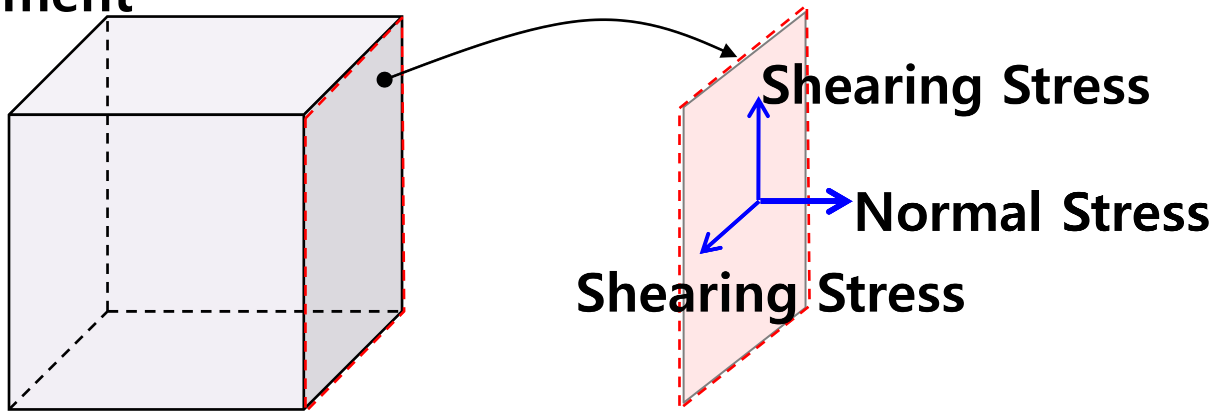
body force of the element

acceleration of the element

mass of the element

Surface Forces : Normal Stress and Shearing Stress

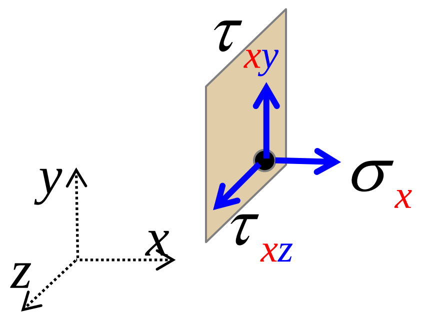
Element



σ **Normal stress : one normal direction**

τ **Shearing stress : two tangential direction**

Notation



τ_{xy}

Orientation of the surface

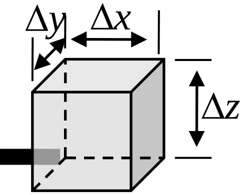
Direction of the force

$\sigma_{xx} \Rightarrow \sigma_x$

CAUTION

$$\sigma_x \neq \frac{\partial \sigma}{\partial x}$$

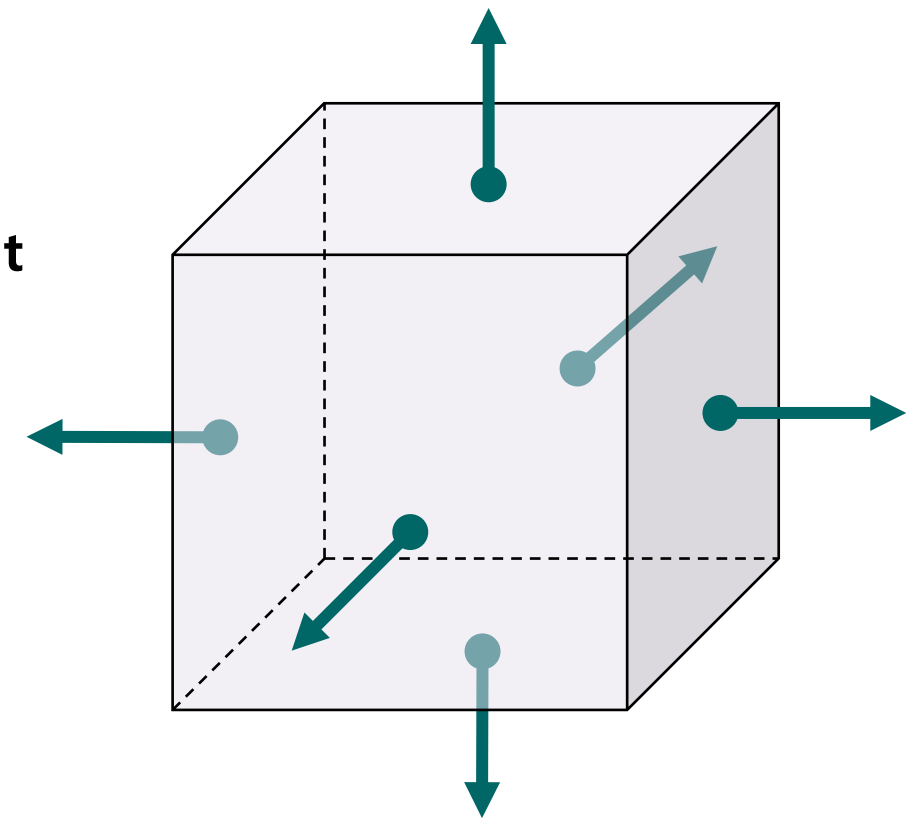
Sign Convention for Normal Stress



Normal Stress Sign Convention*

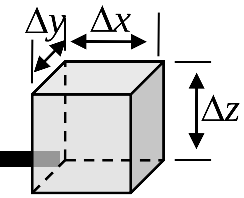
A normal stress is defined as positive if it is a tensile stress, i.e., if it is directed away from the surface upon which it acts

Element



*Wang.C.T , Applied Elasticity , McGRAW-HILL, 1953, p2
*Kundu P.K., Cohen I.M., Fulid Mechanics, Fourth Edition, Academic Press, p31

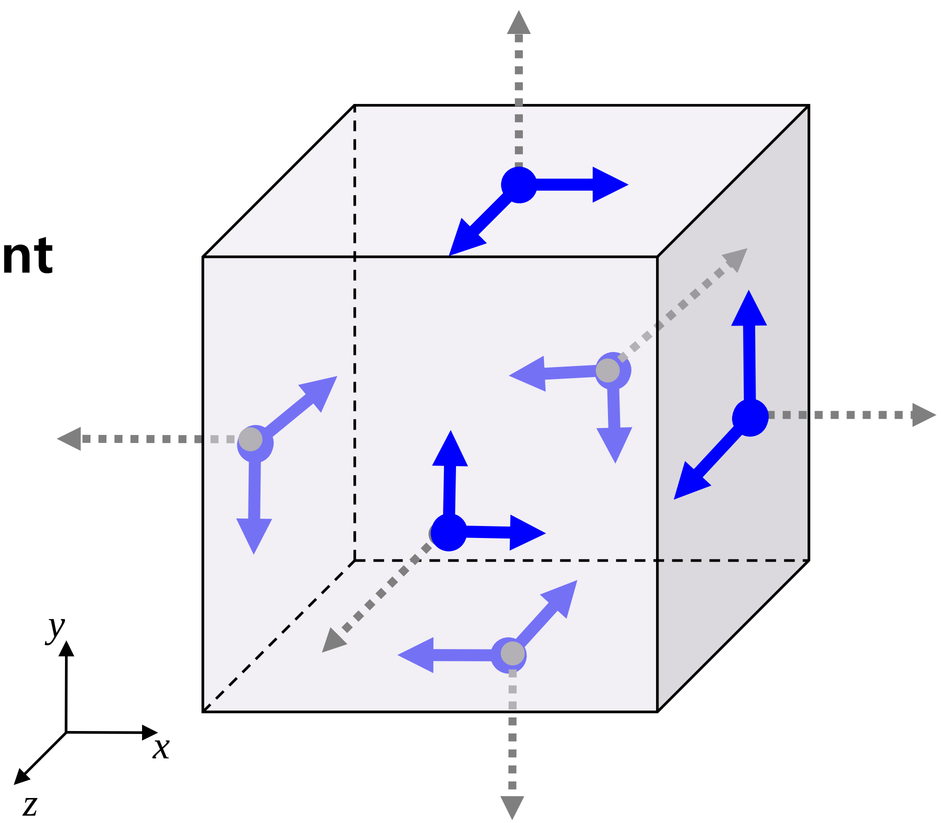
Sign Convention for Normal Stress



Shearing Stress Sign Convention*

A shearing stresses are positive if they are in the positive directions of the other two coordinates axes on any surface where the tensile stress is in the positive direction of the coordinate axis.

Element



If the tensile stress is opposite to the positive axis, the positive directions of the shearing stresses are also opposite to the positive axes.

*Wang.C.T , Applied Elasticity , McGRAW-HILL, 1953, p2

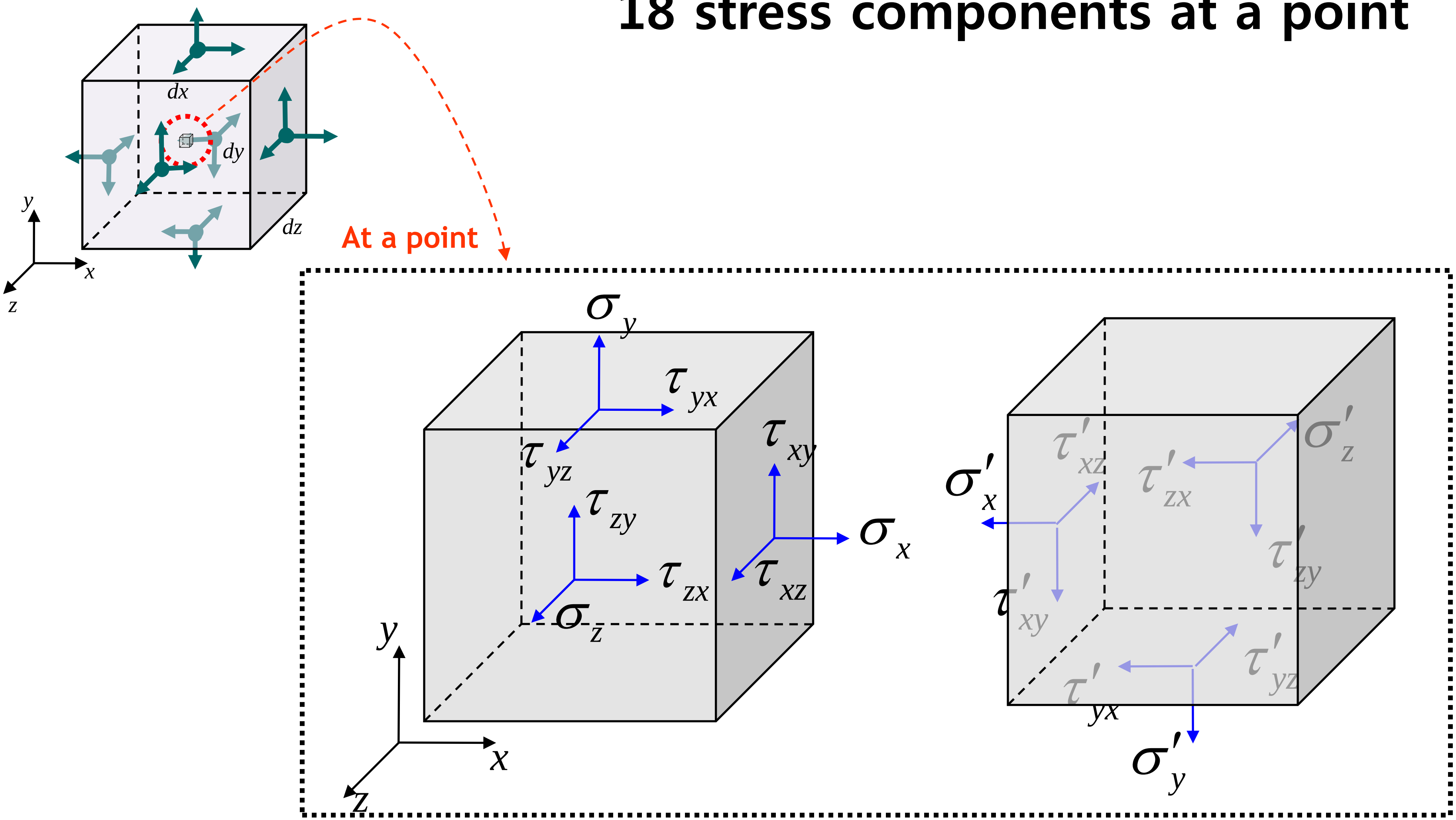
 *Kundu P.K., Cohen I.M., Fulid Mechanics, Fourth Edition, Academic Press, p31

Stress Components

No. of Stress Components at a Point

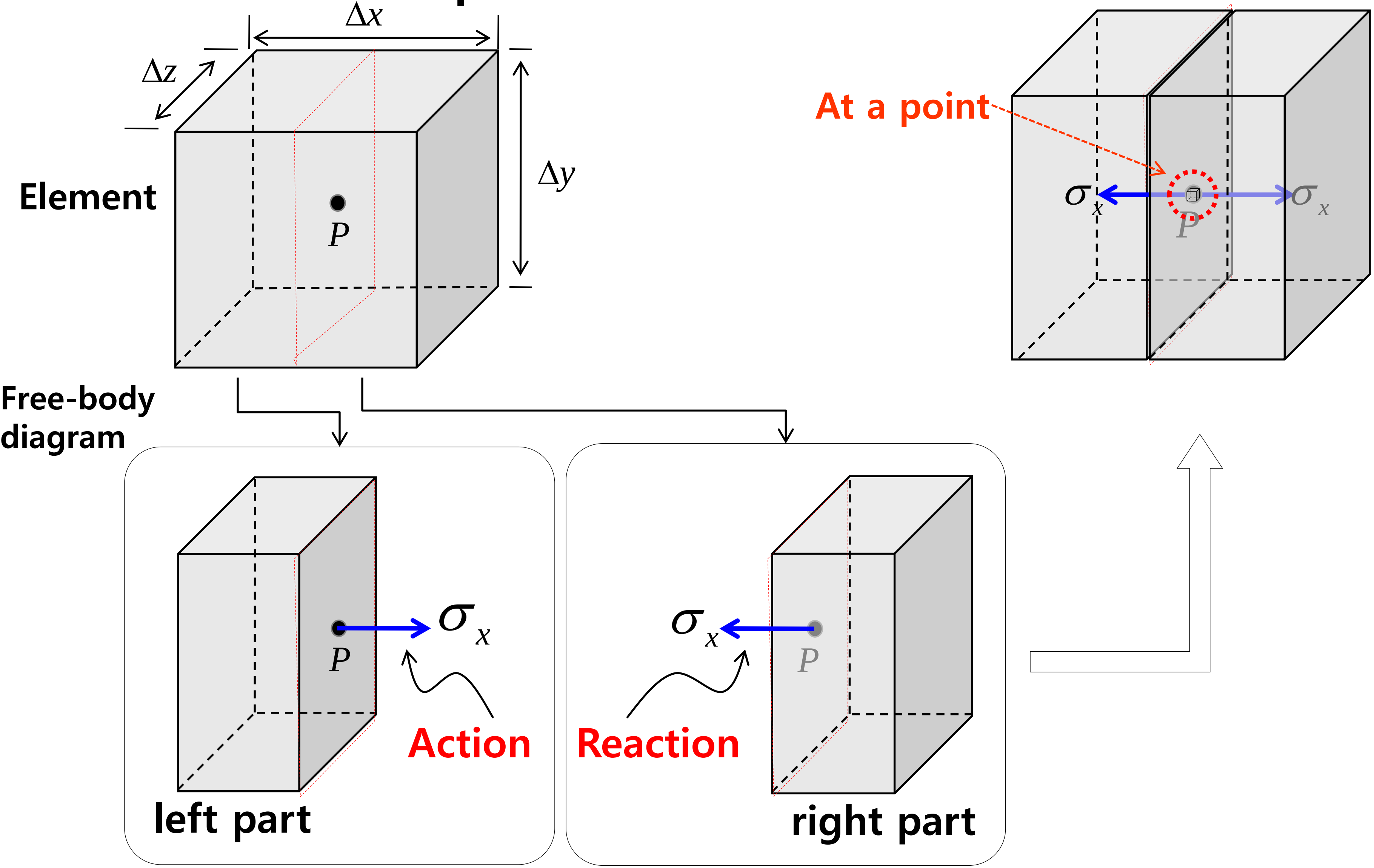
element

18 stress components at a point



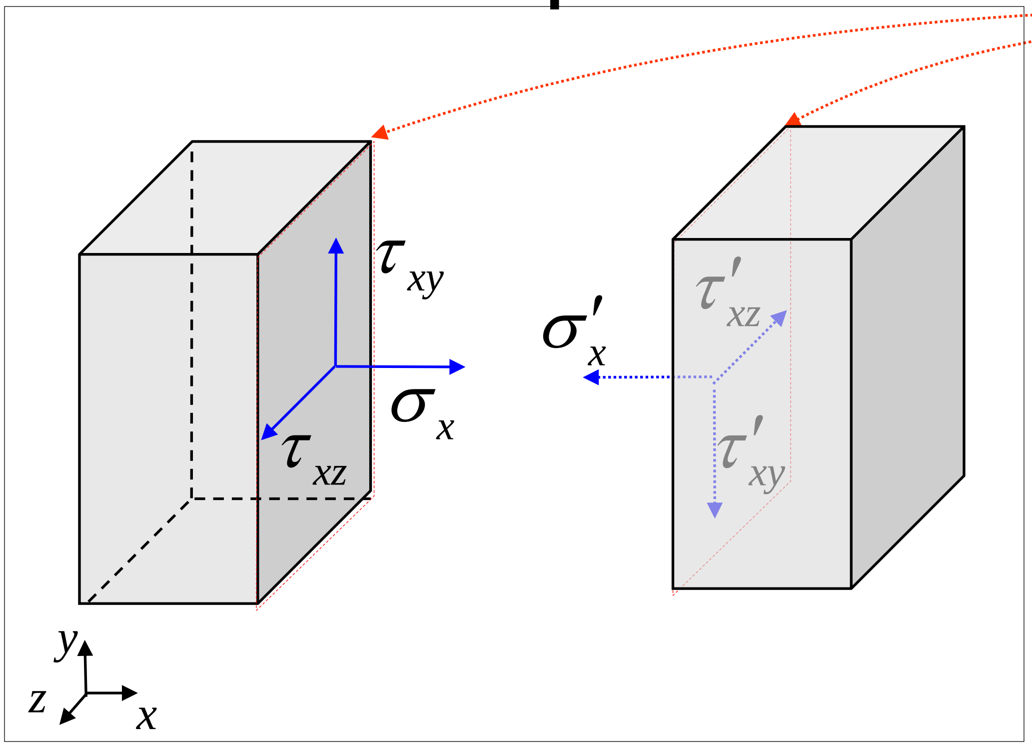
Stress Components

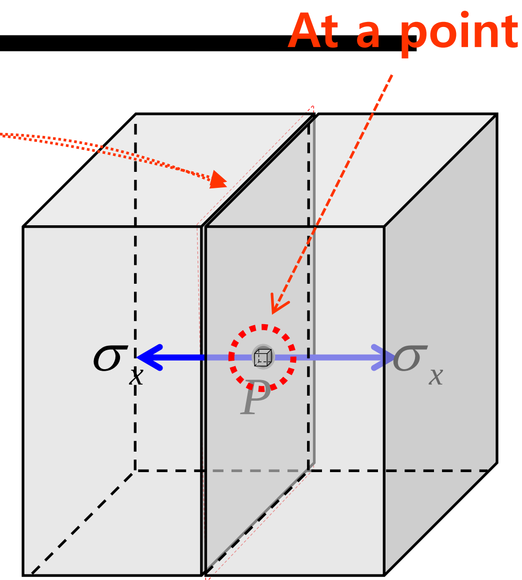
No. of Stress Components at a Point



Stress Components

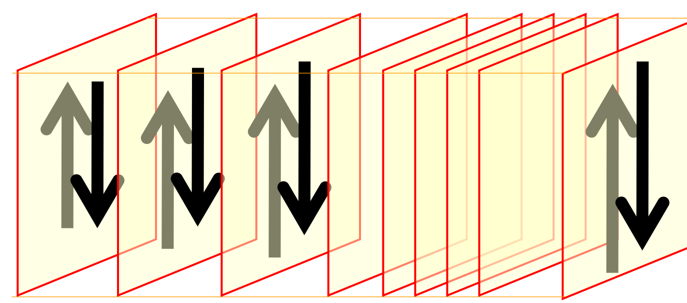
No. of Stress Components at a Point





$$\sigma'_x = \sigma_x,\ \tau'_{xy} = \tau_{xy},\ \tau'_{xz} = \tau_{xz}$$

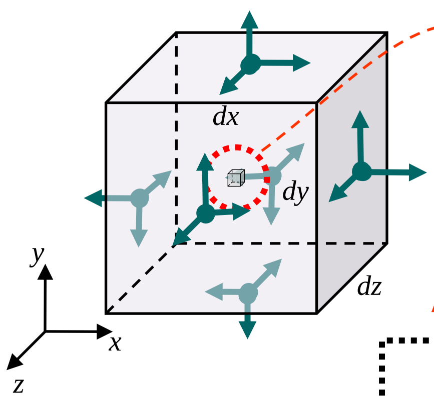
c.f) shear force



Stress Components

No. of Stress Components at a Point

element



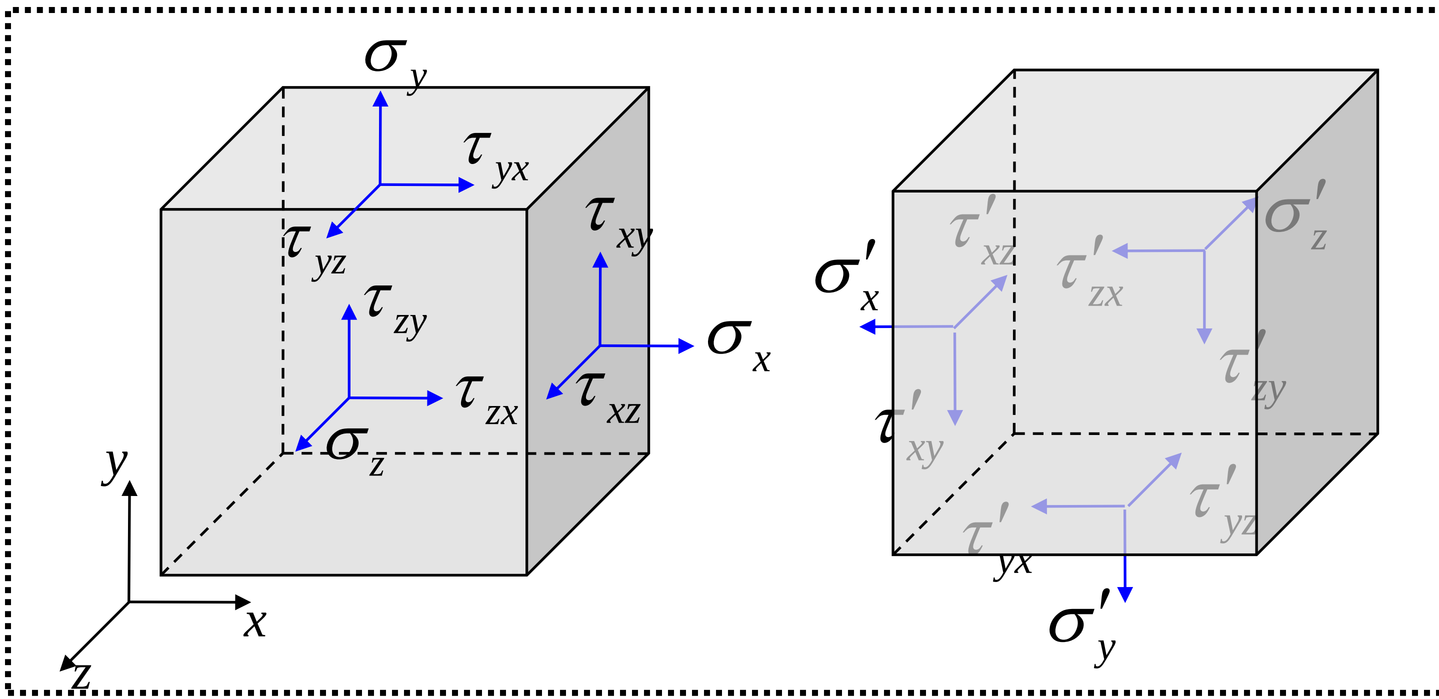
$$\sigma'_x = \sigma_x,\ \tau'_{xy} = \tau_{xy},\ \tau'_{xz} = \tau_{xz}$$

in same way

$$\sigma'_y = \sigma_y,\ \tau'_{yx} = \tau_{yx},\ \tau'_{yz} = \tau_{yz}$$

$$\sigma'_z = \sigma_z,\ \tau'_{zy} = \tau_{zy},\ \tau'_{zx} = \tau_{zx}$$

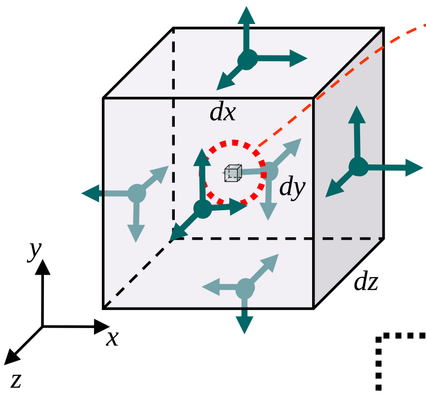
At a point



Stress Components

No. of Stress Components at a Point

element



At a point

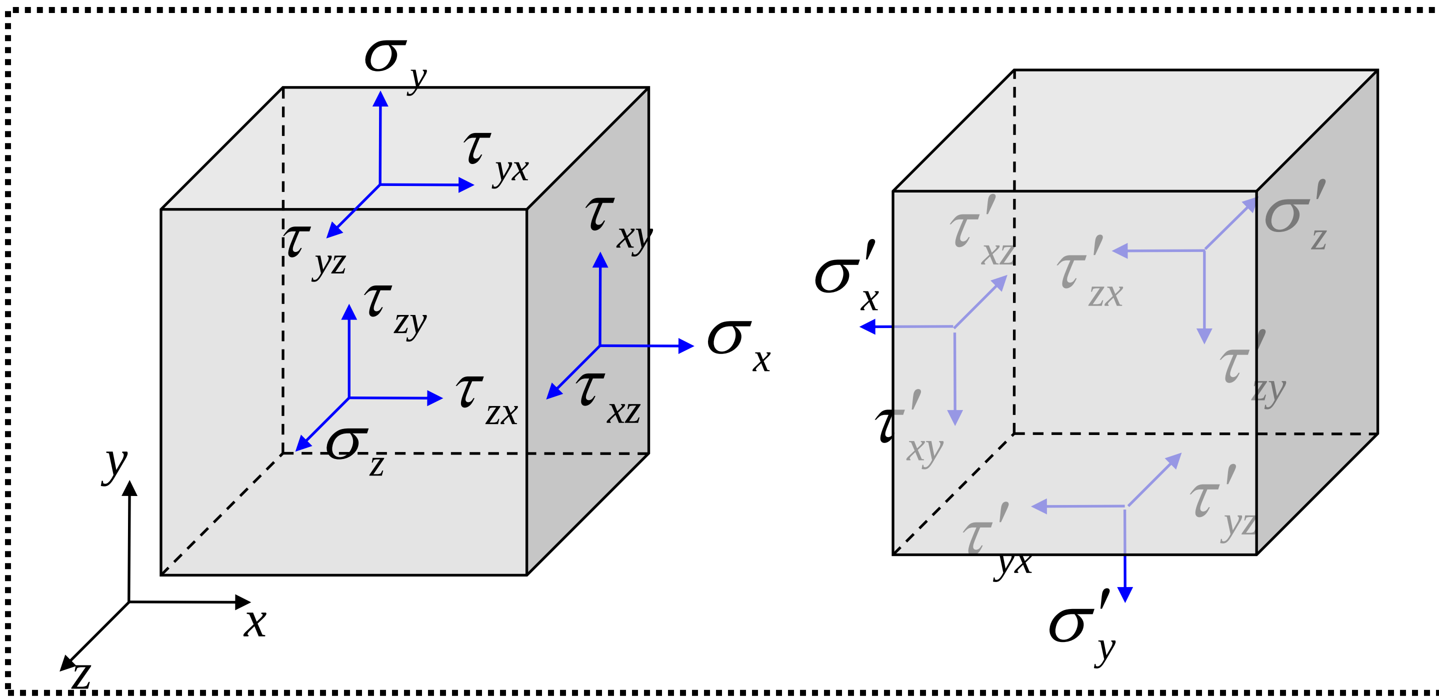
18 stress components at a point

$$\sigma'_x = \sigma_x,\ \tau'_{xy} = \tau_{xy},\ \tau'_{xz} = \tau_{xz}$$

$$\sigma'_y = \sigma_y,\ \tau'_{yx} = \tau_{yx},\ \tau'_{yz} = \tau_{yz}$$

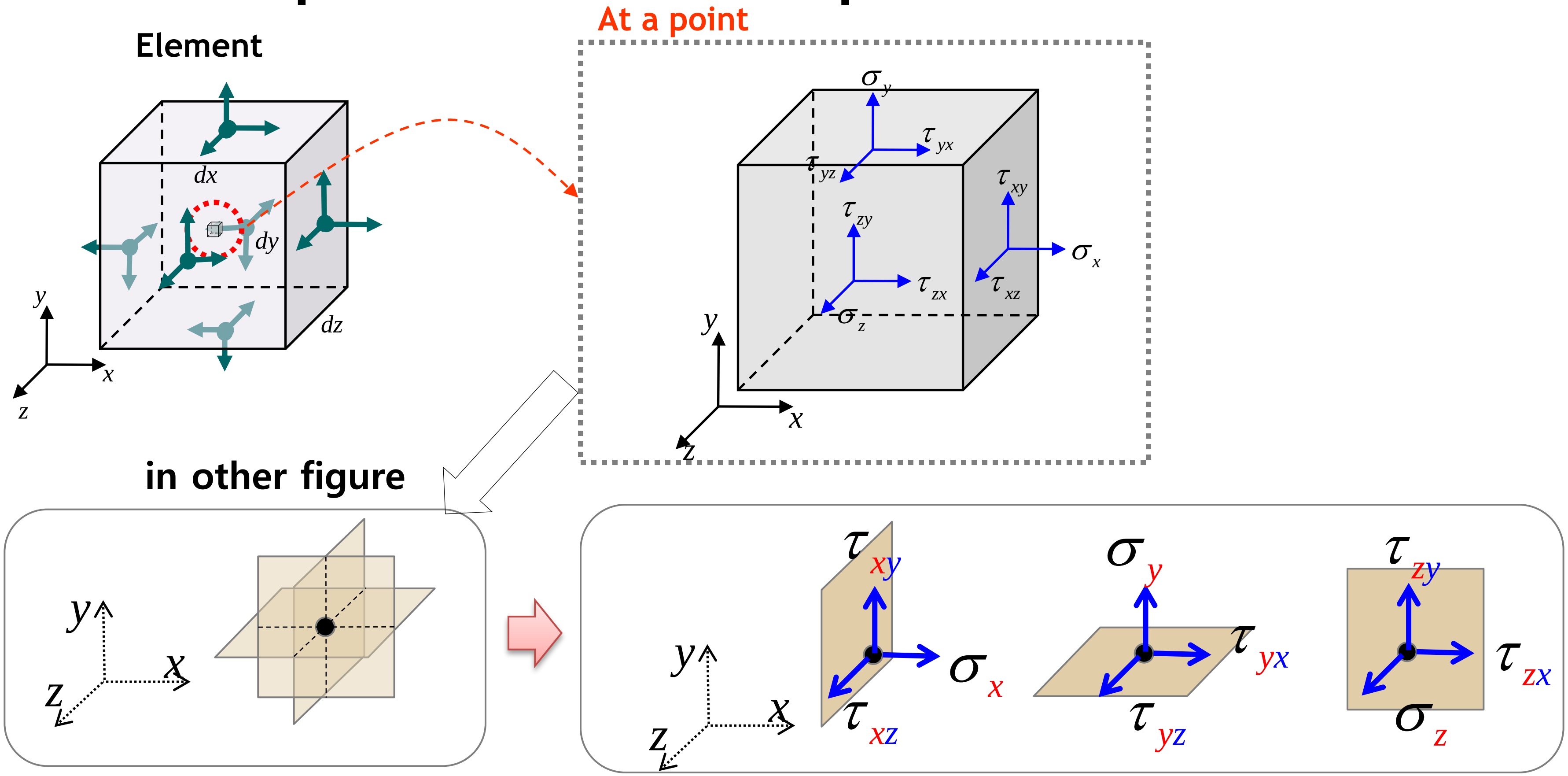
$$\sigma'_z = \sigma_z,\ \tau'_{zy} = \tau_{zy},\ \tau' x = \tau_{zx}$$

9 independent stress components



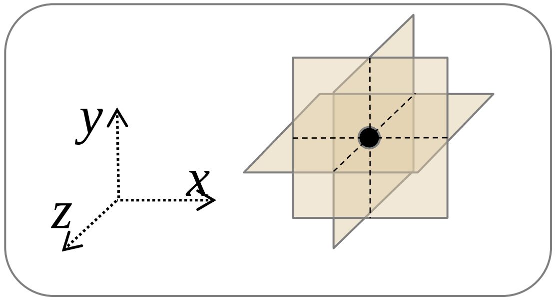
Stress Components

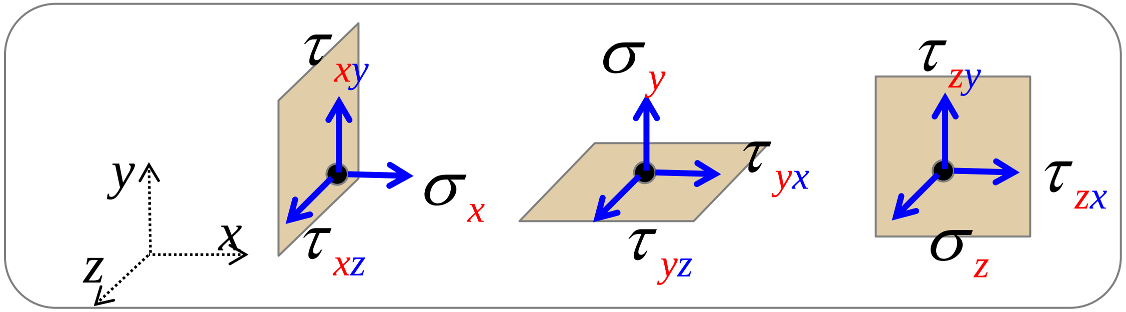
No. of Independent Stress Components at a Point



Through ***a point*** in a body we can construct ***three orthogonal coordinate planes*** on which we have ***9 independent stress components***

Stress Components





Through ***a point*** in a body we can construct ***three orthogonal coordinate planes*** on which we have ***9 independent stress components***

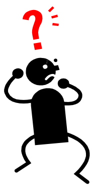
Are all the 9 stress components independent at a point?

Let us consider the moment equations for an element

Moment Equations for Element
- Stresses on the Surface of an Element

Moment Equations $\Leftarrow$ **Force** $\Leftarrow$ **stresses**

Stresses on the Surface of an Element



How we can describe the stresses on the surface of an element, if the stresses at the center of the element are known?

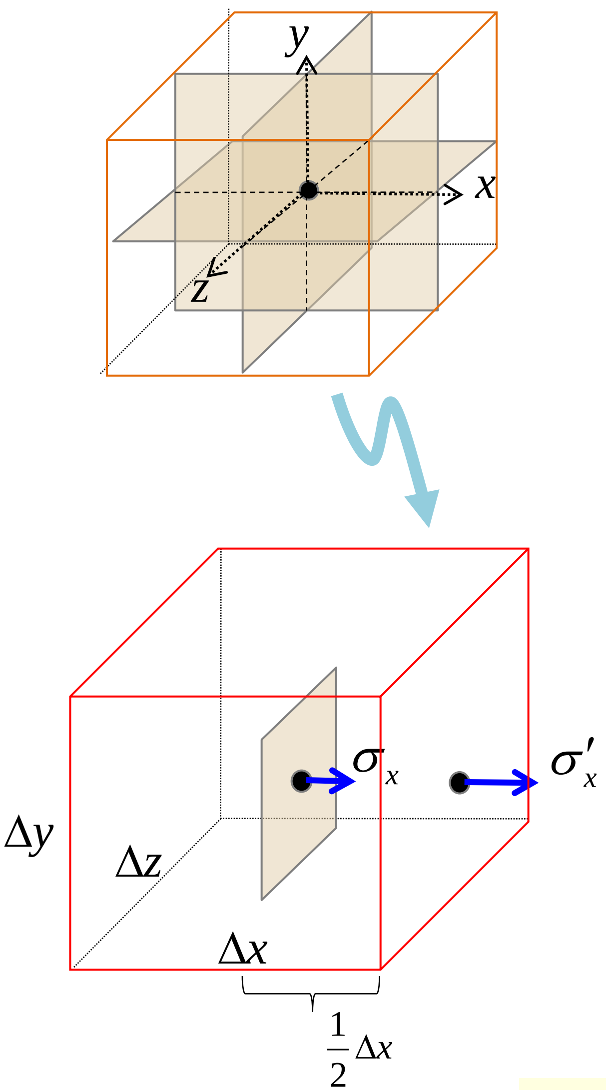
Taylor Series

$$f(x^* + \Delta x) = f(x^*) + f'(x^*)\Delta x + \frac{1}{2} f''(x^*)\Delta x^2 + \dots$$

for example, $f \rightarrow \sigma,\ \Delta x \rightarrow \frac{1}{2}\Delta x$ $\Downarrow$

$$\therefore \sigma'_x = \sigma_x + \frac{\partial \sigma_x}{\partial x}\left(\frac{1}{2}\Delta x\right) + \cancel{\frac{1}{2}\frac{\partial^2 \sigma_x}{\partial x^2}\left(\frac{1}{2}\Delta x\right)^2} + \dots$$

Linearization



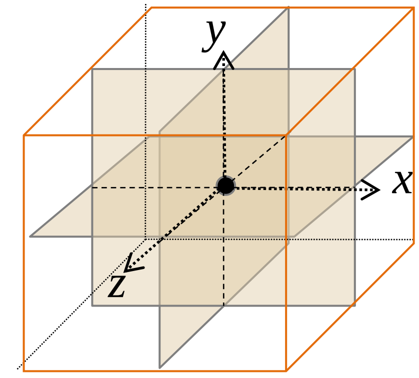
Stresses on the Surface of an Element

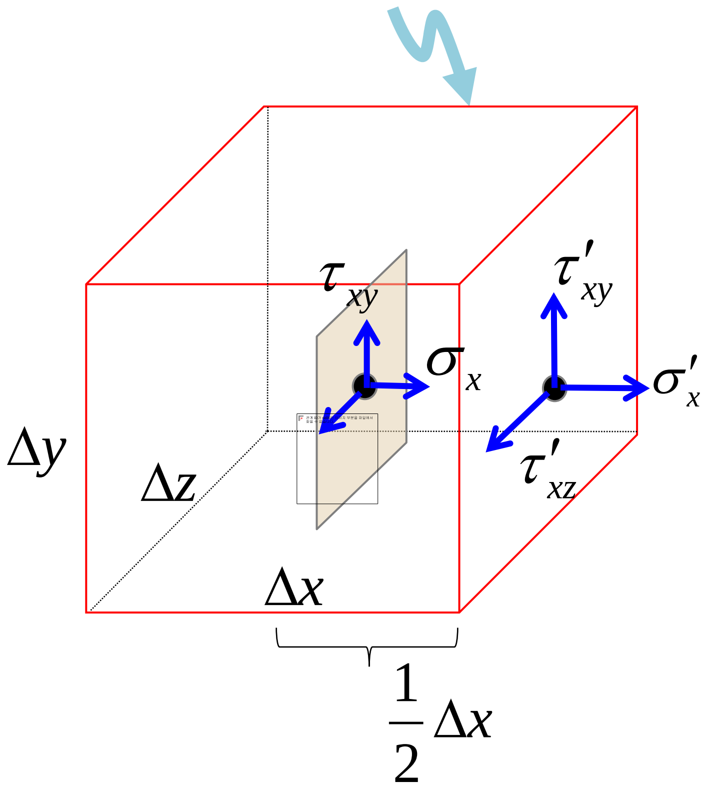
Stresses on the surface of an element

$$\sigma'_x = \sigma_x + \frac{\partial \sigma_x}{\partial x}\frac{1}{2}\Delta x$$

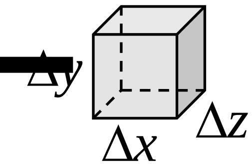
$$\tau'_{xy} = \tau_{xy} + \frac{\partial \tau_{xy}}{\partial x}\frac{1}{2}\Delta x$$

$$\tau'_{xz} = \tau_{xz} + \frac{\partial \tau_{xz}}{\partial x}\frac{1}{2}\Delta x$$

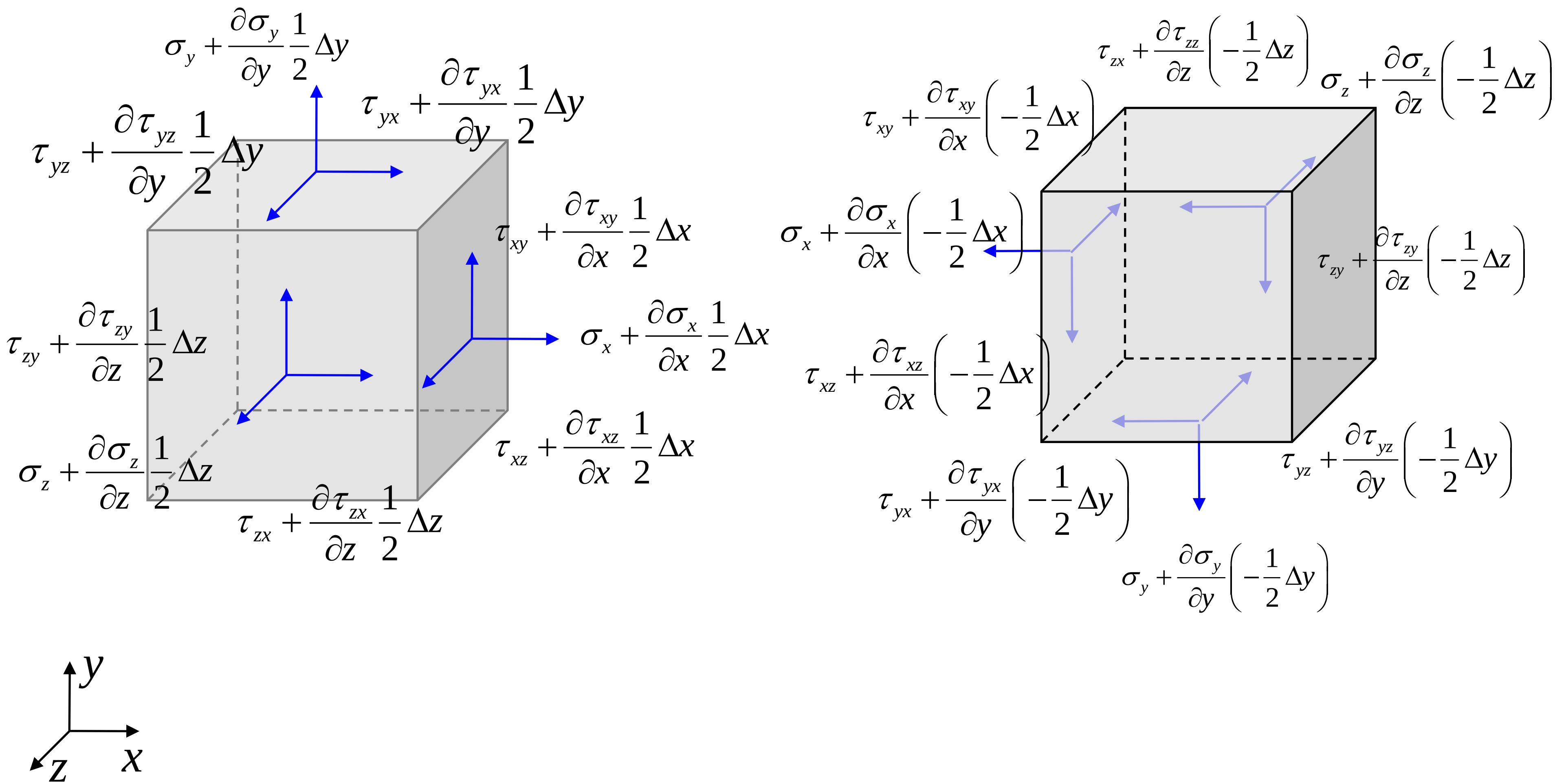




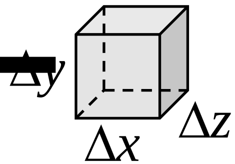
Stresses on the Surface of an Element



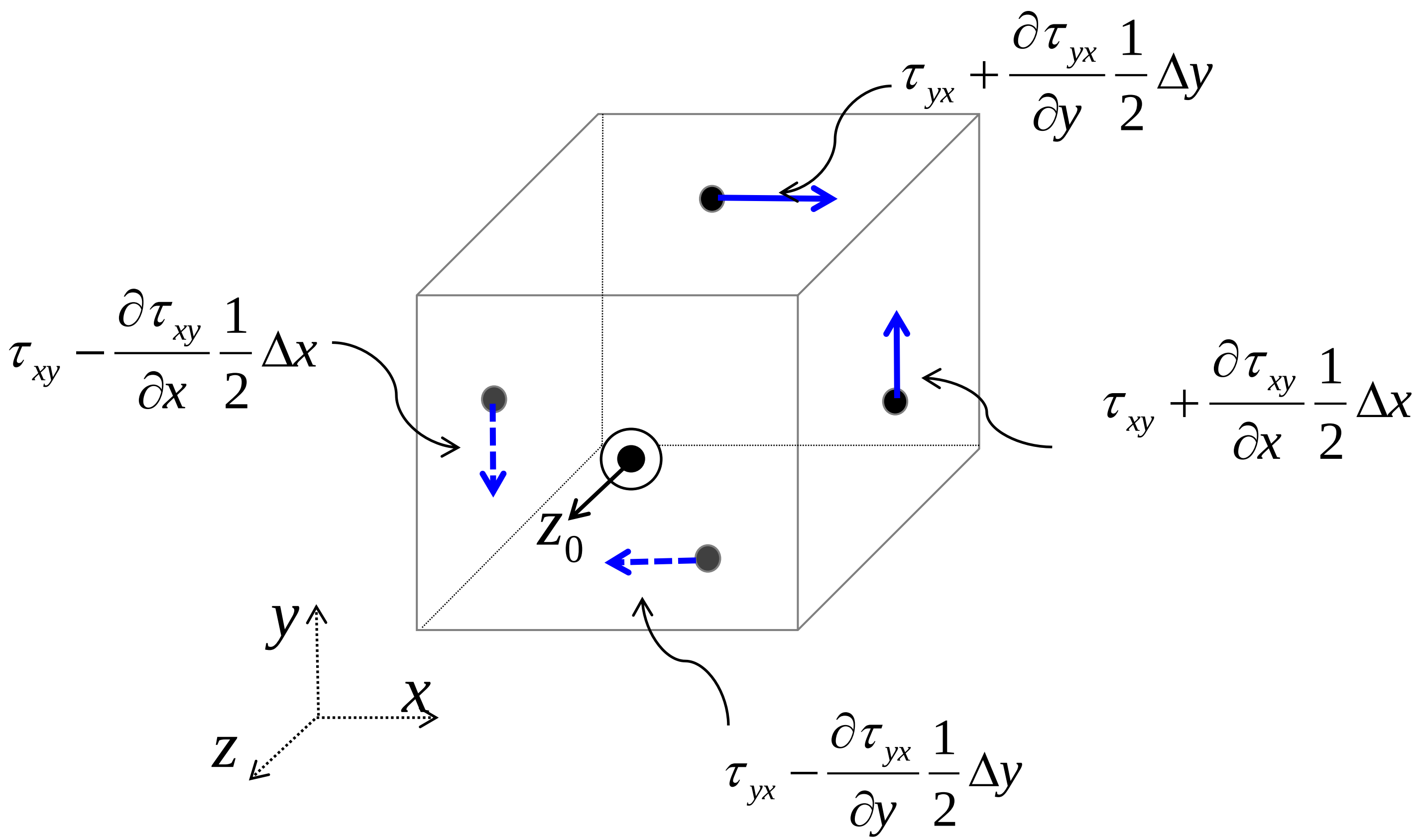
All stress components on the surface of an element



Stresses on the Surface of an Element



Stress components associated with the moment about z_0 axis



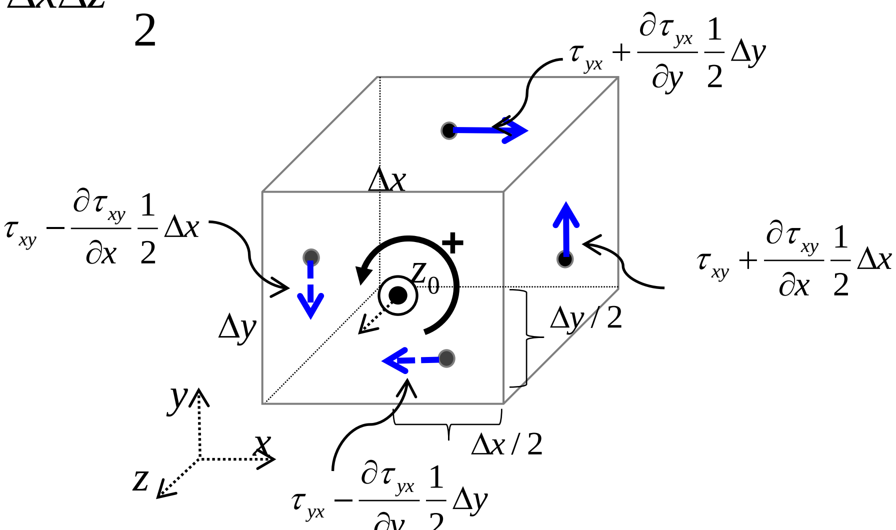
Moment Equations for Element

Moment about z_0 axis

$$M_{z_0} = \left(\tau_{xy} + \cancel{\frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}}\right)\Delta y \Delta z \frac{\Delta x}{2} - \left(\tau_{yx} + \cancel{\frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}}\right)\Delta x \Delta z \frac{\Delta y}{2}$$

$$+ \left(\tau_{xy} - \cancel{\frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}}\right)\Delta y \Delta z \frac{\Delta x}{2} - \left(\tau_{yx} - \cancel{\frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}}\right)\Delta x \Delta z \frac{\Delta y}{2}$$

$$= \left(\tau_{xy} - \tau_{yx}\right)\Delta x \Delta y \Delta z$$



for small volume

$$M_{z_0} = \left(\tau_{xy} - \tau_{yx}\right) dxdydz$$

Moment Equations for Element with Vector Notation

Moment about z_0 axis

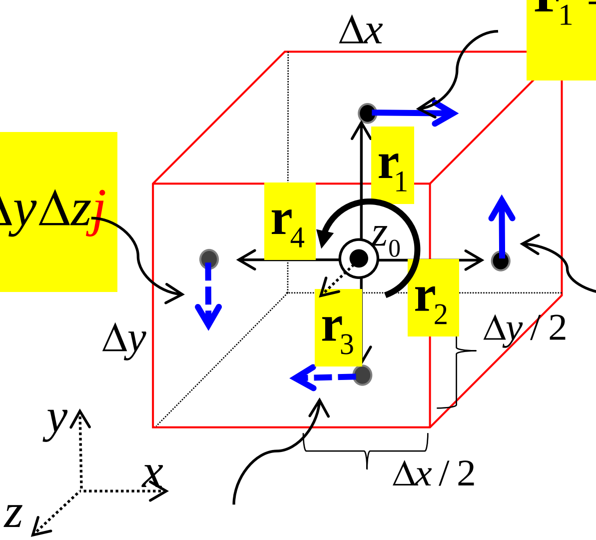
$$\mathbf{M}_{z_0} = \mathbf{r}_1 \times \mathbf{F}_1 + \mathbf{r}_2 \times \mathbf{F}_2 + \mathbf{r}_3 \times \mathbf{F}_3 + \mathbf{r}_4 \times \mathbf{F}_4$$

$$\mathbf{F}_1 = \left(\tau_{yx} + \frac{\partial \tau_{yx}}{\partial y} \frac{1}{2} \Delta y \right) \Delta x \Delta z \, \mathbf{i}$$

$$\mathbf{r}_1 = \frac{1}{2} \Delta y \, \mathbf{j}$$

$$\mathbf{F}_4 = -\left(\tau_{xy} - \frac{\partial \tau_{xy}}{\partial x} \frac{1}{2} \Delta x \right) \Delta y \Delta z \, j$$

$$\mathbf{r}_4 = -\frac{1}{2} \Delta x \, \mathbf{i}$$



$$\mathbf{F}_2 = \left(\tau_{xy} + \frac{\partial \tau_{xy}}{\partial x} \frac{1}{2} \Delta x \right) \Delta y \Delta z \, j$$

$$\mathbf{r}_2 = \frac{1}{2} \Delta x \, \mathbf{i}$$

$$\mathbf{F}_3 = -\left(\tau_{yx} - \frac{\partial \tau_{yx}}{\partial y} \frac{1}{2} \Delta y \right) \Delta x \Delta z \, \mathbf{i}$$

$$\mathbf{r}_3 = -\frac{1}{2} \Delta y \, \mathbf{j}$$

Moment Equations for Element : Vector Notation

$$\mathbf{M}_{z_0} = \mathbf{r}_1 \times \mathbf{F}_1 + \mathbf{r}_2 \times \mathbf{F}_2 + \mathbf{r}_3 \times \mathbf{F}_3 + \mathbf{r}_4 \times \mathbf{F}_4$$

$$= \left(\frac{\Delta y}{2}\mathbf{j}\right) \times \left(\left(\tau_{yx} + \frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}\right)\Delta x \Delta z\,\mathbf{i}\right) + \left(\frac{\Delta x}{2}\mathbf{i}\right) \times \left(\left(\tau_{xy} + \frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}\right)\Delta y \Delta z\,\mathbf{j}\right)$$

$$+ \left(-\frac{\Delta y}{2}\mathbf{j}\right) \times \left(-\left(\tau_{yx} - \frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}\right)\Delta x \Delta z\mathbf{i}\right) + \left(-\frac{\Delta x}{2}\mathbf{i}\right) \times \left(-\left(\tau_{xy} - \frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}\right)\Delta y \Delta z\mathbf{j}\right)$$

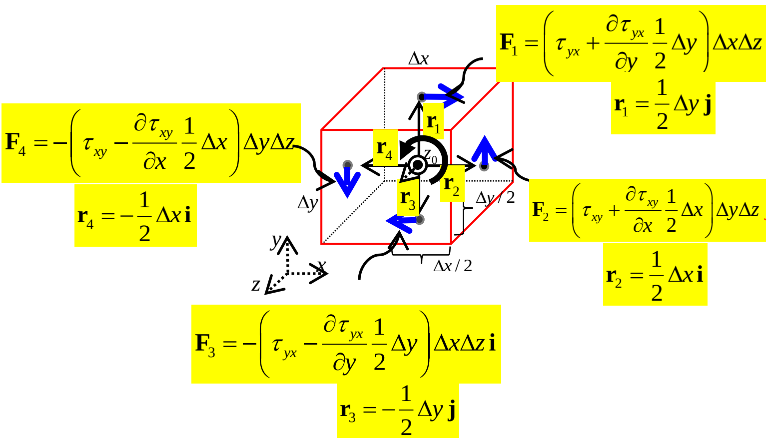
$$= \left(-\left(\tau_{yx} + \cancel{\frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}}\right)\frac{\Delta x \Delta y \Delta z}{2}\mathbf{k}\right) + \left(\left(\tau_{xy} + \cancel{\frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}}\right)\frac{\Delta x \Delta y \Delta z}{2}\mathbf{k}\right)$$

$$+ \left(-\left(\tau_{yx} - \cancel{\frac{\partial \tau_{yx}}{\partial y}\frac{\Delta y}{2}}\right)\frac{\Delta x \Delta y \Delta z}{2}\mathbf{k}\right) + \left(\left(\tau_{xy} - \cancel{\frac{\partial \tau_{xy}}{\partial x}\frac{\Delta x}{2}}\right)\frac{\Delta x \Delta y \Delta z}{2}\mathbf{k}\right)$$

$$= \left(\tau_{xy} - \tau_{yx}\right)\Delta x \Delta y \Delta z\,\mathbf{k}$$

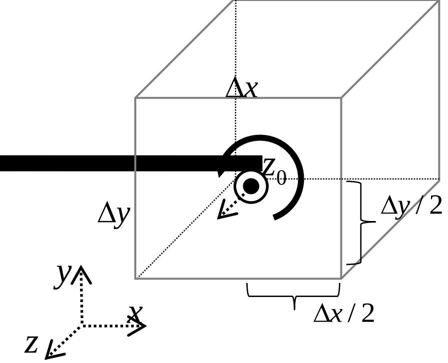
for small volume

$$\mathbf{M}_{z_0} = \left(\tau_{xy} - \tau_{yx}\right) dxdydz\,\mathbf{k}$$



, or $M_{z_0} = \left(\tau_{xy} - \tau_{yx}\right) dxdydz$

Moment Equations for Element



Rotational equilibrium of the element*

$$I_{z_0}\dot{\omega}=M_{z_0}$$

$$\frac{\rho dxdydz}{12}\left(dx^2+dy^2\right)\dot{\omega}=\left(\tau_{xy}-\tau_{yx}\right)dxdydz$$

$$\frac{\rho}{12}\left(dx^2+dy^2\right)\dot{\omega}=\left(\tau_{xy}-\tau_{yx}\right)$$

$$M_{z_0}=\left(\tau_{xy}-\tau_{yx}\right)dxdydz$$

※ **Mass moment of inertia**

$$I_{z_0}=\frac{m}{12}\left(dx^2+dy^2\right)$$

$$=\frac{\rho dxdydz}{12}\left(dx^2+dy^2\right)$$

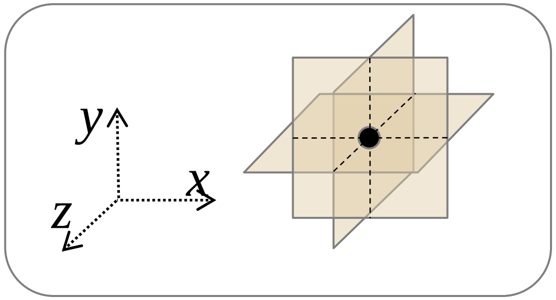
To the point of center , $dx\rightarrow 0,\ dy\rightarrow 0$

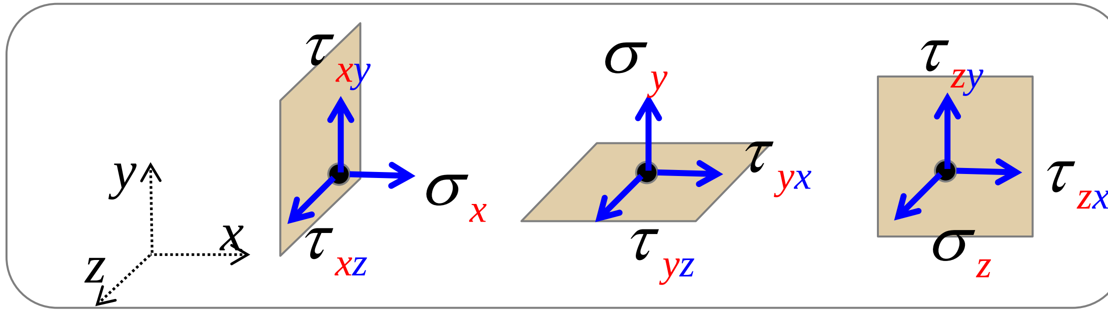
$$\frac{\rho}{12}\cancel{\left(dx^2+dy^2\right)}\dot{\omega}\quad=0$$

$\therefore\tau_{xy}=\tau_{yx}$ **in the same way,** $\left(\tau_{xz}=\tau_{zx},\tau_{yz}=\tau_{zy}\right)$

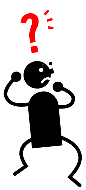
*Kundu,P.K., Fluid Mechanics, Academic Press, 2008, pp.88-93

Stress Components at a point





Through ***a point*** in a body we can construct ***three orthogonal coordinate planes*** on which we have ***9 independent stress components***



Are all the 9 stress components independent at a point?

Let us consider the moment equations for an element

$$\tau_{xy} = \tau_{yx}, \quad \tau_{xz} = \tau_{zx}, \quad \tau_{yz} = \tau_{zy}$$

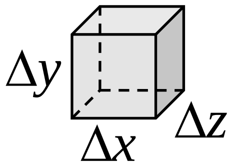
Six independent stress components

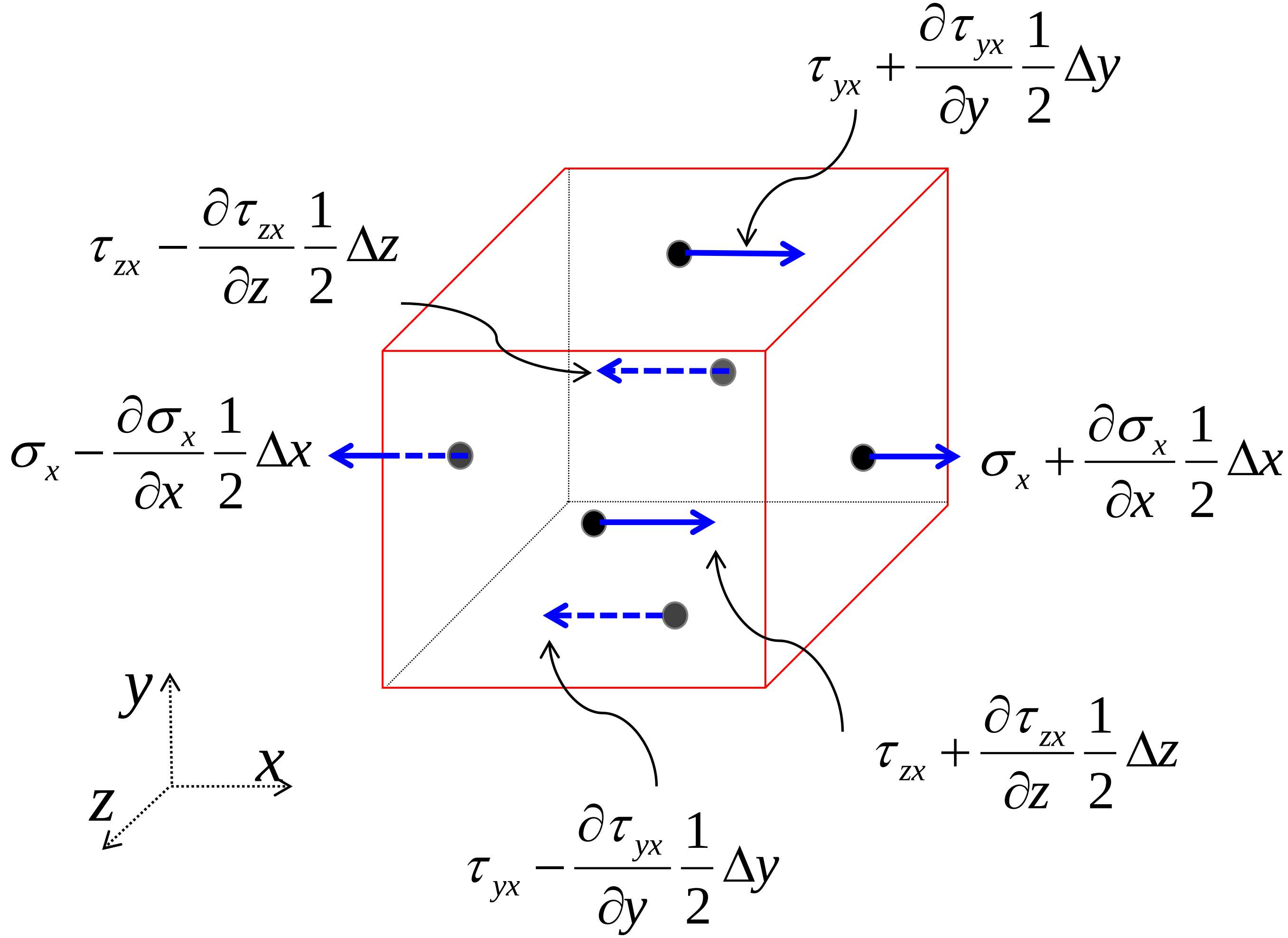
$$\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{xz}, \tau_{yz}$$

Stress Analysis
- Net surface in the *x*-direction

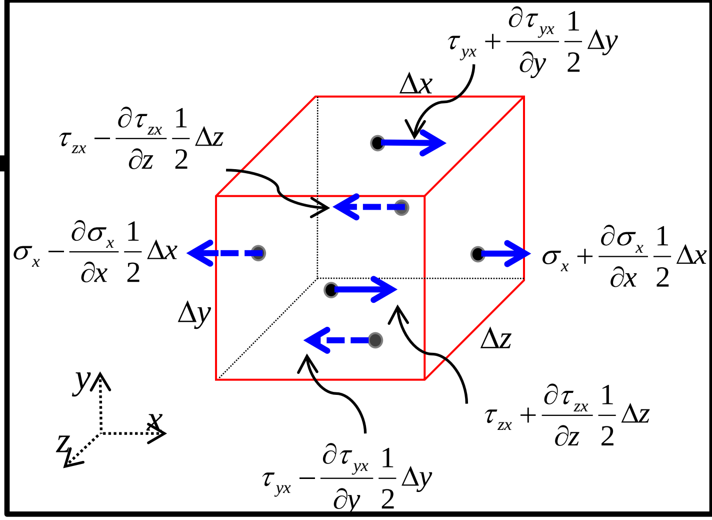
$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \boxed{\mathbf{F}_{Surface}}$$

Stress components associated with the forces in the *x*-direction





Stress Analysis
- Net surface on the element in the *x*-direction



Net surface force acting on the element in the *x*-direction :

$$stress \times area = force$$

$$\left[\sigma_x+\frac{\partial\sigma_x}{\partial x}\frac{1}{2}\Delta x-\left(\sigma_x-\frac{\partial\sigma_x}{\partial x}\frac{1}{2}\Delta x\right)\right]\Delta y\Delta z$$

$$+\left[\tau_{yx}+\frac{\partial\tau_{yx}}{\partial y}\frac{1}{2}\Delta y-\left(\tau_{yx}-\frac{\partial\tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\right]\Delta z\Delta x$$

$$+\left[\tau_{zx}+\frac{\partial\tau_{zx}}{\partial z}\frac{1}{2}\Delta z-\left(\tau_{zx}-\frac{\partial\tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\right]\Delta x\Delta y$$

$$=\left[\frac{\partial\sigma_x}{\partial x}+\frac{\partial\tau_{yx}}{\partial y}+\frac{\partial\tau_{zx}}{\partial z}\right]\Delta x\Delta y\Delta z$$

Stress Analysis with Vector Notation

$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \boxed{\mathbf{F}_{Surface}}$$

Stress components acting on the element in the *x*-direction

$$\mathbf{F}_1 = \left(\tau_{yx} + \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\Delta x \Delta z\,\mathbf{i}$$

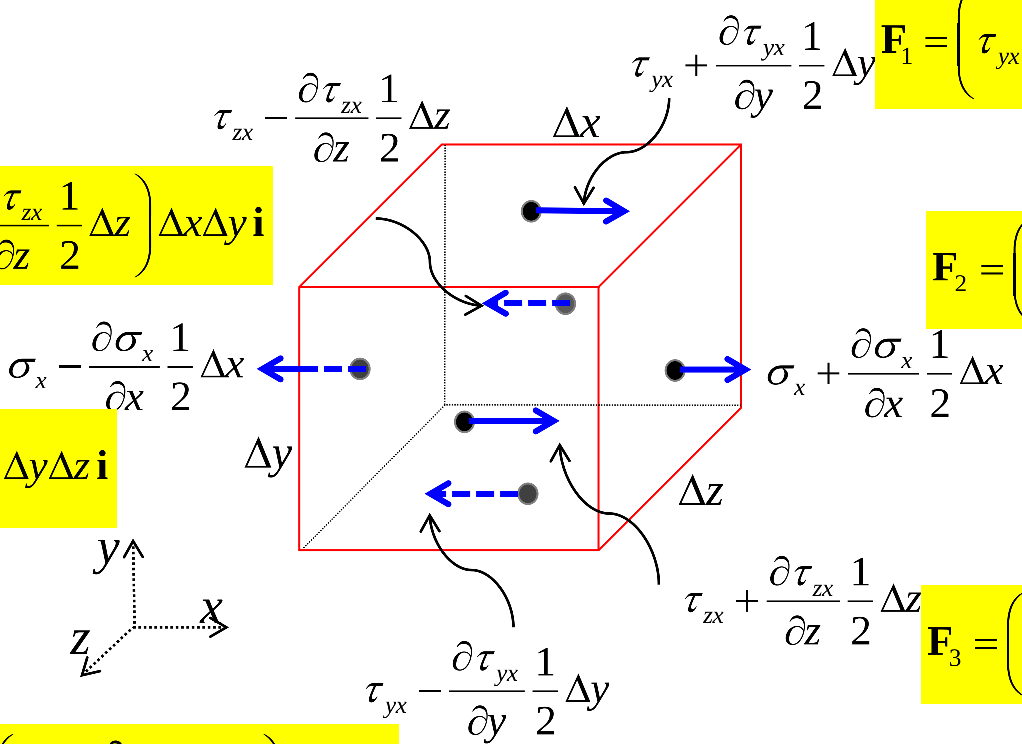
$$\mathbf{F}_2 = \left(\sigma_x + \frac{\partial \sigma_x}{\partial x}\frac{1}{2}\Delta x\right)\Delta y \Delta z\,\mathbf{i}$$

$$\mathbf{F}_3 = \left(\tau_{zx} + \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\Delta x \Delta y\,\mathbf{i}$$

$$\mathbf{F}_4 = -\left(\tau_{zx} - \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\Delta x \Delta y\,\mathbf{i}$$

$$\mathbf{F}_5 = -\left(\sigma_x - \frac{\partial \sigma_x}{\partial x}\frac{1}{2}\Delta x\right)\Delta y \Delta z\,\mathbf{i}$$

$$\mathbf{F}_6 = -\left(\tau_{yx} - \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\Delta x \Delta z\,\mathbf{i}$$



Stress Analysis

Net surface force acting on the element in the *x*-direction :

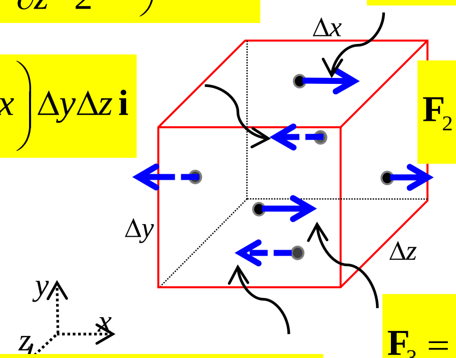
$$\sum \mathbf{F}_{Surface,x} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 + \mathbf{F}_4 + \mathbf{F}_5 + \mathbf{F}_6$$

$$\mathbf{F}_4 = -\left(\tau_{zx} - \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\Delta x \Delta y\,\mathbf{i}$$

$$\mathbf{F}_1 = \left(\tau_{yx} + \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\Delta x \Delta z\,\mathbf{i}$$

$$\mathbf{F}_5 = -\left(\sigma_{x} - \frac{\partial \sigma_{x}}{\partial x}\frac{1}{2}\Delta x\right)\Delta y \Delta z\,\mathbf{i}$$

$$\mathbf{F}_2 = \left(\sigma_{x} + \frac{\partial \sigma_{x}}{\partial x}\frac{1}{2}\Delta x\right)\Delta y \Delta z\,\mathbf{i}$$



$$\mathbf{F}_3 = \left(\tau_{zx} + \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\Delta x \Delta y\,\mathbf{i}$$

$$\mathbf{F}_6 = -\left(\tau_{yx} - \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\Delta x \Delta z\,\mathbf{i}$$

$$stress \times area = force$$

$$\sum \mathbf{F}_{Surface,x} = \left[\tau_{zx} + \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z - \left(\tau_{zx} - \frac{\partial \tau_{zx}}{\partial z}\frac{1}{2}\Delta z\right)\right]\Delta x \Delta y\,\mathbf{i}$$

$$+\left[\tau_{yx} + \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y - \left(\tau_{yx} - \frac{\partial \tau_{yx}}{\partial y}\frac{1}{2}\Delta y\right)\right]\Delta z \Delta x\,\mathbf{i}$$

$$+\left[\sigma_{x} + \frac{\partial \sigma_{x}}{\partial x}\frac{1}{2}\Delta x - \left(\sigma_{x} - \frac{\partial \sigma_{x}}{\partial x}\frac{1}{2}\Delta x\right)\right]\Delta y \Delta z\,\mathbf{i}$$

$$= \left[\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}\right]\Delta x \Delta y \Delta z\,\mathbf{i}$$

Stress Analysis

$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \boxed{\mathbf{F}_{Surface}}$$

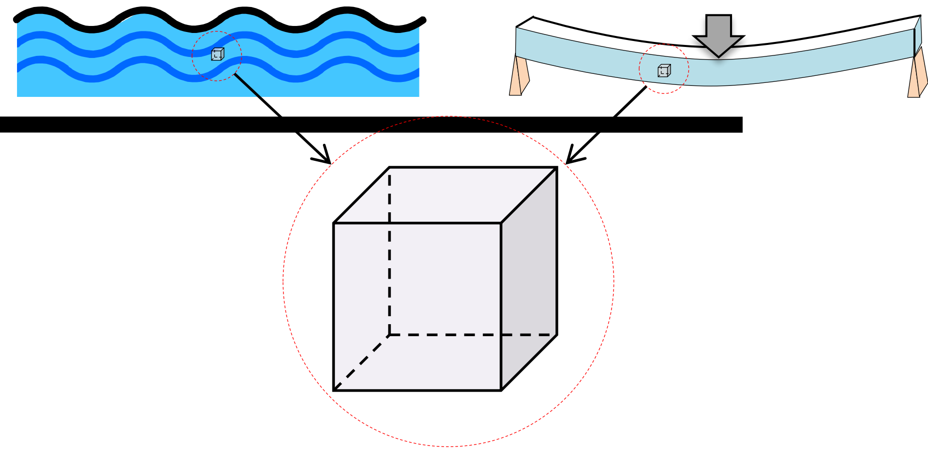
Net surface force acting on the element in the *x,y,* and *z* direction for unit volume

$$F_{Surface,x} = \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$F_{Surface,y} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

$$F_{Surface,z} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

Equations of Motions



$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \mathbf{F}_{Surface}$$

surface force of the element

body force of the element

acceleration of the element

mass of the element

per unit volume

$$F_{Surface,x} = \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$F_{Surface,y} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

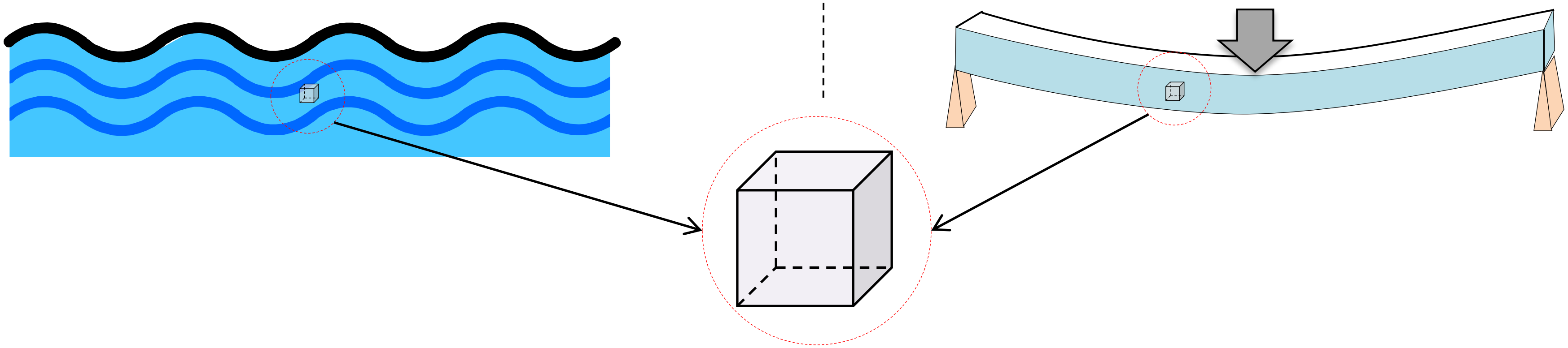
$$F_{Surface,z} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

Surface Force

$$m\frac{d\mathbf{V}}{dt} = \mathbf{F}_{Body} + \mathbf{F}_{Surface}$$

Fluid Mechanics

Elasticity



Fluid Mechanics

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

Elasticity

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

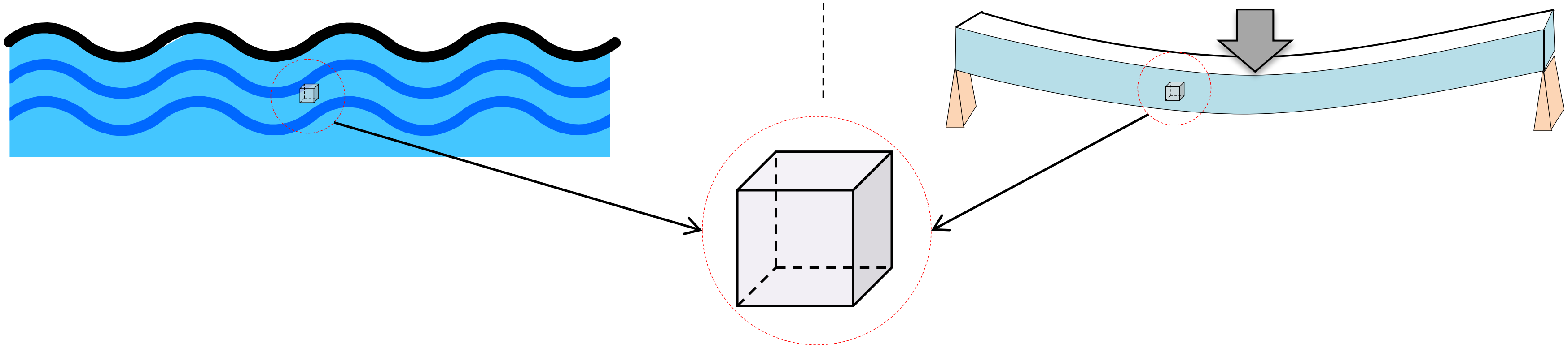
$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

Body Force, Surface Force

$$m\frac{d\mathbf{V}}{dt} = \boxed{\mathbf{F}_{Body} + \mathbf{F}_{Surface}}$$

Fluid Mechanics

Elasticity



$$\rho g_x + \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$\rho g_y + \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

$$\rho g_z + \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

$$X + \frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z}$$

$$Y + \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{zy}}{\partial z}$$

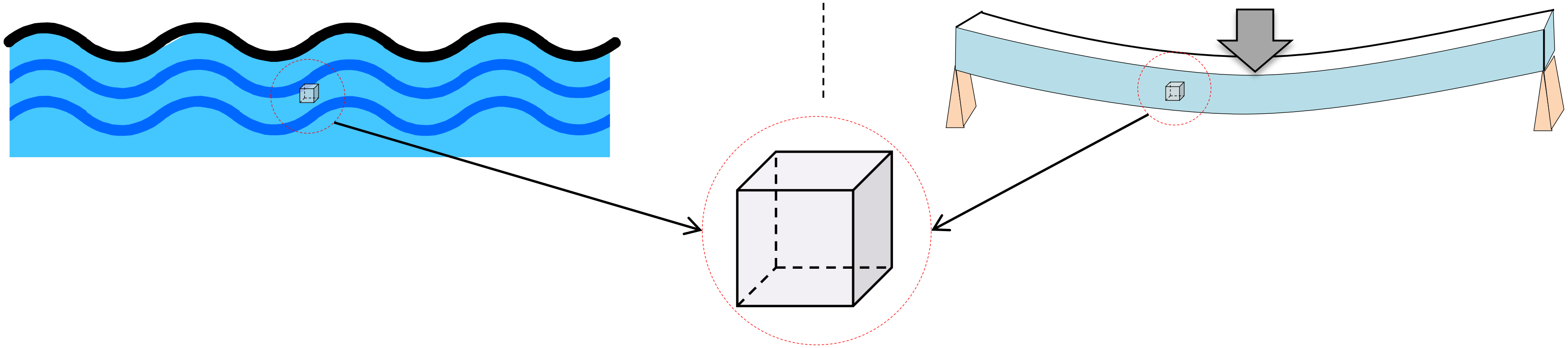
$$Z + \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_z}{\partial z}$$

Acceleration

$$m\boxed{\frac{d\mathbf{V}}{dt}} = \mathbf{F}_{Body} + \mathbf{F}_{Surface}$$

Fluid Mechanics

Elasticity



$$\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right)$$

$$\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right)$$

$$\left(\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right)$$

$$\frac{d^2 u}{dt^2}$$

$$\frac{d^2 v}{dt^2}$$

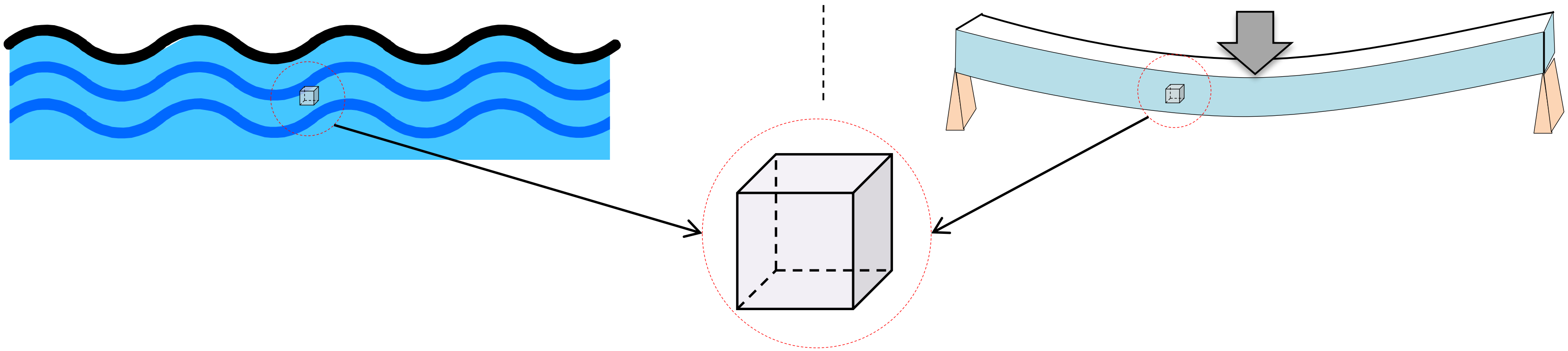
$$\frac{d^2 w}{dt^2}$$

Mass , Acceleration

$$m\frac{d\mathbf{V}}{dt}=\mathbf{F}_{Body}+\mathbf{F}_{Surface}$$

Fluid Mechanics

Elasticity



$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)$$

$$\rho\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}\right)$$

$$\rho\left(\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}\right)$$

$$\rho\frac{d^2u}{dt^2}$$

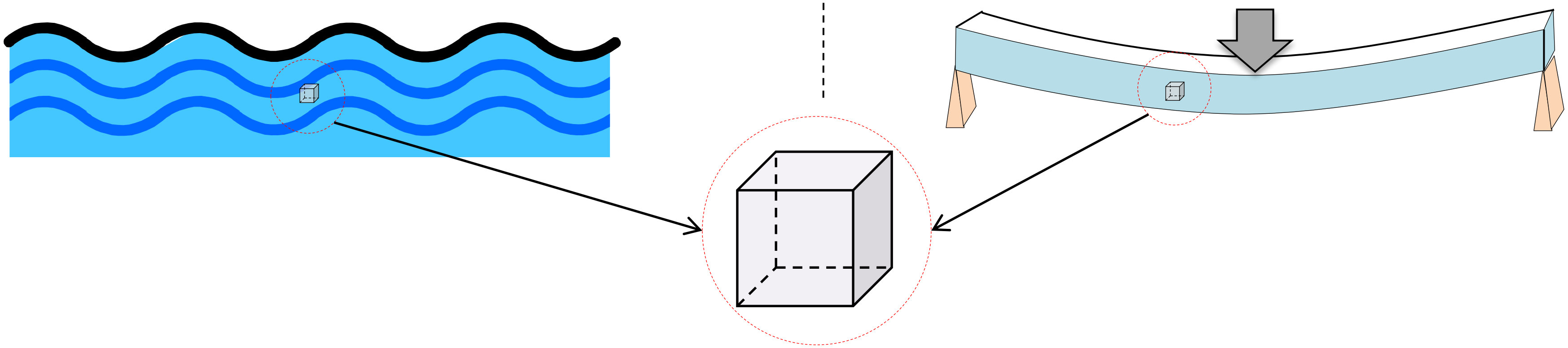
$$\rho\frac{d^2v}{dt^2}$$

$$\rho\frac{d^2w}{dt^2}$$

Equations of Motions

$$m\frac{d\mathbf{V}}{dt}=\mathbf{F}_{Body}+\mathbf{F}_{Surface}$$

Fluid Mechanics | **Elasticity**



Fluid Mechanics

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)=\rho g_x+\frac{\partial\sigma_x}{\partial x}+\frac{\partial\tau_{yx}}{\partial y}+\frac{\partial\tau_{zx}}{\partial z}$$

$$\rho\left(\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}\right)=\rho g_y+\frac{\partial\tau_{xy}}{\partial x}+\frac{\partial\sigma_y}{\partial y}+\frac{\partial\tau_{zy}}{\partial z}$$

$$\rho\left(\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}\right)=\rho g_z+\frac{\partial\tau_{xz}}{\partial x}+\frac{\partial\tau_{yz}}{\partial y}+\frac{\partial\sigma_z}{\partial z}$$

⇨ **Cauchy Equation**

Elasticity

$$\rho\frac{d^2u}{dt^2}=X+\frac{\partial\sigma_x}{\partial x}+\frac{\partial\tau_{yx}}{\partial y}+\frac{\partial\tau_{zx}}{\partial z}$$

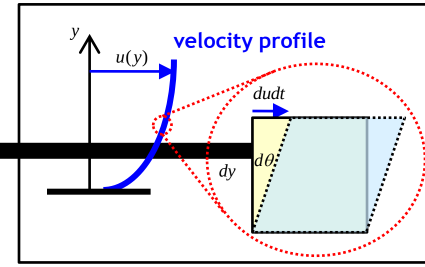
$$\rho\frac{d^2v}{dt^2}=Y+\frac{\partial\tau_{xy}}{\partial x}+\frac{\partial\sigma_y}{\partial y}+\frac{\partial\tau_{zy}}{\partial z}$$

$$\rho\frac{d^2w}{dt^2}=Z+\frac{\partial\tau_{xz}}{\partial x}+\frac{\partial\tau_{yz}}{\partial y}+\frac{\partial\sigma_z}{\partial z}$$

⇨ **Static Equilibrium :** $\frac{d^2u}{dt^2}=\frac{d^2v}{dt^2}=\frac{d^2w}{dt^2}=0$

Equations of Motions

Dynamics and Statics



Equations of Motion

$$\boxed{m\frac{d\mathbf{V}}{dt}} = \sum \mathbf{F} = \mathbf{F}_{Body} + \boxed{\mathbf{F}_{Surface}}$$

→ **dynamics** $\sum \mathbf{F} \rightarrow \text{cause motion} \rightarrow m\frac{d\mathbf{V}}{dt}$

→ **statics** $\sum \mathbf{F} \rightarrow \text{without motion} \rightarrow \text{'internal change'} \rightarrow \text{strain}$

Fluid Equations of Motion
Navier Stokes Equations

Total Derivative of Velocity Vector

$$\boxed{\rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right)} = \rho g_x - \boxed{\frac{\partial P}{\partial x} + \frac{\mu}{3}\frac{\partial}{\partial x}(\Theta) + \mu\nabla^2 u}$$

$\mathbf{V}$: velocity vector

$\mathbf{V} = (u(x,y,z;t), v(x,y,z;t), w(x,y,z;t))$

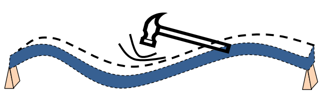
μ : viscosity , $\left[\frac{\mathrm{N}\cdot\mathrm{s}}{m^2}\right]$

$\Theta = \nabla \bullet \mathbf{V} = \frac{du}{dx} + \frac{dv}{dy} + \frac{dw}{dz}$

Newtonian fluid

$\tau \propto \frac{du}{dy}, \ \tau_{xy} = \mu\left(\frac{du}{dy} + \frac{dv}{dx}\right)$

Linear Elastic Equations of Motion*
(also called *the Navier Equations* : principal Equations for the motion)



$$\boxed{\rho\frac{\partial^2 \overline{u}}{\partial t^2}} = \rho F + \boxed{(\lambda+\mu)\frac{\partial^2 \overline{u}}{\partial x^2} + \mu\nabla^2 \overline{u}}$$

$\overline{u} = \overline{u}(x,t)$: displacement

ν : Poisson's Ratio
G : Shear Modulus
E : Young's Modulus

μ, λ : Lame Elastic constant

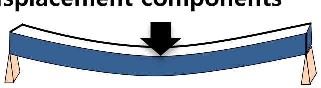
$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, \left[\frac{N}{m^2}\right]$

$\mu = G, \left[\frac{N}{m^2}\right]$

$\tau_{xy} = \mu\gamma_{xy} = \mu\left(\frac{\partial \overline{v}}{\partial x} + \frac{\partial \overline{u}}{\partial y}\right)$

$\varepsilon = \frac{\delta L}{L}, \ \varepsilon_x = \frac{\partial \overline{u}}{\partial x}$

Formulation of Elasticity Problems (Static) in displacement components



$$\boxed{0} = X + \boxed{(\lambda+G)\frac{\partial^2 \overline{u}}{\partial x^2} + G\nabla^2 \overline{u}}$$

$\overline{u} = \overline{u}(x)$: displacement

*Betounes D., Partial Differential Equations for Computational Science, Springer, 1998, p343

2. Grillage Analysis for Midship Structure

Seoul National Univ.
@SDAL
Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

Summary

u : Axial Displacement G : Shear Modulus A : Sectional Area G : Shear Modulus E : Young's Modulus

w : Vertical Displacement θ : Angle of Twist l : Length J : Polar Moment of Inertia I : Moment of Inertia

Element | Differential Equation → Weighted Residual → •Discretization •Approximation → Finite Element Method | •Superposition of Stiffness Matrix •Coordinates Transformation **Behavior → Structure**

$$\mathbf{M}\ddot{\mathbf{x}} = \sum \mathbf{F} \text{ , where } \ddot{\mathbf{x}} = 0$$

$$\mathbf{Kd} = \mathbf{F}$$

Bar (P, P, l)

$$EA\frac{d^2u(x)}{dx^2} = 0 \Rightarrow \int_0^L W_l\left(EA\frac{d^2\hat{u}(x)}{dx^2}\right)dx = 0 \Rightarrow \frac{EA}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -P(0) \\ P(l) \end{bmatrix}$$

$$\hat{u}(x) = a_0 + a_1 x$$

Tension → **Truss**

Beam (m(0), m(l), V(0), V(l), l)

$$EI\frac{d^4y(x)}{dx^4} = 0 \Rightarrow \int_0^L W_l\left(EI\frac{d^4\hat{y}(x)}{dx^4}\right)dx = 0 \Rightarrow \frac{2EI}{l^3}\begin{bmatrix} 6 & 3l & -6 & 3l \\ 3l & 2l^2 & -3l & l^2 \\ -6 & -3l & 6 & -3l \\ 3l & l^2 & -3l & 2l^2 \end{bmatrix}\begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix} = \begin{bmatrix} V(0) \\ -m(0) \\ -V(l) \\ m(l) \end{bmatrix}$$

$$\hat{y}(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$

Bending → **Frame**

Shaft (T, T, l)

$$GJ\frac{d^2\theta(x)}{dx^2} = 0 \Rightarrow \int_0^L W_l\left(GJ\frac{d^2\hat{\theta}(x)}{dx^2}\right)dx = 0 \Rightarrow \frac{GJ}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} -T(0) \\ T(l) \end{bmatrix}$$

$$\hat{\theta}(x) = c_0 + c_1 x$$

Torsion → **Grillage**

Application

Midship Cargo Hold → (Engineering Concept !) → **Grillage Modeling**
- 3D→ 2D
- Equivalent Force & Moment
- Boundary condition

→ **Solution**
- programming
- visulaization

Beam Theory : Sign Convention, Deflection of Beam

Elasticity : Displacement, Strain, Stress, Force Equilibrium, Compatibility, Constitutive Equation

2.1. Element : Bar

- Derivation of the Stiffness Matrix by Applying direct equilibrium approach

Element : Bar (1 element , 2 nodes)
- Direct equilibrium approach

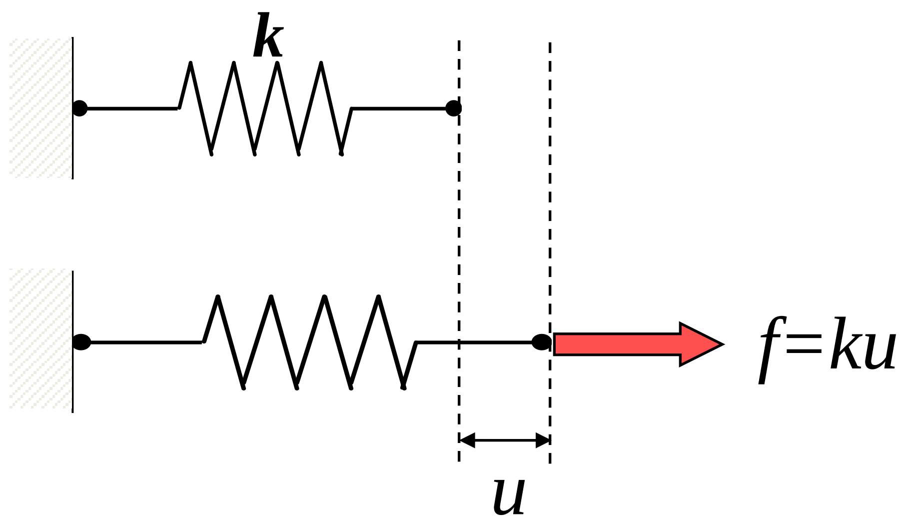
- Hooke's Law

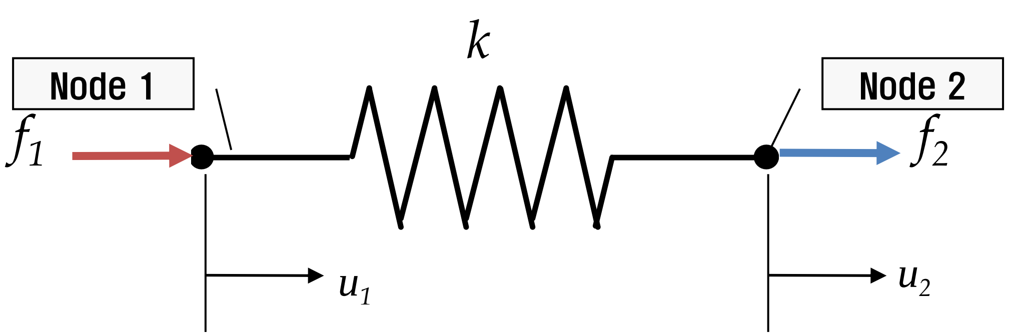
$$f = ku$$

f : External force

k : Stiffness coefficient

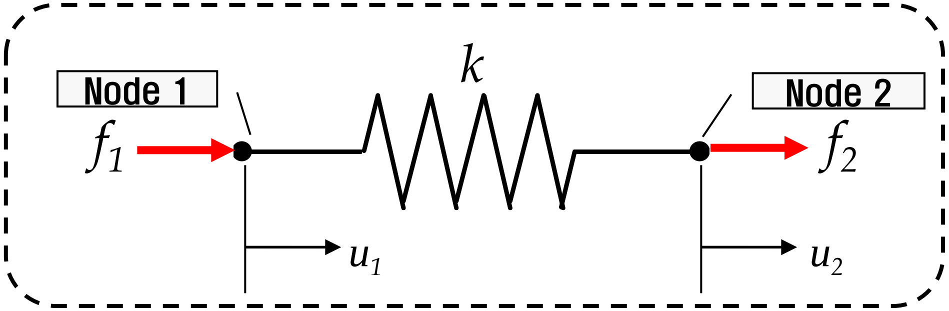
u : Displacement



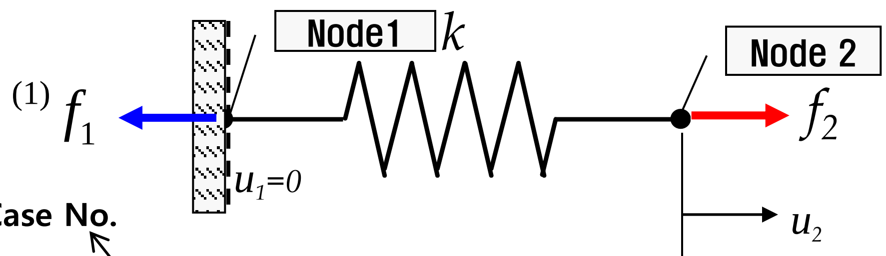


Element : Bar (1 element , 2 nodes)
- Direct equilibrium approach

f_i : External force



① Case #1: the node 1 is fixed (u_1=0)

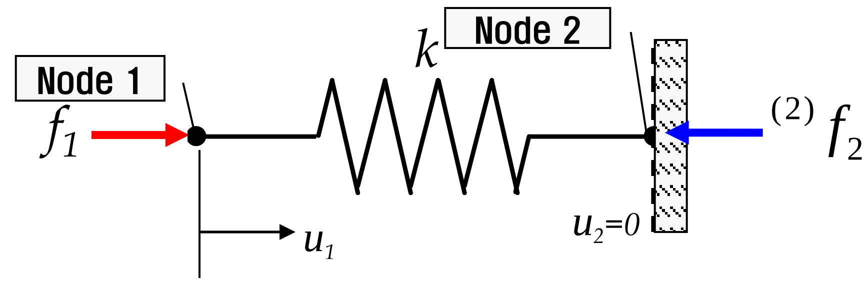


Case No.

$${}^{(1)}f_1 = -ku_2 \qquad {}^{(1)}f_2 = ku_2$$

Node No.

② Case #2: the node 2 is fixed (u_2=0)



$${}^{(2)}f_1 = ku_1 \qquad {}^{(2)}f_2 = -ku_1$$

③ Actual Case: the nodes are not fixed

Stiffness Matrix

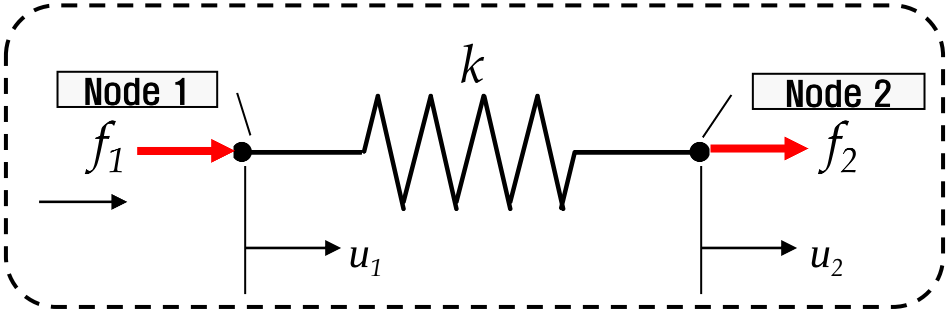
$$f_1 = {}^{(1)}f_1 + {}^{(2)}f_1 = ku_1 - ku_2 = k(u_1 - u_2)$$

$$f_2 = {}^{(1)}f_2 + {}^{(2)}f_2 = -ku_1 + ku_2 = k(u_2 - u_1)$$

Matrix Form

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \boxed{\begin{bmatrix} k & -k \\ -k & k \end{bmatrix}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \longrightarrow [f] = [K][u]$$

Element : Bar (1 element , 2 nodes)
- Direct equilibrium approach



Stiffness Matrix

$$\begin{bmatrix} f_1 \\ \boxed{f_2} \end{bmatrix} = \begin{bmatrix} k & -k \\ -k & k \end{bmatrix} \begin{bmatrix} \overset{=0}{\boxed{u_1}} \\ u_2 \end{bmatrix}$$

This stiffness matrix is Singular, i.e., its inverse does not exist and it cannot be solved!

It means that the structure has not been secured to the ground. As the system stands, no limitation has been placed on any of the displacements u_1 and u_2. Therefore, the application of any form of external loading will **result in the system moving as a rigid body**.

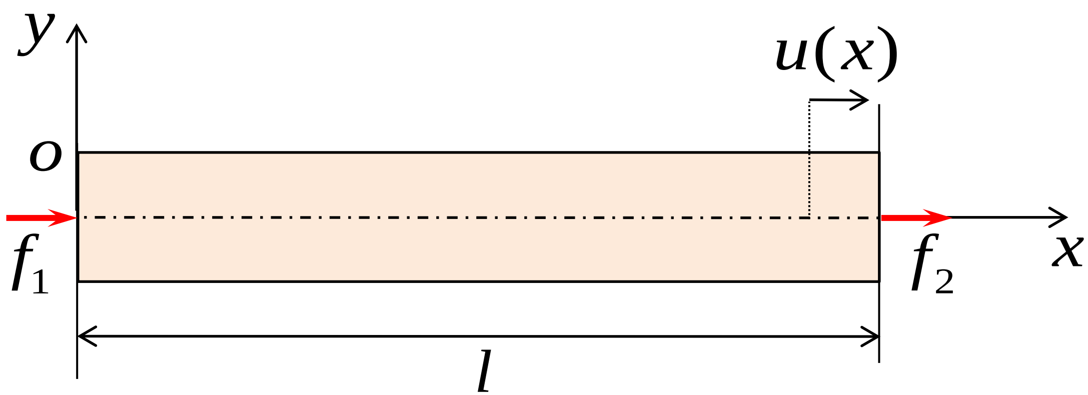
The problem can be rendered solvable simply by **specifying sufficient boundary conditions to prevent the structure moving** as a rigid body. Therefore assume node 1 to be fixed $\boxed{(u_1=0)}$ and the force applied on node 2 $\boxed{f_2}$ is given; then u_2 can be determined and f_1 is found consequently.

Rockey K. C., The Finite Element method, A basic introduction, Crosby Lockwood Staples, 1975, pp. 16~17

2.2 Element : Bar

- Derivation of the Stiffness Matrix by Applying Galerkin's Residual Method

Element : Bar – Problem Definition



f_1, f_2 : concentrated forces exerted on the ends of the bar

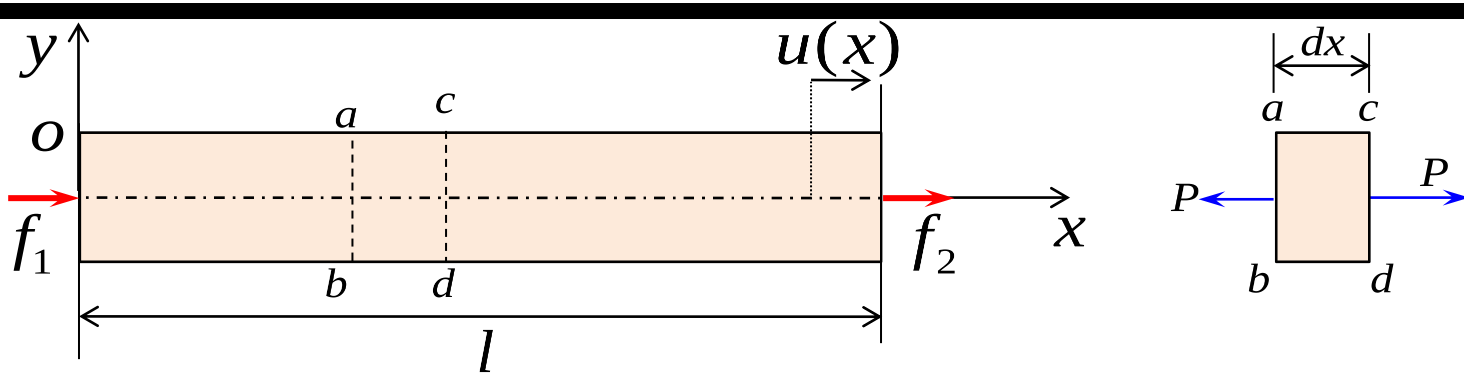
Given:

1) The concentrated forces f_1 and f_2 are exerted on the ends of the bar.
2) There is no distributed force.

Find:

1) The displacement at the ends of the bar u_1, u_2.

Element : Bar - Differential Eq.



P: the (internal) forces acting on the cross sections of a small element of the bar of length dx

f_1, f_2 :concentrated forces exerted on the ends of the bar

The bar element is assumed to have constant cross-section area A, modulus of elasticity E, and initial length l, and there is no distributed force:

$$P = A(x)\sigma$$

$$\sigma = E\varepsilon$$

$$P = EA(x)\varepsilon$$

$$\varepsilon = \frac{du(x)}{dx}$$

$$P = EA(x)\frac{du(x)}{dx}$$

From the force equilibrium, "P" dose not change along the x-axis

$$\frac{dP}{dx} = \frac{d}{dx}\left(EA(x)\frac{du(x)}{dx}\right) = 0$$

If A(x) is constant "A"

$$EA\frac{d^2u(x)}{dx^2} = 0 \Rightarrow$$

Differential Equation (Governing Equation)

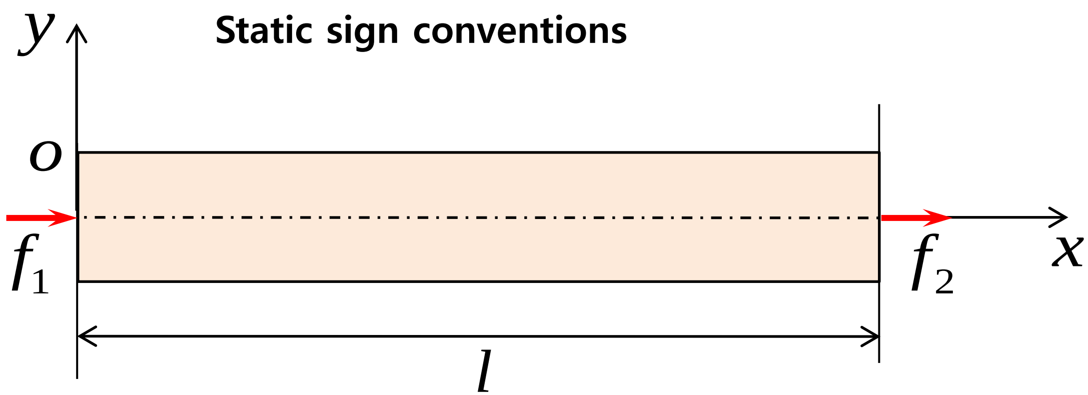
Boundary Condition

$$EA\left.\frac{du}{dx}\right|_{x=0} = P(0) \quad , \quad EA\left.\frac{du}{dx}\right|_{x=l} = P(l)$$

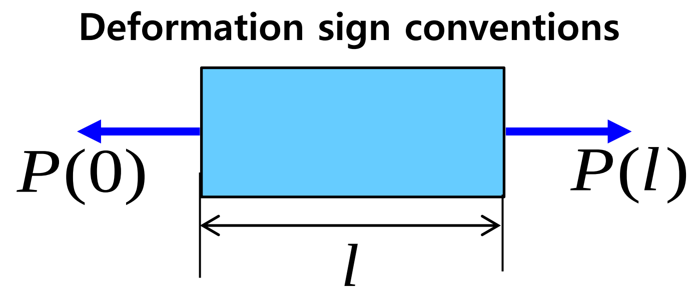
Then, how can we represent the boundary conditions with the given external forces?

Reference) Logan, A first course in the finite element method, 3rd edition, Thomson learning, 2002, p.64

Element : Bar – Determination of the stress resultant



f_1, f_2 : concentrated forces exerted on the ends of the bar



Remember!

The stress resultant, such as a tensile force P at the end of the bar, is not given, but can be represented with the given external forces f_1, f_2.

Boundary Condition

$$EA\left.\frac{du}{dx}\right|_{x=0} = P(0) \quad , \quad EA\left.\frac{du}{dx}\right|_{x=l} = P(l)$$

Tensile force at $x=0$

$$P(0) = -f_1$$

Tensile force at $x=l$

$$P(l) = +f_2$$

The minus sign of tensile force P(0) is the result of opposite sign conventions between the static and deformation at $x=0$.

Function Approximation by Shape Functions (Basis Function of order 1)

Differential Equation without distributed force

$$EA\frac{d^2u(x)}{dx^2}=0$$

Assume an approximation for the displacement in axial direction through the element length to be $\hat{u}(x)=a_0+a_1x$

This displacement function is appropriate because there are two degrees of freedom.

$\hat{u}(0)=u_1,\ \hat{u}(l)=u_2$ should be satisfied.

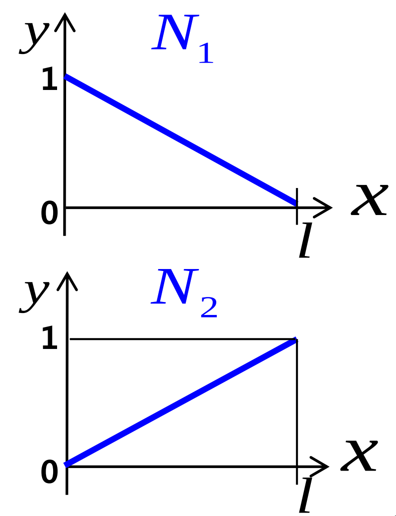
$$a_0=u_1,\ a_1=\frac{1}{l}(u_2-u_1)$$

$$\hat{u}(x)=\left(1-\frac{x}{l}\right)u_1+\frac{x}{l}u_2$$

$$\hat{u}(x)=N_1u_1+N_2u_2$$

$$where\ N_1=1-\frac{x}{l},\ N_2=\frac{x}{l}$$

Shape functions Ni



Galerkin's Residual Method

Differential Equation without distributed force

$$EA\frac{d^2u(x)}{dx^2}=0 \quad \Longrightarrow \quad \hat{u}(x)=N_1u_1+N_2u_2 \quad \text{where } N_1=1-\frac{x}{l},\ N_2=\frac{x}{l} \quad \Longrightarrow$$

since $\hat{u}(x)$ is approximated solution $\quad EA\frac{d^2\hat{u}(x)}{dx^2}\neq 0=R$

Thus substituting the approximated solution, which satisfy the boundary conditions, into the differential equation results in a residual *R* over the whole region of the problem as follows

$$\iiint_V R\,dV$$

In the residual method, we require that a weighted value of the residual be a minimum over the whole region. The weighting functions allow the weighted integral of residuals to go to zero

$$\iiint_V R\,W\,dV=0$$

, where W is an independent *weighting function*

We could require that an "appropriate number" of integrals of the error, "weighted in different ways", be zero

$$\iiint_V R\,W_i\,dV=0$$

, where " i "is i-th weighting function.

Galerkin's Residual Method:

The basis functions N_i are chosen to play the role of the weighting functions W_i

$$\iiint_V R\,N_i\,dV=0 \qquad ,(i=1,2)$$

Element : Bar (1 element , 2 nodes)
- Galerkin's Residual Method

Galerkin's residual method

Differential Equation

$$\boxed{EA\frac{d^2u(x)}{dx^2}=0}$$

$$\int_0^l AE\frac{d^2\hat{u}(x)}{dx^2}N_i\,dx=0 \quad ,(i=1,2)$$

$$where,\ \hat{u}(x)=N_1u_1+N_2u_2\,,N_1=1-\frac{x}{l},\ N_2=\frac{x}{l}$$

integration by parts

$$\left[N_iAE\frac{d\hat{u}}{dx}\right]_0^l-\int_0^l AE\frac{d\hat{u}}{dx}\frac{dN_i}{dx}dx=0$$

since $\frac{d\hat{u}}{dx}=\frac{dN_1}{dx}u_1+\frac{dN_2}{dx}u_2=\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}\begin{bmatrix}u_1\\u_2\end{bmatrix}$

$$AE\int_0^l\frac{dN_i}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\left[N_iAE\frac{d\hat{u}}{dx}\right]_0^l \qquad ,(i=1,2)$$

$$\Rightarrow\begin{cases} i=1: & AE\int_0^l\frac{dN_1}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\boxed{\left[N_1AE\frac{d\hat{u}}{dx}\right]_0^l}\Rightarrow \cancel{N_1}AE\left.\frac{d\hat{u}}{dx}\right|_{x=l}-N_1AE\left.\frac{d\hat{u}}{dx}\right|_{x=0}\Rightarrow\boxed{-P(0)} \\ i=2: & AE\int_0^l\frac{dN_2}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\boxed{\left[N_2AE\frac{d\hat{u}}{dx}\right]_0^l}\Rightarrow N_2AE\left.\frac{d\hat{u}}{dx}\right|_{x=l}-\cancel{N_2}AE\left.\frac{d\hat{u}}{dx}\right|_{x=0}\Rightarrow\boxed{P(l)}\end{cases}$$

since $N_1(0)=1,\ N_1(l)=0$

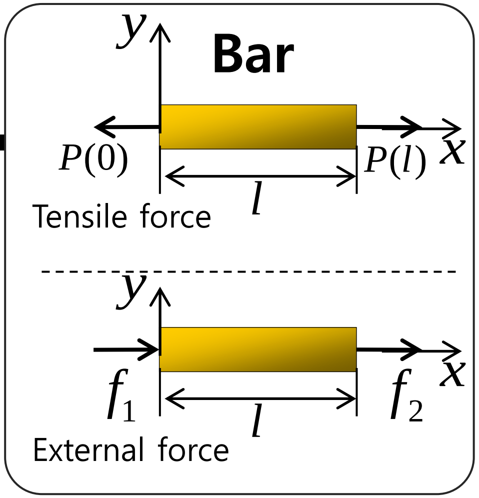
$N_2(0)=0,\ N_2(l)=1$

Element : Bar (1 element , 2 nodes) - Galerkin's Residual Method

$$\begin{bmatrix} AE\int_0^l \frac{dN_1}{dx}\left[-\frac{1}{l}\quad \frac{1}{l}\right]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = -P(0) \\ AE\int_0^l \frac{dN_2}{dx}\left[-\frac{1}{l}\quad \frac{1}{l}\right]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = P(l) \end{bmatrix}$$

Differential Equation

$$EA\frac{d^2u(x)}{dx^2} = 0$$



L.H.S

$$AE\int_0^l \begin{bmatrix}-\frac{1}{l}\end{bmatrix}\left[-\frac{1}{l}\quad \frac{1}{l}\right]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = AE\frac{1}{l^2}\int_0^l [1\quad -1]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = AE\frac{1}{l^2}[l\quad -l]\begin{bmatrix}u_1\\u_2\end{bmatrix} = \frac{AE}{l}[1\quad -1]\begin{bmatrix}u_1\\u_2\end{bmatrix}$$

$$AE\int_0^l \begin{bmatrix}\frac{1}{l}\end{bmatrix}\left[-\frac{1}{l}\quad \frac{1}{l}\right]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = AE\frac{1}{l^2}\int_0^l [-1\quad 1]dx\begin{bmatrix}u_1\\u_2\end{bmatrix} = AE\frac{1}{l^2}[-l\quad l]\begin{bmatrix}u_1\\u_2\end{bmatrix} = \frac{AE}{l}[-1\quad 1]\begin{bmatrix}u_1\\u_2\end{bmatrix}$$

$$\therefore \begin{bmatrix} \frac{AE}{l}[1\quad -1]\begin{bmatrix}u_1\\u_2\end{bmatrix} = -P(0) \\ \frac{AE}{l}[-1\quad 1]\begin{bmatrix}u_1\\u_2\end{bmatrix} = P(l) \end{bmatrix}$$

$$\boxed{\therefore \mathbf{Kd} = \mathbf{F} \quad where\ , \mathbf{K} = \frac{EA}{l}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}, \mathbf{d} = \begin{bmatrix}u_1\\u_2\end{bmatrix}, \mathbf{F} = \begin{bmatrix}-P(0)\\P(l)\end{bmatrix} = \begin{bmatrix}f_1\\f_2\end{bmatrix}}$$

REFERENCE: GALERKIN'S RESIDUAL METHOD

Approximation by Trial Functions
- Function Approximation by Trial Functions

frequently referred *as shape or basis function*

If we can find any function ψ satisfying $\psi|_{\Gamma} = \phi|_{\Gamma}$

and if we introduce a set of independent ***trial functions***

$\{N_m \; ; \; m = 1, 2, 3 \ldots\}$ **such that** $N_m|_{\Gamma} = 0$ **for all** m

then at all points in Ω , we can approximate to ϕ by

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

Approximation by Trial Functions
- Function Approximation by Trial Functions

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

The manner in which ψ and the trial function set are defined automatically ensures that $\hat{\phi}\big|_\Gamma = \phi\big|_\Gamma$

the approximation has the property that whatever the values of the parameters a_m

Approximation by Trial Functions
- Weighted Residual Approximations

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

We shall now attempt to develop a general method for determining the parameters a_m in the approximation

We begin by introducing the error, or *residual* R_Ω in the approximation

$$R_\Omega \equiv \phi - \hat{\phi}$$

which is a function of position in Ω

Approximation by Trial Functions - Weighted Residual Approximations

In an attempt to reduce this residual in some **overall manner** over the whole domain Ω

We could require that an appropriate number of **integrals of the error over** Ω, weighted in different ways, **be zero**

$$\int_{\Omega} W_i \, (\phi - \hat{\phi}) \, d\Omega \equiv \int_{\Omega} W_i \, R_{\Omega} \, d\Omega = 0$$

$$i = 1, 2, \ldots, M$$

where W_i is a set of independent ***weighting functions***

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_{\Gamma} = \phi\big|_{\Gamma}$$

$$N_m\big|_{\Gamma} = 0$$

where, a_m are some parameters which are **computed so as to obtain a good "fit"**

residual

$$R_{\Omega} = \phi - \hat{\phi}$$

[Zienkiewicz 1983] Ch. 2.1

Approximation by Trial Functions
- Weighted Residual Approximations

The general completeness (convergence) requirement

$$\hat{\phi} \rightarrow \phi \quad \textbf{as} \quad M \rightarrow \infty$$

can then be cast in an alternative form by requiring

$$\int_{\Omega} W_i \, R_{\Omega} \, d\Omega = 0 \quad \textbf{for all} \; i \quad \textbf{as} \quad M \rightarrow \infty$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_{\Gamma} = \phi\big|_{\Gamma}$$

$$N_m\big|_{\Gamma} = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

residual

$$R_{\Omega} = \phi - \hat{\phi}$$

Approximation by Trial Functions - Weighted Residual Approximations

alternative form of completeness requirement

$$\int_\Omega W_i R_\Omega \, d\Omega = 0 \quad \text{for all } i \quad \text{as } M \to \infty$$

$$\boxed{\int_\Omega W_i (\phi - \hat{\phi}) \, d\Omega = 0}$$

standard weighted residual statement

Find

$$\int_\Omega W_i \left(\phi - \psi - \sum_{m=1}^{M} \boxed{a_m} N_m \right) d\Omega = 0$$

User defined weighting function

chosen to be zero at the B/C

chosen to satisfy the B/C

the function to be approximated is given

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

residual

$$R_\Omega = \phi - \hat{\phi}$$

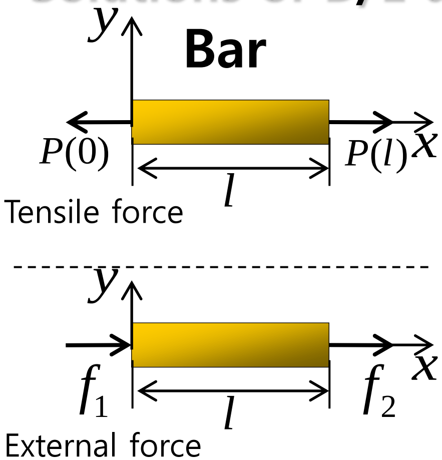
2.3. Element : Bar

- Comparison between "Direct Equilibrium Approach" and "Galerkin's Residual Method"

E: Young's Modulus, A: sectional area

Element : Bar (1 element , 2 nodes)

Solutions of D/E using Galerkin's Residual Method



Differential Equation

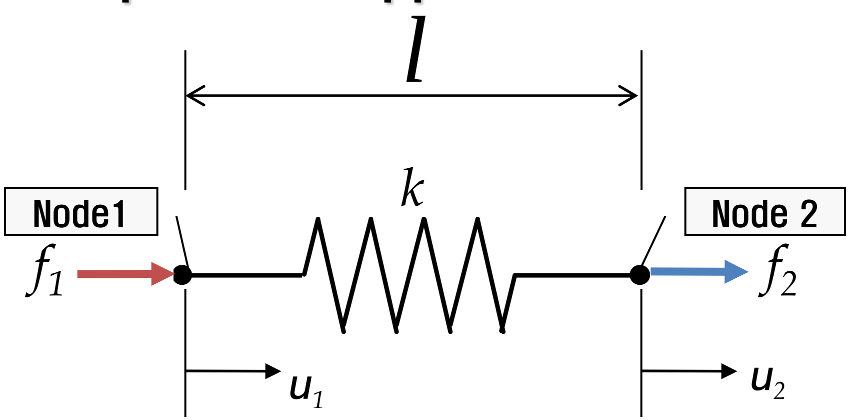
$$EA\frac{d^2u(x)}{dx^2}=0$$

$$\hat{u}(x)=a_0+a_1x$$

Galerkin's residual method

$$\frac{EA}{l}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\begin{bmatrix}-P(0)\\ P(l)\end{bmatrix}=\begin{bmatrix}f_1\\ f_2\end{bmatrix}$$

Direct equilibrium approach



$$k=\frac{EA}{l}$$

$$\begin{bmatrix}f_1\\ f_2\end{bmatrix}=\begin{bmatrix}k & -k\\ -k & k\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

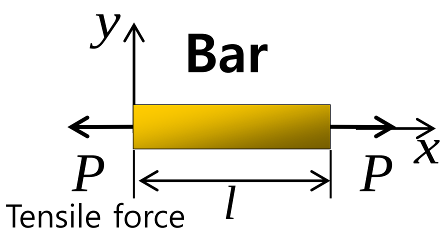
stiffness matrix

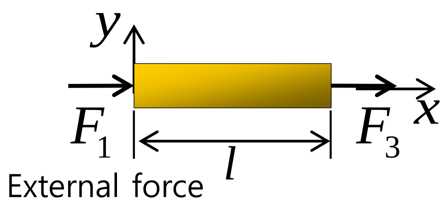
$$[f]=[K][u]$$

2.4. Element : Bar

- Derivation of Stiffness Matrix for a Bar Composed of 2 Elements by Applying Galerkin's Residual Method

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method





Differential Equation

$$EA\frac{d^2u(x)}{dx^2}=0 \qquad 0<x<l$$

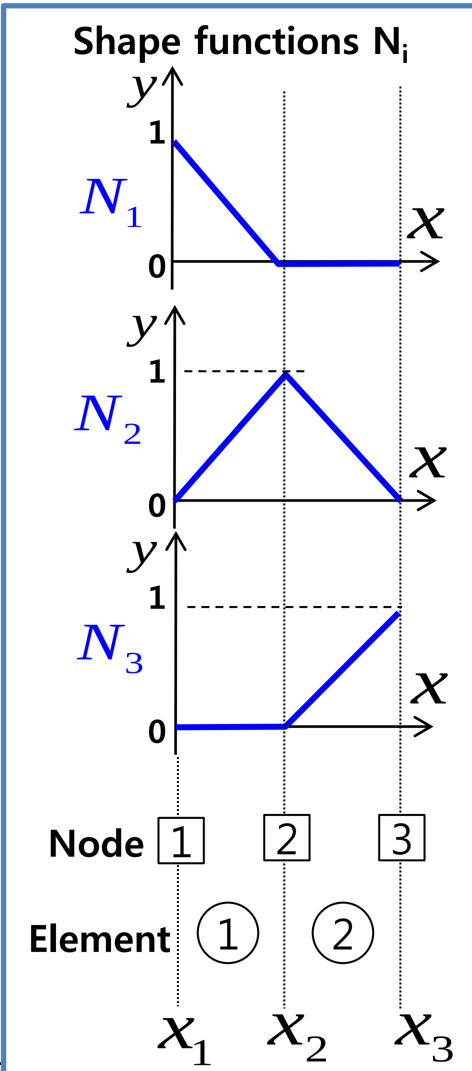
Boundary Condition

$$EA\left.\frac{du}{dx}\right|_{x=0}=P(0) \quad , \quad EA\left.\frac{du}{dx}\right|_{x=l}=P(l)$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

$$\boxed{A(u) = EA\frac{d^2u}{dx^2} = 0 \quad in \;\; 0 < x < l}$$

$$u \approx \hat{u} = N_1u_1 + N_2u_2 + N_3u_3$$



The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{u}) - \cancel{A(u)} = EA\frac{d^2\hat{u}}{dx^2} \quad in \;\; 0 < x < l$$

The weighted residual form:

$$\int_0^l W_i \mathbf{R}_\Omega dx = 0, \; i = 1, 2, 3$$

$$\int_0^l W_i \left(EA\frac{d^2\hat{u}}{dx^2} \right) dx = 0, \; i = 1, 2, 3$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$u \approx \hat{u} = N_1 u_1 + N_2 u_2 + N_3 u_3$$

$$\int_0^l W_i \left(EA \frac{d^2 \hat{u}}{dx^2} \right) dx = 0, \ i = 1, 2, 3$$

$$\downarrow$$

$$\int_0^l W_i EA \frac{d^2 \hat{u}}{dx^2} dx = 0, \ i = 1, 2, 3$$

$$\downarrow$$

$$EA \int_0^l W_i \frac{d^2 \hat{u}}{dx^2} dx = 0, \ i = 1, 2, 3$$

$\downarrow$ **Integration by parts**

$$-EA \int_0^l \frac{dW_i}{dx} \frac{d\hat{u}}{dx} dx + EA \left[W_i \frac{d\hat{u}}{dx} \right]_0^l = 0, \ i = 1, 2, 3$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$u \approx \hat{u} = N_1 u_1 + N_2 u_2 + N_3 u_3$

$$EA\int_0^l \frac{dW_i}{dx}\frac{d\left(N_1u_1 + N_2u_2 + N_3u_3\right)}{dx}dx - EA\left[W_i\frac{d\hat{u}}{dx}\right]_0^l = 0,\ i = 1,2,3$$

$\downarrow$ **Galerkin methods** $W_i = N_i$

$$EA\int_0^l \frac{dN_i}{dx}\frac{d\left(N_1u_1 + N_2u_2 + N_3u_3\right)}{dx}dx - EA\left[N_i\frac{d\hat{u}}{dx}\right]_0^l = 0,\ i = 1,2,3$$

$\downarrow$

$$EA\int_0^l u_1\frac{dN_i}{dx}\frac{dN_1}{dx}dx + EA\int_0^l u_2\frac{dN_i}{dx}\frac{dN_2}{dx}dx + EA\int_0^l u_3\frac{dN_i}{dx}\frac{dN_3}{dx}dx$$

$$= EA\left[N_i\frac{d\hat{u}}{dx}\right]_0^l,\ i = 1,2,3$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^l u_1 \frac{dN_i}{dx}\frac{dN_1}{dx}dx + \int_0^l u_2 \frac{dN_i}{dx}\frac{dN_2}{dx}dx + \int_0^l u_3 \frac{dN_i}{dx}\frac{dN_3}{dx}dx\right) =$$

$$+\left[N_i EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$\downarrow i = 1, 2, 3$$

$$EA\left(\int_0^l u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx + \int_0^l u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx + \int_0^l u_3 \frac{dN_1}{dx}\frac{dN_3}{dx}dx\right) = \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$EA\left(\int_0^l u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx + \int_0^l u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx + \int_0^l u_3 \frac{dN_2}{dx}\frac{dN_3}{dx}dx\right) = \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$EA\left(\int_0^l u_1 \frac{dN_3}{dx}\frac{dN_1}{dx}dx + \int_0^l u_2 \frac{dN_3}{dx}\frac{dN_2}{dx}dx + \int_0^l u_3 \frac{dN_3}{dx}\frac{dN_3}{dx}dx\right) = \left[N_3 EA\frac{d\hat{u}}{dx}\right]_0^l$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^l u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx+\int_0^l u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx+\int_0^l u_3 \frac{dN_1}{dx}\frac{dN_3}{dx}dx\right)=\left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$EA\left(\int_0^l u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx+\int_0^l u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx+\int_0^l u_3 \frac{dN_2}{dx}\frac{dN_3}{dx}dx\right)=\left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$EA\left(\int_0^l u_1 \frac{dN_3}{dx}\frac{dN_1}{dx}dx+\int_0^l u_2 \frac{dN_3}{dx}\frac{dN_2}{dx}dx+\int_0^l u_3 \frac{dN_3}{dx}\frac{dN_3}{dx}dx\right)=\left[N_3 EA\frac{d\hat{u}}{dx}\right]_0^l$$

$$\downarrow$$

$$EA\underbrace{\begin{bmatrix}\int_0^l \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_3}{dx}dx\\ \int_0^l \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_3}{dx}dx\\ \int_0^l \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_3}{dx}dx\end{bmatrix}}_{\mathbf{K}}\underbrace{\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}}_{\mathbf{d}}=\underbrace{\begin{bmatrix}\left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^l\\ \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^l\\ \left[N_3 EA\frac{d\hat{u}}{dx}\right]_0^l\end{bmatrix}}_{\mathbf{F}}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

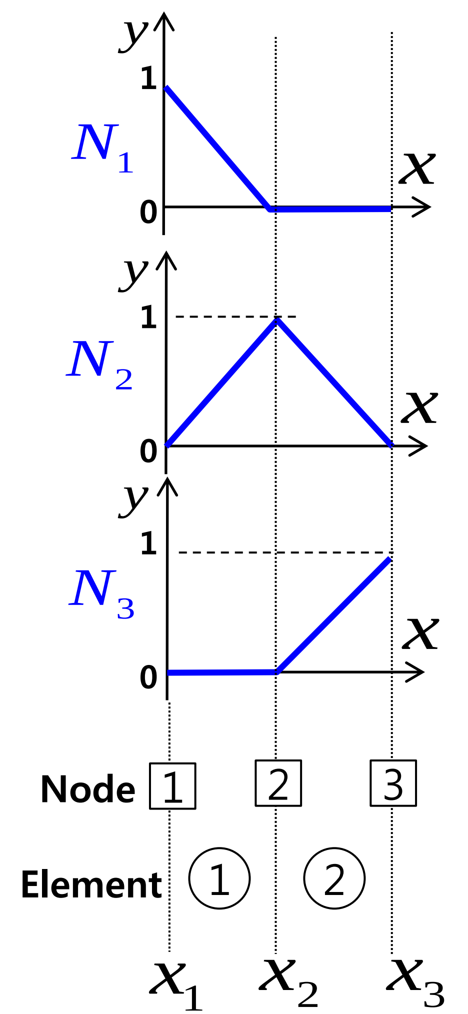
The weighted residual form:

$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = EA\begin{bmatrix} \int_0^l \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_0^l \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_0^l \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad \mathbf{F} = \begin{bmatrix} \left[N_1 EA \frac{d\hat{u}}{dx} \right]_0^l \\ \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^l \\ \left[N_3 EA \frac{d\hat{u}}{dx} \right]_0^l \end{bmatrix}$$

N_i is corresponding to the 1st order B-spline basis functions



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

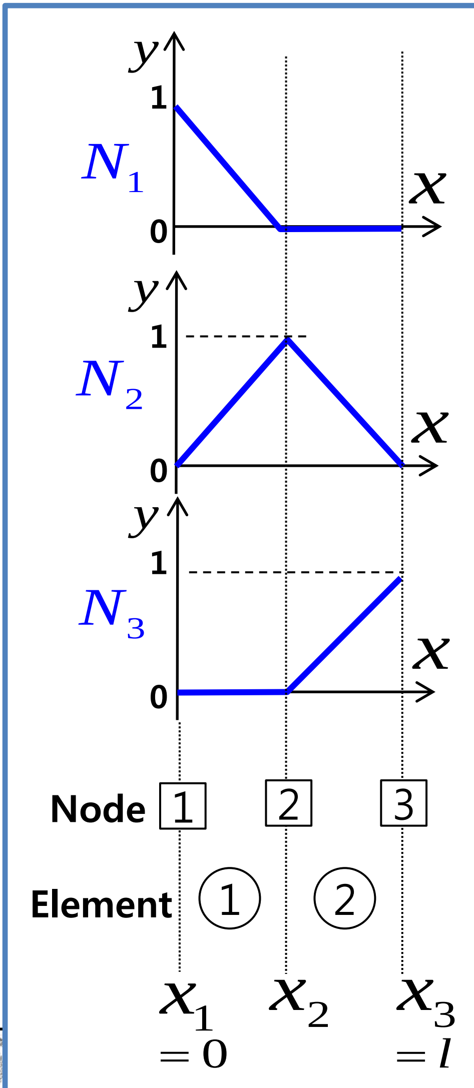
$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = EA\begin{bmatrix} \int_0^l \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_0^l \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_0^l \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \int_0^l \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA\begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \boxed{\int_{x_1}^{x_2}} \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_1}^{x_2}}$$

$$\mathbf{K}_2 = EA\begin{bmatrix} \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_{x_2}^{x_3} \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \boxed{\int_{x_2}^{x_3}} \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_2}^{x_3}}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

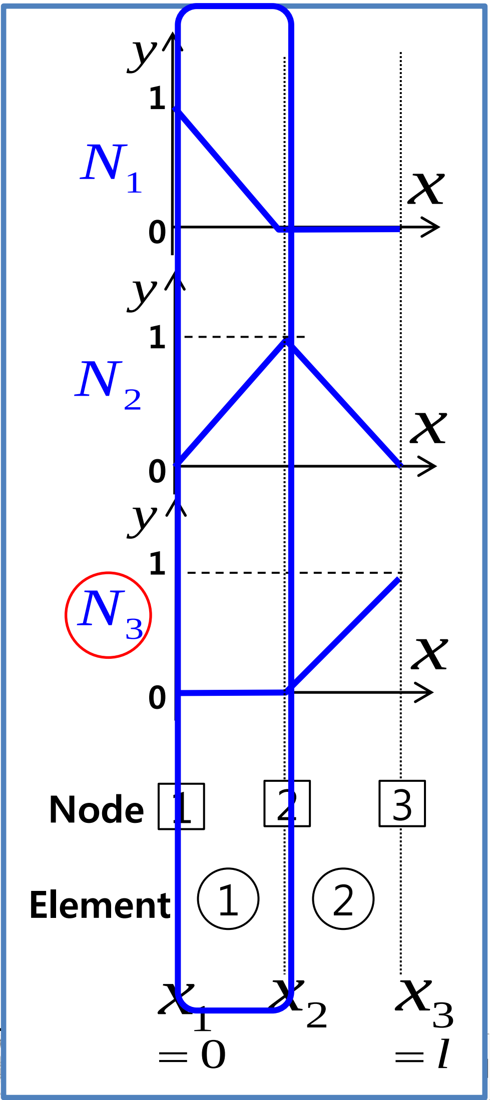
The weighted residual form:

$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA \begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix}$$

$$= EA \begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & 0 \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $\mathbf{K}\mathbf{d} = \mathbf{F}$

$$\frac{dN_1}{dx} = -\frac{1}{l/2}, \quad \frac{dN_2}{dx} = \frac{1}{l/2}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

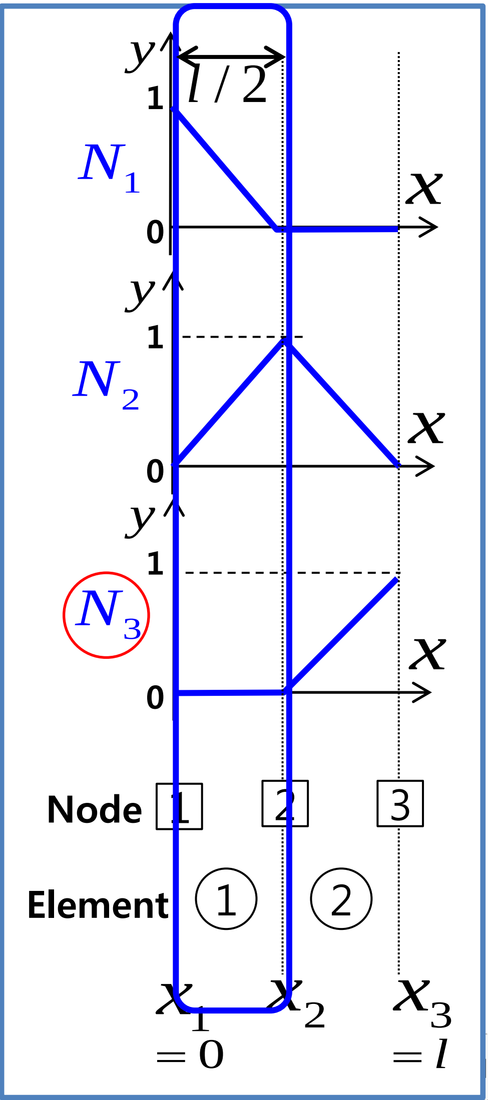
$$\mathbf{K}_1 = EA\begin{bmatrix} \boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx} & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & 0 \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \boxed{\int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx} = \int_0^{l/2}\left(-\frac{1}{l/2}\right)\left(-\frac{1}{l/2}\right)dx$$

$$= \int_0^{l/2}\frac{1}{(l/2)^2}dx = \left[\frac{1}{(l/2)^2}x\right]_0^{l/2} = \frac{1}{l/2}$$

$$\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx = \int_0^{l/2}\left(-\frac{1}{l/2}\right)\left(\frac{1}{l/2}\right)dx$$

$$= -\int_0^{l/2}\frac{1}{(l/2)^2}dx = -\left[\frac{1}{(l/2)^2}x\right]_0^{l/2} = -\frac{1}{l/2}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

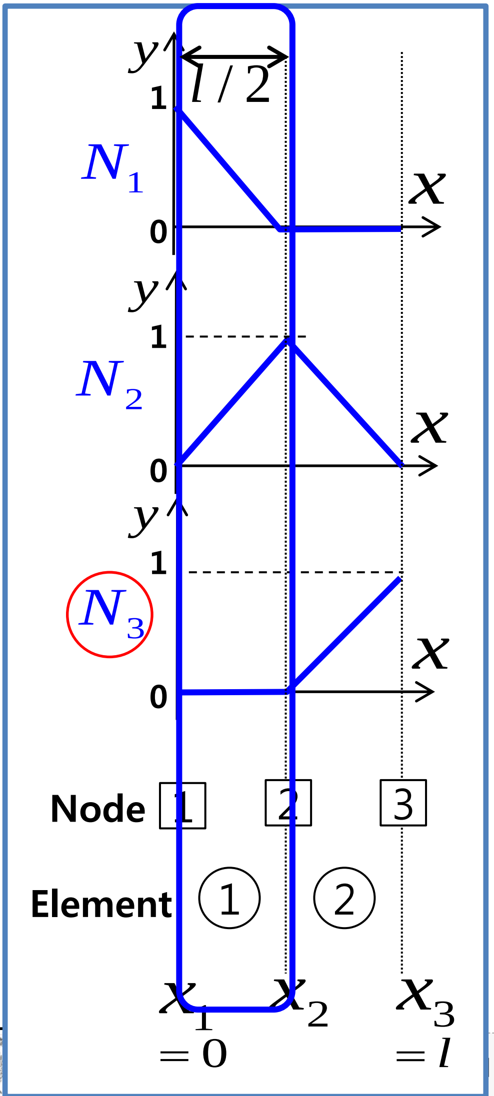
$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA\begin{bmatrix} \frac{1}{l/2} & -\frac{1}{l/2} & 0 \\ -\frac{1}{l/2} & \frac{1}{l/2} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

In the same manner,

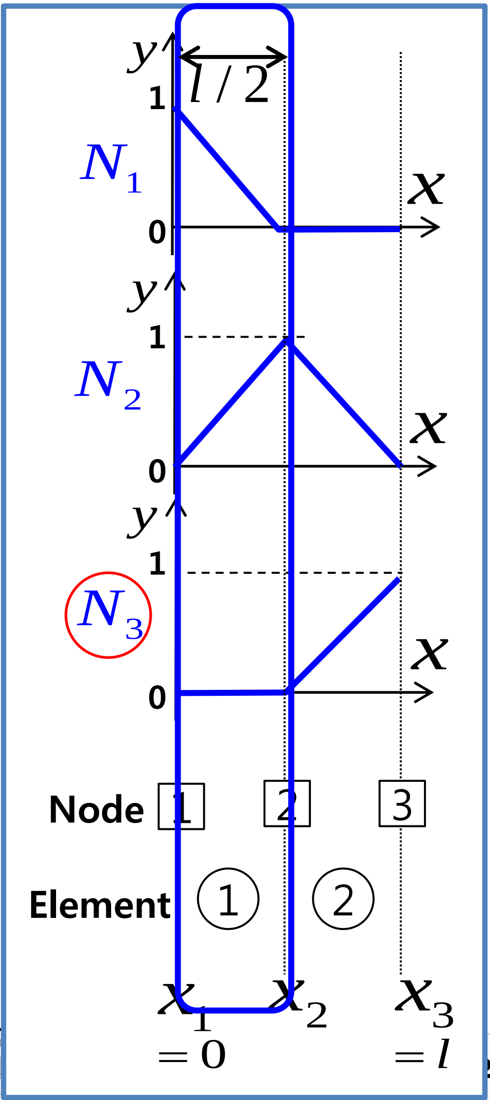
$$\mathbf{K}_2 = EA\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{l/2} & -\frac{1}{l/2} \\ 0 & -\frac{1}{l/2} & \frac{1}{l/2} \end{bmatrix} = \frac{EA}{l/2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $\mathbf{Kd} = \mathbf{F}$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2 = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$



$$\mathbf{K}_1 = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{K}_2 = \frac{EA}{l/2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

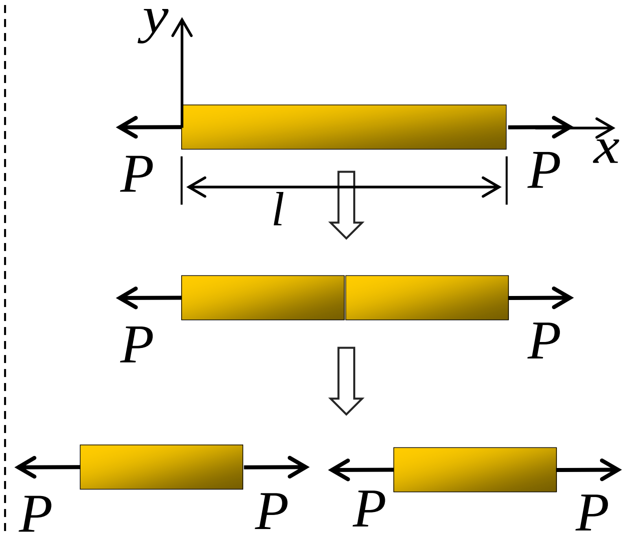
The weighted residual form:

$$\mathbf{Kd} = \mathbf{F}$$

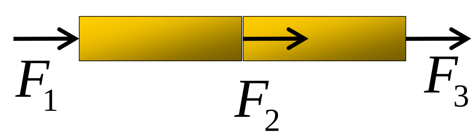
$$\mathbf{K} = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} \left[N_1 EA \dfrac{d\hat{u}}{dx} \right]_0^l \\ \left[N_2 EA \dfrac{d\hat{u}}{dx} \right]_0^l \\ \left[N_3 EA \dfrac{d\hat{u}}{dx} \right]_0^l \end{bmatrix} = \begin{bmatrix} -P(0) \\ 0 \\ P(l) \end{bmatrix}$$

Tensile force



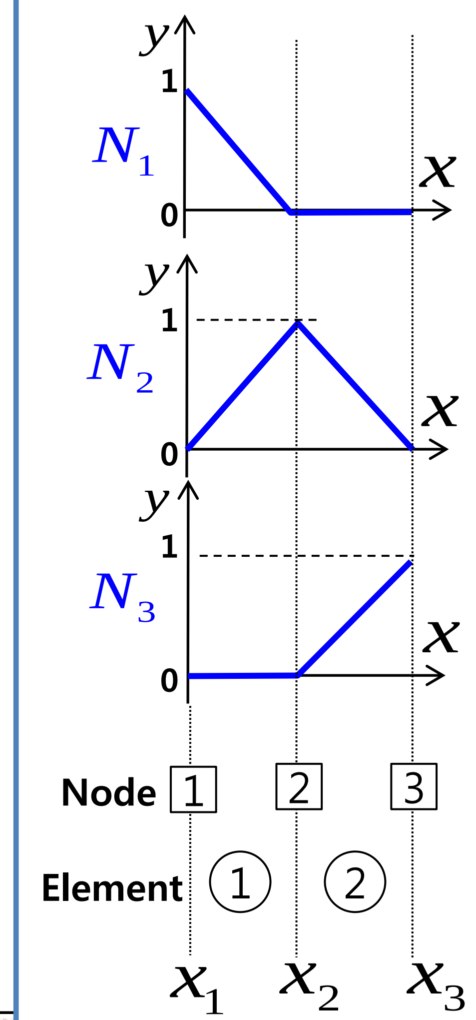
External forces



$$-P(0) = F_1$$

$$F_2 = -P + P = 0$$

$$P(l) = F_3$$



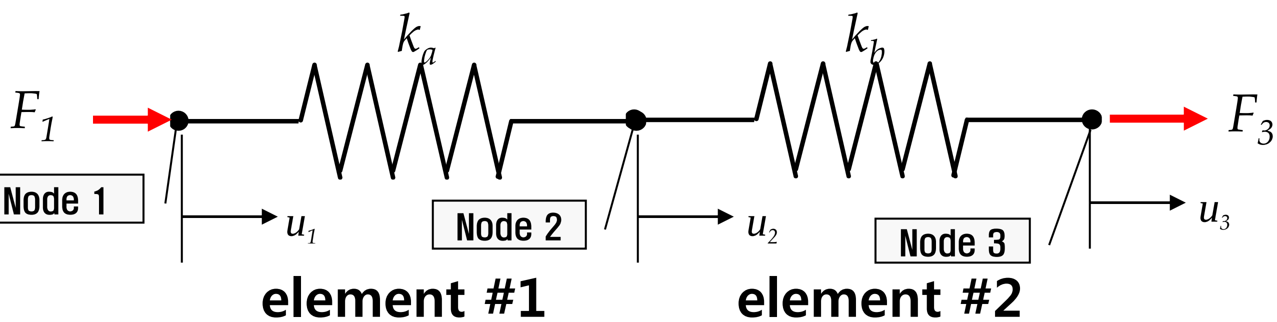
2.5. Element : Bar

- Derivation of Stiffness Matrix for a Bar Composed of 2 Elements by Superposition of Stiffness Matrix

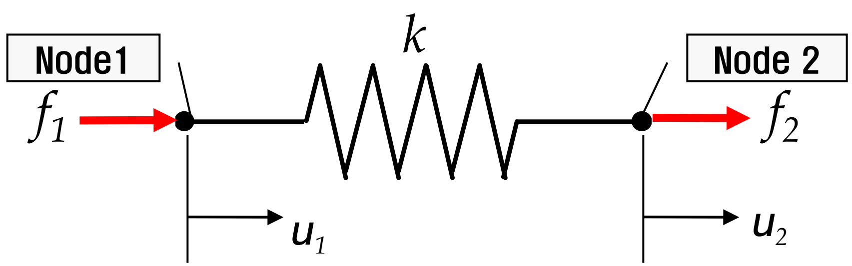
Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach

We will consider two spring assemblage.

F_1 and F_3 are <u>**external forces**</u> which are applied at node 1 and 3, respectively.



Recall stiffness matrix for single bar element

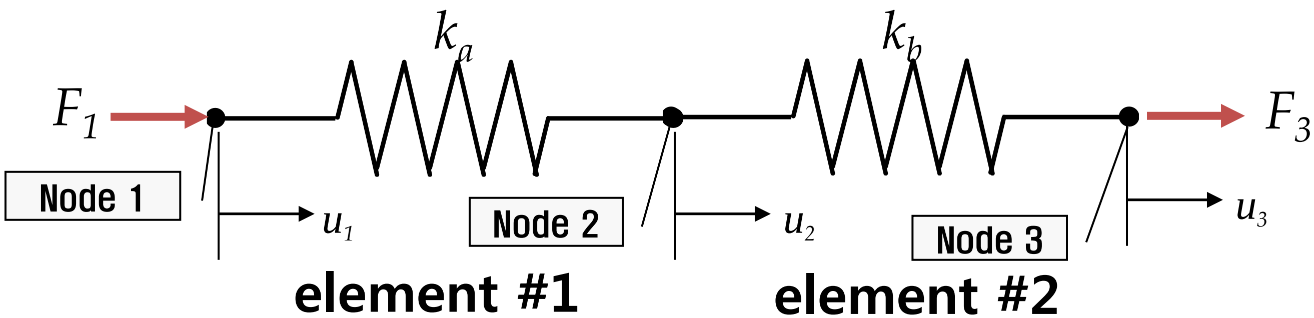


$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \boxed{\begin{bmatrix} k & -k \\ -k & k \end{bmatrix}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

stiffness matrix

$$[f] = [K][u]$$

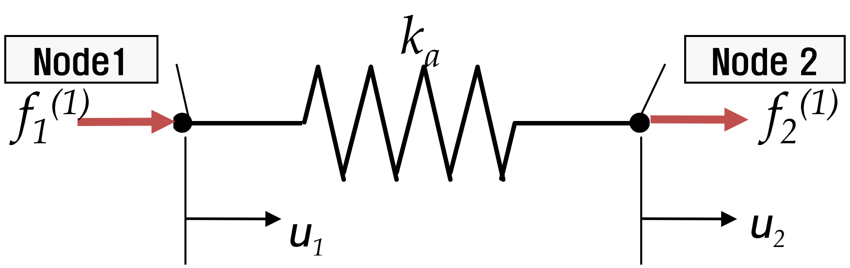
Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach



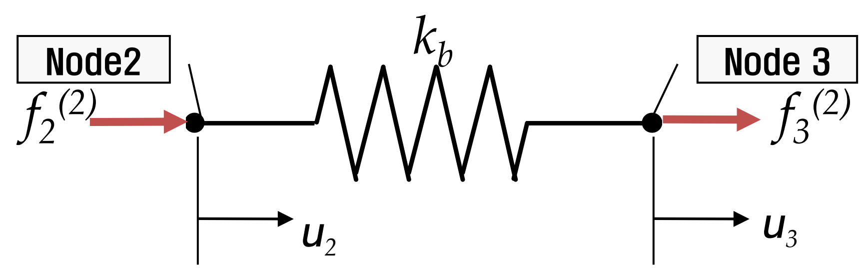
F_i: External forces at the i^{th} node.

Free-body diagrams of each element and nodes are shown as follows:

element #1



element #2

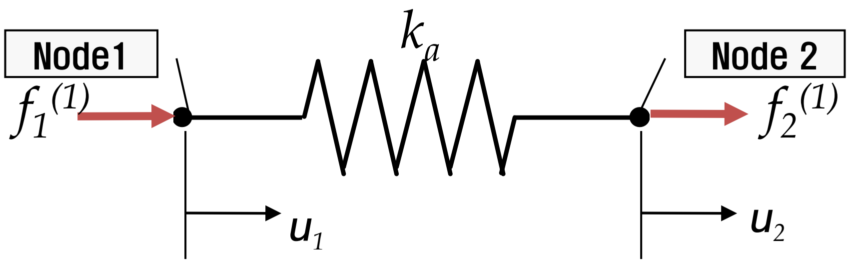


, where forces $f_i^{(j)}$ correspond to internal forces at the i^{th} node of the j^{th} element.

Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach

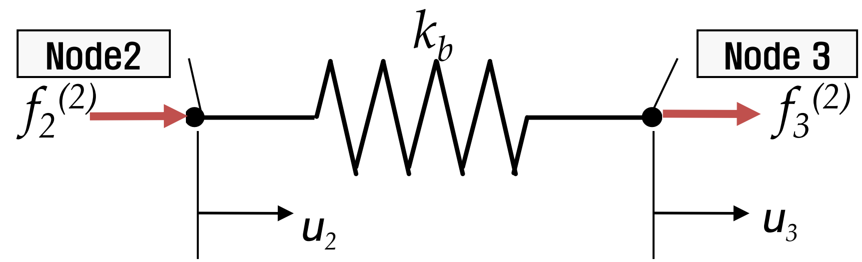
<u>Free-body diagrams</u> of each element and nodes are shown as follows:

element #1



$$\begin{bmatrix} f_1^{(1)} \\ f_2^{(1)} \end{bmatrix} = \begin{bmatrix} k_a & -k_a \\ -k_a & k_a \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

element #2



$$\begin{bmatrix} f_2^{(2)} \\ f_3^{(2)} \end{bmatrix} = \begin{bmatrix} k_b & -k_b \\ -k_b & k_b \end{bmatrix} \begin{bmatrix} u_2 \\ u_3 \end{bmatrix}$$

, where forces $f_i^{(j)}$ correspond to internal forces at the i^th node of the j^th element.

Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach

element #1

$$\begin{bmatrix} f_1^{(1)} \\ f_2^{(1)} \end{bmatrix} = \begin{bmatrix} k_a & -k_a \\ -k_a & k_a \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \longrightarrow \begin{aligned} f_1^{(1)} &= k_a u_1 - k_a u_2 \\ f_2^{(1)} &= -k_a u_1 + k_a u_2 \end{aligned}$$

$$+$$

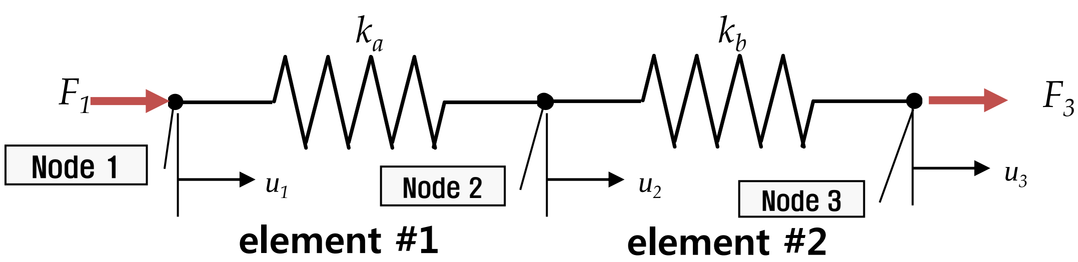
element #2

$$\begin{bmatrix} f_2^{(2)} \\ f_3^{(2)} \end{bmatrix} = \begin{bmatrix} k_b & -k_b \\ -k_b & k_b \end{bmatrix} \begin{bmatrix} u_2 \\ u_3 \end{bmatrix} \longrightarrow \begin{aligned} f_2^{(2)} &= k_b u_2 - k_b u_3 \\ f_3^{(2)} &= -k_b u_2 + k_b u_3 \end{aligned}$$

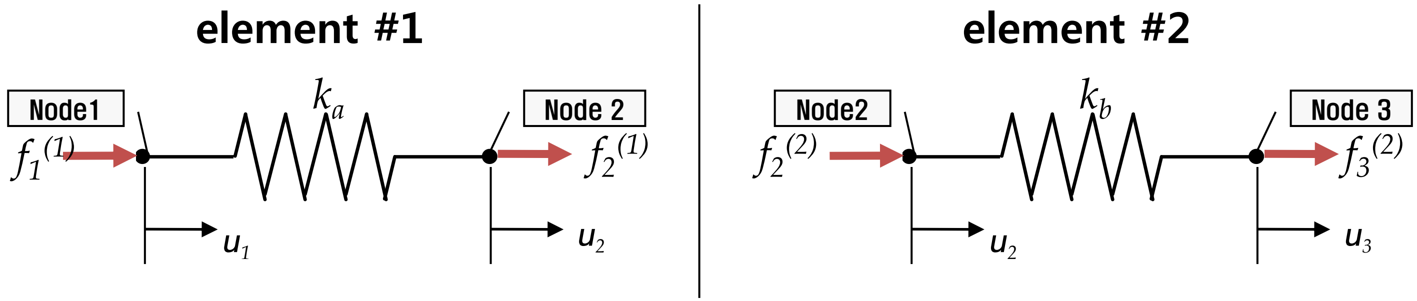
$$=$$

$$\begin{bmatrix} f_1^{(1)} \\ f_2^{(1)} + f_2^{(2)} \\ f_3^{(2)} \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \longleftarrow \begin{aligned} f_1^{(1)} &= k_a u_1 - k_a u_2 \\ f_2^{(1)} + f_2^{(2)} &= -k_a u_1 + \left(k_a + k_b\right) u_2 - k_b u_3 \\ f_3^{(2)} &= -k_b u_2 + k_b u_3 \end{aligned}$$

Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach



F_i : forces at the i^th node.



$f_i^{(j)}$ correspond to internal forces at the i^th node of the j^th element.

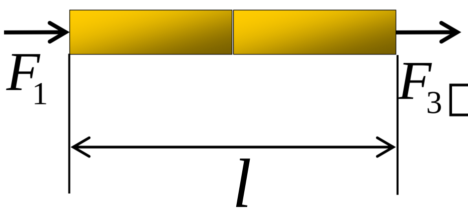
Based on the free-body diagrams, and the fact that external forces must equal internal forces at each node, we can write nodal equilibrium equations at each node as follows.

$$F_1 = f_1^{(1)}, \ \ 0 = f_2^{(1)} + f_2^{(2)}, \ \ F_3 = f_3^{(2)}$$

Therefore

$$\begin{bmatrix} f_1^{(1)} \\ f_2^{(1)} + f_2^{(2)} \\ f_3^{(2)} \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \Rightarrow \begin{bmatrix} F_1 \\ 0 \\ F_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Solutions of D/E using Galerkin's Residual Method



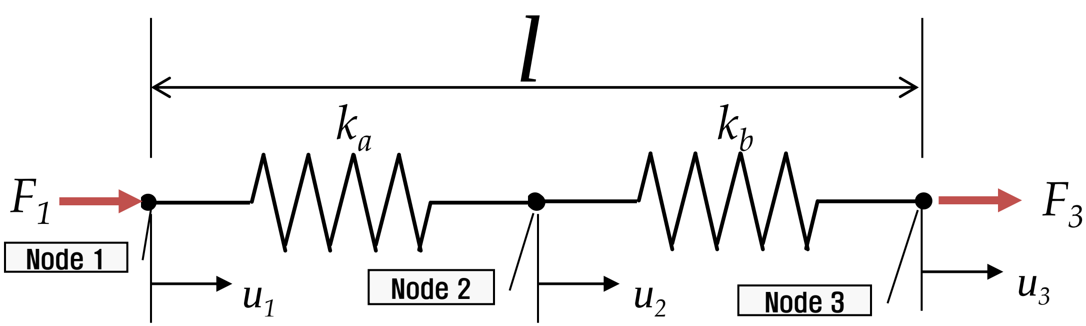
Differential Equation

$$EA\frac{d^2u(x)}{dx^2}=0$$

Galerkin's residual method

$$\begin{bmatrix} F_1 \\ 0 \\ F_3 \end{bmatrix} = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Direct equilibrium approach



$$\begin{bmatrix} F_1 \\ 0 \\ F_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$k_a = k_b = k = \frac{EA}{l/2}$$

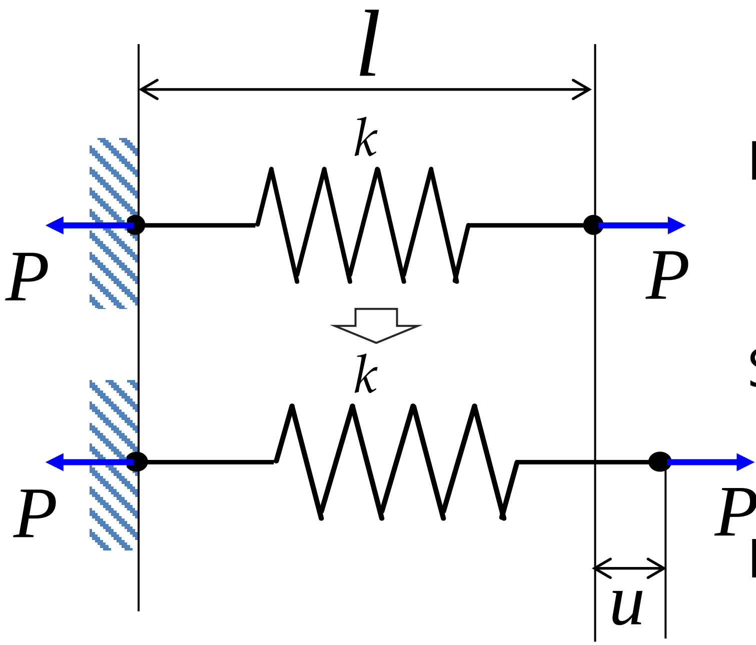
$$\begin{bmatrix} F_1 \\ 0 \\ F_3 \end{bmatrix} = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

If one spring is divided into two elements, then the elements have same spring constants.

$$k_a = k_b = k = \frac{EA}{l/2}$$

※ superposition of stiffness matrix

Reference) Calculation of stiffness constant of spring



Young's Modulus: E

Cross Section Area: A

Stress(axial force per unit cross section area): σ

Strain(elongation per unit length): ε

Constitutive Equation

$$\sigma = E\varepsilon$$

Multiplying cross section area A gives

$$A\sigma = EA\varepsilon$$

Substituting "stress resultant" P for $A\sigma$ gives

$$P = EA\varepsilon$$

Multiplying the length l yields

$$lP = EAl\varepsilon$$

$$= EA \cdot u \quad \text{or} \quad P = \frac{EA}{l} \cdot u$$

Therefore

$$P = ku, \; where \; k = \frac{EA}{l}$$

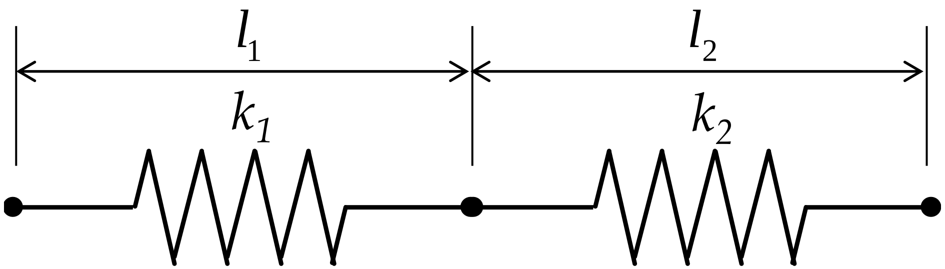
Reference) Calculation of stiffness constant of spring

$$k = \frac{EA}{l}$$

Series assembly of the spring

Assumption: Young's modulus and area are same.

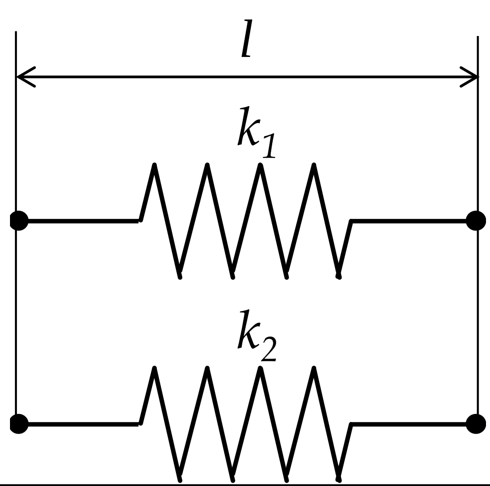
$$k_1 = \frac{EA}{l_1}, \quad k_2 = \frac{EA}{l_2}, \quad k_3 = \frac{EA}{l_1 + l_2}$$



$$\Rightarrow \frac{1}{k_1} + \frac{1}{k_2} = \frac{1}{k_3}$$

Parallel assembly of the spring

Assumption: Young's modulus and length are same.



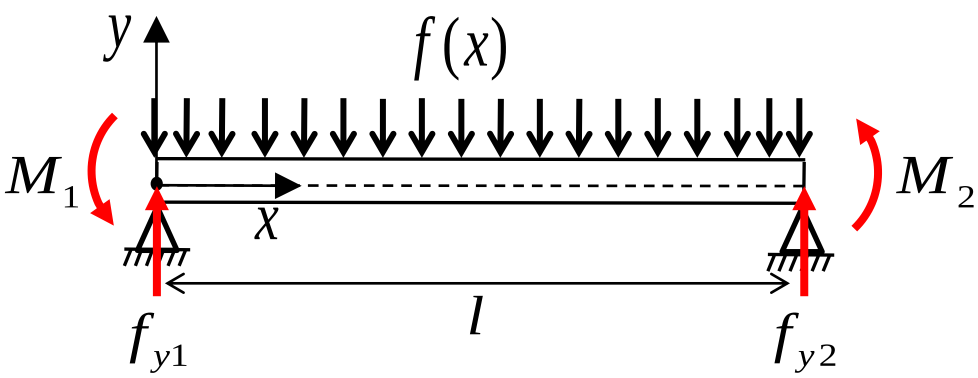
$$k_1 = \frac{EA_1}{l}, k_2 = \frac{EA_2}{l}, k_3 = \frac{E(A_1 + A_2)}{l}$$

$$\Rightarrow k_1 + k_2 = k_3$$

2.6. Element : Beam

- Derivation of the Stiffness Matrix by Applying Direct Equilibrium Approach

Element : Beam – Problem Definition



- Consider a beam element with the simple supports
- Moments and rotations are positive in the counterclockwise direction.
- Force and displacement are positive in the positive y direction.

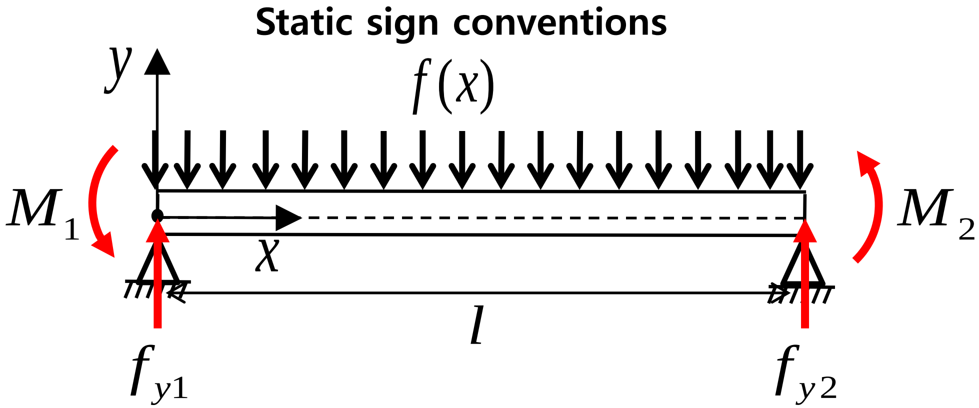
Given:

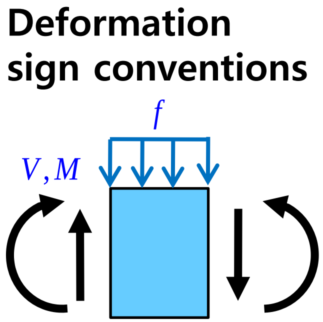
1) The concentrated forces f_{y1} and f_{y2} are exerted on the ends of the bar.
2) The moment M_1 and M_2 are exerted on the ends of the bar.
3) distributed force f(x) is applied to the element

Find:

1) The vertical displacement at the ends of the bar d_{y1}, d_{y2}.
2) The rotation angle at the ends of the bar ϕ_1, ϕ_2.

Element : Beam – Determination of the stress resultant





Remember! The stress resultants, such as a bending moment "m" and shear force "V" at the end of the beam, are not given, but can be represented by the given external forces.

Shear force at x=0

$$V(0) = f_{y1}$$

Bending moment at x=0

$$m(0) = -M_1$$

Shear force at x=l

$$V(l) = -f_{y2}$$

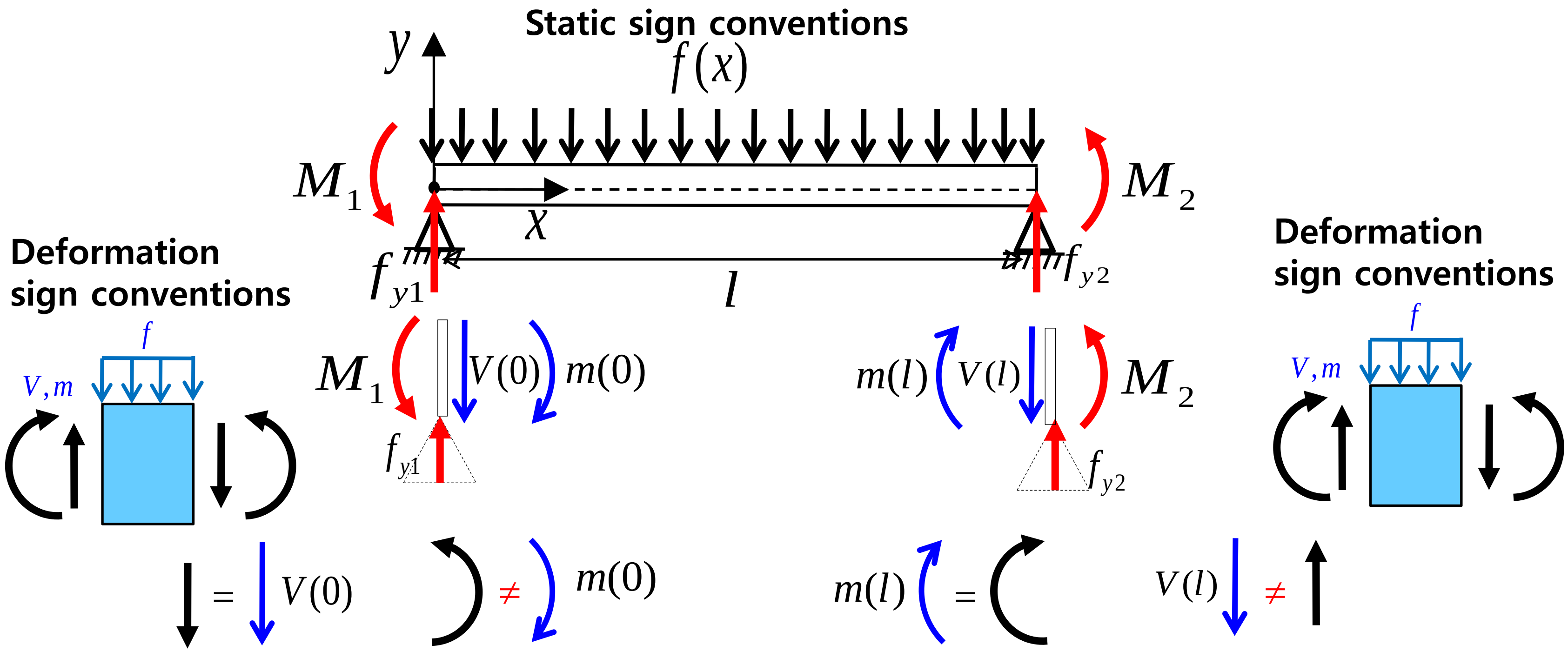
Bending moment at x=l

$$m(l) = M_2$$

The minus sign of bending moment $m(0)$ and shear force $V(l)$ are the result of opposite sign conventions between the static and deformation at x=0 and x=l respectively.

Remember: The bending moment and shear force at the ends of the beam are independent of the distributed force, but only dependent on the boundary conditions.

Element : Beam – Determination of the stress resultant



Remember! The stress resultants, such as a bending moment "m" and shear force "V" at the end of the beam, are not given, but can be represented by the given external forces.

Shear force at x=0

$$V(0) = f_{y1}$$

Bending moment at x=0

$$m(0) = -M_1$$

Shear force at x=l

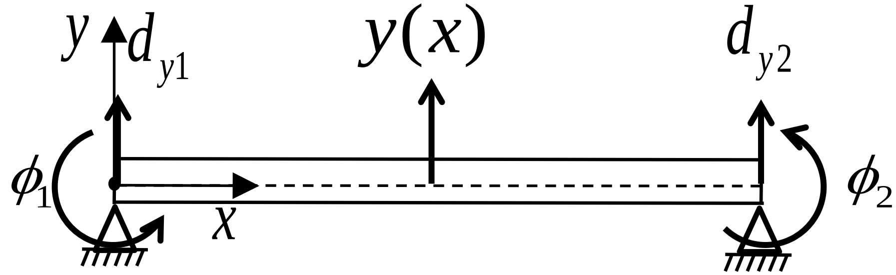
$$V(l) = -f_{y2}$$

Bending moment at x=l

$$m(l) = M_2$$

Remember: The bending moment and shear force at the ends of the beam are independent of the distributed force, but only dependent on the boundary conditions.

Element: Beam
- Derivation of the beam elemental stiffness matrix



$y(x)$: the vertical displacement variation through the element length

$$V(0) = f_{y1} \qquad V(l) = -f_{y2}$$
$$m(0) = -M_1 \qquad m(l) = M_2$$

Assume an approximation function for the vertical displacement variation through the element length to be

$$\hat{y}(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$

The complete cubic displacement function is appropriate because there are four degrees of freedom.

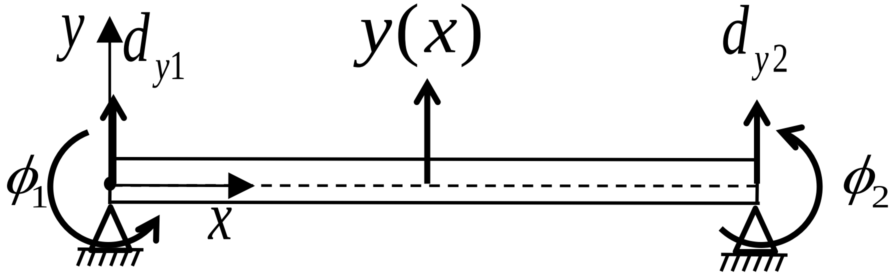
The boundary conditions $\hat{y}(0) = d_{y1},\ \hat{y}(l) = d_{y2},\ \frac{d\hat{y}}{dx}(0) = \phi_1,\ \frac{d\hat{y}}{dx}(l) = \phi_2$ should be satisfied.

$$b_0 = d_{y1} \quad , b_1 = \phi_1 \quad , b_2 = -\frac{3}{l^2}\left(d_{y1} - d_{y2}\right) - \frac{1}{l}\left(2\phi_1 + \phi_2\right)$$
$$b_3 = \frac{2}{l^3}\left(d_{y1} - d_{y2}\right) + \frac{1}{l^2}\left(\phi_1 + \phi_2\right)$$

$$\hat{y}(x) = d_{y1} + \phi_1 x + \left[-\frac{3}{l^2}\left(d_{y1} - d_{y2}\right) - \frac{1}{l}\left(2\phi_1 + \phi_2\right)\right]x^2 + \left[\frac{2}{l^3}\left(d_{y1} - d_{y2}\right) + \frac{1}{l^2}\left(\phi_1 + \phi_2\right)\right]x^3$$

$$or\ \ \hat{y}(x) = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)d_{y1} + \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)\phi_1 + \frac{1}{l^3}(-2x^3 + 3x^2 l)d_{y2} + \frac{1}{l^3}(x^3 l - x^2 l^2)\phi_2$$

Element: Beam
- Reference) Derivation of shape function Ni



$y(x)$: the vertical displacement variation through the element length

$$V(0) = f_{y1} \qquad V(l) = -f_{y2}$$
$$m(0) = -M_1 \qquad m(l) = M_2$$

$$\hat{y}(x) = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)d_{y1} + \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)\phi_1 + \frac{1}{l^3}(-2x^3 + 3x^2 l)d_{y2} + \frac{1}{l^3}(x^3 l - x^2 l^2)\phi_2$$

$$\hat{y}(x) = N_1 d_{y1} + N_2 \phi_1 + N_3 d_{y2} + N_4 \phi_2$$

⇩

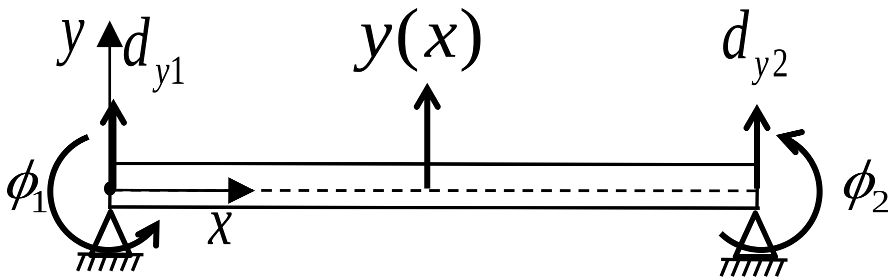
$$\hat{y}(x) = \begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix}$$

$$N_1 = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)$$
$$N_2 = \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)$$
$$N_3 = \frac{1}{l^3}(-2x^3 + 3x^2 l)$$
$$N_4 = \frac{1}{l^3}(x^3 l - x^2 l^2)$$

The shape functions and their coefficients, i.e., displacement of deformation, can be regarded as basis functions and control points of B-spline curve, respectively.

Reference) Logan, A first course in the finite element method, 3rd edition, Thomson learning, 2002, pp.138~158

Element: Beam
- Derivation of the beam elemental stiffness matrix



$v(x)$: the vertical displacement variation through the element length

$$V(0) = f_{y1} \qquad V(l) = -f_{y2}$$
$$m(0) = -M_1 \qquad m(l) = M_2$$

$$\hat{y}(x) = d_{y1} + \phi_1 x + \left[-\frac{3}{l^2}\left(d_{y1} - d_{y2}\right) - \frac{1}{l}\left(2\phi_1 + \phi_2\right) \right] x^2 + \left[\frac{2}{l^3}\left(d_{y1} - d_{y2}\right) + \frac{1}{l^2}\left(\phi_1 + \phi_2\right) \right] x^3$$

Recall that the approximation function $\hat{y}(x)$ is so constructed that **the boundary conditions of $y(x)$ are satisfied.** -> $EI\dfrac{d^3\hat{y}}{dx^3} = V(x), EI\dfrac{d^2\hat{y}}{dx^2} = m(x)$

Derive the element stiffness matrix and equations using a **direct equilibrium approach**.

$$f_{y1} = V(0) = EI\frac{d^3\hat{y}(0)}{dx^3} = \frac{EI}{l^3}(12d_{y1} + 6l\phi_1 - 12d_{y2} + 6\phi_2)$$

$$M_1 = -m(0) = -EI\frac{d^2\hat{y}(0)}{dx^2} = \frac{EI}{l^3}(6ld_{y1} + 4l^2\phi_1 - 6ld_{y2} + 2l^2\phi_2)$$

$$f_{y2} = -V(l) = -EI\frac{d^3\hat{y}(l)}{dx^3} = \frac{EI}{l^3}(-12d_{y1} - 6l\phi_1 + 12d_{y2} - 6\phi_2)$$

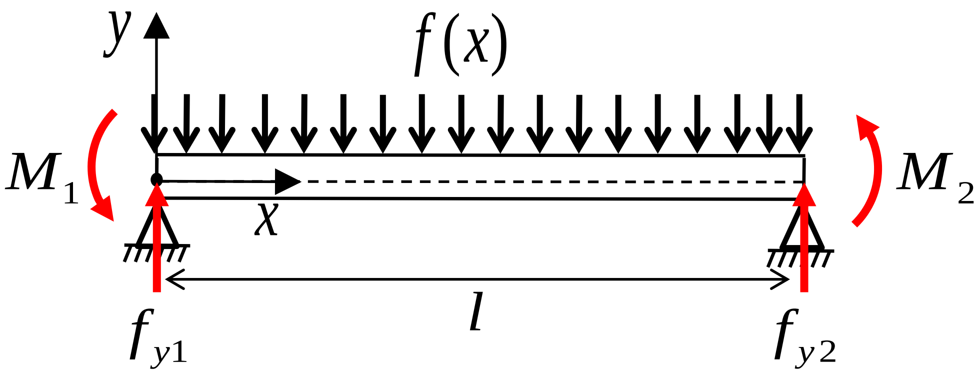
$$M_2 = m(l) = EI\frac{d^2\hat{y}(l)}{dx^2} = \frac{EI}{l^3}(6ld_{y1} + 2l^2\phi_1 - 6ld_{y2} + 4l^2\phi_2)$$

Matrix form ⇒

$$\begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} & -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} & -\frac{6EI}{l^2} & \frac{2EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} & \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} & -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix} = \begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix}$$

Reference) Logan, A first course in the finite element method, 3rd edition, Thomson learning, 2002, pp.138~158

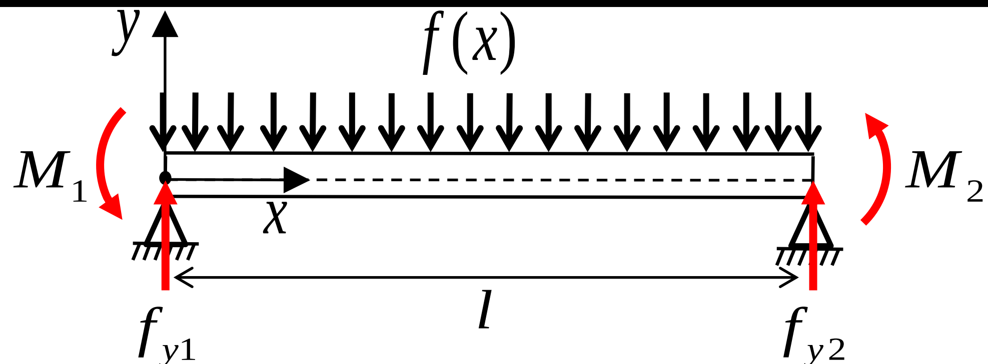
Element: Beam
- Equivalent nodal forces



$$\begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} & -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} & -\frac{6EI}{l^2} & \frac{2EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} & \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} & -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix} = \begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} + \boxed{\begin{bmatrix} f'_{y1,f(x)} \\ M'_{1,f(x)} \\ f'_{y2,f(x)} \\ M'_{2,f(x)} \end{bmatrix}}$$

equivalent nodal forces

Element: Beam
- Work-Equivalence Method



We can use the work-equivalence method to replace a distributed force by a set of discrete forces. This method is based on the concept that the work of the distributed force f(x) in going through the displacement field *y(x)* is equal to the work done by nodal force $f'_{yi,f(x)}$ and $M'_{i,f(x)}$ in going through the nodal displacement d_{y1}, and ϕ_1 for arbitrary nodal displacements.

Work done by distributed force

$$W_{distibuted} = -\int_0^l f(x) \cdot y(x)\, dx$$

In this problem, the distributed force is acting in negative direction of the static sign conventions.

Work done by nodal forces

$$W_{discrete} = M'_{1,f(x)} \cdot \phi_1 + M'_{2,f(x)} \cdot \phi_2 + f'_{y1,f(x)} \cdot d_{y1} + f'_{y2,f(x)} \cdot d_{y2}$$

Element: Beam
- Work-Equivalence Method

Work done by distributed force

$$W_{distibuted} = -\int_0^L f(x)\cdot y(x)\,dx$$

$$\hat{y}(x) = \begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix} = \mathbf{N}^T \mathbf{d}$$

$$W_{distibuted} = -\int_0^l f(x)\cdot \mathbf{N}^T\,dx \cdot \mathbf{d} = \begin{bmatrix} -\int_0^l f(x)\cdot N_1\,dx \\ -\int_0^l f(x)\cdot N_2\,dx \\ -\int_0^l f(x)\cdot N_3\,dx \\ -\int_0^l f(x)\cdot N_4\,dx \end{bmatrix}^T \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix}$$

$$N_1 = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3) \qquad N_3 = \frac{1}{l^3}(-2x^3 + 3x^2 l)$$

$$N_2 = \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3) \qquad N_4 = \frac{1}{l^3}(x^3 l - x^2 l^2)$$

Work done by nodal forces

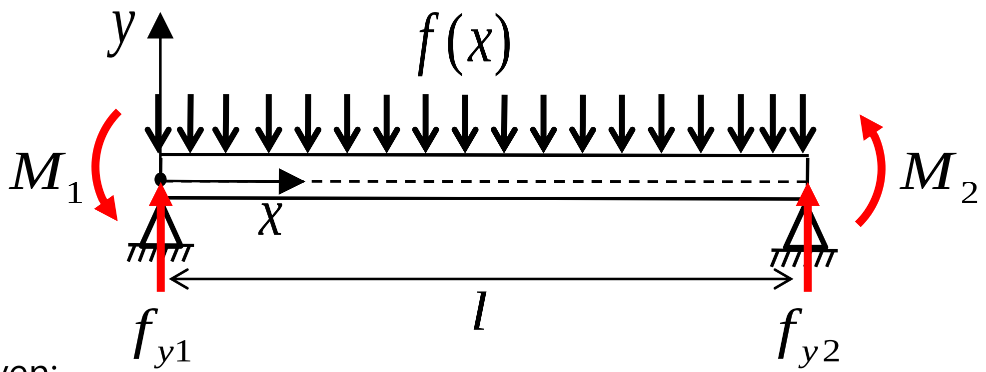
$$W_{discrete} = M'_{1,f(x)}\cdot\phi_1 + M'_{2,f(x)}\cdot\phi_2 + f'_{y1,f(x)}\cdot d_{y1} + f'_{y2,f(x)}\cdot d_{y2}$$

$$W_{discrete} = \begin{bmatrix} f'_{y1,f(x)} \\ M'_{1,f(x)} \\ f'_{y2,f(x)} \\ M'_{2,f(x)} \end{bmatrix}^T \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix}$$

Work-Equivalence

$$\begin{bmatrix} f'_{y1,f(x)} \\ M'_{1,f(x)} \\ f'_{y2,f(x)} \\ M'_{2,f(x)} \end{bmatrix} = \begin{bmatrix} -\int_0^l f(x)\cdot N_1\,dx \\ -\int_0^l f(x)\cdot N_2\,dx \\ -\int_0^l f(x)\cdot N_3\,dx \\ -\int_0^l f(x)\cdot N_4\,dx \end{bmatrix}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix



Given:

1) The concentrated forces f_{y1} and f_{y2} are exerted on the ends of the bar.
2) The moment M_1 and M_2 are exerted on the ends of the bar.
3) distributed force f(x) is applied to the element

Find:

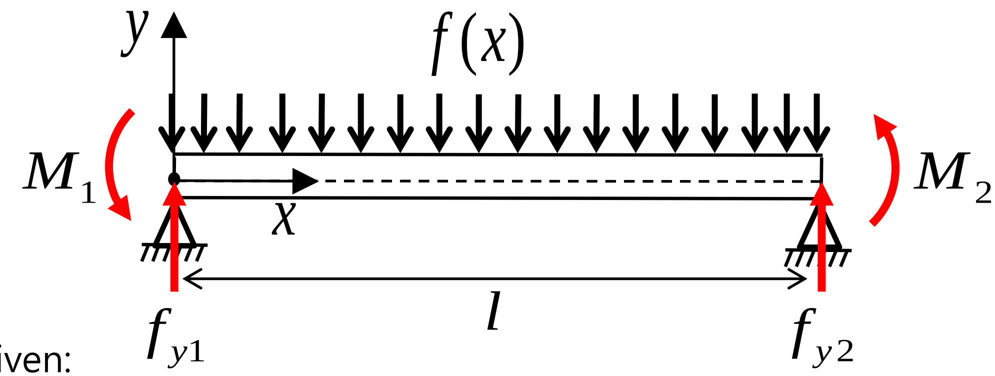
1) The vertical displacement at the ends of the bar d_{y1}, d_{y2}.
2) The rotation angle at the ends of the bar ϕ_1, ϕ_2.

$$\begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} & -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} & -\frac{6EI}{l^2} & \frac{2EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} & \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} & -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix} = \begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} + \begin{bmatrix} -\int_0^l f(x)\cdot N_1\,dx \\ -\int_0^l f(x)\cdot N_2\,dx \\ -\int_0^l f(x)\cdot N_3\,dx \\ -\int_0^l f(x)\cdot N_4\,dx \end{bmatrix}$$

2.7. Element : Beam

- Derivation of Stiffness Matrix by Applying Galerkin's Residual Method

Element : Beam – Problem Definition



Given:

1) The concentrated forces f_{y1} and f_{y2} are exerted on the ends of the bar.
2) The moment M_1 and M_2 are exerted on the ends of the bar.
3) distributed force f(x) is applied to the element

Find:

1) The vertical displacement at the ends of the bar d_{y1}, d_{y2}.
2) The rotation angle at the ends of the bar ϕ_1, ϕ_2.

Differential Equation: Deflection Curve of a Beam is derived as follows.

$V(x)$: Shear force $m(x)$: Bending moment

"Deflection Curve of a Beam"

$$\boxed{\frac{d^2y}{dx^2} = \frac{m(x)}{EI}} \Rightarrow \boxed{EI\frac{d^2y}{dx^2} = m(x)} \Rightarrow \boxed{EI\frac{d^3y}{dx^3} = -V(x)} \Rightarrow \boxed{EI\frac{d^4y}{dx^4} = -f(x)}$$

Boundary Condition

$$EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=0} = V(0) = f_{y1} \quad , EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=0} = m(0) = -M_1 \quad , EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=l} = V(l) = -f_{y2} \quad , EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=l} = m(l) = M_2$$

Element : Beam - Galerkin's Residual Method

Beam - Galerkin's Residual Method

$$\int_0^l \left[EI\frac{d^4\hat{y}(x)}{dx^4} + f(x) \right] N_i \, dx = 0 \quad ,(i=1,2,3,4)$$

$$, where \ \hat{y}(x) = N_1 d_{y1} + N_2\phi_1 + N_3 d_{y2} + N_4\phi_2 = \mathbf{N}^T \mathbf{d}$$

$$\hat{y}(x) = \begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix} \begin{bmatrix} d_{y1} \\ \phi_1 \\ d_{y2} \\ \phi_2 \end{bmatrix}$$

$$= \mathbf{N}^T \mathbf{d}$$

, where

$$N_1 = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)$$

$$N_2 = \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)$$

$$N_3 = \frac{1}{l^3}(-2x^3 + 3x^2 l)$$

$$N_4 = \frac{1}{l^3}(x^3 l - x^2 l^2)$$

Galerkin Method

the test functions N_i are chosen to play the role of the weighting functions *W*

$$\iiint_V R \, N_i \, dV = 0 \qquad ,(i=1,2,3,4)$$

→weighting function

→residual (test function $\mathbf{N}$ used)

Differential Equation

$$EI\frac{d^4 y(x)}{dx^4} + f(x) = 0$$

Boundary Conditions

$$EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=0} = V(0) = f_{y1} \quad , EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=l} = V(l) = -f_{y2}$$

$$, EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=0} = m(0) = -M_1 \quad , EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=l} = m(l) = M_2$$

Element : Beam - Galerkin's Residual Method

Beam - Galerkin's Residual Method

$$\int_0^l \left[EI \frac{d^4\hat{y}(x)}{dx^4} + f(x) \right] N_i \, dx = 0 \quad ,(i=1,2,3,4)$$

$$, where \ \hat{y}(x) = \mathbf{N}^T \mathbf{d}$$

$$N_1 = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)$$

$$N_2 = \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)$$

$$N_3 = \frac{1}{l^3}(-2x^3 + 3x^2 l)$$

$$N_4 = \frac{1}{l^3}(x^3 l - x^2 l^2)$$

integration by parts

$$\left[N_i EI \frac{d^3\hat{y}}{dx^3} \right]_0^l - \int_0^l EI \frac{d^3\hat{y}}{dx^3} \frac{dN_i}{dx} dx + \int_0^l f(x) N_i \, dx = 0 \quad ,(i=1,2,3,4)$$

integration by parts again

$$EI \int_0^l \frac{d^2 N_i}{dx^2} \frac{d^2\hat{y}}{dx^2} dx + EI \left[N_i \frac{d^3\hat{y}}{dx^3} - \frac{dN_i}{dx} \frac{d^2\hat{y}}{dx^2} \right]_0^l + \int_0^l f(x) N_i \, dx = 0 \quad ,(i=1,2,3,4)$$

$$EI \int_0^l \frac{d^2 N_i}{dx^2} \mathbf{B} \, dx \, \mathbf{d} + EI \left[N_i V - \frac{dN_i}{dx} m \right]_0^l + \int_0^l f(x) N_i \, dx = 0 \quad ,(i=1,2,3,4)$$

In matrix form,

$$EI \int_0^l \mathbf{B}^T \mathbf{B} \, dx \, \mathbf{d} = EI \left[\frac{d\mathbf{N}^T}{dx} m - \mathbf{N}^T V \right]_0^l - \int_0^l \mathbf{N}^T f(x) \, dx = 0$$

Differential Equation

$$\boxed{EI \frac{d^4 v(x)}{dx^4} + f(x) = 0}$$

Boundary Conditions

$$EI \left. \frac{d^3\hat{y}}{dx^3} \right|_{x=0} = V(0) = f_{y1}$$

$$, EI \left. \frac{d^2\hat{y}}{dx^2} \right|_{x=0} = m(0) = -M_1$$

$$, EI \left. \frac{d^3\hat{y}}{dx^3} \right|_{x=l} = V(l) = -f_{y2}$$

$$, EI \left. \frac{d^2\hat{y}}{dx^2} \right|_{x=l} = m(l) = M_2$$

$V(x)$ and $m(x)$ are substituted for d^3v/dx^3 and d^2v/dx^2 when the "x" is 0 or l, since $V(0)$, $m(0)$, $V(l)$, and $m(l)$ are the boundary conditions.

$$\frac{d^2\hat{y}}{dx^2} = \mathbf{B} \cdot \mathbf{d}$$

$$B_1 = \frac{1}{l^3}(12x - 6l)$$

$$B_2 = \frac{1}{l^3}(6xl - 4l^2)$$

$$B_3 = \frac{1}{l^3}(-12x + 6l)$$

$$B_4 = \frac{1}{l^3}(6xl - 2l^2)$$

Element : Beam - Galerkin's Residual Method

$$EI\int_0^l \mathbf{B}^T\mathbf{B}\,dx\,\mathbf{d} = \boxed{EI\left[\frac{d\mathbf{N}^T}{dx}m - \mathbf{N}^T V\right]_0^l - \int_0^l \mathbf{N}^T f(x)\,dx}$$

R.H.S

$$\boxed{EI\left[\frac{d\mathbf{N}^T}{dx}m - \mathbf{N}^T V\right]_0^l - \int_0^l \mathbf{N}^T f(x)\,dx} = m(l)\frac{d\mathbf{N}^T}{dx}(l) - V(l)\mathbf{N}^T(l) - m(0)\frac{d\mathbf{N}^T}{dx}(0) + V(0)\mathbf{N}^T(0) - \int_0^l \mathbf{N}^T f(x)\,dx$$

$$= m(l)\begin{bmatrix}0\\0\\0\\1\end{bmatrix} - V(l)\begin{bmatrix}0\\0\\1\\0\end{bmatrix} - m(0)\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + V(0)\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + \begin{bmatrix}-\int_0^l N_1 f(x)\,dx\\ -\int_0^l N_2 f(x)\,dx\\ -\int_0^l N_3 f(x)\,dx\\ -\int_0^l N_4 f(x)\,dx\end{bmatrix} = \boxed{\begin{bmatrix}V(0) - \int_0^l N_1 f(x)\,dx\\ -m(0) - \int_0^l N_2 f(x)\,dx\\ -V(l) - \int_0^l N_3 f(x)\,dx\\ m(l) - \int_0^l N_4 f(x)\,dx\end{bmatrix}}$$

, where

$$\mathbf{N}(x) = \frac{1}{l^3}\begin{bmatrix}2x^3 - 3x^2 l + l^3 & x^3 l - 2x^2 l^2 + x l^3 & -2x^3 + 3x^2 l & x^3 l - x^2 l^2\end{bmatrix}$$

$$\frac{d\mathbf{N}(x)}{dx} = \frac{1}{l^3}\begin{bmatrix}6x^2 - 6xl & 3x^2 l - 4xl^2 + l^3 & -6x^2 + 6xl & 3x^2 l - 2xl^2\end{bmatrix}$$

$$\mathbf{N}(0) = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix} \quad ,\mathbf{N}(l) = \begin{bmatrix}0 & 0 & 1 & 0\end{bmatrix}$$

$$,\left.\frac{d\mathbf{N}(x)}{dx}\right|_{x=0} = \begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix} \;,\left.\frac{d\mathbf{N}(x)}{dx}\right|_{x=l} = \begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}$$

Element : Beam - Galerkin's Residual Method

$$\overbrace{\boxed{EI\int_0^l \mathbf{B}^T\mathbf{B}\,dx}}^{\mathbf{K}}\,\mathbf{d} = \boxed{\begin{bmatrix} V(0)-\int_0^l N_1 f(x)\,dx \\ -m(0)-\int_0^l N_2 f(x)\,dx \\ -V(l)-\int_0^l N_3 f(x)\,dx \\ m(l)-\int_0^l N_4 f(x)\,dx \end{bmatrix}} = \mathbf{F}$$

Applying Galerkin's residual method and integration by parts

Differential Equation

$$\boxed{EI\frac{d^4 v(x)}{dx^4} + f = 0}$$

Boundary Conditions

$$EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=0} = V(0) = f_{y1}$$

$$, EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=0} = m(0) = -M_1$$

$$, EI\left.\frac{d^3\hat{y}}{dx^3}\right|_{x=l} = V(l) = -f_{y2}$$

$$, EI\left.\frac{d^2\hat{y}}{dx^2}\right|_{x=l} = m(l) = M_2$$

$$\mathbf{K} = EI\int_0^l \left(\mathbf{B}^T\mathbf{B}\right)dx$$

$$= \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}$$

$$\boxed{\therefore \mathbf{Kd} = \mathbf{F} \quad where,\ \mathbf{K} = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} V(0)-\int_0^l N_1 f(x)\,dx \\ -m(0)-\int_0^l N_2 f(x)\,dx \\ -V(l)-\int_0^l N_3 f(x)\,dx \\ m(l)-\int_0^l N_4 f(x)\,dx \end{bmatrix}}$$

If the "f(x)" is constant "f", then

$$\boxed{\therefore \mathbf{Kd} = \mathbf{F} \quad where,\ \mathbf{K} = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} V(0)-\frac{l}{2}f \\ -m(0)-\frac{l^2}{12}f \\ -V(l)-\frac{l}{2}f \\ m(l)+\frac{l^2}{12}f \end{bmatrix}}$$

(derivation)

$$\mathbf{K} = EI\int_0^l \left(\mathbf{B}^T\mathbf{B}\right)dx$$

$$= EI\int_0^l \begin{bmatrix} \frac{1}{l^3}(12x-6l)\frac{1}{l^3}(12x-6l) & \frac{1}{l^3}(12x-6l)\frac{1}{l^3}(6xl-4l^2) & \frac{1}{l^3}(12x-6l)\frac{1}{l^3}(-12x+6l) & \frac{1}{l^3}(12x-6l)\frac{1}{l^3}(6xl-2l^2) \\ \frac{1}{l^3}(6xl-4l^2)\frac{1}{l^3}(12x-6l) & \frac{1}{l^3}(6xl-4l^2)\frac{1}{l^3}(6xl-4l^2) & \frac{1}{l^3}(6xl-4l^2)\frac{1}{l^3}(-12x+6l) & \frac{1}{l^3}(6xl-4l^2)\frac{1}{l^3}(6xl-2l^2) \\ \frac{1}{l^3}(-12x+6l)\frac{1}{l^3}(12x-6l) & \frac{1}{l^3}(-12x+6l)\frac{1}{l^3}(6xl-4l^2) & \frac{1}{l^3}(-12x+6l)\frac{1}{l^3}(-12x+6l) & \frac{1}{l^3}(-12x+6l)\frac{1}{l^3}(6xl-2l^2) \\ \frac{1}{l^3}(6xl-2l^2)\frac{1}{l^3}(12x-6l) & \frac{1}{l^3}(6xl-2l^2)\frac{1}{l^3}(6xl-4l^2) & \frac{1}{l^3}(6xl-2l^2)\frac{1}{l^3}(-12x+6l) & \frac{1}{l^3}(6xl-2l^2)\frac{1}{l^3}(6xl-2l^2) \end{bmatrix} dx$$

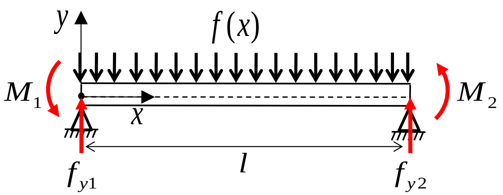
$$= EI\int_0^l \begin{bmatrix} \frac{1}{l^6}(144x^2-144xl+36l^2) & \frac{1}{l^6}(72x^2l-84xl^2+24l^3) & \frac{1}{l^6}(-144x^2+144xl-36l^2) & \frac{1}{l^6}(72x^2l-60xl^2+12l^3) \\ \frac{1}{l^6}(72x^2l-84xl^2+24l^3) & \frac{1}{l^6}(36x^2l^2-48xl^3+16l^4) & \frac{1}{l^6}(-72x^2l+84xl^2-24l^3) & \frac{1}{l^6}(36x^2l^2-36xl^3+8l^4) \\ \frac{1}{l^6}(-144x^2+144xl-36l^2) & \frac{1}{l^6}(-72x^2l+84xl^2-24l^3) & \frac{1}{l^6}(144x^2-144xl+36l^2) & \frac{1}{l^6}(-72x^2l+60xl^2-12l^3) \\ \frac{1}{l^6}(72x^2l-60xl^2+12l^3) & \frac{1}{l^6}(36x^2l^2-36xl^3+8l^4) & \frac{1}{l^6}(-72x^2l+60xl^2-12l^3) & \frac{1}{l^3}(36x^2l^2-24xl^3+4l^4) \end{bmatrix} dx$$

$$= \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}$$

2.8. Element : Beam

- Comparison between "Direct Equilibrium Approach" and "Galerkin's Residual Method"

Solutions of D/E using Galerkin's Residual Method



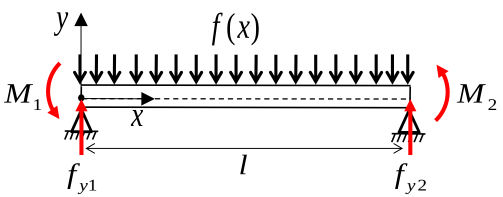
Differential Equation

$$EI\frac{d^4y(x)}{dx^4}+f(x)=0 \Rightarrow \frac{2EI}{l^3}\begin{bmatrix}6 & 3l & -6 & 3l\\ 3l & 2l^2 & -3l & l^2\\ -6 & -3l & 6 & -3l\\ 3l & l^2 & -3l & 2l^2\end{bmatrix}\begin{bmatrix}d_{y1}\\ \phi_1\\ d_{y2}\\ \phi_2\end{bmatrix}=\begin{bmatrix}f_{y1}\\ M_1\\ f_{y2}\\ M_2\end{bmatrix}+\begin{bmatrix}-\int_0^l f(x)\cdot N_1\,dx\\ -\int_0^l f(x)\cdot N_2\,dx\\ -\int_0^l f(x)\cdot N_3\,dx\\ -\int_0^l f(x)\cdot N_4\,dx\end{bmatrix}$$

$$\hat{y}(x)=b_0+b_1x+b_2x^2+b_3x^3$$

Galerkin's residual method

Direct equilibrium approach



$$\frac{2EI}{l^3}\begin{bmatrix}6 & 3l & -6 & 3l\\ 3l & 2l^2 & -3l & l^2\\ -6 & -3l & 6 & -3l\\ 3l & l^2 & -3l & 2l^2\end{bmatrix}\begin{bmatrix}d_{y1}\\ \phi_1\\ d_{y2}\\ \phi_2\end{bmatrix}=\begin{bmatrix}f_{y1}\\ M_1\\ f_{y2}\\ M_2\end{bmatrix}+\begin{bmatrix}-\int_0^l f(x)\cdot N_1\,dx\\ -\int_0^l f(x)\cdot N_2\,dx\\ -\int_0^l f(x)\cdot N_3\,dx\\ -\int_0^l f(x)\cdot N_4\,dx\end{bmatrix}$$

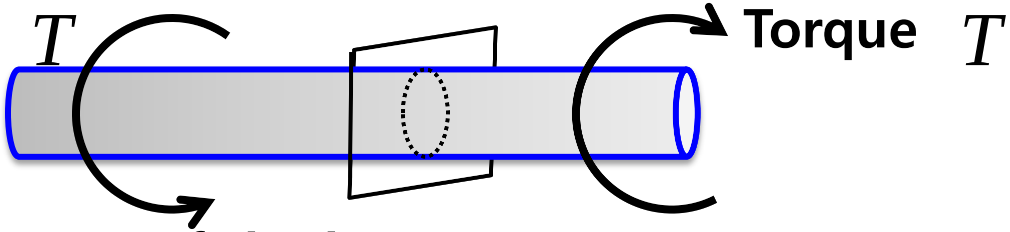
※ superposition of stiffness matrix

2.9. Element : Shaft

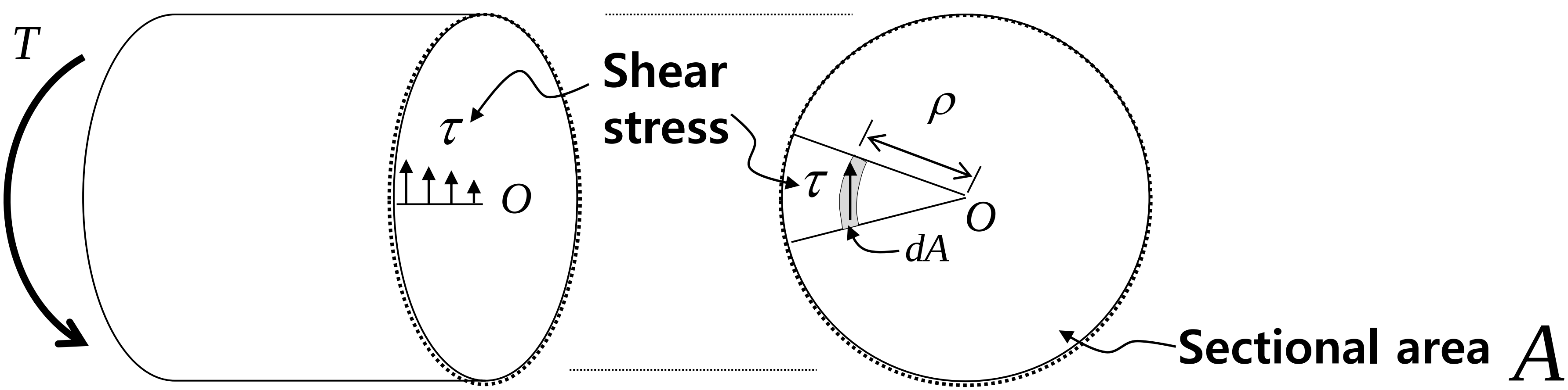
- Derivation of Stiffness Matrix by Applying Galerkin's Residual Method

Shear Stress in torsion

Deformation of a circular bar in pure torsion



Free-body diagram of the bar



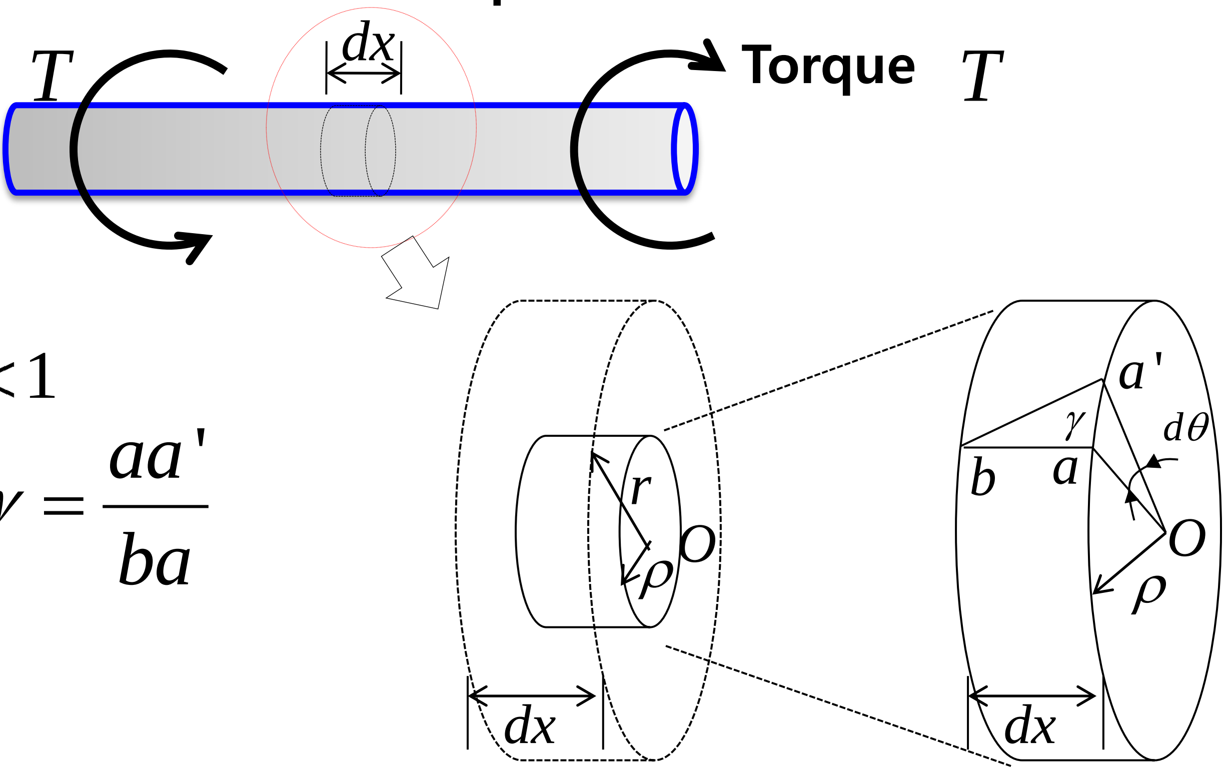
Shear force acting on the area dA : τdA

Resultant moment about a longitudinal axis through point O is equal to the torque :

$$T = \int_A \rho \tau dA$$

Shear Strain in torsion

Deformation of a circular bar in pure torsion



$$aa' = \rho\, d\theta$$

$$ba = dx$$

Assume, $\gamma \ll 1$

$$\gamma \approx \tan\gamma = \frac{aa'}{ba}$$

Shear strain

$$\gamma = \rho \frac{d\theta}{dx}$$

Relation between the torque and the angle of twist

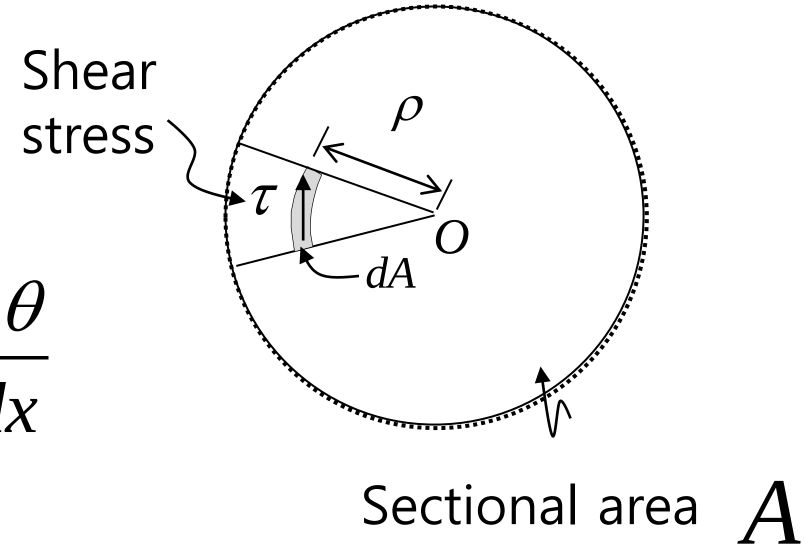
Shear force acting on area dA

$$\tau\, dA$$

Hooke's law in shear deformation

$$\tau = G\gamma$$

Shear strain $\gamma = \rho \dfrac{d\theta}{dx}$



$$\tau\, dA = G\rho \frac{d\theta}{dx} dA$$

Resultant moment about a longitudinal axis through the point O is equal to the torque:

$$T = \int_A \rho\, \tau\, dA = \int_A G \frac{d\theta}{dx} \rho^2 dA = G \frac{d\theta}{dx} \int_A \rho^2 dA$$

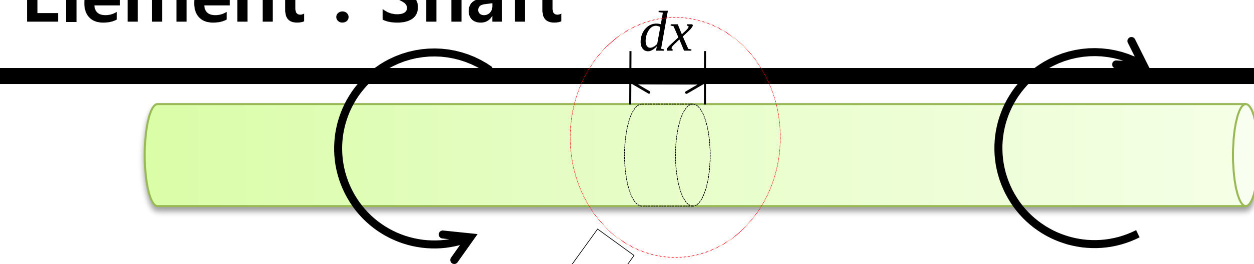
Relation between the torque and the angle of twist

$$T = GJ \frac{d\theta}{dx}$$

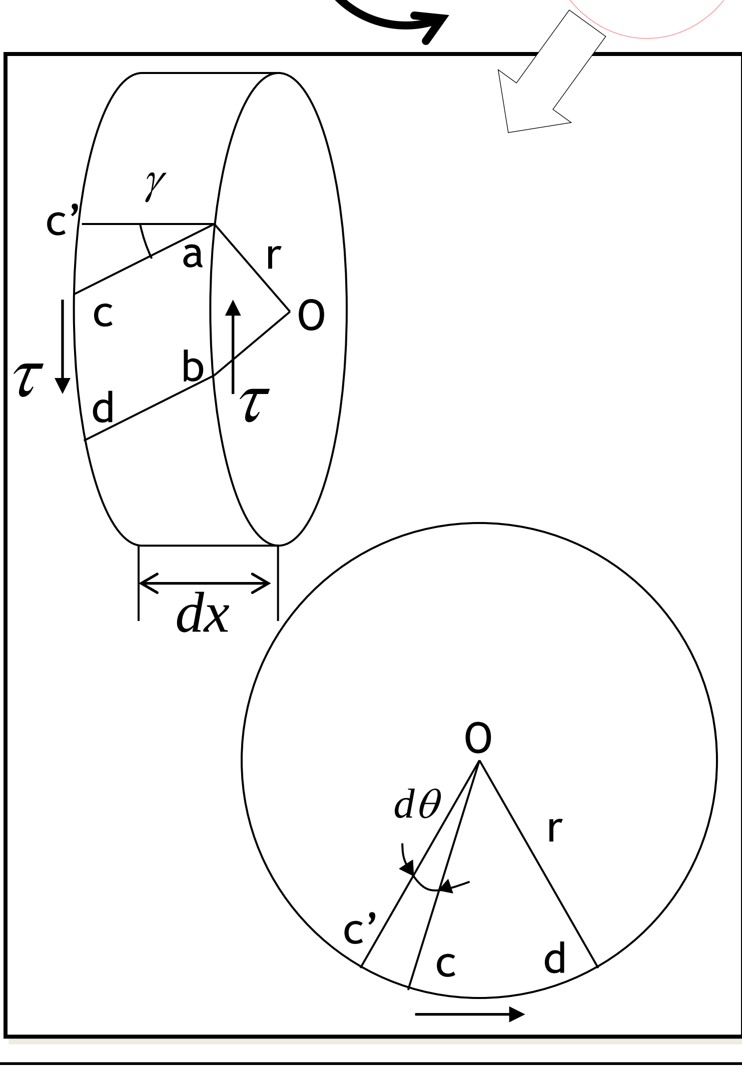
Polar moment of inertia $J = \int_A \rho^2 dA$

Element : Shaft*

τ : shear stress



Shaft



① $cc' = rd\theta$

② **assume,** $\gamma \ll 1$

$$\gamma \approx \tan\gamma = \frac{cc'}{ac'} = \frac{rd\theta}{dx}$$

③ **let** $\dfrac{d\theta}{dx} = \phi$ **, (** ϕ **: angle of twist per unit length)**

$$\gamma = r\phi$$

④ Hooke's law in shear for the linearly elastic material

$$\tau = G\gamma = Gr\phi$$

G : shear modulus of elasticity

γ : shear strain

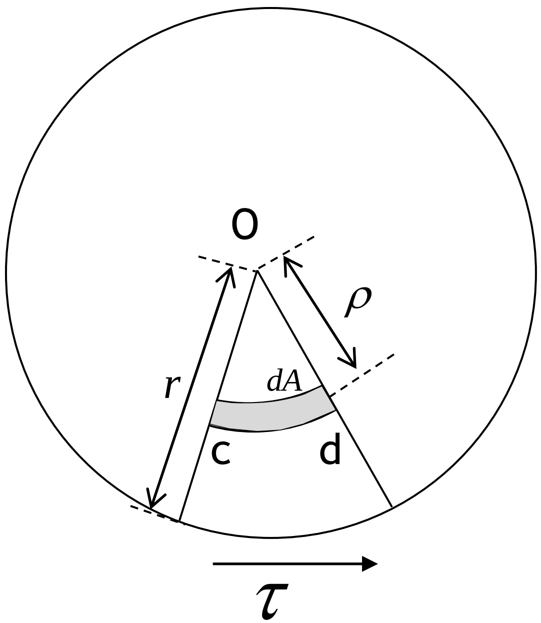
Cf.) $\sigma = E\varepsilon$

*Gere, J.M., Mechanics of Materials, 6th Edition, Thomson, 2006, pp 189~195

Element : Shaft*

$\frac{d\theta}{dx} = \phi \quad , \tau = G\gamma = Gr\phi$

G : Shear Modulus



⑤ the shear stress at an interior point (radius ρ)

$$\tau = G\rho\phi$$

⑥ we consider an element of area dA located at radial distance ρ

☞ shear force acting on the element : $\tau\, dA = G\rho\phi\, dA$

☞ moment : $\rho \times \tau\, dA = G\phi\rho^2\, dA$

⑦ the resultant moment equal to the torque

$$T = \int_A G\phi\rho^2 dA = G\phi\int_A \rho^2 dA = G\phi\, J \quad \Rightarrow \quad T = GJ\frac{d\theta}{dx}$$

※ J : **Polar Moment of Inertia** $\left(J = \int_A \rho^2 dA\right)$

⑧ for a bar in pure torsion, the total angle of twist θ, equal to the rate of twist times the length of the bar

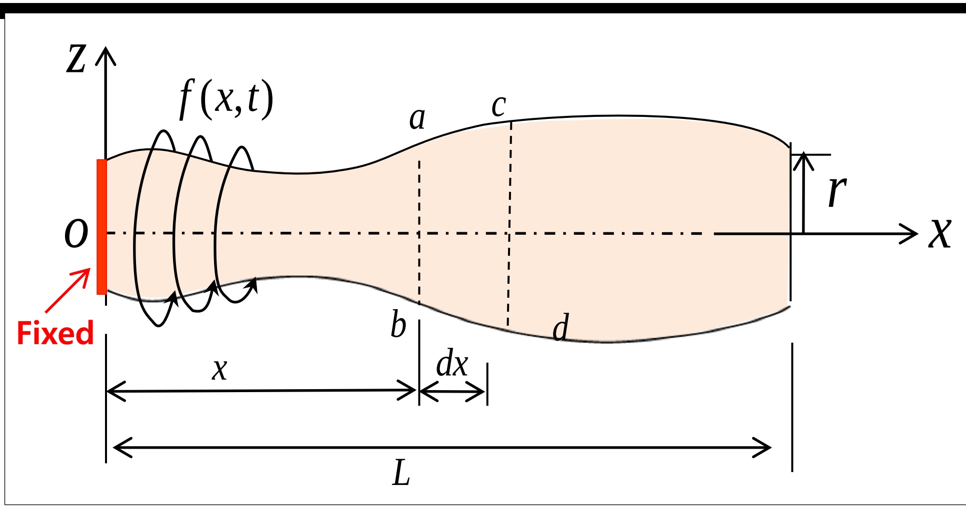
$$\theta = \phi l$$

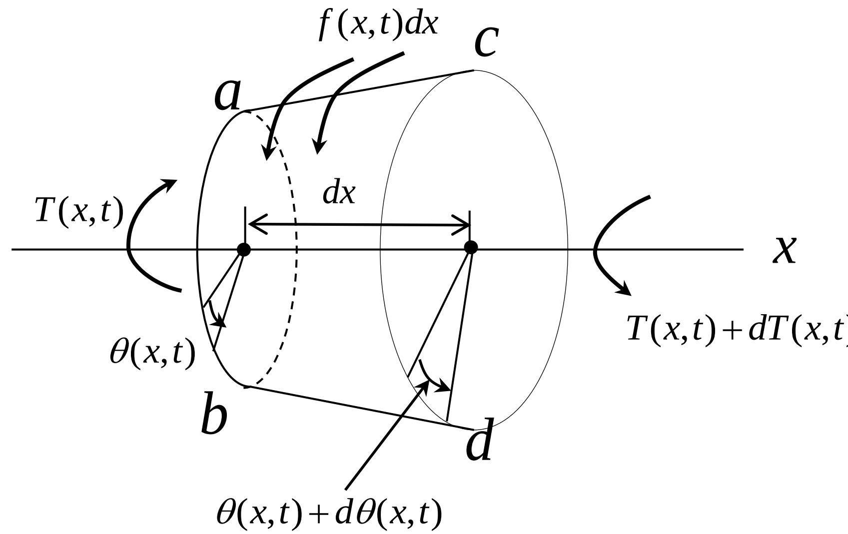
$$T = G\phi\, J = \frac{GJ}{l}\theta$$

$$\boxed{\theta = \frac{T\,l}{GJ}}$$

*Gere, J.M., Mechanics of Materials, 6th Edition, Thomson, 2006, pp 189~195

Equation of Motion for Torsional Vibration of a Shaft





Inertia torque action on an element of length dx

$$I\,dx\frac{\partial^2\theta}{\partial t^2} \qquad ,I: \text{ mass polar moment of inertia of the shaft per unit length}$$

The application of Newton's second law yields the equation of motion

$$(T+dT)+fdx-T=I\,dx\frac{\partial^2\theta}{\partial t^2}$$

$$dT=\frac{\partial T}{\partial x}dx \quad \hookrightarrow \quad \frac{\partial T(x,t)}{\partial x}dx+f(x,t)dx=I\,dx\frac{\partial^2\theta(x,t)}{\partial t^2}$$

divided by $\quad \hookrightarrow \quad \dfrac{\partial T(x,t)}{\partial x}+f(x,t)=I\dfrac{\partial^2\theta(x,t)}{\partial t^2}$

Rao S.S., Mechanical Vibrations, Fourth Edition, Prentice-Hall, 2004, pp.606-607

Equation of Motion for Torsional Vibration of a Shaft

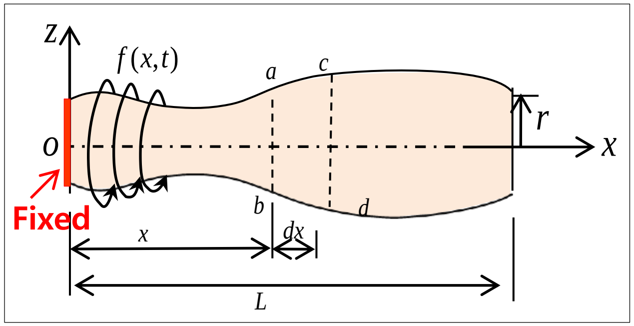
Equation of motion

$$\frac{\partial T(x,t)}{\partial x} + f(x,t) = I\frac{\partial^2 \theta(x,t)}{\partial t^2} \quad \leftarrow \textbf{nonhomogeneous}$$

homogeneous

$$c.f. \quad U_{tt} = c^2 U_{xx}$$

$$J = \int_A \rho^2 dA \qquad T = GJ\frac{d\theta}{dx}$$



Relation between torsional deflection and the twisting moment

$$T(x,t) = GJ(x)\frac{\partial \theta(x,t)}{\partial x}$$

G : Shear modulus

GJ : Torsional stiffness

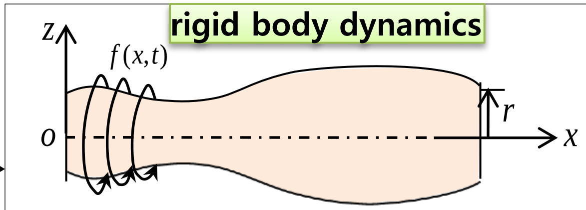
$$\frac{\partial}{dx}\left[GJ(x)\frac{\partial \theta(x,t)}{\partial x}\right] + f(x,t) = I\frac{\partial^2 \theta(x,t)}{\partial t^2}$$

if J is constant

$$GJ\frac{\partial^2 \theta(x,t)}{\partial x^2} + f(x,t) = I\frac{\partial^2 \theta(x,t)}{\partial t^2}$$

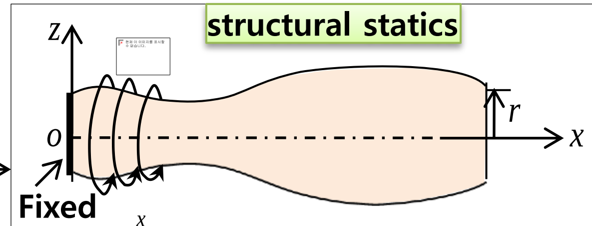
structural vibration

rigid body dynamics



$$f(x,t) = I\frac{\partial^2 \theta(x,t)}{\partial t^2}$$

structural statics



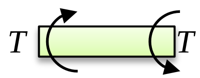
$$GJ\frac{\partial^2 \theta(x)}{\partial x^2} + f(x) = 0$$

Rao S.S., Mechanical Vibrations, Fourth Edition, Prentice-Hall, 2004, pp.606-607

Element : Shaft

T : Torque
l : length
G : Shear Modulus
J : Polar Moment of Inertia

Shaft



$$GJ\frac{d^2\theta(x)}{dx^2}=0 \quad \Rightarrow \quad \int_0^L W_l\left(GJ\frac{d^2\theta(x)}{dx^2}\right)dx=0 \quad \xrightarrow{\theta(x)=c_0+c_1x} \quad \frac{GJ}{l}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}=\begin{bmatrix}-T\\ T\end{bmatrix}$$

Torsion

- Reference : Chapter.1 Bar

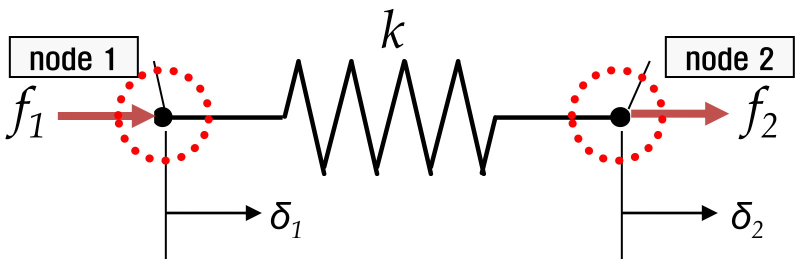
Bar

$$EA\frac{d^2u(x)}{dx^2}=0 \quad \Rightarrow \quad \int_0^L W_l\left(EA\frac{d^2u(x)}{dx^2}\right)dx=0 \quad \xrightarrow{u(x)=a_0+a_1x} \quad \frac{EA}{l}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\begin{bmatrix}-P\\ P\end{bmatrix}$$

Tension

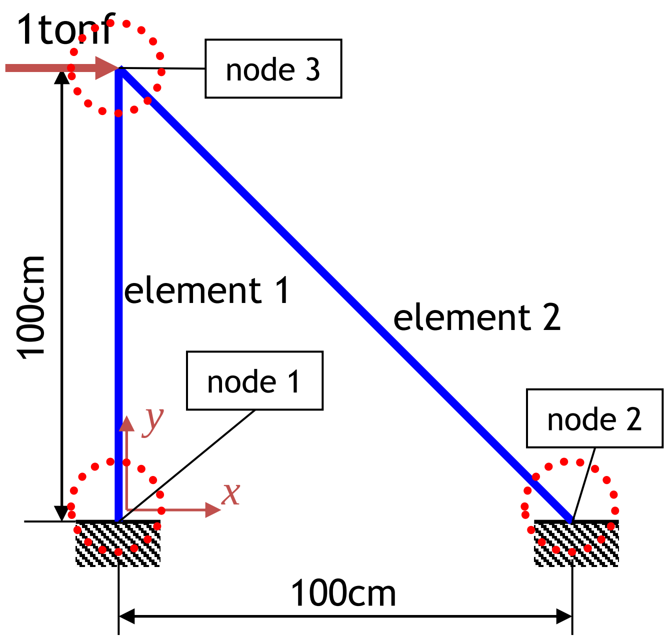
2.10. Superposition of Stiffness Matrix and Coordinate Transformation

Node

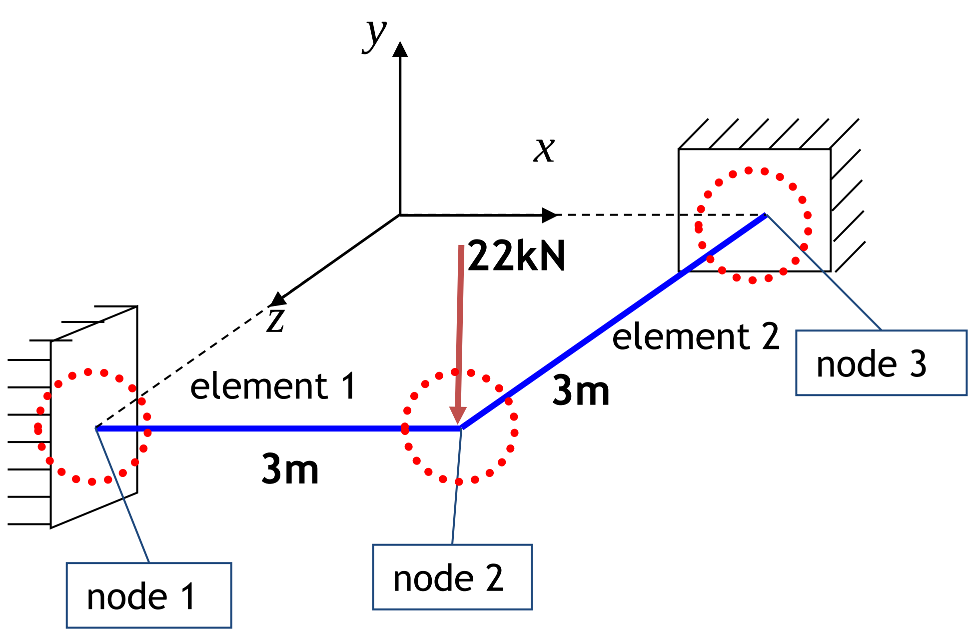


The nodes are points at which equilibrium will be enforced and displacement found. They are generally located at the ends of the elements for most common structural shapes such as bars and beams.*

ex.1)



ex.2)



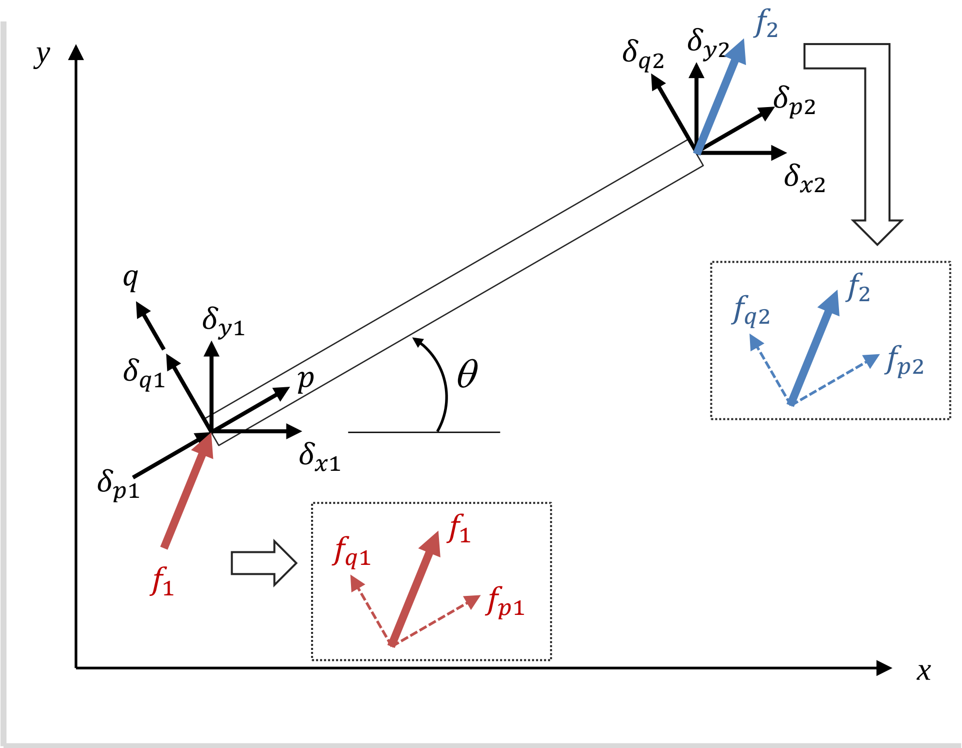
*Sennet, R.E., Matrix Analysis of Structres, Prentice-Hall, 1994, p.4

Representation of the elemental displacements and forces in terms of the global displacements and forces

- The *elemental displacement* $\delta_{p1}, \delta_{q1}, \delta_{p2}, \delta_{q2}$ are parallel and perpendicular to the member coordinate system p and q

- The *elemental forces* which act in the elemental coordinate system p and q are denoted by $f_{p1}, f_{q1}, f_{p2}, f_{q2}$ in order to distinguish them from the global force f_1 and f_2

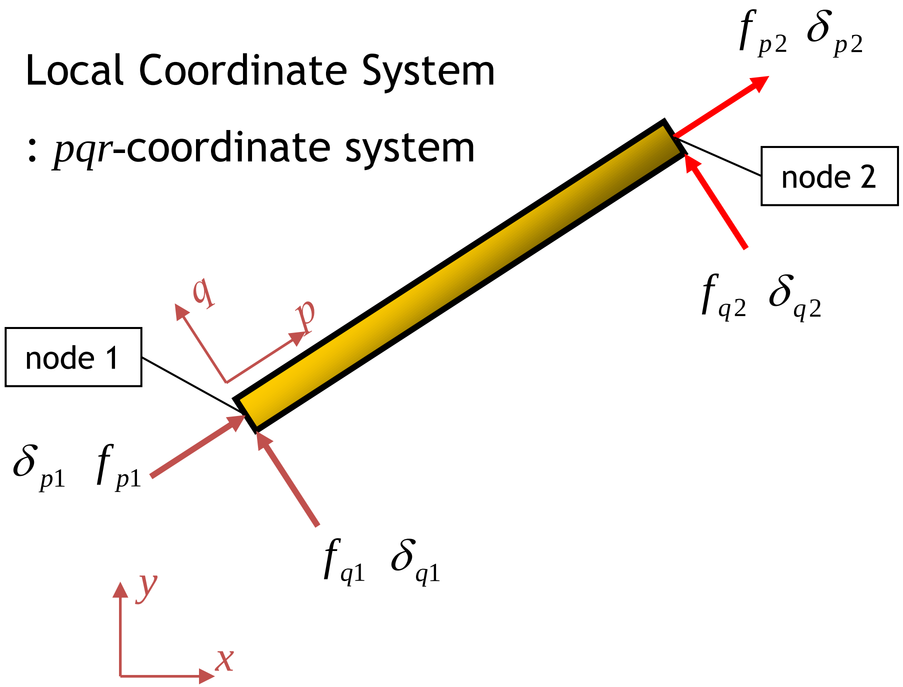
- Express the elemental displacements in terms of the global displacements.



Coordinate System

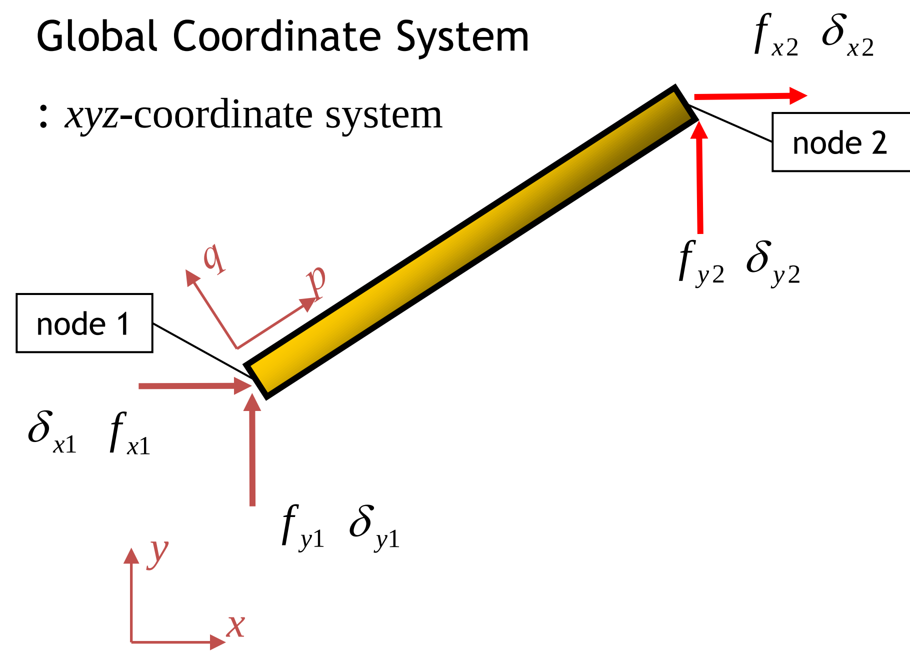
Local Coordinate System

: *pqr*-coordinate system

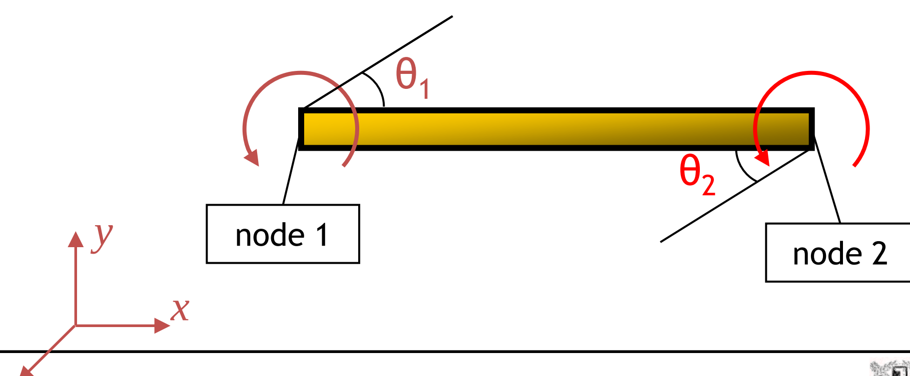


Global Coordinate System

: *xyz*-coordinate system

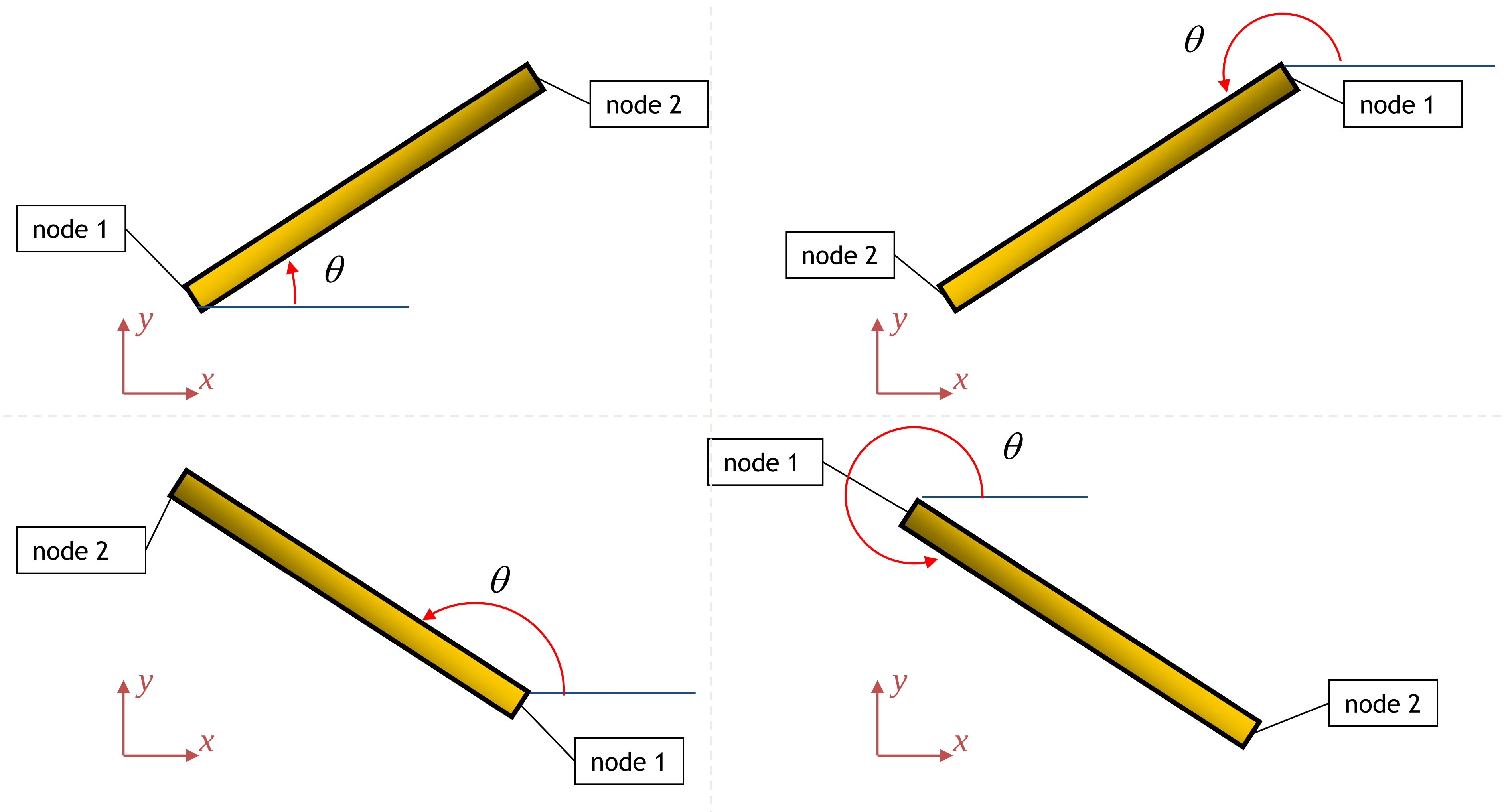


Sign Convention : positive moment



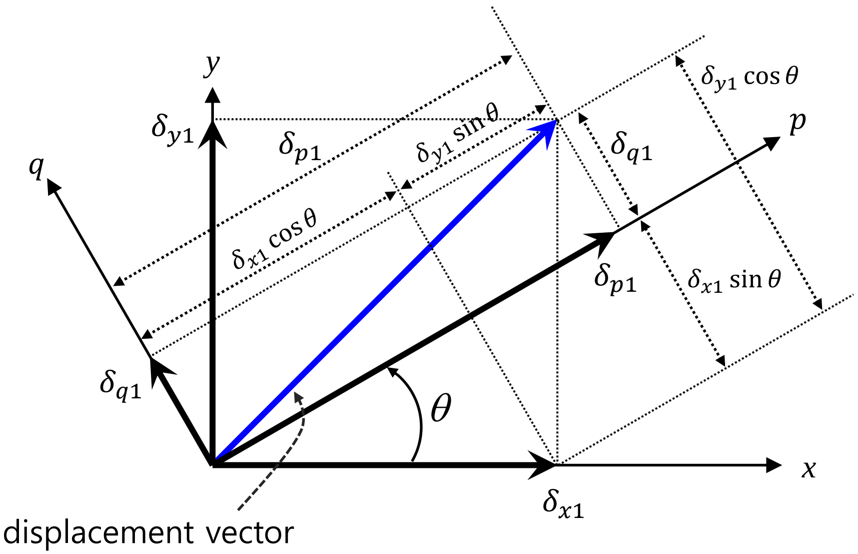
Angle

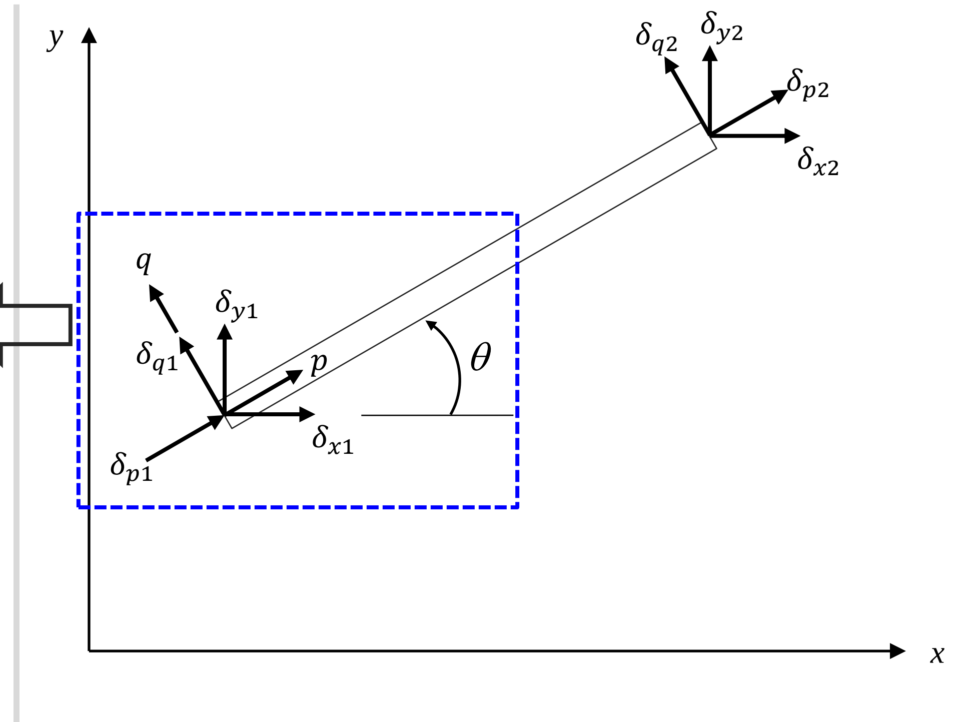
- Angles in global coordinate system : counterclockwise at the lower number of node



Representation of the elemental displacements in terms of the global displacements

- Consider the vector displacement of the left end of the member.





$$\delta_{p1} = \delta_{x1}\cos\theta + \delta_{y1}\sin\theta$$

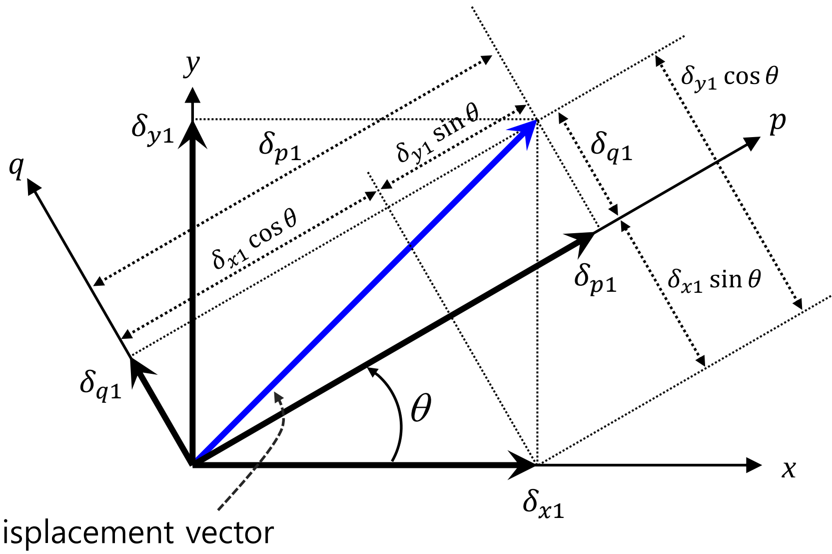
$$\delta_{q1} = \delta_{y1}\cos\theta - \delta_{x1}\sin\theta$$

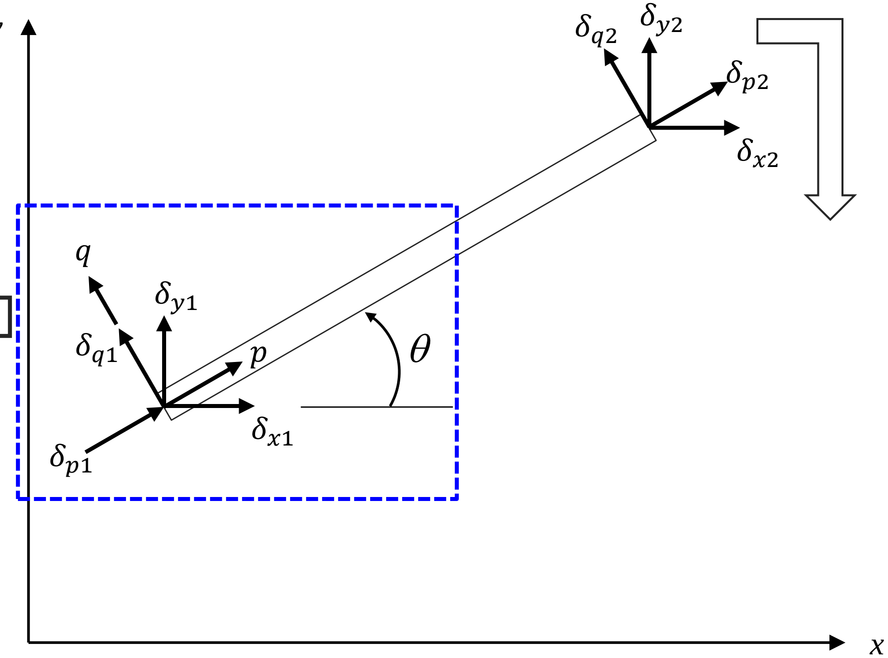
- The same relationships between displacements will also exist at the right end of the member

$$\delta_{p2} = \delta_{x2}\cos\theta + \delta_{y2}\sin\theta$$

$$\delta_{q2} = \delta_{y2}\cos\theta - \delta_{x2}\sin\theta$$

Representation of the elemental displacements in terms of the global displacements





$$\delta_{p1} = \delta_{x1}\cos\theta + \delta_{y1}\sin\theta$$

$$\delta_{q1} = \delta_{y1}\cos\theta - \delta_{x1}\sin\theta$$

$$\delta_{p2} = \delta_{x2}\cos\theta + \delta_{y2}\sin\theta$$

$$\delta_{q2} = \delta_{y2}\cos\theta - \delta_{x2}\sin\theta$$

transformation matrix

$$\begin{bmatrix} \delta_{p1} \\ \delta_{q1} \\ \delta_{p2} \\ \delta_{q1} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x1} \\ \delta_{y1} \end{bmatrix}$$

$$\boxed{[\delta_{pq}] = [\mathbf{T}][\delta_{xy}]}$$

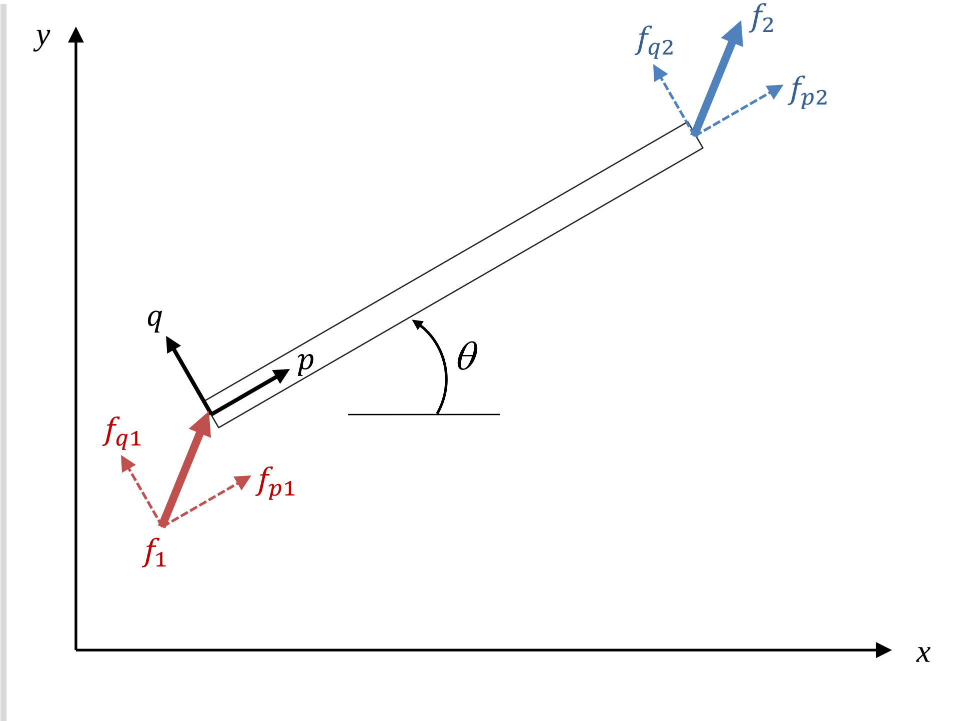
Representation of the elemental forces in terms of the global forces

- Forces in 2 Dimension $(xy \rightarrow pq)$:

In the same way,

$$\begin{bmatrix} f_{p1} \\ f_{q1} \\ f_{p2} \\ f_{q2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix}$$

$$\boxed{[\mathbf{F}_{pq}] = [\mathbf{T}][\mathbf{F}_{xy}]}$$

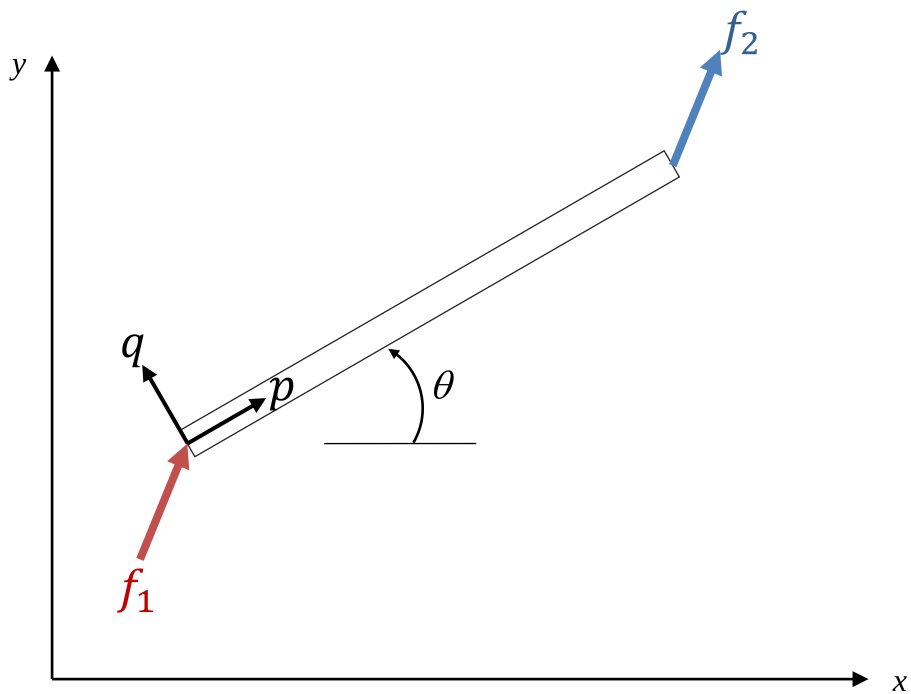


- Forces and Moment in 3 Dimension $(xyz \rightarrow pqr)$

$$\begin{bmatrix} f_{p1} \\ f_{q1} \\ M_{z1} \\ f_{p2} \\ f_{q2} \\ M_{z2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & \sin\theta & 0 \\ 0 & 0 & 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{x1} \\ f_{y1} \\ M_{z1} \\ f_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix}$$

Element : 2-Dimensional Bar

- Solution of 2-Dimensional Bar



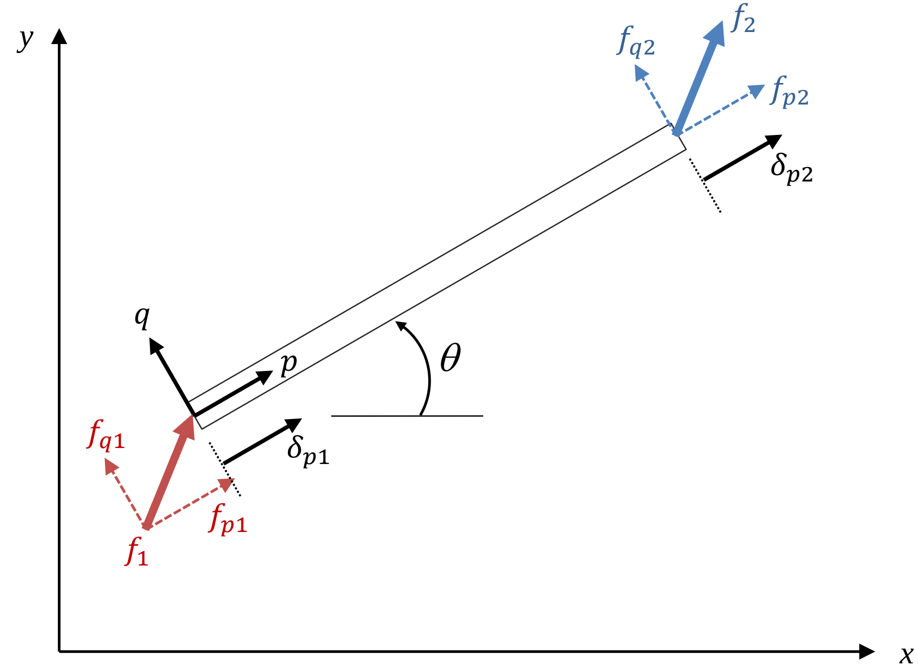
Step1. Find stiffness matrix in the local coordinate system (***pq-coordinate system)***

Step2. Find transformation matrix between the local and the global coordinate system

Step3. Find stiffness matrix in the global coordinate system (***xy-coordinate system)***

Element : 2-Dimensional Bar

Step1. Find stiffness matrix in the local coordinate system (***pq-coordinate system)***



Notation

δ_{pi} : displacement parallel to the p *axis* at node i

δ_{qi} : displacement parallel to the q *axis* at node i

f_{pi} : force parallel to the p *axis* at node i

f_{qi} : force parallel to the q *axis* at node i

$$\begin{bmatrix} f_{p1} \\ f_{q1} \\ f_{p2} \\ f_{q2} \end{bmatrix} = \begin{bmatrix} k & 0 & -k & 0 \\ 0 & 0 & 0 & 0 \\ -k & 0 & k & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{p1} \\ \delta_{q1} \\ \delta_{p2} \\ \delta_{q2} \end{bmatrix} \Rightarrow$$

①

$$\boxed{[\mathbf{F}_{pq}] = [\mathbf{K}_{pq}][\delta_{pq}]}$$

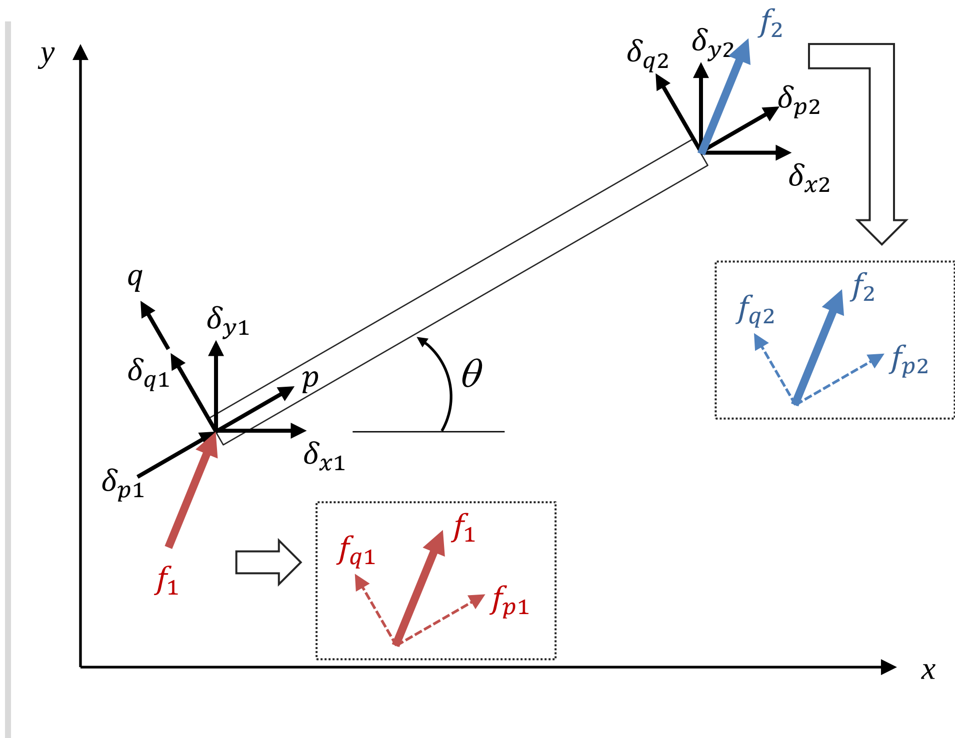
Element : 2-Dimensional Bar

Step2. Find transformation matrix between the local and the global coordinate system

(1) the forces with respect to the global coordinate system

$$\begin{bmatrix} f_{p1} \\ f_{q1} \\ f_{p2} \\ f_{q2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix}$$

② $\boxed{[\mathbf{F}_{pq}] = [\mathbf{T}][\mathbf{F}_{xy}]}$



(2) the displacements with respect to the global coordinate system

$$\begin{bmatrix} \delta_{p1} \\ \delta_{q1} \\ \delta_{p2} \\ \delta_{q2} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

③ $\boxed{[\delta_{pq}] = [\mathbf{T}][\delta_{xy}]}$

Element : 2-Dimensional Bar

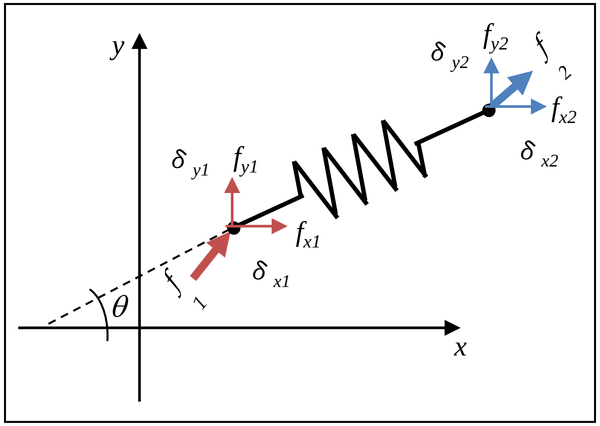
Step3. Find stiffness matrix in the global coordinate system (***xy-coordinate system)***

① $[\mathbf{F}_{pq}] = [\mathbf{K}_{pq}][\delta_{pq}]$ $\Rightarrow$ $[\mathbf{T}][\mathbf{F}_{xy}] = [\mathbf{K}_{pq}][\mathbf{T}][\delta_{xy}]$

② $[\mathbf{F}_{pq}] = [\mathbf{T}][\mathbf{F}_{xy}]$

③ $[\delta_{pq}] = [\mathbf{T}][\delta_{xy}]$

multiply $[\mathbf{T}]^{-1} = [\mathbf{T}]^T$



$$[\mathbf{F}_{xy}] = [\mathbf{T}]^T[\mathbf{K}_{pq}][\mathbf{T}][\delta_{xy}]$$

C:cos
S:sin

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = \begin{bmatrix} C & -S & 0 & 0 \\ S & C & 0 & 0 \\ 0 & 0 & C & -S \\ 0 & 0 & S & C \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} C & S & 0 & 0 \\ -S & C & 0 & 0 \\ 0 & 0 & C & S \\ 0 & 0 & -S & C \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

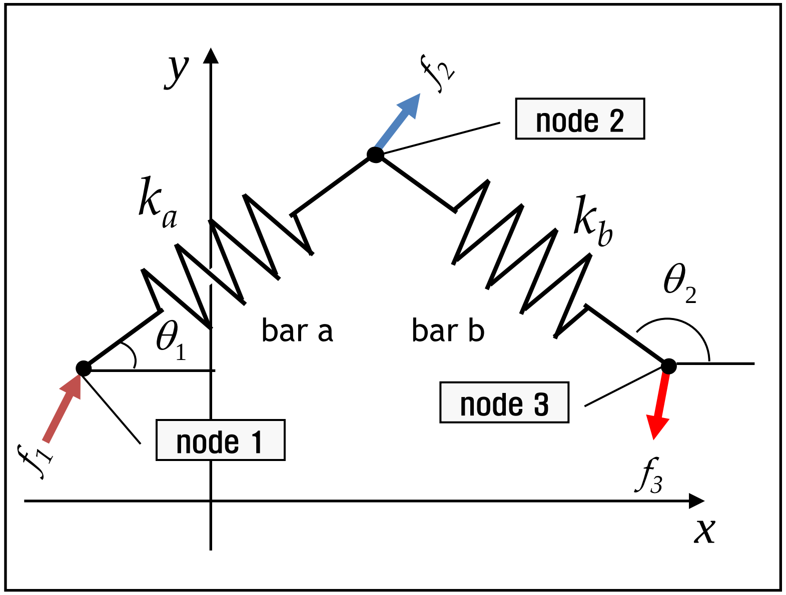
$$\therefore \begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

$$\boxed{[\mathbf{F}_{xy}] = [\mathbf{K}_{xy}][\delta_{xy}]}$$ stiffness equation

Element : 2-Dimensional Bar

C:cos , S:sin

ex.) Find a stiffness equation of the following system:



(1) Stiffness equation of bar a

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k_a \begin{bmatrix} C^2\theta_1 & C\theta_1 S\theta_1 & -C^2\theta_1 & -C\theta_1 S\theta_1 \\ C\theta_1 S\theta_1 & S^2\theta_1 & -C\theta_1 S\theta_1 & -S^2\theta_1 \\ -C^2\theta_1 & -C\theta_1 S\theta_1 & C^2\theta_1 & C\theta_1 S\theta_1 \\ -C\theta_1 S\theta_1 & -S^2\theta_1 & C\theta_1 S\theta_1 & S^2\theta_1 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

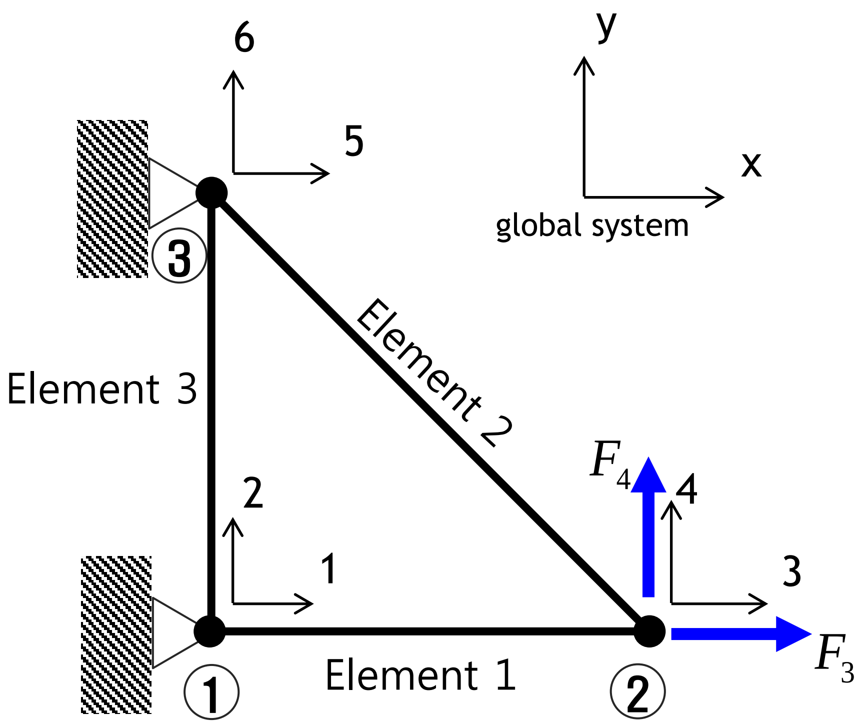
(2) Stiffness equation of bar b

$$\begin{bmatrix} f_{x2} \\ f_{y2} \\ f_{x3} \\ f_{y3} \end{bmatrix} = k_b \begin{bmatrix} C^2\theta_2 & C\theta_2 S\theta_2 & -C^2\theta_2 & -C\theta_2 S\theta_2 \\ C\theta_2 S\theta_2 & S^2\theta_2 & -C\theta_2 S\theta_2 & -S^2\theta_2 \\ -C^2\theta_2 & -C\theta_2 S\theta_2 & C^2\theta_2 & C\theta_2 S\theta_2 \\ -C\theta_2 S\theta_2 & -S^2\theta_2 & C\theta_2 S\theta_2 & S^2\theta_2 \end{bmatrix} \begin{bmatrix} \delta_{x2} \\ \delta_{y2} \\ \delta_{x3} \\ \delta_{y3} \end{bmatrix}$$

(3) Superposition

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \\ f_{x3} \\ f_{y3} \end{bmatrix} = \begin{bmatrix} k_a C^2\theta_1 & k_a C\theta_1 S\theta_1 & -k_a C^2\theta_1 & -k_a C\theta_1 S\theta_1 & 0 & 0 \\ k_a C\theta_1 S\theta_1 & k_a S^2\theta_1 & -k_a C\theta_1 S\theta_1 & -k_a S^2\theta_1 & 0 & 0 \\ -k_a C^2\theta_1 & -k_a C\theta\ S\theta_1 & k_a C^2\theta_1 + k_b C^2\theta_2 & k_a C\theta_1 S\theta_1 + k_b C\theta_2 S\theta_2 & -k_b C^2\theta_2 & -k_b C\theta_2 S\theta_2 \\ -k_a C\theta_1 S\theta_1 & -k_a S^2\theta_1 & k_a C\theta_1 S\theta_1 + k_b C\theta_2 S\theta_2 & k_a S^2\theta_1 + k_b S^2\theta_2 & -k_b C\theta\ S\theta & -k_b S^2\theta_2 \\ 0 & 0 & -k_b C^2\theta_2 & -k_b C\theta_2 S\theta_2 & k_b C^2\theta_2 & k_b C\theta_2 S\theta_2 \\ 0 & 0 & -k_b C\theta_2 S\theta_2 & -k_b S^2\theta_2 & k_b C\theta_2 S\theta_2 & k_b S^2\theta_2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \\ \delta_{x3} \\ \delta_{y3} \end{bmatrix}$$

Truss Examples



The system displacements with respect to the global coordinate system are as follows:

Node	System displacements
1	u_1 , u_2
2	u_3 , u_4
3	u_5 , u_6

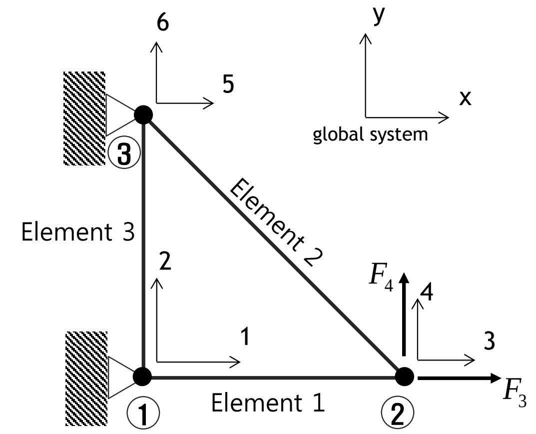
EA = constant

Truss Examples

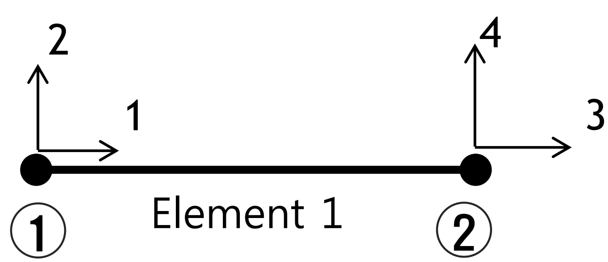
$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

For element number 1, consider nodes number 1 and 2 as the left and right ends of the member, respectively.

Thus, for element number 1, with $\theta_{1\text{-}2} = 0°$ and C = 1, S = 0, we have



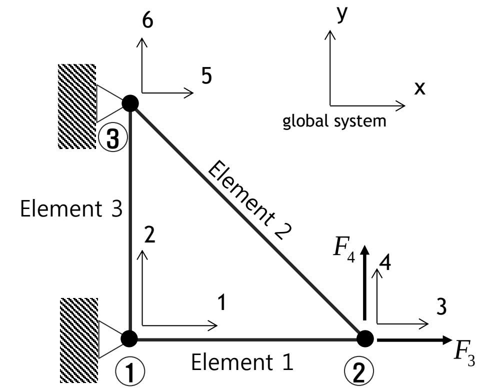
$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EAL \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$



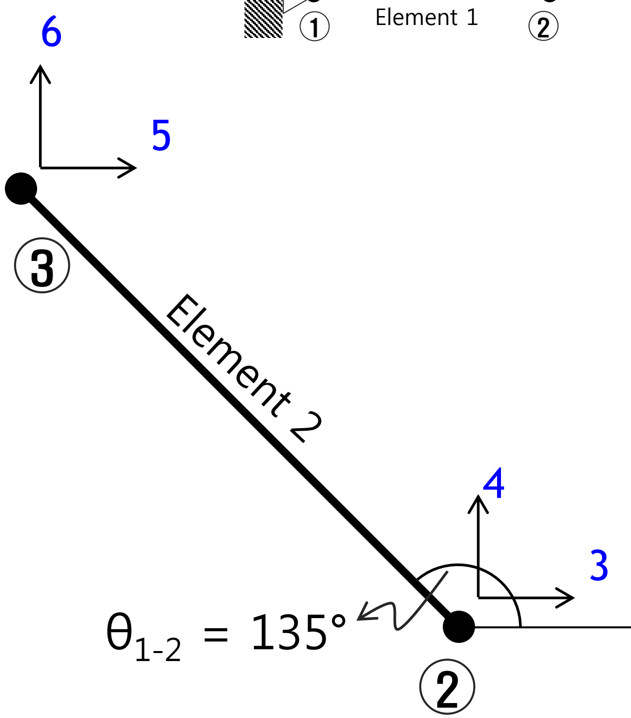
Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

For element number 2, nodes 2 and 3 locate the left and right ends of the member.



Therefore $\theta_{2\text{-}3}$ = 135°, and $C = -1/\sqrt{2}$, $S = 1/\sqrt{2}$
Thus,



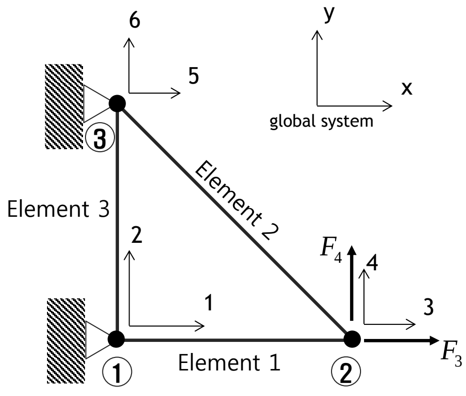
$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$

Truss Examples

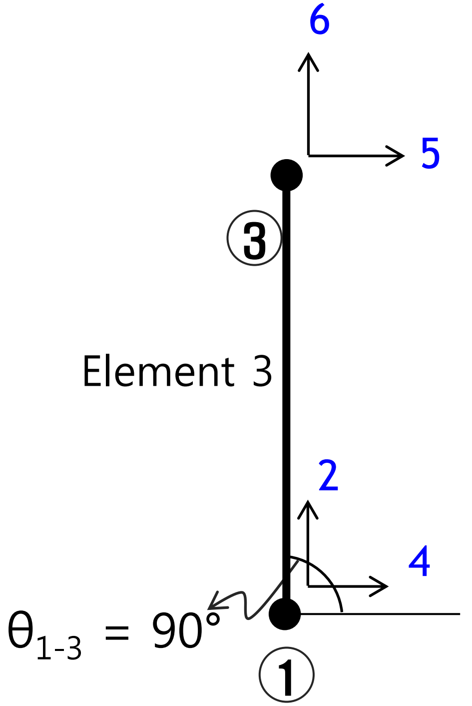
$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

For element 3, nodes 1 and 3 are located at the left and right ends of the member,

So $\theta_{1\text{-}3}$ = 90° and S = 1, C = 0.



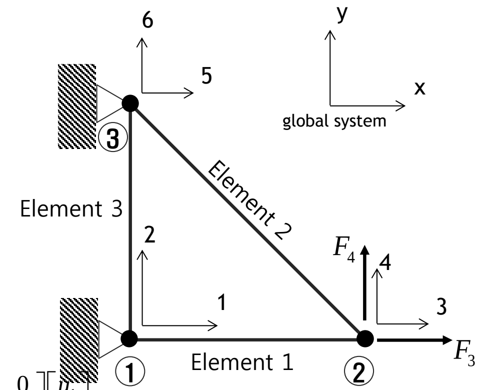
$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$



Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

Since we now have all elemental stiffnesses expressed in terms of the global coordinate system, we can now construct the system stiffness matrix. The structure has three nodes and therefore six degrees of freedom. The structural stiffness matrix will be a 6 x 6 matrix. Accumulating elements of the elemental stiffness matrices using the global codes noted above and to the right of the matrices we find

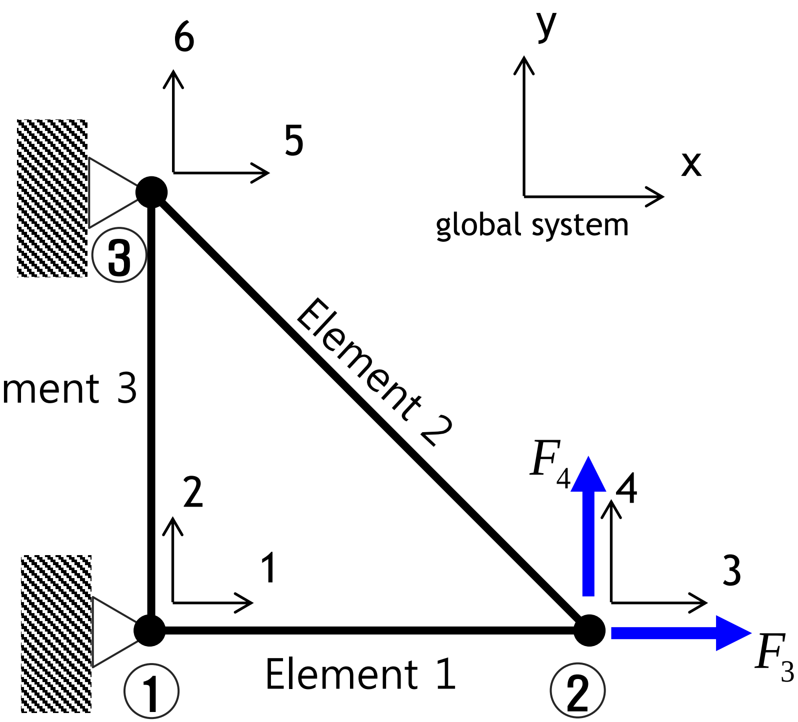


$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EAL \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} + \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} + \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1+1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & -1 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1+1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$

Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$



Both nodes 1 and 3 are pinned. Thus

$$u_1, u_2, u_5, u_6 = 0$$

Eliminating the rows and columns associated with these zero displacements results in the reduced stiffness matrix shown in equation

$$\begin{bmatrix} \cancel{F_1} \\ \cancel{F_2} \\ F_3 \\ F_4 \\ \cancel{F_5} \\ \cancel{F_6} \end{bmatrix} = EA/L \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1+1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & -1 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1+1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} \cancel{u_1} \\ \cancel{u_2} \\ u_3 \\ u_4 \\ \cancel{u_5} \\ \cancel{u_6} \end{bmatrix}$$

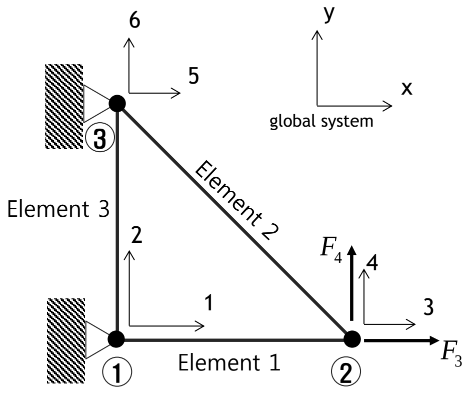
$$\Rightarrow \begin{bmatrix} F_3 \\ F_4 \end{bmatrix} = EA/L \begin{bmatrix} 1+1/2\sqrt{2} & -1/2\sqrt{2} \\ -1/2\sqrt{2} & 1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix}$$

Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

$$\begin{bmatrix} F_3 \\ F_4 \end{bmatrix} = EA/L \begin{bmatrix} 1+1/2\sqrt{2} & -1/2\sqrt{2} \\ -1/2\sqrt{2} & 1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix}$$

$$\Rightarrow \begin{aligned} u_3 &= (F_3 + F_4)L/EA \\ u_4 &= \left\lfloor (1+2\sqrt{2})F_4 + F_3 \right\rfloor L/EA \end{aligned}$$



$$u_1, u_2, u_5, u_6 = 0$$

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \end{bmatrix} = EA/L \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1+1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & 0 & -1/2\sqrt{2} & 1/2\sqrt{2} & 1/2\sqrt{2} & -1/2\sqrt{2} \\ 0 & -1 & 1/2\sqrt{2} & -1/2\sqrt{2} & -1/2\sqrt{2} & 1+1/2\sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix}$$

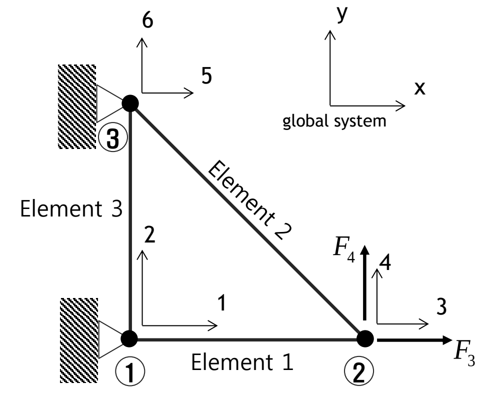
The reactions F_1, F_2, F_5, and F_6 are found by substituting displacements u_1 ~ u_6 into the equation . We find,

$$F_1 = -(F_3 + F_4), F_2 = 0, F_5 = F_4, and \quad F_6 = -F_4$$

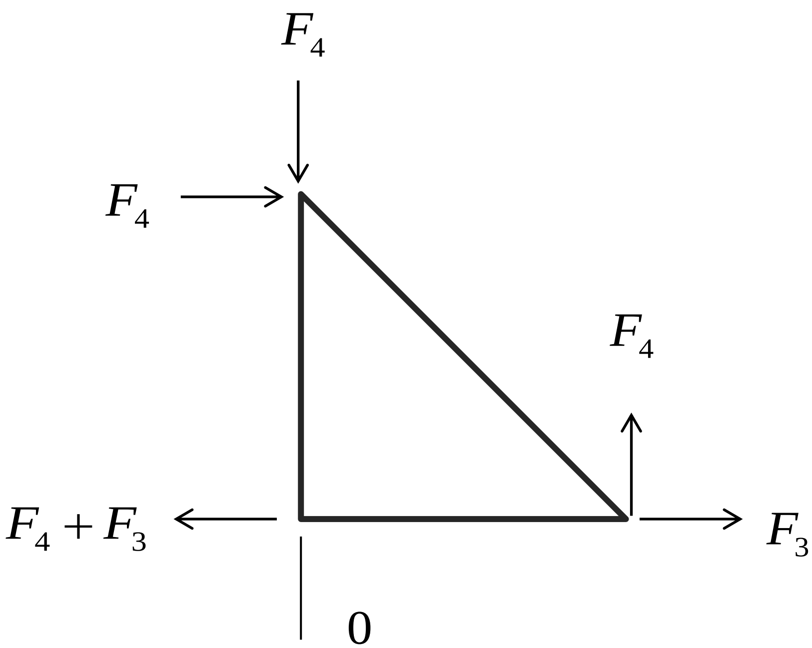
Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

$$F_1 = -(F_3 + F_4), F_2 = 0, F_5 = F_4, and \quad F_6 = -F_4$$



Sketching these reactions and the applied loads F_3 and F_4 on the structure as shown below, we verify **overall equilibrium**.



Truss Examples

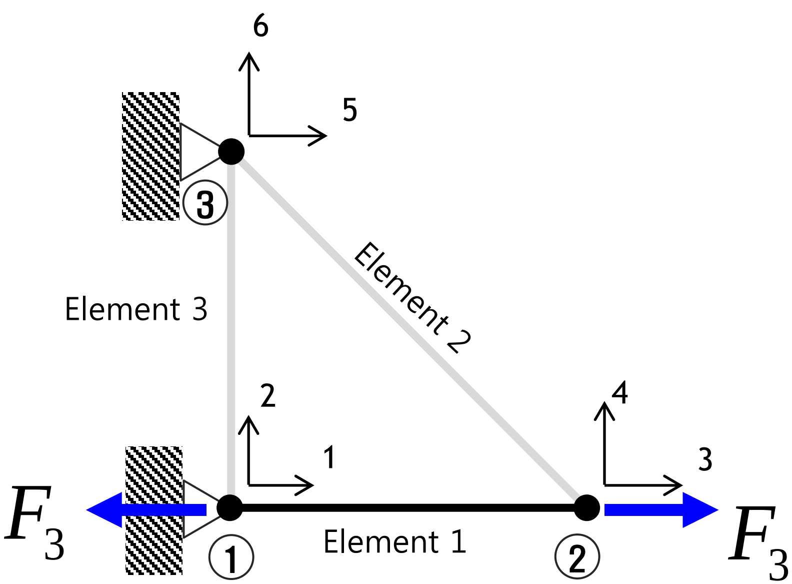
$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

Member force for the element 1

$$\boxed{\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix}} = EA/L \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} -(F_3+F_4) \\ 0 \\ (F_3+F_4) \\ 0 \end{bmatrix}$$

This forces are described in **local coordinate of each element!**

If $F_4 = 0$

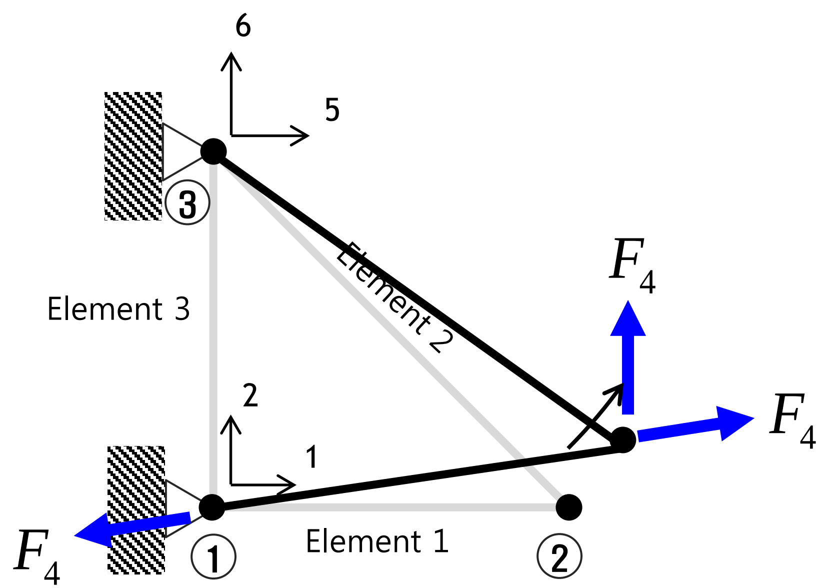


F_3는 1번 element가 모두 지지하고 있음

F_3는 2, 3 element에 어떠한 영향도 주지 않음

즉, 2, 3 element가 없다고 가정 하여도 정적 평형상태임

If $F_3 = 0$



F_4로 인하여 2번 element는 pin 3을 기준으로 회전함

이로 인하여 1번 element는 인장을 하게 되므로

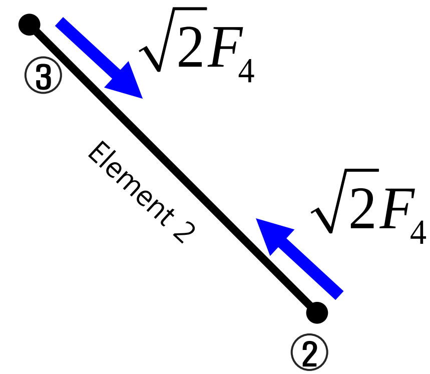
tensile force를 받음

Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

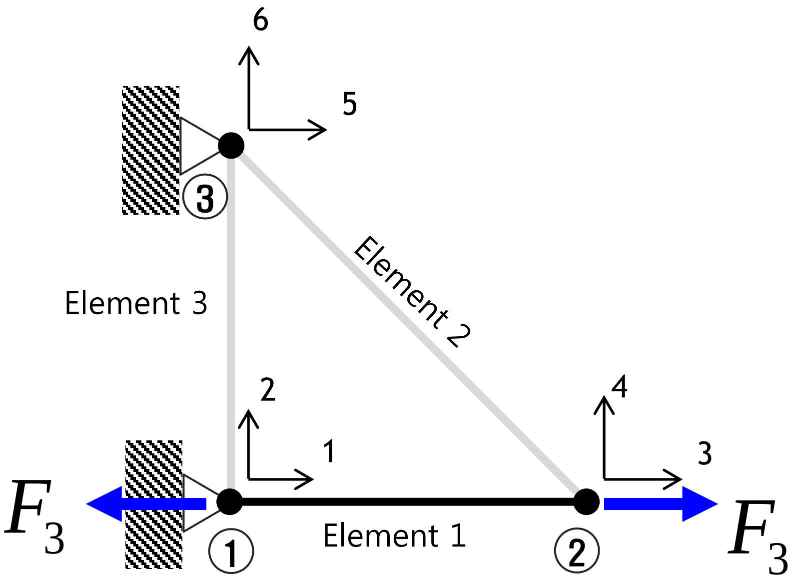
Member force for the element 2

$$\boxed{\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix}} = EA/\sqrt{2}L \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ -1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & -1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} = 1/\sqrt{2} \begin{bmatrix} 2F_4 \\ 0 \\ -2F_4 \\ 0 \end{bmatrix}$$

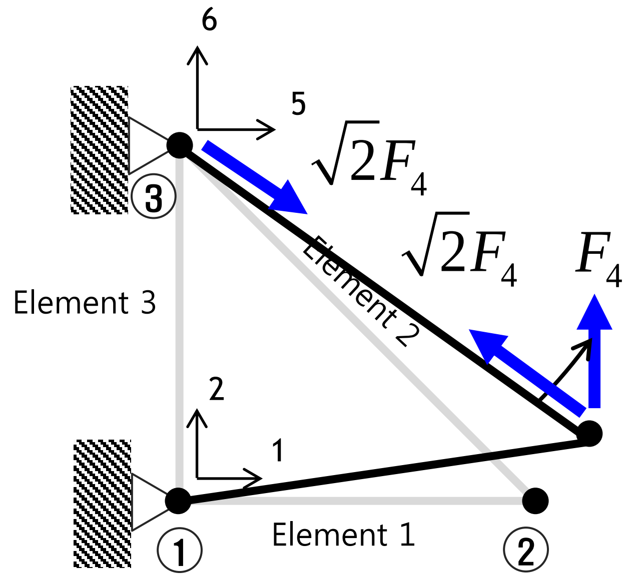


This forces are described in **local coordinate of each element!**

If $F_4 = 0$



If $F_3 = 0$



F_3는 2, 3 element에 어떠한 영향도 주지 않음
즉, 2, 3 element가 없다고 가정 하여도 정적 평형상태임
따라서 2번 element는 F_3로 인하여 어떠한 힘도 받지 않음

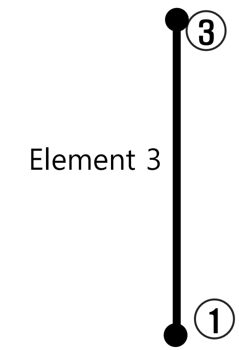
F_4로 인하여 2번 element는 pin 3을 기준으로 회전 함
하지만 1번 element가 회전 운동을 하지 못하도록 막고 있음
이로 인하여 4번 element는 압축력을 받게 됨

Truss Examples

$$\begin{bmatrix} f_{x1} \\ f_{y1} \\ f_{x2} \\ f_{y2} \end{bmatrix} = k \begin{bmatrix} C^2 & CS & -C^2 & -CS \\ CS & S^2 & -CS & -S^2 \\ -C^2 & -CS & C^2 & CS \\ -CS & -S^2 & CS & S^2 \end{bmatrix} \begin{bmatrix} \delta_{x1} \\ \delta_{y1} \\ \delta_{x2} \\ \delta_{y2} \end{bmatrix}$$

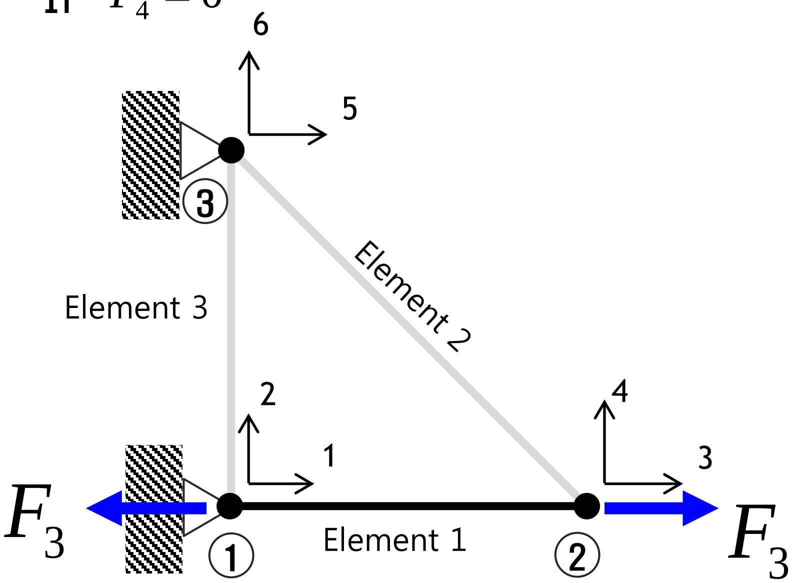
Member force for the element 3

$$\boxed{\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix}} = EA/L \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_5 \\ u_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

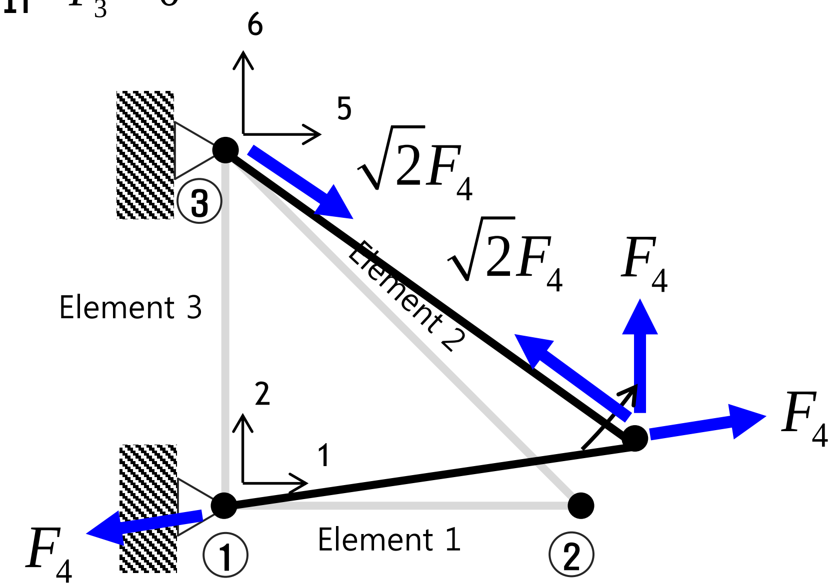


This forces are described in **local coordinate of each element!**

If $F_4 = 0$



If $F_3 = 0$



F_3가 0일 때나 F_4가 0일 때, 어떠한 경우에도 element 3에는 아무런 힘이 가해지지 않음

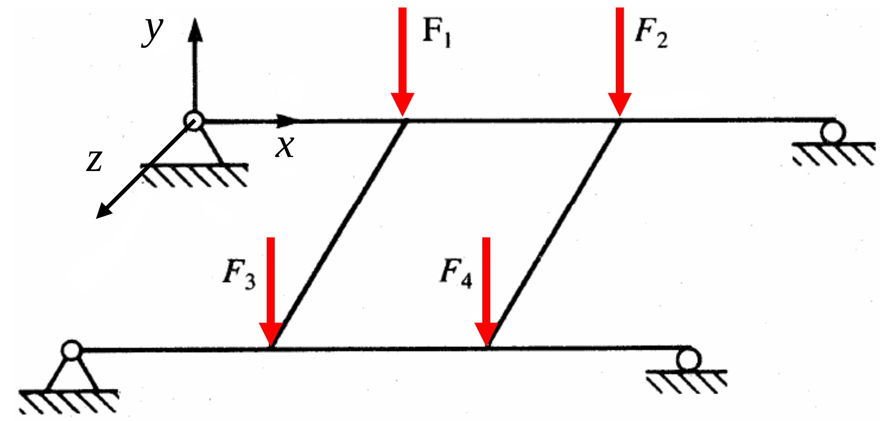
즉, element 3가 없다고 생각해 보아도, 정적 평형상태에 어떠한 영향도 주지 않음

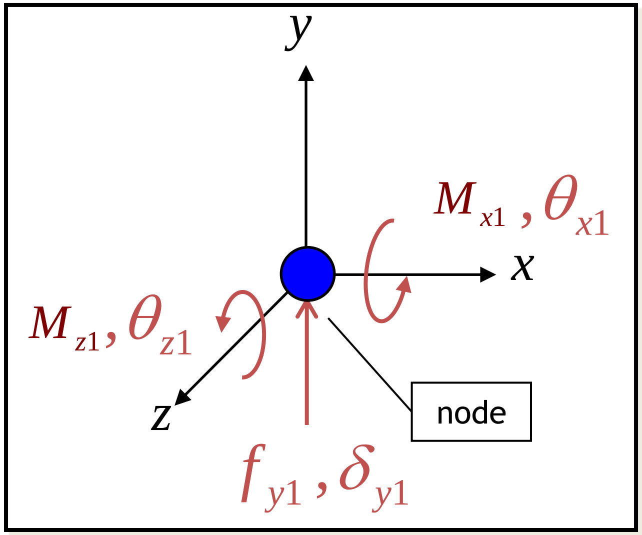
2.11. Stiffness Matrix for Grillage

Grillage

- Grillage* (Grid Structure) : A structure that has loads applied perpendicular to its plane. The elements are assumed to be rigidly connected at the joints. The floor system shown in the figure is an example of a very common grillage.

<example of grillage : typical floor system>



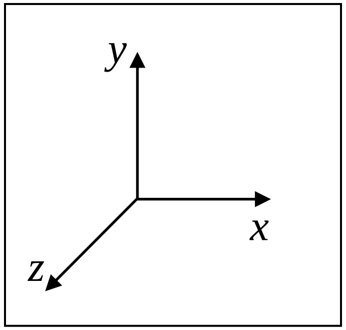


- As in the case of the beam element, we assume that axial deformation is neglected. However, in addition to bending about the horizontal axis of the cross section, the elements will also resist the loads by twisting about the axis of the element, thus developing torsional moments. Therefore, at each joint we will have a vertical displacement, a rotation about the horizontal axis of the cross section due to bending, and a rotation about the axis of the element due to torsion. There are three degrees of freedom at each node.

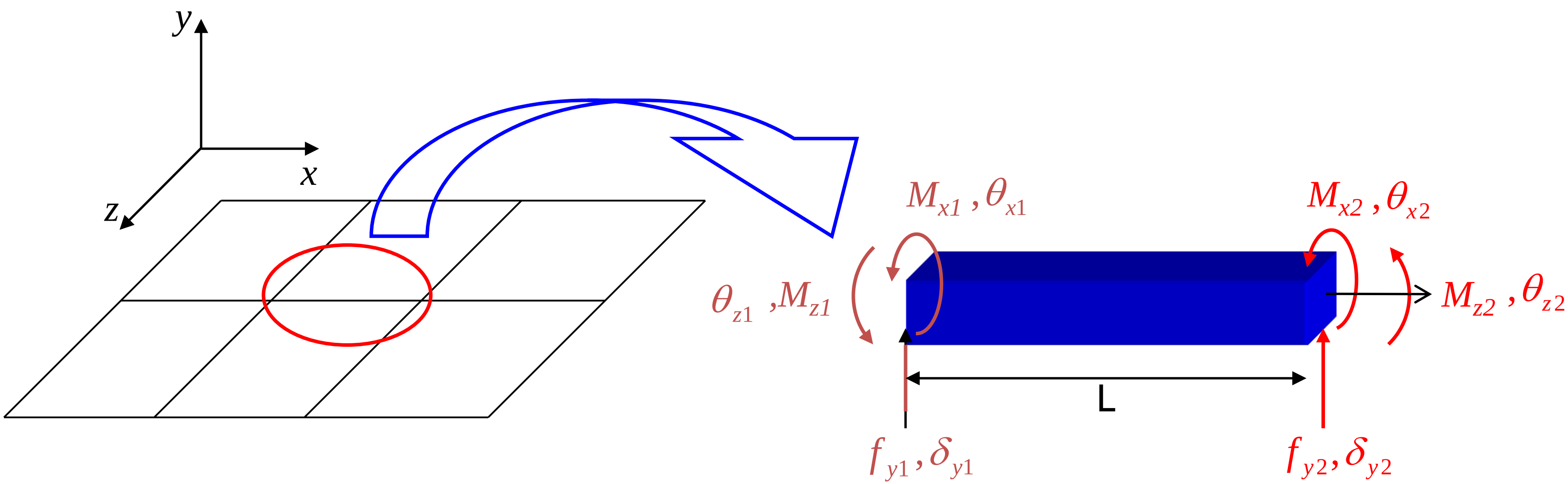
*Daryl L. Logan, A first Course in the Finite Element Method, 2nd edition, PWS Publishing,1993, p.87

Grillage : Stiffness Equations

step1. Coordinate System

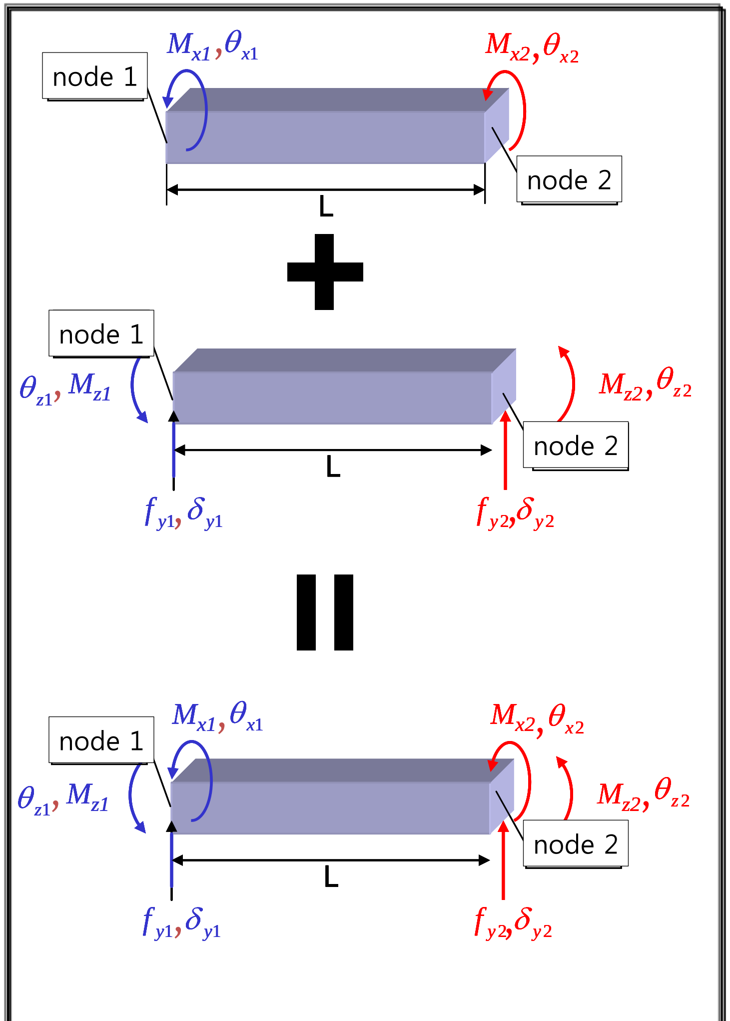


step2. Variables at each nodes



Grillage : Stiffness Equations

Step3. Stiffness Equations



shaft

$$\begin{bmatrix} M_{x1} \\ M_{x2} \end{bmatrix} = \begin{bmatrix} \frac{GJ}{L} & -\frac{GJ}{L} \\ -\frac{GJ}{L} & \frac{GJ}{L} \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \theta_{x2} \end{bmatrix}$$

+

beam

$$\begin{bmatrix} f_{y1} \\ M_{z1} \\ f_{y2} \\ M_{z2} \end{bmatrix} = \begin{bmatrix} \frac{12EI}{L^3} & \frac{6EI}{L^2} & -\frac{12EI}{L^3} & \frac{6EI}{L^2} \\ \frac{6EI}{L^2} & \frac{4EI}{L} & -\frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{12EI}{L^3} & -\frac{6EI}{L^2} & \frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ \frac{6EI}{L^2} & \frac{2EI}{L} & -\frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_{z1} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

=

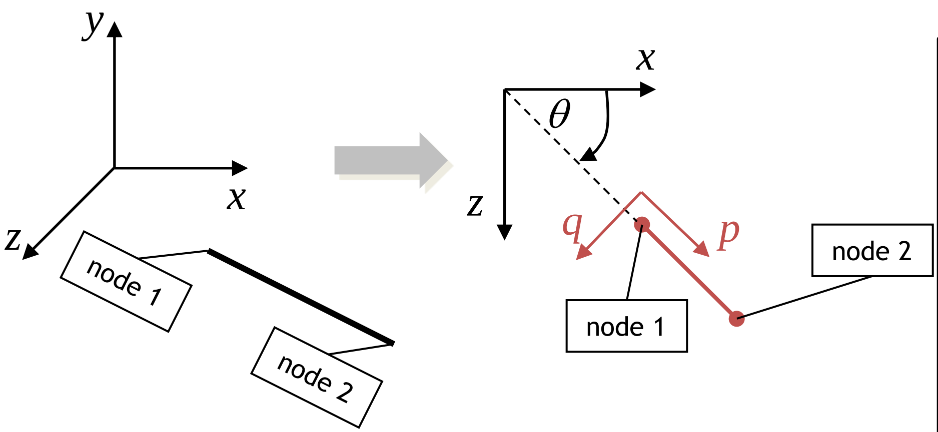
Grillage

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = \begin{bmatrix} \frac{GJ}{L} & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & \frac{12EI}{L^3} & \frac{6EI}{L^2} & 0 & -\frac{12EI}{L^3} & \frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{4EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{GJ}{L} & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} & 0 & \frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{2EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

① $[\mathbf{F}_{pqr}] = [\mathbf{K}_{pqr}][\delta_{pqr}]$

Grillage : Stiffness Equations

- transformation matrix between ***pq*** and ***xy*** coordinate system



rotation transformation along with y axis

$$[\mathbf{T}] = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

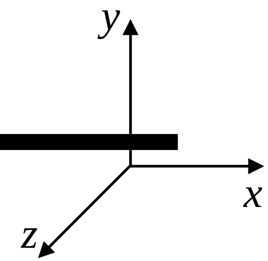
② $[\delta_{pqr}] = [\mathbf{T}][\delta_{xyz}]$

$$\begin{bmatrix} \theta_{p1} \\ \delta_{q1} \\ \theta_{r1} \\ \theta_{p2} \\ \delta_{q2} \\ \theta_{r2} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

③ $[f_{pqr}] = [\mathbf{T}][f_{xyz}]$

$$\begin{bmatrix} M_{p1} \\ f_{q1} \\ M_{r1} \\ M_{p2} \\ f_{q2} \\ M_{r2} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix}$$

Grillage : Stiffness Equations



- Stiffness Equations

① $[f_{pq}] = [\mathbf{K}_{pq}][\delta_{pq}]$ $\Rightarrow$ $[\mathbf{T}][f_{xy}] = [\mathbf{K}_{pq}][\mathbf{T}][\delta_{xy}]$

② $[f_{pq}] = [\mathbf{T}][f_{xy}]$

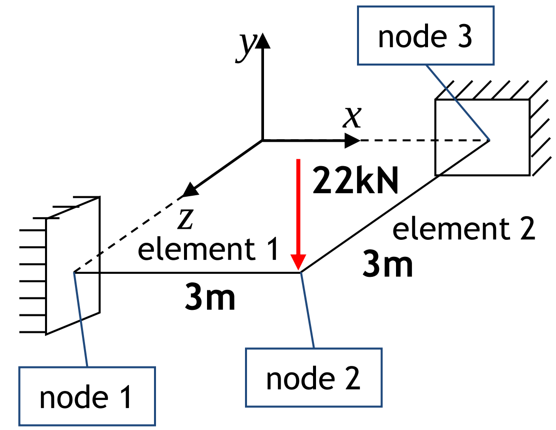
③ $[\delta_{pq}] = [\mathbf{T}][\delta_{xy}]$

multiply $[\mathbf{T}]^{-1} = [\mathbf{T}]^T$

$$[f_{xy}] = [\mathbf{T}]^T[\mathbf{K}_{pq}][\mathbf{T}][\delta_{xy}] = [\mathbf{K}_{xy}][\delta_{xy}]$$

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \frac{GJ}{L} & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & \frac{12EI}{L^3} & \frac{6EI}{L^2} & 0 & -\frac{12EI}{L^3} & \frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{4EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{GJ}{L} & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} & 0 & \frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{2EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

Ex.) Grillage



ex.) Find displacements and reaction force at each nodes of frame in the following figure.

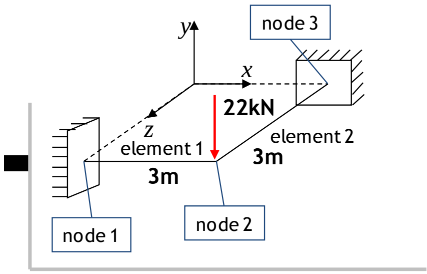
Step1. Input Data

- constants ($\theta_1 = 0$, $\theta_2 = 270°$)

element	cos θ	sin θ	length (m)	moment of inertia (m⁴)	Young's moldulus (kN/m²)	shear modulus (kN/m²)	polar moment of inertia (m⁴)
1	1	0	3	$I=16.6\times10^{-5}$	$E=210\times10^{6}$	$G=84\times10^{6}$	$J=4.6\times10^{-5}$
2	0	−1	3				

Ex.) Grillage

$$
[\mathbf{K}_{pqr}] = \begin{bmatrix} \frac{GJ}{L} & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & \frac{12EI}{L^3} & \frac{6EI}{L^2} & 0 & -\frac{12EI}{L^3} & \frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{4EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{GJ}{L} & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} & 0 & \frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{2EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix}
$$



Step2. Stiffness Equation

$$
[\mathbf{F}_{xyz}] = [\mathbf{K}_{xyz}][\delta_{xyz}] = [\mathbf{T}]^T[\mathbf{K}_{pqr}][\mathbf{T}][\delta_{xyz}]
$$

☞ **element 1**

$$
[\mathbf{T}] = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
[\mathbf{K}_{pqr}] = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 \\ -0.128 & 0 & 0 & 0.128 & 0 & 0 \\ 0 & -1.55 & -2.32 & 0 & 1.55 & -2.32 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65 \end{bmatrix}
$$

$$
\frac{GJ}{L} = \frac{(84\times10^6)\cdot(4.6\times10^{-5})}{3} = 0.128\times10^4
$$

$$
\frac{4EI}{L} = \frac{4\cdot(210\times10^6)(16.6\times10^{-5})}{3} = 4.65\times10^4
$$

$$
\frac{6EI}{L^2} = \frac{6\cdot(210\times10^6)(16.6\times10^{-5})}{3^2} = 2.32\times10^4
$$

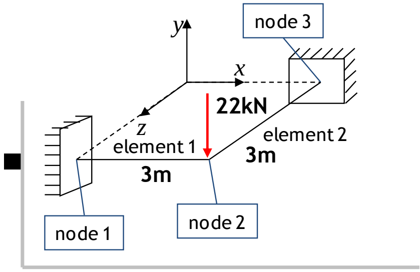
$$
\frac{12EI}{L^3} = \frac{12\cdot(210\times10^6)(16.6\times10^{-5})}{3^3} = 1.55\times10^4
$$

$$
\boxed{[\mathbf{F}_{xyz}] = [\mathbf{T}]^T[\mathbf{K}_{pqr}][\mathbf{T}][\delta_{xyz}]}
$$

$$
\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 \\ -0.128 & 0 & 0 & 0.128 & 0 & 0 \\ 0 & -1.55 & -2.32 & 0 & 1.55 & -2.32 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65 \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}
$$

Ex.) Grillage

$$[\mathbf{K}_{pqr}]=\begin{bmatrix} \frac{GJ}{L} & 0 & 0 & -\frac{GJ}{L} & 0 & 0 \\ 0 & \frac{12EI}{L^3} & \frac{6EI}{L^2} & 0 & -\frac{12EI}{L^3} & \frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{4EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{2EI}{L} \\ -\frac{GJ}{L} & 0 & 0 & \frac{GJ}{L} & 0 & 0 \\ 0 & -\frac{12EI}{L^3} & -\frac{6EI}{L^2} & 0 & \frac{12EI}{L^3} & -\frac{6EI}{L^2} \\ 0 & \frac{6EI}{L^2} & \frac{2EI}{L} & 0 & -\frac{6EI}{L^2} & \frac{4EI}{L} \end{bmatrix}$$



Step2. Stiffness Equation

$$[\mathbf{F}_{xyz}]=[\mathbf{K}_{xyz}][\delta_{xyz}]=[\mathbf{T}]^T[\mathbf{K}_{pqr}][\mathbf{T}][\delta_{xyz}]$$

☞ element 2

$$[\mathbf{T}]=\begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix}=\begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$[\mathbf{K}_{pqr}]=10^4\times\begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 \\ -0.128 & 0 & 0 & 0.128 & 0 & 0 \\ 0 & -1.55 & -2.32 & 0 & 1.55 & -2.32 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65 \end{bmatrix}$$

$$\frac{GJ}{L}=\frac{(84\times10^6)\cdot(4.6\times10^{-5})}{3}=0.128\times10^4$$

$$\frac{4EI}{L}=\frac{4\cdot(210\times10^6)(16.6\times10^{-5})}{3}=4.65\times10^4$$

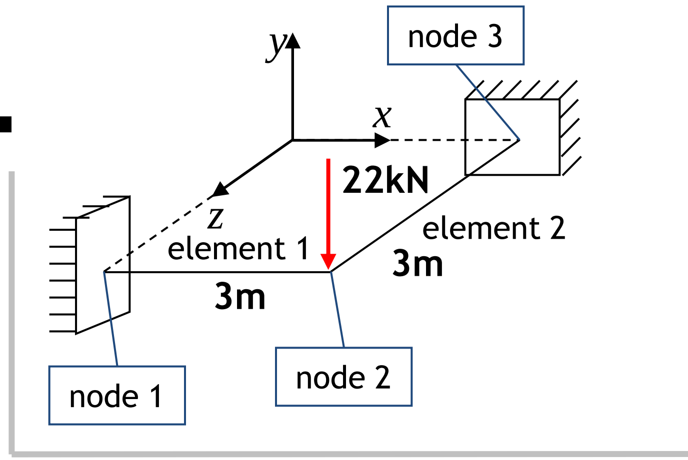
$$\frac{6EI}{L^2}=\frac{6\cdot(210\times10^6)(16.6\times10^{-5})}{3^2}=2.32\times10^4$$

$$\frac{12EI}{L^3}=\frac{12\cdot(210\times10^6)(16.6\times10^{-5})}{3^3}=1.55\times10^4$$

$$[\mathbf{F}_{xyz}]=[\mathbf{T}]^T[\mathbf{K}_{pqr}][\mathbf{T}][\delta_{xyz}]$$

$$\begin{bmatrix} M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix}=10^4\times\begin{bmatrix} 4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 2.32 & 1.55 & 0 & 2.32 & -1.55 & 0 \\ 0 & 0 & 0.128 & 0 & 0 & -0.128 \\ 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix}\begin{bmatrix} \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \\ \theta_{x3} \\ \delta_{y3} \\ \theta_{z3} \end{bmatrix}$$

Ex.) Grillage



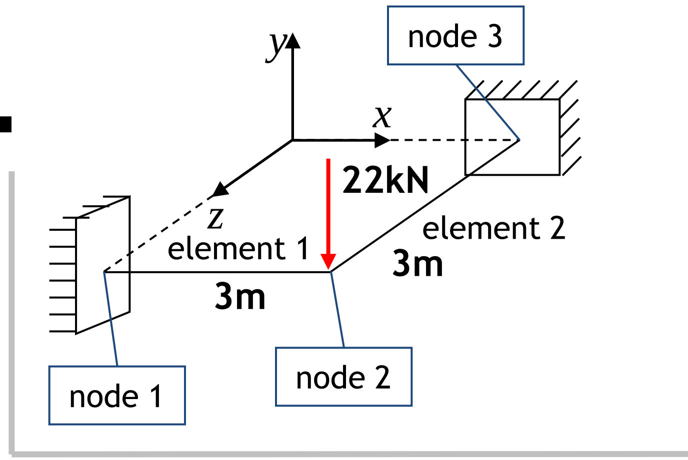
Step3. Find Displacements

- known/unknown displacements
  - ✓ known : $\theta_{x1}, \delta_{y1}, \theta_{z1}, \theta_{x3}, \delta_{y3}, \theta_{z3}(=0)$
  - ✓ unknown : $\theta_{x2}, \delta_{y2}, \theta_{z2}$
- known/unknown forces
  - ✓ known : $M_{x2}(=0)$, $f_{y2}(=-22\text{kN})$, $M_{z2}(=0)$
  - ✓ unknown : $M_{x1}, f_{y1}, M_{z1}, M_{x3}, f_{y3}, M_{z3}$

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 \\ -0.128 & 0 & 0 & 0.128 & 0 & 0 \\ 0 & -1.55 & -2.32 & 0 & 1.55 & -2.32 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65 \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

$$\begin{bmatrix} M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = 10^4 \times \begin{bmatrix} 4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 2.32 & 1.55 & 0 & 2.32 & -1.55 & 0 \\ 0 & 0 & 0.128 & 0 & 0 & -0.128 \\ 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix} \begin{bmatrix} \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \\ \theta_{x3} \\ \delta_{y3} \\ \theta_{z3} \end{bmatrix}$$

Ex.) Grillage



Step3. Find Displacements

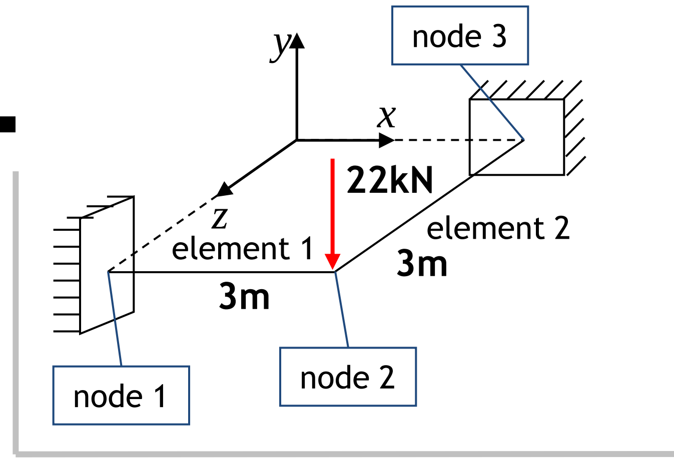
- known/unknown displacements
  - ✓ known : $\theta_{x1}, \delta_{y1}, \theta_{z1}, \theta_{x3}, \delta_{y3}, \theta_{z3}(=0)$
  - ✓ unknown : $\theta_{x2}, \delta_{y2}, \theta_{z2}$
- known/unknown forces
  - ✓ known : $M_{x2}(=0)$, $f_{y2}(=-22\text{kN})$, $M_{z2}(=0)$
  - ✓ unknown : M_{x1}, f_{y1}, M_{z1}, M_{x3}, f_{y3}, M_{z3}

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 & 0 & 0 & 0 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 & 0 & 0 & 0 \\ -0.128 & 0 & 0 & 0.128+4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 0 & -1.55 & -2.32 & 2.32 & 1.55+1.55 & -2.32 & 2.32 & -1.55 & 0 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65+0.128 & 0 & 0 & -0.128 \\ 0 & 0 & 0 & 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ 0 & 0 & 0 & -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \\ \theta_{x3} \\ \delta_{y3} \\ \theta_{z3} \end{bmatrix}$$

-Chapter 7. Grillage

Ex.) Grillage

Step3. Find Displacements



- known/unknown displacements
  - ✓ known : $\theta_{x1},\delta_{y1},\theta_{z1},\theta_{x3},\delta_{y3},\theta_{z3}(=0)$
  - ✓ unknown : $\theta_{x2},\delta_{y2},\theta_{z2}$
- known/unknown forces
  - ✓ known : $M_{x2}(=0)$, $f_{y2}(=-22kN)$, $M_{z2}(=0)$
  - ✓ unknown : M_{x1}, f_{y1}, M_{z1}, M_{x3}, f_{y3}, M_{z3}

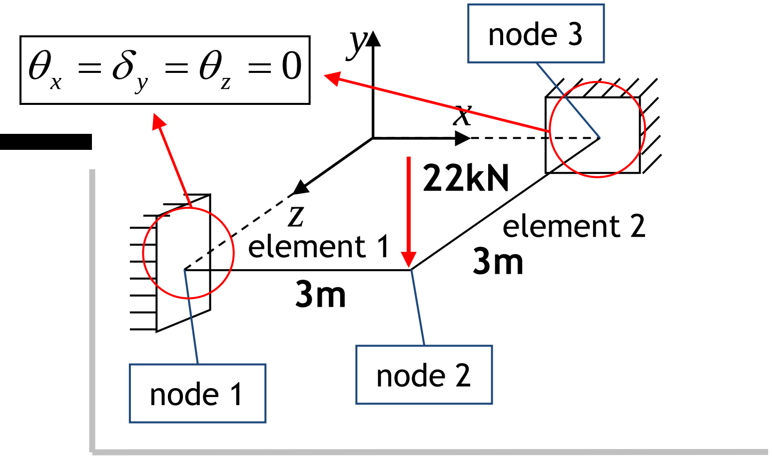
$$\begin{bmatrix} M_{x2}=0 \\ f_{y2}=-22kN \\ M_{2}=0 \end{bmatrix} = 10^4 \times \begin{bmatrix} 4.778 & 2.32 & 0 \\ 2.32 & 3.10 & -2.32 \\ 0 & -2.32 & 4.778 \end{bmatrix} \begin{bmatrix} \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

given

find

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2}=0 \\ f_{y2}=-22kN \\ M_{z2}=0 \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 & 0 & 0 & 0 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 & 0 & 0 & 0 \\ -0.128 & 0 & 0 & 0.128+4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 0 & -1.55 & -2.32 & 2.32 & 1.55+1.55 & -2.32 & 2.32 & -1.55 & 0 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65+0.128 & 0 & 0 & -0.128 \\ 0 & 0 & 0 & 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ 0 & 0 & 0 & -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix} \begin{bmatrix} \theta_{x1} \\ \delta_{y1} \\ \theta_{z1} \\ \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \\ \theta_{x3} \\ \delta_{y3} \\ \theta_{z3} \end{bmatrix}$$

Ex.) Grillage



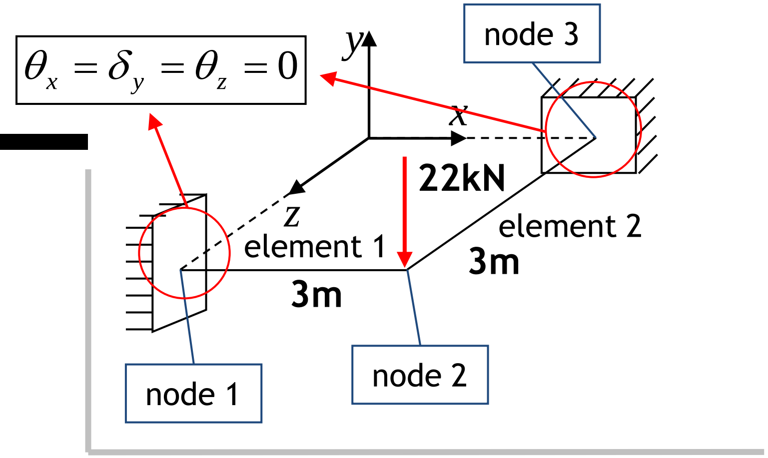
Step3. Find Displacements

- known/unknown displacements
  - ✓ known : $\theta_{x1}, \delta_{y1}, \theta_{z1}, \theta_{x3}, \delta_{y3}, \theta_{z3}(=0)$
  - ✓ unknown : $\theta_{x2}, \delta_{y2}, \theta_{z2}$
- known/unknown forces
  - ✓ known : $M_{x2}(=0)$, $f_{y2}(=-22kN)$, $M_{z2}(=0)$
  - ✓ unknown : M_{x1}, f_{y1}, M_{z1}, M_{x3}, f_{y3}, M_{z3}

$$\begin{bmatrix} M_{x2} = 0 \\ f_{y2} = -22kN \\ M_2 = 0 \end{bmatrix} = 10^4 \times \begin{bmatrix} 4.778 & 2.32 & 0 \\ 2.32 & 3.10 & -2.32 \\ 0 & -2.32 & 4.778 \end{bmatrix} \begin{bmatrix} \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix}$$

$$\therefore \begin{bmatrix} \theta_{x2} \\ \delta_{y2} \\ \theta_{z2} \end{bmatrix} = \frac{1}{10^4} \begin{bmatrix} 4.778 & 2.32 & 0 \\ 2.32 & 3.10 & -2.32 \\ 0 & -2.32 & 4.778 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ -22 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.126 \times 10^{-2} rad \\ -0.259 \times 10^{-2} cm \\ -0.126 \times 10^{-2} rad \end{bmatrix}$$

Ex.) Grillage



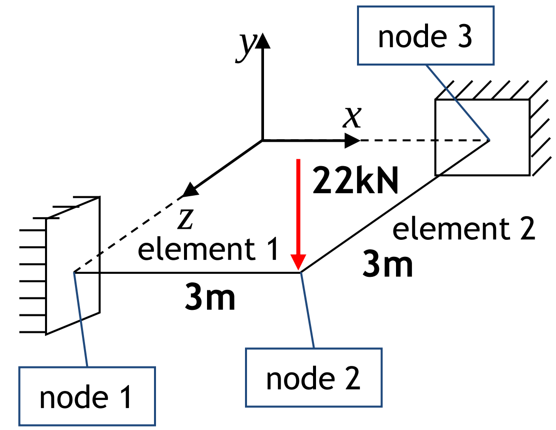
Step3. Find Displacements

- known/unknown displacements
  - ✓ known : $\theta_{x1}, \delta_{y1}, \theta_{z1}, \theta_{x3}, \delta_{y3}, \theta_{z3}(=0)$
  - ✓ unknown : $\theta_{x2}, \delta_{y2}, \theta_{z2}$
- known/unknown forces
  - ✓ known : $M_{x2}(=0)$, $f_{y2}(=-22kN)$, $M_{z2}(=0)$
  - ✓ unknown : M_{x1}, f_{y1}, M_{z1}, M_{x3}, f_{y3}, M_{z3}

find **given**

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2}=0 \\ f_{y2}=-22kN \\ M_{z2}=0 \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 & 0 & 0 & 0 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 & 0 & 0 & 0 \\ -0.128 & 0 & 0 & 0.128+4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 0 & -1.55 & -2.32 & 2.32 & 1.55+1.55 & -2.32 & 2.32 & -1.55 & 0 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65+0.128 & 0 & 0 & -0.128 \\ 0 & 0 & 0 & 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ 0 & 0 & 0 & -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix} \begin{bmatrix} \theta_{x1}=0 \\ \delta_{y1}=0 \\ \theta_{z1}=0 \\ \theta_{x2}=0.126\times10^{-2}rad \\ \delta_{y2}=-0.259\times10^{-2}cm \\ \theta_{z2}=-0.126\times10^{-2}rad \\ \theta_{x3}=0 \\ \delta_{y3}=0 \\ \theta_{z3}=0 \end{bmatrix}$$

Ex.) Grillage



Step4. Find Reaction Forces

superposition of reaction forces of each elements

- reaction forces for element 1

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = 10^4 \times \begin{bmatrix} 0.128 & 0 & 0 & -0.128 & 0 & 0 \\ 0 & 1.55 & 2.32 & 0 & -1.55 & 2.32 \\ 0 & 2.32 & 4.65 & 0 & -2.32 & 2.33 \\ -0.128 & 0 & 0 & 0.128 & 0 & 0 \\ 0 & -1.55 & -2.32 & 0 & 1.55 & -2.32 \\ 0 & 2.32 & 2.33 & 0 & -2.32 & 4.65 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.126\times10^{-2}\,rad \\ -0.259\times10^{-2}\,cm \\ -0.126\times10^{-2}\,rad \end{bmatrix} = \begin{bmatrix} -1.65kN\cdot m \\ 11kN \\ 31kN\cdot m \\ 1.65kN\cdot m \\ -11kN \\ 1.65kN\cdot m \end{bmatrix}$$

- reaction forces for element 2

$$\begin{bmatrix} M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = 10^4 \times \begin{bmatrix} 4.65 & 2.32 & 0 & 2.32 & -2.32 & 0 \\ 2.32 & 1.55 & 0 & 2.32 & -1.55 & 0 \\ 0 & 0 & 0.128 & 0 & 0 & -0.128 \\ 2.32 & 2.32 & 0 & 4.65 & -2.32 & 0 \\ -2.32 & -1.55 & 0 & -2.32 & 1.55 & 0 \\ 0 & 0 & -0.128 & 0 & 0 & 0.128 \end{bmatrix} \begin{bmatrix} 0.126\times10^{-2}\,rad \\ -0.259\times10^{-2}\,cm \\ -0.126\times10^{-2}\,rad \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1.65kN\cdot m \\ -11kN \\ -1.65kN\cdot m \\ -31kN\cdot m \\ 11kN \\ 1.65kN\cdot m \end{bmatrix}$$

Ex.) Grillage

- reaction forces for element 1

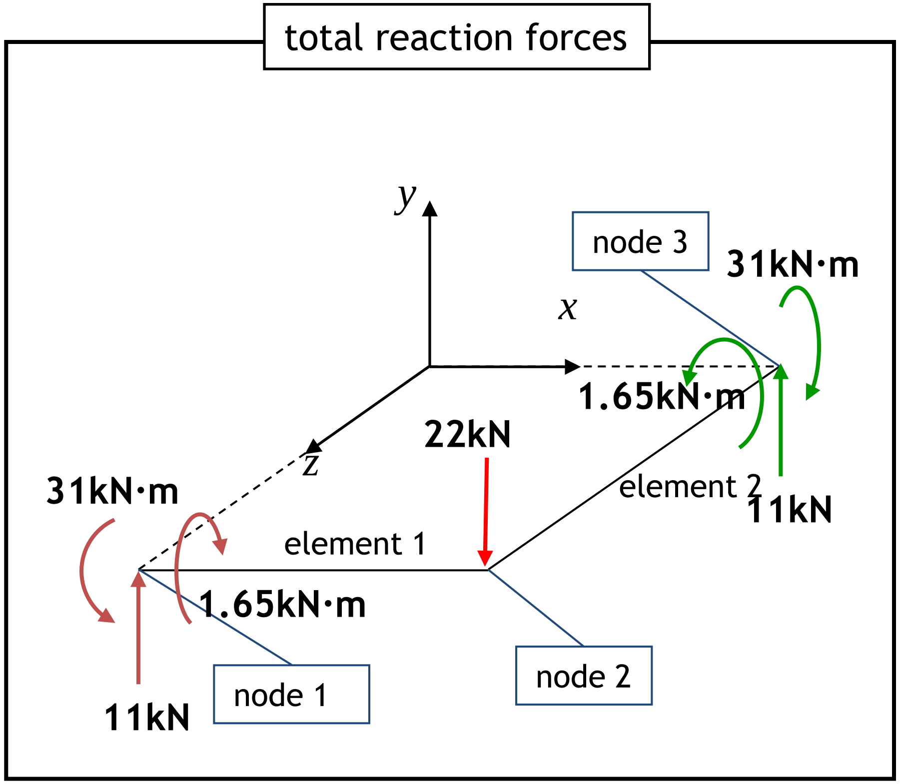
$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \end{bmatrix} = \begin{bmatrix} -1.65kN\cdot m \\ 11kN \\ 31kN\cdot m \\ 1.65kN\cdot m \\ -11kN \\ 1.65kN\cdot m \end{bmatrix}$$

- reaction forces for element 2

$$\begin{bmatrix} M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = \begin{bmatrix} -1.65kN\cdot m \\ -11kN \\ -1.65kN\cdot m \\ -31kN\cdot m \\ 11kN \\ 1.65kN\cdot m \end{bmatrix}$$

- total reaction forces

$$\begin{bmatrix} M_{x1} \\ f_{y1} \\ M_{z1} \\ M_{x2} \\ f_{y2} \\ M_{z2} \\ M_{x3} \\ f_{y3} \\ M_{z3} \end{bmatrix} = \begin{bmatrix} -1.65kN\cdot m \\ 11kN \\ 31kN\cdot m \\ 0 \\ -22kN \\ 0 \\ -31kN\cdot m \\ 11kN \\ 1.65kN\cdot m \end{bmatrix}$$



3. Finite Difference Method and Finite Element Method

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab.
http://asdal.snu.ac.kr

3. Finite Difference Method and Finite Element Method

3.1 Introduction to FDM and FEM

Reference

[Zienkiewicz 1983]
Zienkiewicz,O.C. and Morgan,K.,
Finite Elements and Approximation,
John Wiley & Sons, 1983

Introduction

Physical Phenomena

⇩ quantitative description

Ordinary or Partial Differential Equations with Boundary and Initial Conditions

⇩ available mathematical method

Exact Solution

Major difficulty

only the very simplest forms of equations, within geometrically trivial boundaries

While searching for a quantitative description of physical phenomena, the engineer or the scientist establishes generally a system of ordinary or partial differential equations valid in a certain region (or domain) and imposes on this system suitable boundary and initial conditions

Here come the major difficulties, as only the very simplest forms of equations, within geometrically trivial boundaries, are capable of being solved exactly with available mathematical method

[Zienkiewicz 1983] pp.1

Introduction

O.D.E. : Ordinary Differential Equation
P.D.E. : Partial Differential Equation
B/C : Boundary Condition
I/C : Initial Condition

O.D.E or P.D.E with B/C and I/C → Exact Solution

Difficult

Discretization
Approximation

Algebraic form → Approximated Solution

To overcome such difficulties, it is necessary to recast the problem in a purely algebraic form, involving only basic arithmetic operations. To achieve this, various forms of discretization of the continuum problem defined by the differential equations can be used.

In such a discretization, the infinite set of numbers representing the unknown function or functions is replaced by a finite number of unknown parameters, and this process, in general, requires some form of approximation

[Zienkiewicz 1983] pp.1-2

Introduction

O.D.E or P.D.E with B/C and I/C

Discretization
Approximation

Algebraic form

Finite Difference Method

Finite Element Method

Of the various forms of discretization which are possible, one of the simplest is the *finite difference process* and the others are various trial function approximations falling under the general classification of *finite element methods*.

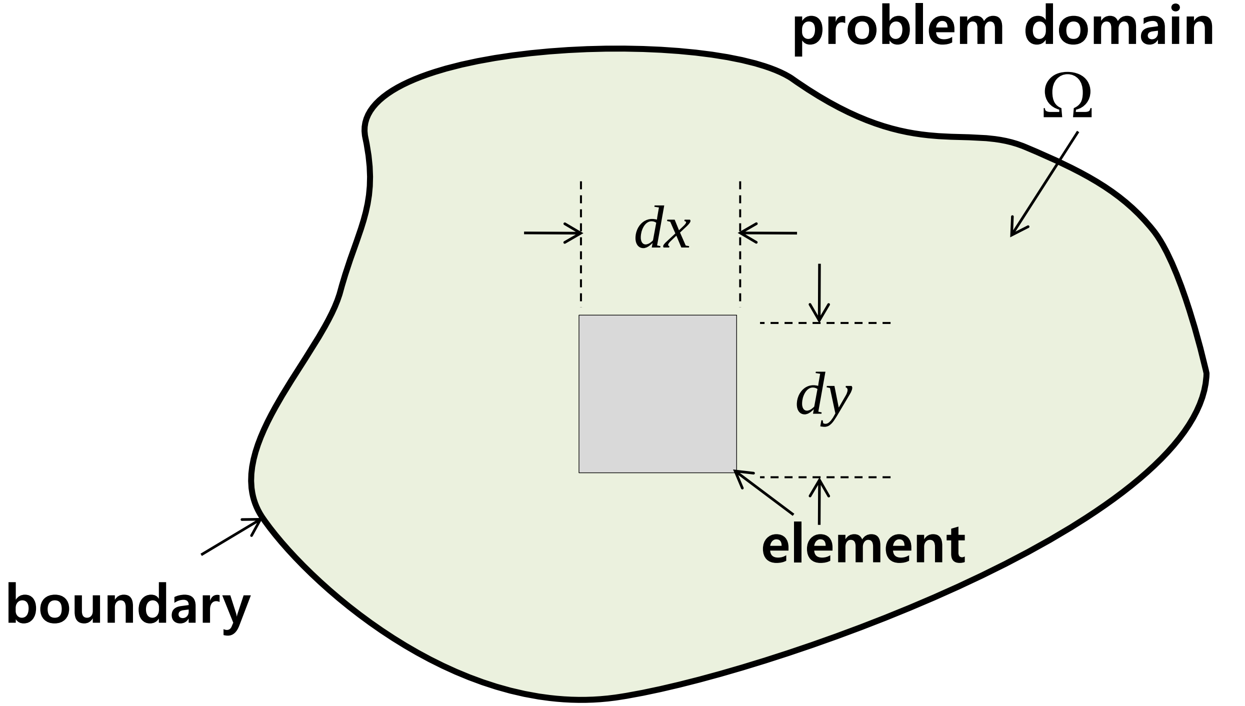
[Zienkiewicz 1983] p.2

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

$\phi(x, y, t)$ **temperature distribution**

$q_x,\ q_y$ **the heat flowing in the direction of the x and y per unit length and in unit time**

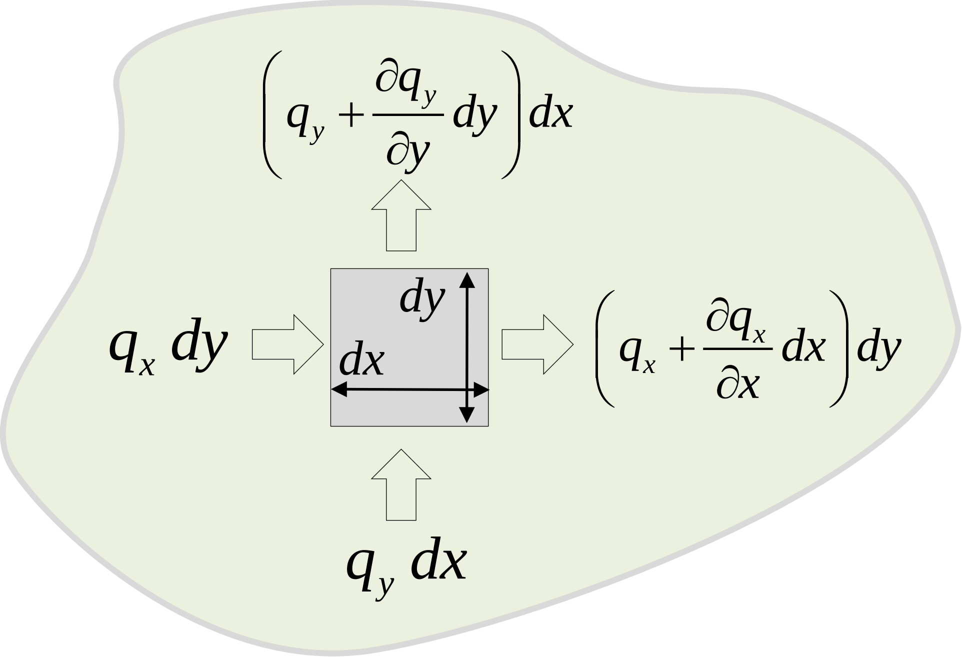


Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

D **the difference between outflow and inflow for an element size** $dx\,dy$

$$D=\left(q_x+\frac{\partial q_x}{\partial x}dx-q_x\right)dy+\left(q_y+\frac{\partial q_y}{\partial y}dy-q_y\right)dx$$



Some Examples of Continuum Problems

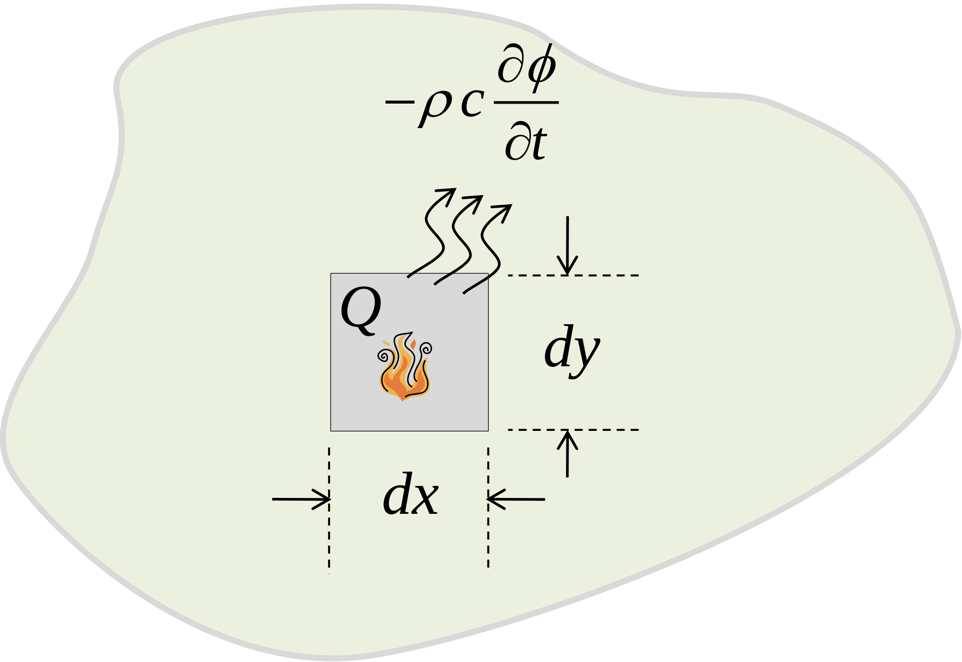
A problem of heat flow in a two-dimensional domain Ω

The heat generated in the element

$$Q\,dx\,dy$$

The heat released in unit time due to the temperature change

$$-\rho c \frac{\partial \phi}{\partial t} dx\,dy$$



where c is the specific heat
and ρ id the density

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

For conservation of heat,
the difference
must be equal to
the sum
of the heat generated
and released in the element

$$\left(q_x + \frac{\partial q_x}{\partial x}dx - q_x\right)dy + \left(q_y + \frac{\partial q_y}{\partial y}dy - q_y\right)dx = Q\,dx\,dy - \rho c \frac{\partial \phi}{\partial t}dx\,dy$$

$$\frac{\partial q_x}{\partial x}dxdy + \frac{\partial q_y}{\partial y}dxdy - Q\,dx\,dy + \rho c \frac{\partial \phi}{\partial t}dx\,dy = 0$$

$$\boxed{\therefore \frac{\partial \boxed{q_x}}{\partial x} + \frac{\partial \boxed{q_y}}{\partial y} - Q + \rho c \frac{\partial \boxed{\phi}}{\partial t} = 0}$$

1 equation
3 variables

Can we solve this problem?

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

Introducing a physical law governing the heat flow in an isometric medium,

$$q_n = -k\frac{\partial \phi}{\partial n}$$ **where, k is a property of the medium known as the conductivity**

specifically, $q_x = -k\dfrac{\partial \phi}{\partial x},\ q_y = -k\dfrac{\partial \phi}{\partial y}$

The heat conservation, therefore, leads to

$$\frac{\partial}{\partial x}\left(-k\frac{\partial \phi}{\partial x}\right)+\frac{\partial}{\partial y}\left(-k\frac{\partial \phi}{\partial y}\right)-Q+\rho c\frac{\partial \phi}{\partial t}=0$$

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right)+Q-\rho c\frac{\partial \phi}{\partial t}=0$$

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

Differential Equation governing the problem at hand

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right)+Q-\rho c\frac{\partial \phi}{\partial t}=0$$

Such a solution needs the specification
of *initial conditions at time*,
and
of *boundary conditions on the boundary*

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

Initial Condition

e.g. the distribution of temperature given everywhere in Ω at $t = t_0$

Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

Boundary Conditions

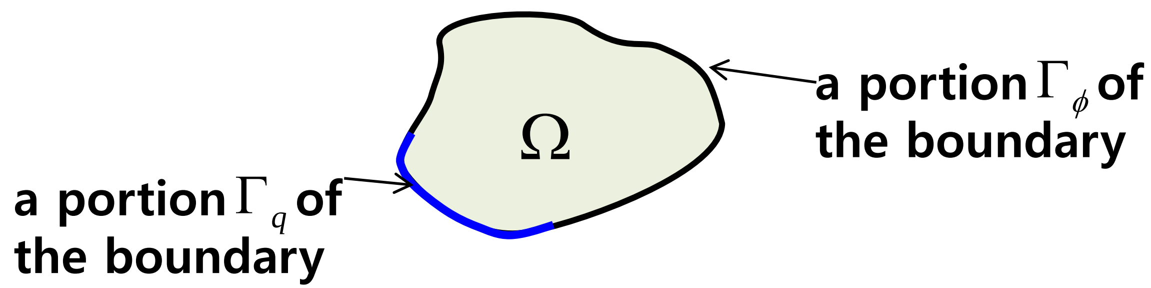
Typically two different kinds of boundary condition may be involved

$\phi - \bar{\phi} = 0 \quad on \quad \Gamma_\phi$ **the values of the temperature are specified**

⇨ **"Dirichlet" B/C**

$-k\dfrac{d\phi}{dn} - \bar{q} = 0 \quad on \quad \Gamma_q$ **the values of the temperature derivative are specified**

⇨ **"Neumman" B/C**



Some Examples of Continuum Problems

A problem of heat flow in a two-dimensional domain Ω

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right)+Q-\rho c\frac{\partial \phi}{\partial t}=0 \quad ,\phi-\bar{\phi}=0 \;\; on \;\; \Gamma_{\phi} \quad ,-k\frac{d\phi}{dn}-\bar{q}=0 \;\; on \;\; \Gamma_{q}$$

if steady-state conditions are assumed $\frac{\partial}{\partial t}=0$

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right)+Q=0$$

for one dimensional problem ⇨ **if** k **is constant**

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right)+Q=0$$

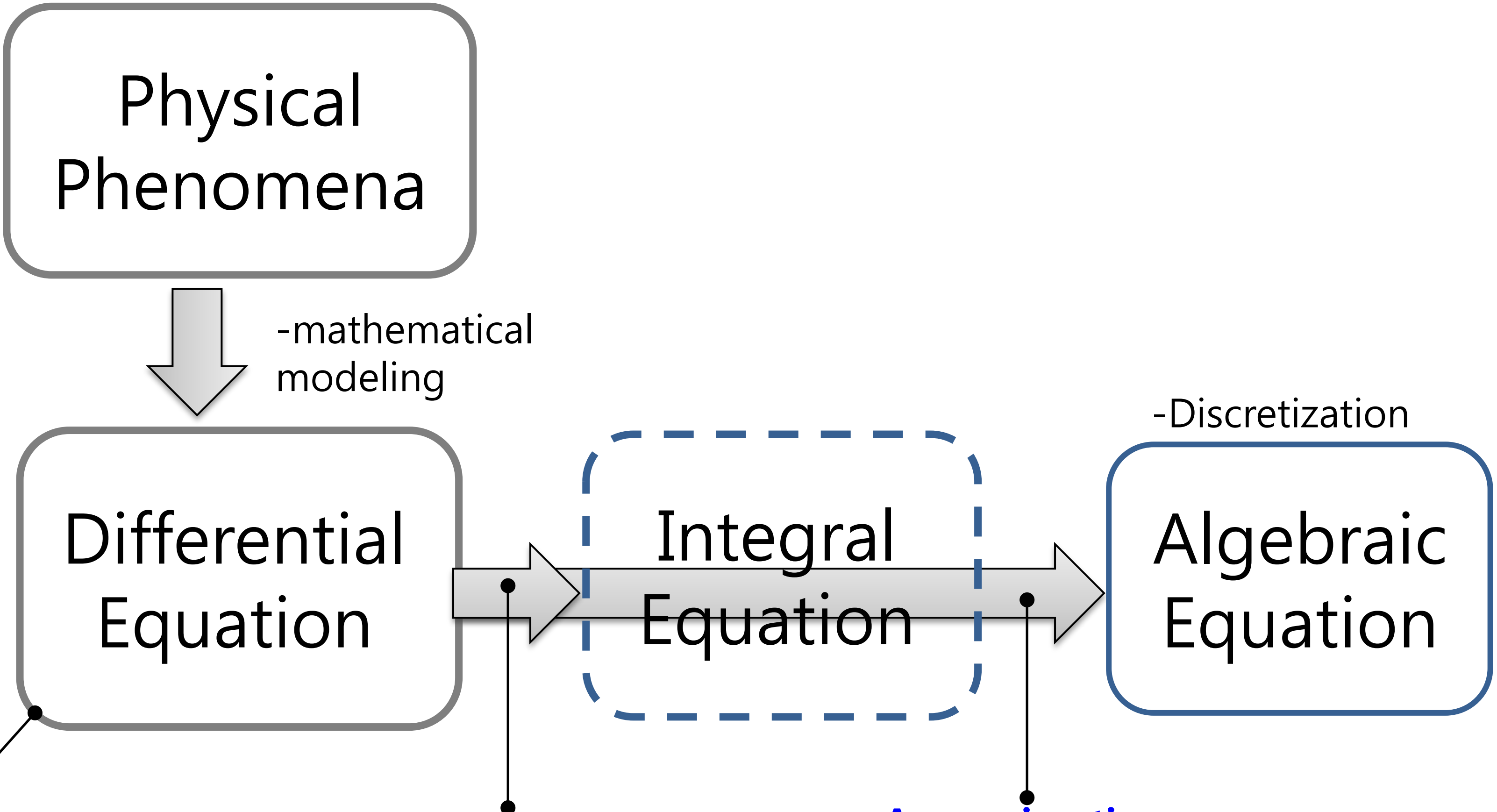
$$k\frac{\partial^2 \phi}{\partial x^2}+Q=0$$

3. Finite Difference Method and Finite Element Method

3.2 Finite Difference Method(FDM)

[Zienkiewicz 1983] Ch. 1.3

Finite Difference Method



Differential Form
- Linear or Nonlinear D.E
- Ordinary or Partial D.E.
- Initial and/or Boundary Value Problem

Integral Form
-Weighted Residual Method
-Calculus of Variation
-Integral Equation

Approximation
- Collocation Method
- Subdomain Method
- Galerkin Method
- Least Square Method

Finite Difference in One Dimension

A Simple one-dimensional boundary value problem :

Problem Definition of Temperature Distribution

We wish to determine a function $\phi(x)$

which satisfies a given differential equation

$$k\frac{d^2\phi}{dx^2} = -Q(x) \quad \text{in the region} \quad 0 < x < L$$

with the associated boundary conditions

$$\phi(0) = \bar{\phi}_0, \quad \phi(L) = \bar{\phi}_L$$

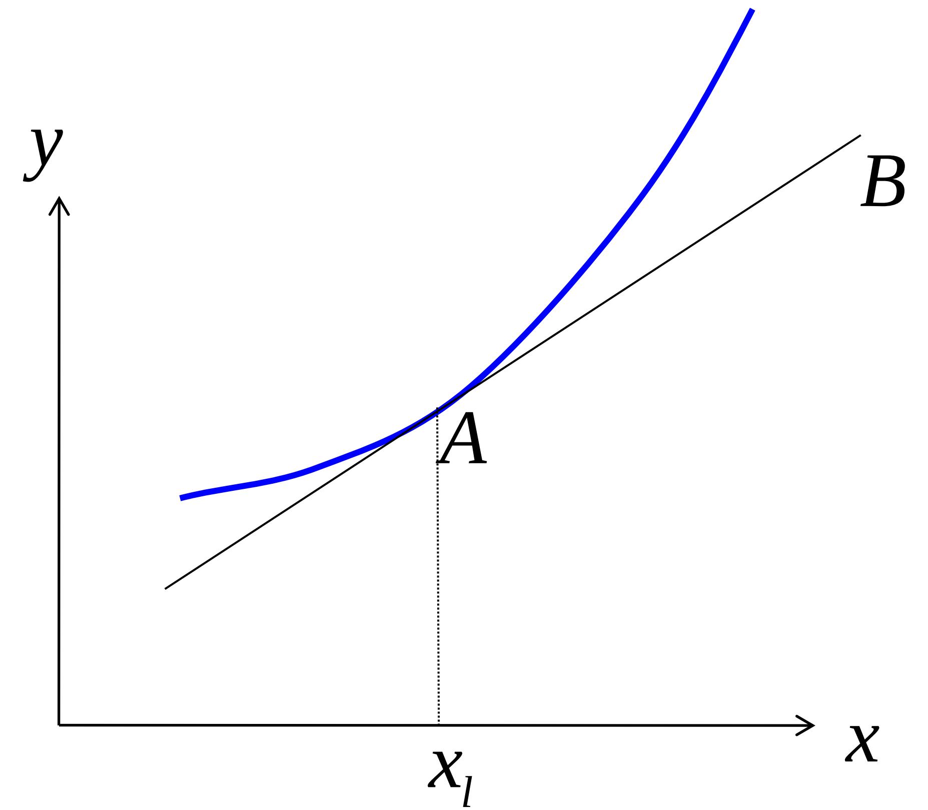
where k is the material thermal conductivity (assumed to be constant)

Finite Difference in One Dimension

The Finite Difference Approximation of Derivatives

A derivative of the function $\phi(x)$ at x_l : slope AB , or $d\phi / dx\big|_l$

$y = \phi(x)$



$(x = l\ )$

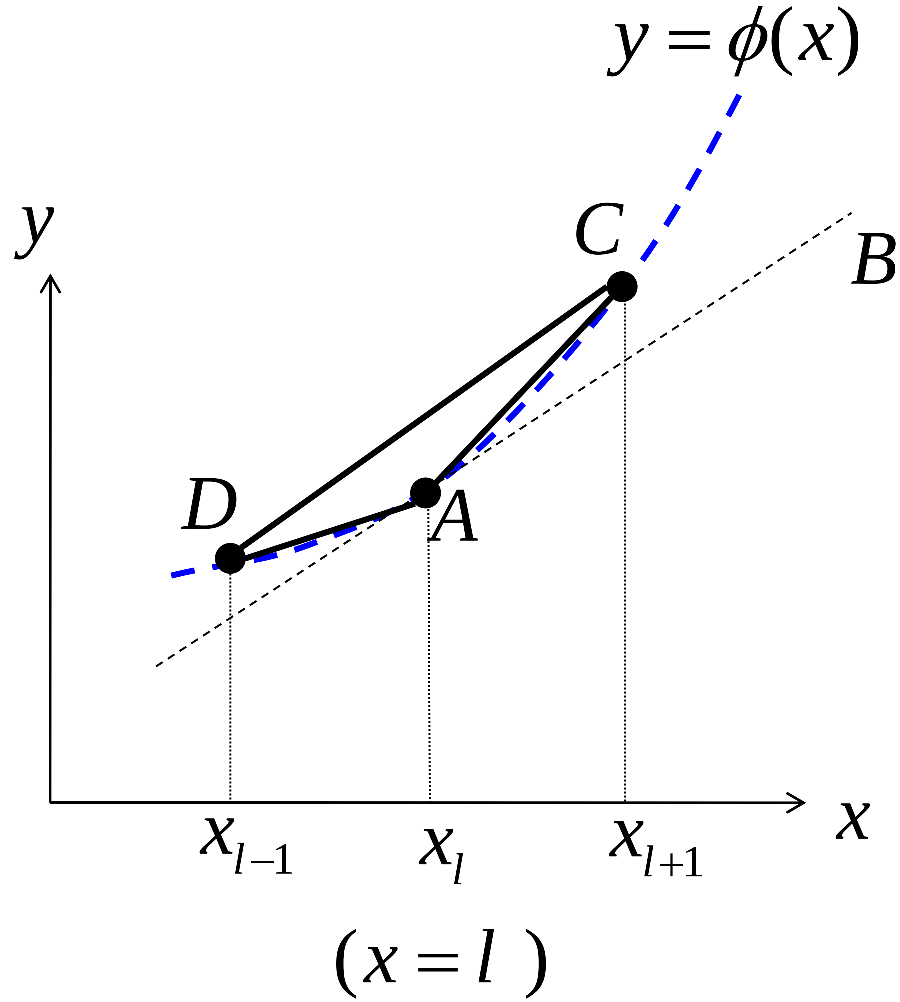
The Finite Difference Approximation of Derivatives

A graphical interpretation of some finite difference approximations to $d\phi / dx|_l$

Forward difference : slope of AC

Backward difference : slope of DA

Central difference : slope of DC



The Finite Difference Approximation of Derivatives

Using Taylor's theorem

$$\phi(x_{l+1}) \equiv \phi(x_l + \Delta x) = \phi(x_l) + \Delta x \left.\frac{d\phi}{dx}\right|_{x=l} + \frac{(\Delta x)^2}{2} \left.\frac{d^2\phi}{dx^2}\right|_{x=l} + \cdots$$

or we can rewrite

$$\phi_{l+1} = \phi_l + \Delta x \left.\frac{d\phi}{dx}\right|_l + \frac{(\Delta x)^2}{2} \left.\frac{d^2\phi}{dx^2}\right|_l + \cdots$$

therefore

$$\left.\frac{d\phi}{dx}\right|_l = \frac{\phi_{l+1} - \phi_l}{\Delta x} - \frac{\Delta x}{2} \left.\frac{d^2\phi}{dx^2}\right|_l - \cdots \quad \Rightarrow \quad \left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_{l+1} - \phi_l}{\Delta x}$$

"forward difference"

The Finite Difference Approximation of Derivatives

In a similar manner by using Taylor's theorem

$$\phi_l = \phi_{l-1} + \Delta x \left.\frac{d\phi}{dx}\right|_l + \frac{(\Delta x)^2}{2} \left.\frac{d^2\phi}{dx^2}\right|_l + \cdots$$

Rewriting the equation for ϕ_{l-1} gives

$$\phi_{l-1} = \phi_l - \Delta x \left.\frac{d\phi}{dx}\right|_l - \frac{(\Delta x)^2}{2} \left.\frac{d^2\phi}{dx^2}\right|_l - \cdots$$

therefore

$$\left.\frac{d\phi}{dx}\right|_l = \frac{\phi_l - \phi_{l-1}}{\Delta x} + \frac{\Delta x}{2} \left.\frac{d^2\phi}{dx^2}\right|_l - \cdots$$

$$\Rightarrow \quad \left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_l - \phi_{l-1}}{\Delta x}$$

"backward difference"

The Finite Difference Approximation of Derivatives

In a similar manner by using Taylor's theorem

$$\phi_{l+1} = \phi_l + \Delta x \left.\frac{d\phi}{dx}\right|_l + \frac{(\Delta x)^2}{2}\left.\frac{d^2\phi}{dx^2}\right|_l + \frac{(\Delta x)^3}{6}\left.\frac{d^3\phi}{dx^3}\right|_l + \cdots \qquad \textbf{(1)}$$

$$\phi_{l-1} = \phi_l - \Delta x \left.\frac{d\phi}{dx}\right|_{l-1} + \frac{(\Delta x)^2}{2}\left.\frac{d^2\phi}{dx^2}\right|_{l-1} - \frac{(\Delta x)^3}{6}\left.\frac{d^3\phi}{dx^3}\right|_l \cdots \qquad \textbf{(2)}$$

(1)-(2) : $$\phi_{l+1} - \phi_{l-1} = 2\Delta x \left.\frac{d\phi}{dx}\right|_l + \frac{(\Delta x)^3}{3}\left.\frac{d^3\phi}{dx^3}\right|_l \cdots$$

$$\Downarrow$$

$$\left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_{l+1} - \phi_{l-1}}{2\Delta x}$$ **"central difference"**

The Finite Difference Approximation of Derivatives

Forward difference : slope of AC

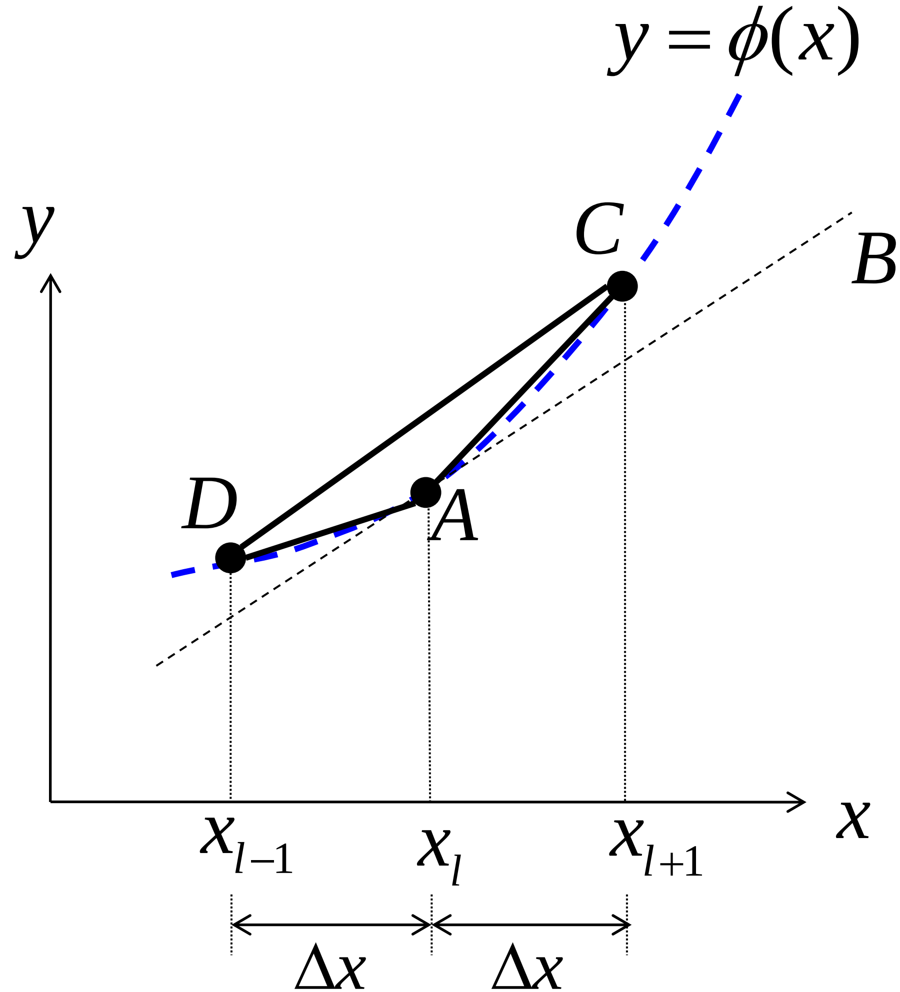
$$\left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_{l+1}-\phi_l}{\Delta x}$$

Backward difference : slope of DA

$$\left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_l-\phi_{l-1}}{\Delta x}$$

Central difference : slope of DC

$$\left.\frac{d\phi}{dx}\right|_l \approx \frac{\phi_{l+1}-\phi_{l-1}}{2\Delta x}$$



The Finite Difference Approximation of Derivatives

In a similar manner by using Taylor's theorem

$$\phi_{l+1} = \phi_l + \Delta x \left.\frac{d\phi}{dx}\right|_l + \frac{(\Delta x)^2}{2}\left.\frac{d^2\phi}{dx^2}\right|_l + \frac{(\Delta x)^3}{6}\left.\frac{d^3\phi}{dx^3}\right|_l + \cdots \quad \textbf{(1)}$$

$$\phi_{l-1} = \phi_l - \Delta x \left.\frac{d\phi}{dx}\right|_{l-1} + \frac{(\Delta x)^2}{2}\left.\frac{d^2\phi}{dx^2}\right|_{l-1} - \frac{(\Delta x)^3}{6}\left.\frac{d^3\phi}{dx^3}\right|_l \cdots \quad \textbf{(2)}$$

(1)+(2) : $$\phi_{l+1} + \phi_{l-1} = 2\phi_l + (\Delta x)^2 \left.\frac{d^2\phi}{dx^2}\right|_l \cdots$$

$$\left.\frac{d^2\phi}{dx^2}\right|_l \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{(\Delta x)^2}$$

"the second derivative of central difference"

The Finite Difference Approximation of Derivatives

The first derivative of central difference at $x = x_{l-1}$

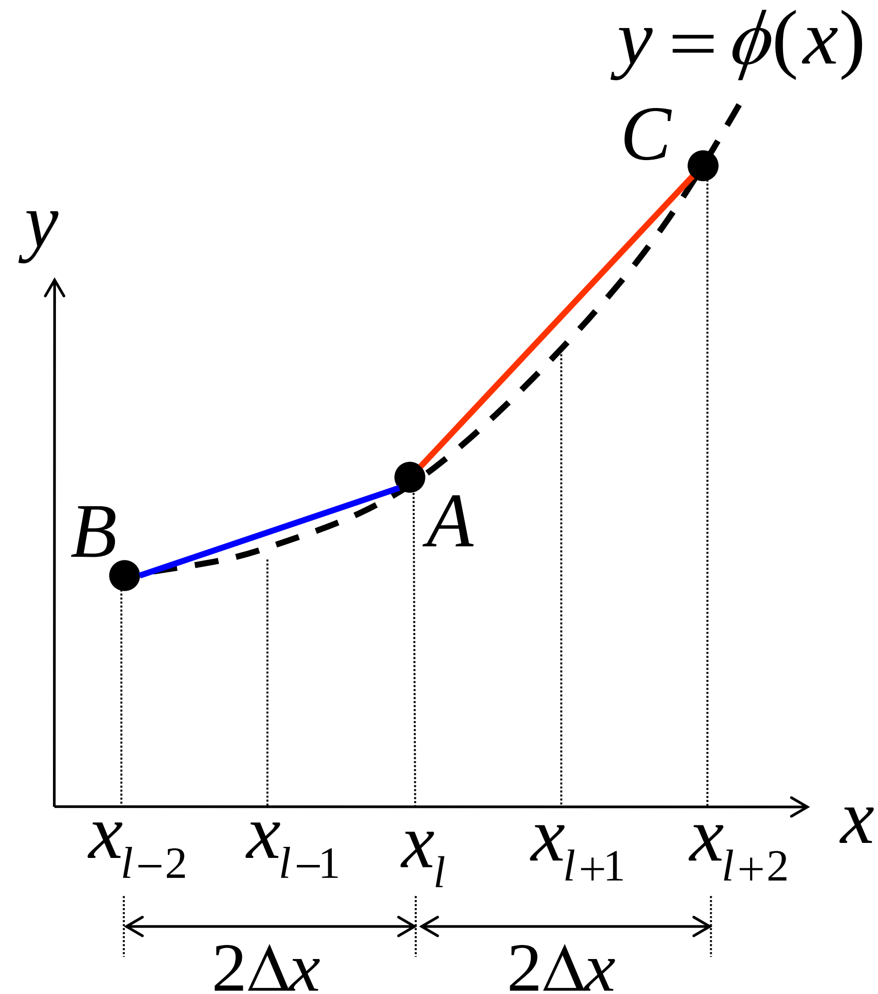
$$\left.\frac{d\phi}{dx}\right|_{l-1} \approx \frac{\phi_l - \phi_{l-2}}{2\Delta x}$$

→ Slop of BA

The first derivative of central difference at $x = x_{l+1}$

$$\left.\frac{d\phi}{dx}\right|_{l+1} \approx \frac{\phi_{l+2} - \phi_l}{2\Delta x}$$

→ Slop of AC



The Finite Difference Approximation of Derivatives

The first derivative of central difference:

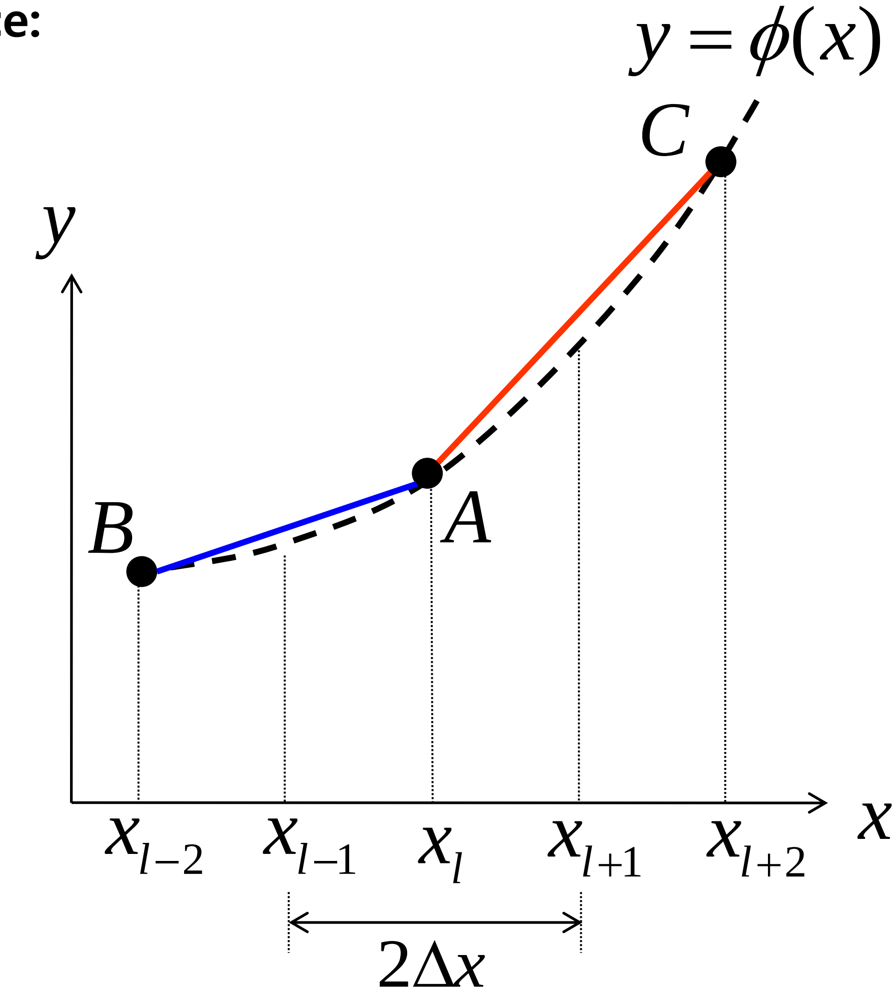
$$\left.\frac{d\phi}{dx}\right|_{l-1} \approx \frac{\phi_l - \phi_{l-2}}{2\Delta x}, \quad \left.\frac{d\phi}{dx}\right|_{l+1} \approx \frac{\phi_{l+2} - \phi_l}{2\Delta x}$$

The second derivative of central difference at $x = x_l$

$$\left.\frac{d^2\phi}{dx^2}\right|_l \approx \frac{\left.\frac{d\phi}{dx}\right|_{l+1} - \left.\frac{d\phi}{dx}\right|_{l-1}}{2\Delta x}$$

$$= \frac{\frac{\phi_{l+2} - \phi_l}{2\Delta x} - \frac{\phi_l - \phi_{l-2}}{2\Delta x}}{2\Delta x}$$

$$= \frac{\phi_{l+2} - 2\phi_l + \phi_{l-2}}{(2\Delta x)^2}$$



The Finite Difference Approximation of Derivatives

The second derivative of central difference at $x = x_l$

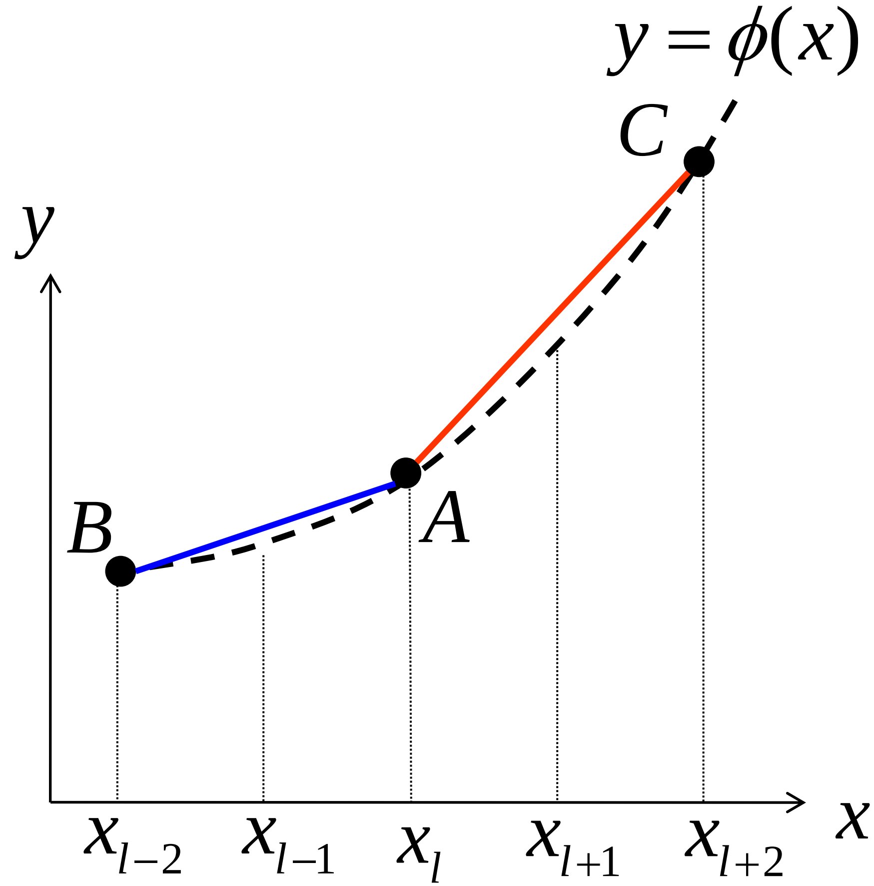
$$\left.\frac{d^2\phi}{dx^2}\right|_l \approx \frac{\phi_{l+2} - 2\phi_l + \phi_{l-2}}{(2\Delta x)^2}$$

$$2\Delta x \to \Delta x$$

$$\phi_{l-2} \to \phi_{l-1}$$

$$\phi_{l+2} \to \phi_{l+1}$$

$$\left.\frac{d^2\phi}{dx^2}\right|_l \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{(\Delta x)^2}$$



Finite Difference in One Dimension
-Solution of a Differential Equation by the Finite Difference Method

$$k\frac{d^2\phi}{dx^2} = -Q(x) \qquad ,0 < x < L \qquad ,\phi(0) = \bar{\phi}_0, \quad \phi(L) = \bar{\phi}_L$$

second derivatives approximated by the central difference method

$$\left.\frac{d^2\phi}{dx^2}\right|_l \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2}$$

the approximation produces the equation

$k\dfrac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = -Q(x_l)$ **at each of the interior grid points** x_l

$$, l = 1, 2, \ldots, L-1$$

$$,\phi(0) \equiv \phi_0 = \bar{\phi}_0, \quad \phi(L) \equiv \phi_L = \bar{\phi}_L$$

Finite Difference in One Dimension
-Solution of a Differential Equation by the Finite Difference Method

$$Q_l = Q(x_l)$$

$$k\frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = -Q(x) \quad , l = 1, 2, \ldots, L-1 \quad , \phi_0 = \bar{\phi}_0, \quad \phi_L = \bar{\phi}_L$$

An equation of this form arises at each of the interior grid points on the finite different mesh.
Writing down these equation separately gives

$$\phi_2 - 2\phi_1 + \phi_0 = -\frac{\Delta x^2 Q_1}{k} \quad \Rightarrow \quad \phi_2 - 2\phi_1 = -\frac{\Delta x^2 Q_1}{k} - \bar{\phi}_0$$

$$\phi_3 - 2\phi_2 + \phi_1 = -\frac{\Delta x^2 Q_2}{k}$$

$$\vdots$$

$$\phi_{L-1} - 2\phi_{L-2} + \phi_{L-3} = -\frac{\Delta x^2 Q_{L-2}}{k}$$

$$\phi_{L-2} - 2\phi_{L-1} + \phi_L = -\frac{\Delta x^2 Q_{L-1}}{k} \quad \Rightarrow \quad -2\phi_{L-1} + \phi_{L-2} = -\frac{\Delta x^2 Q_{L-1}}{k} - \bar{\phi}_L$$

this set of equations may be written as a single matrix equation

$$\mathbf{K}\boldsymbol{\phi} = \mathbf{f}$$

where,

$$\mathbf{K} = \begin{bmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}, \boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_{L-2} \\ \phi_{L-1} \end{bmatrix}, \mathbf{f} = \begin{bmatrix} \dfrac{\Delta x^2 Q_1}{k} + \bar{\phi}_0 \\ \dfrac{\Delta x^2 Q_2}{k} \\ \vdots \\ \dfrac{\Delta x^2 Q_{L-1}}{k} + \bar{\phi}_L \end{bmatrix}$$

multiply -1 both side of the equations in the previous slide

$$-\phi_2 + 2\phi_1 = \frac{\Delta x^2 Q_1}{k} + \bar{\phi}_0$$

$$-\phi_3 + 2\phi_2 - \phi_1 = \frac{\Delta x^2 Q_2}{k}$$

$$\vdots$$

$$-\phi_{L-1} + 2\phi_{L-2} - \phi_{L-3} = \frac{\Delta x^2 Q_{L-2}}{k}$$

$$2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2 Q_{L-1}}{k} + \bar{\phi}$$

The original problem of determining an unknown continuous function $\phi(x)$

$\Downarrow$

has been replaced by the problem of solving a matrix equation for the discrete set of values $\phi_1, \phi_2, \ldots, \phi_{L-1}$

The finite difference method will, therefore, give information about the function values at the mesh points

but no information about the functions values between these points

Differential Equation

$$k\frac{d^2\phi}{dx^2} = -Q(x) \quad ,0 < x < L$$

$$,\phi(0) = \bar{\phi}_0, \quad \phi(L) = \bar{\phi}_L$$

Solution by the Finite Difference Method

given

$$\mathbf{K}\boxed{\boldsymbol{\phi}} = \boxed{\mathbf{f}}$$

find

Example 1.1

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$

Boundary Condition $\phi = 0$ at $x = 0$ and $\phi = 1$ at $x = 1$

A mesh spacing $\Delta x = \dfrac{1}{3}$ is chosen

Solution)

The left side of the governing equation can be approximated as follows

$$\frac{d^2\phi_l}{dx^2} \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2}$$

$$\frac{d^2\phi_l}{dx^2} \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2}$$

Example 1.1

governing equation $\frac{d^2\phi}{dx^2} = \phi$ **A mesh spacing:** $\Delta x = \frac{1}{3}$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

Solution)

$\Delta x = \frac{1}{3}$

$\phi_0 = 0$ $\boxed{\phi_1}$ $\boxed{\phi_2}$ $\phi_3 = 1$ **2 unknowns**

$x_0 = 0$ $x_1 = \frac{1}{3}$ $x_2 = \frac{2}{3}$ $x_3 = 1$

$$\frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = \phi_l \Rightarrow \phi_{l+1} - 2\phi_l + \phi_{l-1} = \Delta x^2 \phi_l$$

when $l = 1,\quad \phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$

$l = 2,\quad \phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$

2 equations

Solution)

when

$$l=1, \quad \phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$$

$$l=2, \quad \phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$$

$$\Delta x = \frac{1}{3}$$

$$\phi_2 - 2\phi_1 - \frac{1}{9}\phi_1 = -\phi_0$$

$$-2\phi_2 - \frac{1}{9}\phi_2 + \phi_1 = -\phi_3$$

$$\phi_0 = 0$$

$$\phi_3 = 1$$

$$-\phi_2 + 2\frac{1}{9}\phi_1 = 0$$

$$2\frac{1}{9}\phi_2 - \phi_1 = 1$$

$$\phi_1 = 0.2893, \phi_2 = 0.6107$$

Analytic Solution)

Suppose that $\phi = e^{\lambda x}$

Substituting $\phi = e^{\lambda x}$ **into the governing equation gives**

$$\lambda^2 e^{\lambda x} = e^{\lambda x} \quad \Rightarrow \quad \lambda^2 = 1 \quad \Rightarrow \quad \lambda = \pm 1$$

General solution: $c_1 e^{x} + c_2 e^{-x}$

Analytic Solution)

General solution: $\phi = c_1 e^x + c_2 e^{-x}$

From B/C: $\phi(0) = 0$ $\qquad\qquad$ $\phi(1) = 1$

$$\boxed{c_1 + c_2 = 0 \qquad\qquad c_1 e + c_2 \frac{1}{e} = 1}$$

$$\Downarrow$$

$$c_1 = \frac{1}{e - 1/e}, \quad c_2 = \frac{1}{-e + 1/e}$$

$\phi_1 = 0.2893, \phi_2 = 0.6107$

Analytic Solution)

$$\phi = \frac{1}{e - 1/e} e^{x} + \frac{1}{-e + 1/e} e^{-x}$$

$$\phi_1 = \phi\left(\frac{1}{3}\right) = 0.2889, \phi_2 = \phi\left(\frac{2}{3}\right) = 0.6102$$

FDM) $\phi_1 = 0.2893, \phi_2 = 0.6107$

Derivative Boundary Conditions
-Boundary condition in terms of a derivative

If the gradient of the temperature is specified for the previous heat conduction example

$$-k\frac{d\phi}{dx} = \bar{q} \qquad \text{at } x = L$$

then

$$-\phi_L + 2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2\, Q_{L-1}}{k}$$

unknown

$$-\phi_2 + 2\phi_1 = \frac{\Delta x^2\, Q_1}{k} + \bar{\phi}_0$$

$$-\phi_3 + 2\phi_2 - \phi_1 = \frac{\Delta x^2\, Q_2}{k}$$

$$\vdots$$

$$-\phi_{L-1} + 2\phi_{L-2} - \phi_{L-3} = \frac{\Delta x^2\, Q_{L-2}}{k}$$

$$-\phi_L + 2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2\, Q_{L-1}}{k}$$

a set of $L-1$ equations

L unknowns

we need one more equation

Derivative Boundary Conditions
-Boundary condition in terms of a derivative

unknown

$$-\phi_L + 2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2 \, Q_{L-1}}{k}$$

$$-k \left. \frac{d\phi}{dx} \right|_L = \bar{q}$$

$$\left. \frac{d\phi}{dx} \right|_L = -\frac{\bar{q}}{k}$$

one more equation
by the backward difference approximation

$$\frac{\phi_L - \phi_{L-1}}{\Delta x} = -\frac{\bar{q}}{k}$$

If we want to use the central difference approximation,

first, we introduce a fictitious mesh point

$$x_{L+1} \equiv x_L + \Delta x$$

with the associated "temperature" ϕ_{L+1}

then we have

$$-\phi_2 + 2\phi_1 = \frac{\Delta x^2 Q_1}{k} + \bar{\phi}_0$$

$$\vdots$$

$$-\phi_L + 2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2 Q_{L-1}}{k}$$

$$-\phi_{L+1} + 2\phi_L - \phi_{L-1} = \frac{\Delta x^2 Q_L}{k}$$

⇨ a set of L equations

$L+1$ unknowns

⇩

we need one more equation

ϕ_{L+1} has no physical significance as the point x_{L+1} lies outside of the boundary

one more equation
by the central difference approximation

$$\frac{\phi_{L+1} - \phi_{L-1}}{2\,\Delta x} = -\frac{\bar{q}}{k}$$

$$-k \left.\frac{d\phi}{dx}\right|_{L} = \bar{q}$$

$$-\phi_2 + 2\phi_1 = \frac{\Delta x^2\, Q_1}{k} + \bar{\phi}_0$$

$$\vdots$$

$$-\phi_L + 2\phi_{L-1} - \phi_{L-2} = \frac{\Delta x^2\, Q_{L-1}}{k}$$

$$-\phi_{L+1} + 2\phi_L - \phi_{L-1} = \frac{\Delta x^2\, Q_L}{k}$$

a set of L equations

$L+1$ unknowns

$\Downarrow$

we need one more equation

Example 1.2

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\frac{d^2\phi}{dx^2} = \phi$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

A mesh spacing $\Delta x = \frac{1}{3}$ is chosen

Solution)

The left side of the governing equation can be approximated as follows

$$\frac{d^2\phi_l}{dx^2} \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2}$$

Example 1.2

governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ **A mesh spacing:** $\Delta x = \dfrac{1}{3}$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

Solution)

$d\phi / dx\big|_3 = 1$

$\phi_0 = 0$ $\quad \boxed{\phi_1} \quad \boxed{\phi_2} \quad \boxed{\phi_3}$ **3 unknowns**

$\Delta x = \dfrac{1}{3}$

$x_0 = 0 \qquad x_1 = \dfrac{1}{3} \qquad x_2 = \dfrac{2}{3} \qquad x_3 = 1$

$$\frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = \phi_l \;\Rightarrow\; \phi_{l+1} - 2\phi_l + \phi_{l-1} = \Delta x^2 \phi_l$$

when $l = 1$, $\phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$

$l = 2$, $\phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$

2 equations

Solution)

Using a backward difference representation of the derivative boundary condition at $x_3 = 1$ produces

$$\frac{\phi_l - \phi_{l-1}}{\Delta x} = \left.\frac{d\phi}{dx}\right|_{x_l} \Rightarrow \frac{\phi_3 - \phi_2}{\Delta x} = 1 \Rightarrow \phi_3 - \phi_2 = \frac{1}{3}$$

$$\phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$$

$$\phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$$

$$\xrightarrow{\phi_0 = 0,\ \Delta x = \frac{1}{3}}$$

$$\boxed{\begin{aligned} \phi_3 - \phi_2 &= \frac{1}{3} \\ \phi_2 - 2\frac{1}{9}\phi_1 &= 0 \\ \phi_1 - 2\frac{1}{9}\phi_2 + \phi_3 &= 0 \end{aligned}}$$

$$\boxed{\phi_1 = 0.2477, \phi_2 = 0.5229, \phi_3 = 0.8563}$$

Analytic Solution)

Suppose that $\phi = e^{\lambda x}$

Substituting $\phi = e^{\lambda x}$ into the governing equation gives

$$\lambda^2 e^{\lambda x} = e^{\lambda x} \quad \Rightarrow \quad \lambda^2 = 1 \quad \Rightarrow \quad \lambda = \pm 1$$

General solution: $c_1 e^{x} + c_2 e^{-x}$

Analytic Solution)

General solution: $\phi = c_1 e^{x} + c_2 e^{-x} \Rightarrow \dfrac{d\phi}{dx} = c_1 e^{x} - c_2 e^{-x}$

From B.C.: $\phi(0) = 0$ $\quad d\phi / dx\big|_{x=1} = 1$

$$\boxed{c_1 + c_2 = 0 \qquad c_1 e - c_2 \frac{1}{e} = 1}$$

$$\Downarrow$$

$$c_1 = \frac{1}{e + 1/e}, \qquad c_2 = \frac{-1}{e + 1/e}$$

$$\phi_1 = 0.2477, \phi_2 = 0.5229, \phi_3 = 0.8563$$

Analytic Solution)

$$\phi = \frac{1}{e+1/e} e^x - \frac{1}{e+1/e} e^{-x}$$

$$\phi_1 = \phi\left(\frac{1}{3}\right) = 0.2200, \phi_2 = \phi\left(\frac{2}{3}\right) = 0.4648, \phi_3 = \phi(1) = 0.7616$$

FDM) $\phi_1 = 0.2477, \phi_2 = 0.5229, \phi_3 = 0.8563$

Example 1.3

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

A mesh spacing $\Delta x = \dfrac{1}{3}$ **is chosen**

Solve this problem using central difference approximation for B.C.

Solution)

The left side of the governing equation can be approximated as follows

$$\frac{d^2\phi_l}{dx^2} \approx \frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2}$$

Example 1.3

governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ **A mesh spacing:** $\Delta x = \dfrac{1}{3}$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

Solution)

$d\phi / dx\big|_3 = 1$

$\phi_0 = 0$ $\boxed{\phi_1}$ $\boxed{\phi_2}$ $\boxed{\phi_3}$ $\boxed{\phi_4}$ **4 unknowns**

$\Delta x = \dfrac{1}{3}$

$x_0 = 0$ $x_1 = \dfrac{1}{3}$ $x_2 = \dfrac{2}{3}$ $x_3 = 1$ $x_4 = \dfrac{4}{3}$

The fictitious mesh point

$$\frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = \phi_l \Rightarrow \phi_{l+1} - 2\phi_l + \phi_{l-1} = \Delta x^2 \phi_l$$

when $l = 1,$ $\phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$

$l = 2,$ $\phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$ **3 equations**

$l = 3,$ $\phi_4 - 2\phi_3 + \phi_2 = \Delta x^2 \phi_3$

Solution) **Central differencing of the derivative boundary condition at $x_3 = 1$ produces**

$$\frac{\phi_{l+1} - \phi_{l-1}}{2\Delta x} = \left.\frac{d\phi}{dx}\right|_{x_l} \Rightarrow \frac{\phi_4 - \phi_2}{2\Delta x} = 1 \Rightarrow$$

$$\phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$$

$$\phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$$

$$\phi_4 - 2\phi_3 + \phi_2 = \Delta x^2 \phi_3$$

$$\xrightarrow{\phi_0 = 0,\ \Delta x = \frac{1}{3}}$$

$$\boxed{\begin{aligned} \phi_4 - \phi_2 &= \frac{2}{3} \\ \phi_2 - 2\frac{1}{9}\phi_1 &= 0 \\ \phi_1 - 2\frac{1}{9}\phi_2 + \phi_3 &= 0 \\ \phi_2 - 2\frac{1}{9}\phi_3 + \phi_4 &= 0 \end{aligned}}$$

$$\boxed{\phi_1 = 0.2168, \phi_2 = 0.4576, \phi_3 = 0.7493}$$

Example 1.2, 1.3

governing equation $\dfrac{d^2\phi}{dx^2}=\phi$ **A mesh spacing:** $\Delta x=\dfrac{1}{3}$

Boundary Condition $\phi=0$ **at** $x=0$ **and** $d\phi/dx=1$ **at** $x=1$

Solution using backward difference representation of the derivative boundary condition

$$\boxed{\phi_1=0.2477,\phi_2=0.5229,\phi_3=0.8563}$$

Solution using central differencing of the derivative boundary condition

$$\boxed{\phi_1=0.2168,\phi_2=0.4576,\phi_3=0.7493}$$

Exact Solution

$$\boxed{\phi_1=0.2200,\phi_2=0.4648,\phi_3=0.7616}$$

Solution using central differencing can be seen to be considerably more accurate than the solution calculated using the backward difference representation of the derivative B.C

Nonlinear Problems

Physical Phenomena

⇩ mathematical modeling

Nonlinear Differential Equation and/or B.C.

⇩ usually fail

Exact Solution

The mathematical modeling of physical problems frequently produces governing differential equations and/or boundary conditions that are nonlinear in character

whereas analytical methods of solution for linear equations normally fail to cope with nonlinear differential equations

Nonlinear Differential Equation

Discretization
Approximation

Nonlinear Algebraic equations

When the boundary value problem is nonlinear, application of the finite difference method produces a set of nonlinear algebraic equations

we can consider the physically realistic problem where the thermal conductivity k is a given function of the temperature ϕ

then, the governing equation is nonlinear

$$\frac{d}{dx}\left[k(\phi)\frac{d\phi}{dx}\right]=-Q(x)$$

$$\frac{d}{dx}\left[k(\phi)\frac{d\phi}{dx}\right] = -Q(x)$$

by using a central difference approximation, we can write

$$\frac{k(\phi)\left.\frac{d\phi}{dx}\right|_{l+\frac{1}{2}} - k(\phi)\left.\frac{d\phi}{dx}\right|_{l-\frac{1}{2}}}{\Delta x} = -Q_l$$

or, $$k(\phi)\left.\frac{d\phi}{dx}\right|_{l+\frac{1}{2}} - k(\phi)\left.\frac{d\phi}{dx}\right|_{l-\frac{1}{2}} = -\Delta x\, Q_l$$

where the subscript $l+\frac{1}{2}$ **indicates an evaluation at the point midway between** x_l **and** x_{l+1}

$$k(\phi)\left.\frac{d\phi}{dx}\right|_{l+\frac{1}{2}} - k(\phi)\left.\frac{d\phi}{dx}\right|_{l-\frac{1}{2}} = -\Delta x\, Q_l$$

$$\left.\frac{d\phi}{dx}\right|_{l+\frac{1}{2}} = \frac{\phi_{l+1} - \phi_l}{\Delta x}$$

by using a central difference approximation again,

$$k(\phi_{l+\frac{1}{2}})\frac{\phi_{l+1} - \phi_l}{\Delta x} - k(\phi_{l-\frac{1}{2}})\frac{\phi_l - \phi_{l-1}}{\Delta x} = -\Delta x\, Q_l$$

$$k(\phi_{l+\frac{1}{2}})\left(\phi_{l+1} - \phi_l\right) - k(\phi_{l-\frac{1}{2}})\left(\phi_l - \phi_{l-1}\right) = -\left(\Delta x\right)^2 Q_l$$

thus application of the finite difference method to the original nonlinear differential equation has produced the set of nonlinear algebraic equations

$$k(\phi_{l+\frac{1}{2}})\phi_{l+1} - \left[k(\phi_{l+\frac{1}{2}}) + k(\phi_{l-\frac{1}{2}})\right]\phi_l + k(\phi_{l-\frac{1}{2}})\phi_{l-1} = -\left(\Delta x\right)^2 Q_l \quad , l = 1, 2, \ldots, L-1$$

it should be noted that this equation reduces to linear equations when k is constant

nonlinear algebraic equations

$$-k(\phi_{l+\frac{1}{2}})\phi_{l+1}+\left[k(\phi_{l+\frac{1}{2}})+k(\phi_{l-\frac{1}{2}})\right]\phi_l-k(\phi_{l-\frac{1}{2}})\phi_{l-1}=\left(\Delta x\right)^2 Q_l \quad , l=1,2,\ldots,L-1$$

it may be conveniently expressed in the form

$$\mathbf{K}(\boldsymbol{\phi})\boldsymbol{\phi}=\mathbf{f}$$

Simple iteration in which the system of equations is solved repeatedly with successively improved values of $\mathbf{K}(\phi)$

If we start from some initial guess $\phi = \phi_0$

and evaluate the matrix $\mathbf{K}(\phi_0) = \mathbf{K}_0$

an improved approximation for ϕ_1 can be obtained as

$$\phi_1 = \mathbf{K}_0^{-1}\mathbf{f}$$

This process can be obviously continued writing

$$\phi_n = \mathbf{K}_{n-1}^{-1}\mathbf{f}$$

$$\phi_n = \mathbf{K}_{n-1}^{-1}\mathbf{f}$$

This process is proceeding until the difference between

ϕ_n and ϕ_{n-1} is within a suitable tolerance

Example 1.4

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\frac{d}{dx}\left[k\frac{d\phi}{dx}\right]=-10x$, where $k=1+0.1\phi$

Boundary Condition $\phi=0$ at $x=0$ and $\phi=0$ at $x=1$

A mesh spacing $\Delta x=\frac{1}{3}$ is chosen

Solution)

Using central difference representation,
the governing equation can be approximated as follows

$$-k_{l+1/2}\,\phi_{l+1}+\left(k_{l+1/2}+k_{l-1/2}\right)\phi_l-k_{l-1/2}\,\phi_{l-1}=10x_l\Delta x^2$$

$$\Delta x=\frac{1}{3}$$

Example 1.4

$$x_1=\frac{1}{3}$$

Solution)

$$x_2=\frac{2}{3}$$

One methods of obtaining this value is to use the approximation

$$\phi_{1/2}\approx\frac{\phi_0+\phi_1}{2}\qquad\phi_{3/2}\approx\frac{\phi_1+\phi_2}{2}\qquad\phi_{5/2}\approx\frac{\phi_2+\phi_3}{2}$$

$\phi_{1/2},\phi_{3/2},\phi_{5/2}$ **can be represented with** $\phi_0,\phi_1,\phi_2,\phi_3$

$$k_{1/2}=1+0.1\phi_{1/2}=1+0.05\left(\phi_0+\phi_1\right)=1+0.05\phi_1$$

$$k_{3/2}=1+0.1\phi_{3/2}=1+0.05\left(\phi_1+\phi_2\right)$$

$$k_{5/2}=1+0.1\phi_{5/2}=1+0.05\left(\phi_2+\phi_3\right)=1+0.05\phi_2$$

Example 1.4

Solution)

$\Delta x = \frac{1}{3}$

$\boxed{\phi_0 = 0}$	$\boxed{\phi_1}$	$\boxed{\phi_2}$	$\boxed{\phi_3 = 0}$
$x_0 = 0$	$x_1 = \frac{1}{3}$	$x_2 = \frac{2}{3}$	$x_3 = 1$

2 knowns

2 unknowns

The central difference representation of the governing equation

$$-k_{l+1/2}\,\phi_{l+1} + \left(k_{l+1/2} + k_{l-1/2}\right)\phi_l - k_{l-1/2}\,\phi_{l-1} = 10 x_l \Delta x^2$$

when $l = 1,$

$$-k_{3/2}\,\phi_2 + \left(k_{3/2} + k_{1/2}\right)\phi_1 - k_{1/2}\,\phi_0 = 10 x_1 \Delta x^2$$

$l = 2,$

$$-k_{5/2}\,\phi_3 + \left(k_{5/2} + k_{3/2}\right)\phi_2 - k_{3/2}\,\phi_1 = 10 x_2 \Delta x^2$$

$$\Delta x = \frac{1}{3}$$

Example 1.4

$$\phi_{1/2} \approx \frac{\phi_0 + \phi_1}{2},\ \phi_{3/2} \approx \frac{\phi_1 + \phi_2}{2},\ \phi_{5/2} \approx \frac{\phi_2 + \phi_3}{2}$$

$$x_1 = \frac{1}{3}$$

Solution) $\boxed{\phi_0 = 0}$ $\boxed{\phi_1}$ $\boxed{\phi_2}$ $\boxed{\phi_3 = 0}$ **2 knowns, 2 unknowns**

$$x_2 = \frac{2}{3}$$

$l = 1,$

$$-k_{3/2}\,\phi_2 + \left(k_{3/2} + k_{1/2}\right)\phi_1 - k_{1/2}\,\cancel{\phi_0} = 10 x_1 \Delta x^2$$

$l = 2,$

$$-k_{5/2}\,\cancel{\phi_3} + \left(k_{5/2} + k_{3/2}\right)\phi_2 - k_{3/2}\,\phi_1 = 10 x_2 \Delta x^2$$

$$\downarrow$$

$$-k_{3/2}\,\phi_2 + \left(k_{3/2} + k_{1/2}\right)\phi_1 = 10 x_1 \Delta x^2$$

$$\left(k_{5/2} + k_{3/2}\right)\phi_2 - k_{3/2}\,\phi_1 = 10 x_2 \Delta x^2$$

$$\Delta x = \frac{1}{3},\ x_1 = \frac{1}{3},\ x_2 = \frac{2}{3}$$

$$-k_{3/2}\,\phi_2 + \left(k_{3/2} + k_{1/2}\right)\phi_1 = \frac{10}{27}$$

$$\left(k_{5/2} + k_{3/2}\right)\phi_2 - k_{3/2}\,\phi_1 = \frac{20}{27}$$

$$\Delta x = \frac{1}{3}$$

Example 1.4

$$\phi_{1/2} \approx \frac{\phi_0 + \phi_1}{2},\ \ \phi_{3/2} \approx \frac{\phi_1 + \phi_2}{2},\ \phi_{5/2} \approx \frac{\phi_2 + \phi_3}{2}$$

$$x_1 = \frac{1}{3}$$

$$x_2 = \frac{2}{3}$$

Solution)

$$-k_{3/2}\,\phi_2 + \left(k_{3/2} + k_{1/2}\right)\phi_1 = \frac{10}{27}$$

$$\left(k_{5/2} + k_{3/2}\right)\phi_2 - k_{3/2}\,\phi_1 = \frac{20}{27}$$

$$\begin{bmatrix} \left(k_{3/2} + k_{1/2}\right) & -k_{3/2} \\ -k_{3/2} & \left(k_{5/2} + k_{3/2}\right) \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} \frac{10}{27} \\ \frac{20}{27} \end{bmatrix}$$

$$k_{1/2} = 1 + 0.05\phi_1,$$

$$k_{3/2} = 1 + 0.05\left(\phi_1 + \phi_2\right),$$

$$k_{5/2} = 1 + 0.05\phi_2$$

$$\underbrace{\begin{bmatrix} 2 + 0.05\left(2\phi_1 + \phi_2\right) & -1 - 0.05\left(\phi_1 + \phi_2\right) \\ -1 - 0.05\left(\phi_1 + \phi_2\right) & \left(2 + 0.05\left(\phi_1 + 2\phi_2\right)\right) \end{bmatrix}}_{\mathbf{K}(\boldsymbol{\phi})} \underbrace{\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}}_{\boldsymbol{\phi}} = \underbrace{\begin{bmatrix} \frac{10}{27} \\ \frac{20}{27} \end{bmatrix}}_{\mathbf{f}}$$

$$\Delta x = \frac{1}{3}$$

Example 1.4

$$x_1 = \frac{1}{3} \qquad x_2 = \frac{2}{3}$$

Solution)

$$\mathbf{K}(\boldsymbol{\phi})\boldsymbol{\phi} = \mathbf{f} \qquad \mathbf{K}(\boldsymbol{\phi}) = \begin{bmatrix} 2+0.05(2\phi_1+\phi_2) & -1-0.05(\phi_1+\phi_2) \\ -1-0.05(\phi_1+\phi_2) & (2+0.05(\phi_1+2\phi_2)) \end{bmatrix}, \boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}, \mathbf{f} = \begin{bmatrix} \frac{10}{27} \\ \frac{20}{27} \end{bmatrix}$$

Initial guess

$$\text{Step 0: } \boldsymbol{\phi}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \mathbf{K}_0 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \xrightarrow{\boldsymbol{\phi}_1 = \mathbf{K}_0^{-1}\mathbf{f}} \boldsymbol{\phi}_1 = \begin{bmatrix} 0.49383 \\ 0.61728 \end{bmatrix}$$

$$\text{Step 1: } \boldsymbol{\phi}_1 = \begin{bmatrix} 0.49383 \\ 0.61728 \end{bmatrix} \longrightarrow \mathbf{K}_1 = \begin{bmatrix} 2.08025 & -1.05556 \\ -1.05556 & 2.08642 \end{bmatrix} \xrightarrow{\boldsymbol{\phi}_2 = \mathbf{K}_1^{-1}\mathbf{f}} \boldsymbol{\phi}_2 = \begin{bmatrix} 0.48190 \\ 0.59883 \end{bmatrix}$$

$$\text{Step 2: } \boldsymbol{\phi}_2 = \begin{bmatrix} 0.48190 \\ 0.59883 \end{bmatrix} \longrightarrow \mathbf{K}_2 = \begin{bmatrix} 2.07813 & -1.05404 \\ -1.05404 & 2.08398 \end{bmatrix} \xrightarrow{\boldsymbol{\phi}_3 = \mathbf{K}_2^{-1}\mathbf{f}} \boldsymbol{\phi}_3 = \begin{bmatrix} 0.48221 \\ 0.59934 \end{bmatrix}$$

$$\text{Step 3: } \boldsymbol{\phi}_3 = \begin{bmatrix} 0.48221 \\ 0.59934 \end{bmatrix} \longrightarrow \mathbf{K}_3 = \begin{bmatrix} 2.07819 & -1.05408 \\ -1.05408 & 2.08404 \end{bmatrix} \xrightarrow{\boldsymbol{\phi}_4 = \mathbf{K}_3^{-1}\mathbf{f}} \boldsymbol{\phi}_4 = \begin{bmatrix} 0.48220 \\ 0.59932 \end{bmatrix}$$

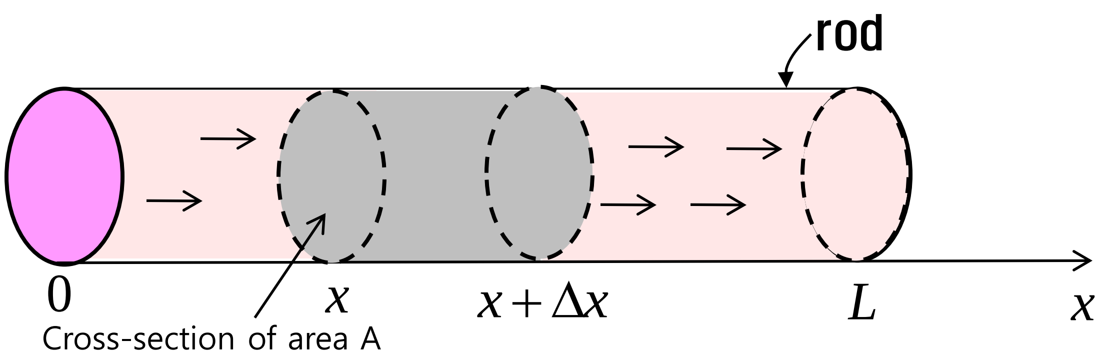
[REVIEW] HEAT EQUATION

Some Examples of Continuum Problems

☑1-D Heat Equation

■ Assumptions

- The flow of heat within the rod takes place only in the x-direction
- The lateral, or curved, surface of the rod is insulated; that is , no heat escapes from this surface
- No heat is being generated within rod
- The rod is homogeneous; that is, its mass per unit volume ρ is constant
- The specific heat(비열)* γ and thermal conductivity(열전도도)** K of material of the rod are constants



One dimension flow of heat

* Oxtoby, Principles of Modern Chemistry, Sixth Edition, Thomson, Index 1.25, “Specific heat capacity : The amount of heat required to raise the temperature of one gram of a substance by one kelvin at constant pressure”

** 여상도, 열역학 개념의 해설, 청문각, 2006, p18 “온도가 동일한 두 물체와 우리의 손이 닿았을 때 그 차갑고 뜨거운 정도가 다른 이유는, 두 물체의 온도가 다르기 때문이 아니라 우리 손에서 물체로 이동하는 열의 전달 속도가 다르기 때문이다. 열전도도가 큰 철판이 열전도도가 작은 나무판에 비해 훨씬 빨리 손으로부터 열을 빼앗아 간다.”

1-D Heat Equation

Two empirical laws of heat conduction

(i) The quantity of heat(열량) Q in an element of mass(질량소) m is

$$Q = c m \phi$$

specific heat(비열) c
temperature of the element q

(ii) The rate of heat flow(열흐름율) Q_t through the cross-section indicated in Figure is proportional to the area A of the cross-section and the partial derivative with respect to x of the temperature

$$Q_t = -kA \frac{d\phi}{dx}$$

thermal conductivity(열전도도) k

Heat flows in the direction of decreasing temperature

1-D Heat Equation

$$(i)\ Q = c\ m\phi$$

$$\Big\downarrow \text{ Substitute } m = \rho A \Delta x$$

$$Q = c\,\rho A \Delta x\,\phi$$

$$\Big\downarrow \text{ Differentiate respect to time}$$

$$\frac{dQ}{dt} = c\,\rho A \Delta x \frac{dq}{dt} \cdots (1)$$

1-D Heat Equation

$$(i)\ Q = c\ m\phi$$

$\downarrow$ Substitute $m = \rho A \Delta x$

$$Q = c\,\rho A \Delta x\,\phi$$

$\downarrow$ Differentiate respect to time

$$\frac{dQ}{dt} = c\,\rho A\,\Delta x\,\frac{dq}{dt} \cdots (1)$$

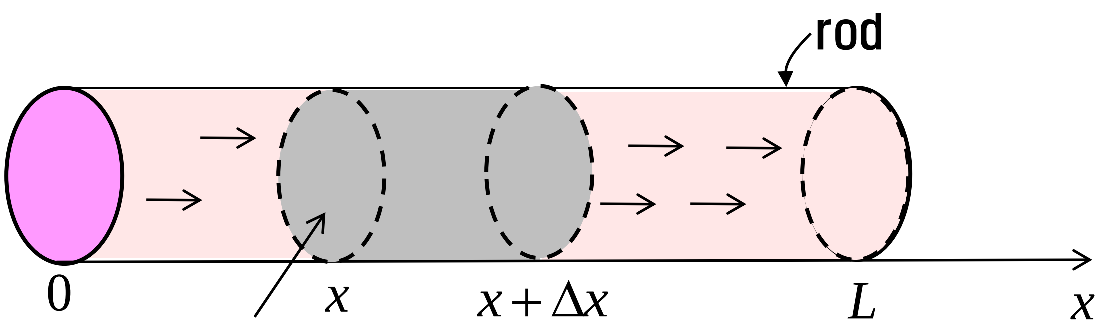
1-D Heat Equation

$$(ii)\ \frac{dQ}{dt} = -kA\frac{d\phi}{dx}$$ The rate of heat flow(열흐름율) through the cross-section

Net flow rate in $(x, x+\Delta x)$

$$\frac{dQ}{dt} = -kA\frac{d\phi}{dx}(x,t) - [-kA\frac{d\phi}{dx}(x+\Delta x,t)]$$

$$= kA[\frac{d\phi}{dx}(x+\Delta x,t) - \frac{d\phi}{dx}(x,t)] \quad \cdots(2)$$



One dimension flow of heat

1-D Heat Equation

$$\frac{dQ}{dt} = c\,\rho A \Delta x \frac{d\phi}{dt} \cdots (1)$$

From (1) and (2)

$$\frac{dQ}{dt} = kA[\frac{d\phi}{dx}(x+\Delta x, t) - \frac{d\phi}{dx}(x,t)] \;\cdots (2)$$

$$kA[\frac{d\phi}{dx}(x+\Delta x, t) - \frac{d\phi}{dx}(x,t)] = c\,\rho A \Delta x \frac{d\phi}{dt}$$

$$\frac{1}{\Delta x} k \left[\frac{d\phi}{dx}(x+\Delta x, t) - \frac{d\phi}{dx}(x,t) \right] = c\rho \frac{d\phi}{dt}$$

$$q_x(x+\Delta x, t) - q_x(x,t)$$

$$= \frac{q(x+\Delta x, t) - q(x,t)}{\Delta x}$$

As $\Delta x \to 0,$

$$\frac{1}{dx}\left[k \frac{d\phi}{dx} \right] = c\rho \frac{d\phi}{dt}$$

O.D.E or P.D.E with B/C and I/C

Discretization
Approximation

↓

Algebraic form

3. Finite Difference Method and Finite Element Method

3.3 Finite Element Method(FEM)

Of the various forms of discretization which are possible, one of the simplest is the *finite difference process* and the others are various trial function approximations falling under the general classification of *finite element methods*.

[Zienkiewicz 1983] p.2

The "Finite element method"
is
a tool
for the approximate solution
of differential equation (with B/C*)
governing mathematical modeling
diverse physical phenomena

[Zienkiewicz 1983] vii

Function Approximation by trial function

Function Approximation by Trial Functions

Introduction

In the finite difference method we have concentrated on defining the value of the unknown function Φ(x) at a finite number of values x

Alternative methods for determining numerically the solution to differential equations can ,however, be developed by making the process of function approximation more systematic and general

Function Approximation by Trial Functions

We wish to approximate a given function ϕ in some region bounded by a closed curve Γ

In problems involving differential equations,
it is required to find the solution satisfying certain boundary conditions.

We ,therefore, attempt initially to construct approximations which are exact equal to prescribed values of ϕ on the boundary curve Γ

frequently referred *as shape or basis function*

If we can find any function ψ satisfying $\psi\big|_{\Gamma} = \phi\big|_{\Gamma}$

and if we introduce a set of independent *trial functions*

$\{N_m \;;\; m = 1, 2, 3 \ldots\}$ such that $N_m\big|_{\Gamma} = 0$ for all m

then at all points in Ω , we can approximate to ϕ by

$$\phi \;\approx\; \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

The manner in which ψ and the trial function set are defined automatically ensures that $\hat{\phi}\big|_\Gamma = \phi\big|_\Gamma$ the approximation has the property that whatever the values of the parameters a_m

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

The trial function set should clearly be chosen so as to ensure that improvement in the approximation occurs with increase in the number M of trial functions used

⇩

Completeness (convergence) requirement

$$\hat{\phi} \to \phi \quad \text{as} \quad M \to \infty$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

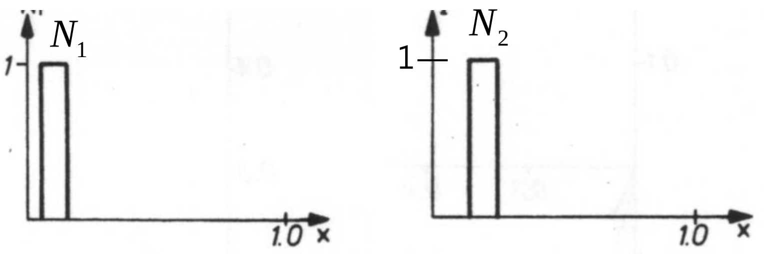
Completeness

$$\hat{\phi} \to \phi$$

as $M \to \infty$

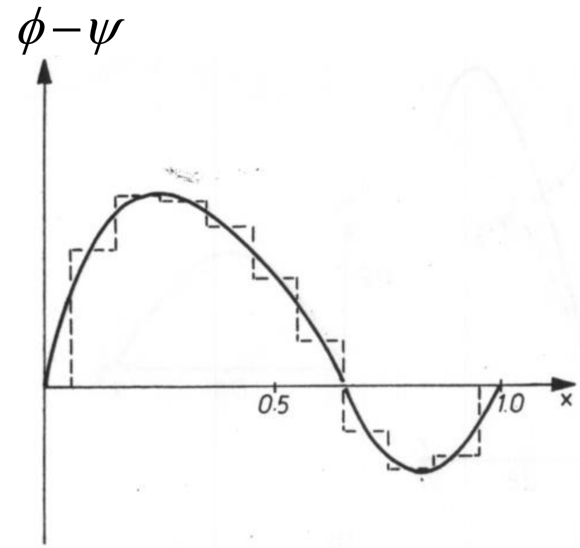
Example)

the chosen functions N_m **are of a discontinuous form, shown to have value unity on a suitable interval and the value zero elsewhere**



...

any function can be approximated as closely as desired by dividing the total domain into ever smaller intervals



We shall now attempt to develop a general method for determining the parameters a_m in the approximation

We begin by introducing the error, or *residual* R_Ω in the approximation

$$R_\Omega \equiv \phi - \hat{\phi}$$

which is a function of position in Ω

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi|_\Gamma = \phi|_\Gamma$$

$$N_m|_\Gamma = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

Weighted Residual Approximations

In an attempt to reduce this residual in some overall manner over the whole domain Ω

we could require that an appropriate number of integrals of the error over Ω, weighted in different ways, be zero

$$\int_{\Omega} W_l \, (\phi - \hat{\phi}) \, d\Omega \equiv \int_{\Omega} W_l \, R_{\Omega} \, d\Omega = 0$$

$$l = 1, 2, \ldots, M$$

where W_l is a set of independent *weighting functions*

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_{\Gamma} = \phi\big|_{\Gamma}$$

$$N_m\big|_{\Gamma} = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

residual

$$R_{\Omega} = \phi - \hat{\phi}$$

The general completeness (convergence) requirement

$$\hat{\phi} \rightarrow \phi \quad \text{as} \quad M \rightarrow \infty$$

can then be cast in an alternative form by requiring

$$\int_{\Omega} W_l \, R_{\Omega} \, d\Omega = 0 \quad \text{for all } l \quad \text{as} \quad M \rightarrow \infty$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_{\Gamma} = \phi\big|_{\Gamma}$$

$$N_m\big|_{\Gamma} = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

residual

$$R_{\Omega} = \phi - \hat{\phi}$$

alternative form of completeness requirement

$$\int_\Omega W_l \, R_\Omega \, d\Omega = 0 \quad \text{for all } l \quad \text{as} \quad M \to \infty$$

$$\boxed{\int_\Omega W_l \, (\phi - \hat{\phi}) \, d\Omega = 0}$$

standard weighted residual statement

Find

$$\int_\Omega W_l \, (\phi - \psi - \sum_{m=1}^{M} \boxed{a_m} N_m) \, d\Omega = 0$$

User defined weighting function

chosen to be zero at the B/C

chosen to satisfy the B/C

the function to be approximated is given

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi\big|_\Gamma = \phi\big|_\Gamma$$

$$N_m\big|_\Gamma = 0$$

where, a_m are some parameters which are computed so as to obtain a good "fit"

residual

$$R_\Omega = \phi - \hat{\phi}$$

Expansion of the equation of weighted residual

$$\int_{\Omega} W_l \left(\phi - \psi - \sum_{m=1}^{M} \boxed{a_m} N_m \right) d\Omega = 0$$

Given: $\square$ **Find:** $\boxed{}$

$$\int_{\Omega} W_l (\phi - \psi) d\Omega = \int_{\Omega} \sum_{m=1}^{M} a_m W_l N_m d\Omega$$

$$\int_{\Omega} W_l (\phi - \psi) d\Omega = \sum_{m=1}^{M} \left[a_m \int_{\Omega} W_l N_m d\Omega \right]$$

$m = 1, 2, .., M$

$$\int_{\Omega} W_l (\phi - \psi) d\Omega = \boxed{a_1 \int_{\Omega} W_l N_1 d\Omega + a_2 \int_{\Omega} W_l N_2 d\Omega + \cdots a_M \int_{\Omega} W_l N_M d\Omega}$$

$l = 1, 2, .., M$

$$\int_{\Omega} W_1 (\phi - \psi) d\Omega = a_1 \int_{\Omega} W_1 N_1 d\Omega + a_2 \int_{\Omega} W_1 N_2 d\Omega + \cdots a_M \int_{\Omega} W_1 N_M d\Omega$$

$$\int_{\Omega} W_2 (\phi - \psi) d\Omega = a_1 \int_{\Omega} W_2 N_1 d\Omega + a_2 \int_{\Omega} W_2 N_2 d\Omega + \cdots a_M \int_{\Omega} W_2 N_M d\Omega$$

$$\vdots$$

$$\int_{\Omega} W_M (\phi - \psi) d\Omega = a_1 \int_{\Omega} W_M N_1 d\Omega + a_2 \int_{\Omega} W_M N_2 d\Omega + \cdots a_M \int_{\Omega} W_M N_M d\Omega$$

Given: Find:

(Derivation)

$$\int_\Omega W_1(\phi-\psi)d\Omega = a_1\int_\Omega W_1 N_1 d\Omega + a_2\int_\Omega W_1 N_2 d\Omega + \cdots a_M\int_\Omega W_1 N_M d\Omega$$

$$\int_\Omega W_2(\phi-\psi)d\Omega = a_1\int_\Omega W_2 N_1 d\Omega + a_2\int_\Omega W_2 N_2 d\Omega + \cdots a_M\int_\Omega W_2 N_M d\Omega$$

$$\vdots$$

$$\int_\Omega W_M(\phi-\psi)d\Omega = a_1\int_\Omega W_M N_1 d\Omega + a_2\int_\Omega W_M N_2 d\Omega + \cdots a_M\int_\Omega W_M N_M d\Omega$$

in matrix form

$$\begin{bmatrix} \int_\Omega W_1(\phi-\psi)d\Omega \\ \int_\Omega W_2(\phi-\psi)d\Omega \\ \vdots \\ \int_\Omega W_M(\phi-\psi)d\Omega \end{bmatrix} = \begin{bmatrix} \int_\Omega W_1 N_1 d\Omega & \int_\Omega W_1 N_2 d\Omega & \cdots & \int_\Omega W_1 N_M d\Omega \\ \int_\Omega W_2 N_1 d\Omega & \int_\Omega W_2 N_2 d\Omega & \cdots & \int_\Omega W_2 N_M d\Omega \\ & \vdots & & \\ \int_\Omega W_M N_1 d\Omega & \int_\Omega W_M N_2 d\Omega & \cdots & \int_\Omega W_M N_M d\Omega \end{bmatrix} \boxed{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}}$$

(Derivation)

$$\int_\Omega W_1(\phi-\psi)d\Omega = a_1\int_\Omega W_1 N_1 d\Omega + a_2\int_\Omega W_1 N_2 d\Omega + \cdots a_M\int_\Omega W_1 N_M d\Omega$$

$$\int_\Omega W_2(\phi-\psi)d\Omega = a_1\int_\Omega W_2 N_1 d\Omega + a_2\int_\Omega W_2 N_2 d\Omega + \cdots a_M\int_\Omega W_2 N_M d\Omega$$

$$\vdots$$

$$\int_\Omega W_M(\phi-\psi)d\Omega = a_1\int_\Omega W_M N_1 d\Omega + a_2\int_\Omega W_M N_2 d\Omega + \cdots a_M\int_\Omega W_M N_M d\Omega$$

in matrix form

$$\begin{bmatrix} \int_\Omega W_1(\phi-\psi)d\Omega \\ \int_\Omega W_2(\phi-\psi)d\Omega \\ \vdots \\ \int_\Omega W_M(\phi-\psi)d\Omega \end{bmatrix} = \begin{bmatrix} \int_\Omega W_1 N_1 d\Omega & \int_\Omega W_1 N_2 d\Omega & \cdots & \int_\Omega W_1 N_M d\Omega \\ \int_\Omega W_2 N_1 d\Omega & \int_\Omega W_2 N_2 d\Omega & \cdots & \int_\Omega W_2 N_M d\Omega \\ & \vdots & & \\ \int_\Omega W_M N_1 d\Omega & \int_\Omega W_M N_2 d\Omega & \cdots & \int_\Omega W_M N_M d\Omega \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}$$

$$[\mathbf{f}] = [\mathbf{K}]\,\boxed{[\mathbf{a}]}$$

Weighted Residual Approximations
- Matrix representation

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

(Derivation)

It can be written quite generally as

given

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

find

the function to be approximated is given

$$\mathbf{f} = [\,f_l\,], \quad f_l = \int_\Omega W_l\,(\phi - \psi)\,d\Omega$$

chosen to satisfy the B/C

chosen to be zero at the B/C

$$\mathbf{K} = [K_{lm}], \quad K_{lm} = \int_\Omega W_l\,N_m\,d\Omega$$

find

$$\mathbf{a}^T = [a_1 \; a_2 \; \cdots \; a_M\,]$$

various forms of weighting functions sets can be used in practice

☑ Unit Impulse

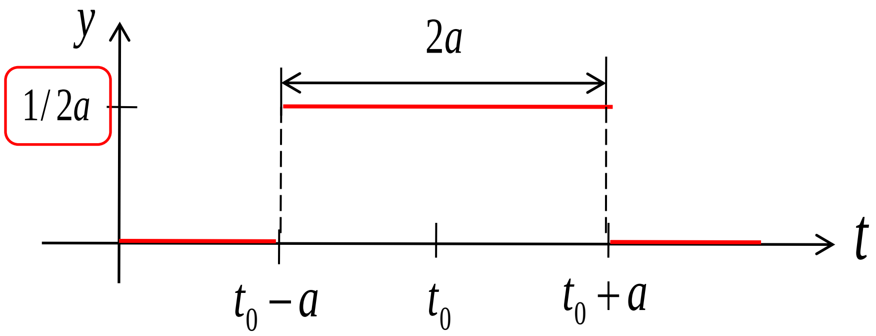
- External force of large magnitude that acts only for a very short period of time

$$\delta_a(t-t_0)=\begin{cases} 0, & 0\le t<t_0-a \\ \dfrac{1}{2a}, & t_0-a\le t<t_0+a \\ 0, & t\ge t_0+a \end{cases}$$

$$\int_0^{\infty}\delta_a(t-t_0)dt=1$$

→'Unit' impulse

Magnitude of the unit impulse function



For small a, $\delta_a(t-t_0)$ have large magnitude.

Weighted Residual Method
- Weighting function: Dirac delta function

The Dirac delta function is chosen as weighting functions sets.

☑ The Dirac Delta Function

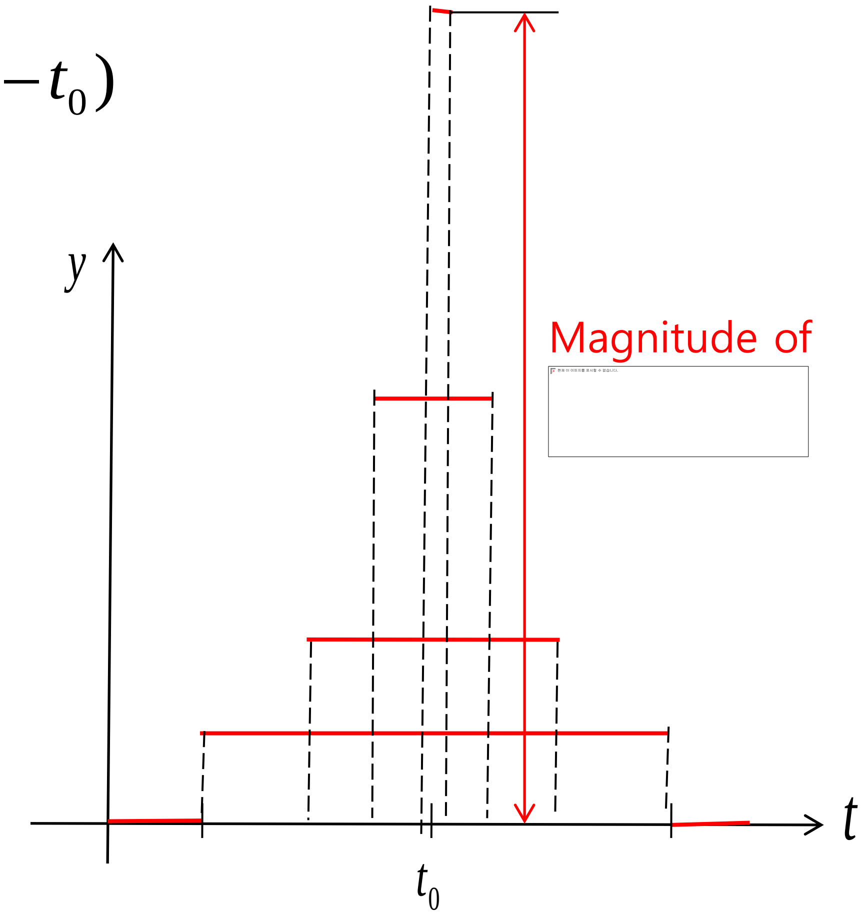
$$\delta(t-t_0)=\lim_{a\to 0}\delta_a(t-t_0)$$

For small *a*, $\delta_a(t-t_0)$ have large magnitude.

$$(i)\ \delta(t-t_0)=\begin{cases}\boxed{\infty,\quad t=t_0}\\ 0,\quad t\neq t_0\end{cases}$$

$$(ii)\ \int_0^x \delta(t-t_0)dt=1$$

$$(iii)\ \int_0^x G(t)\cdot\delta(t-t_0)dt=G(t_0)$$



Weighted Residual Method
- Point Collocation

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$W_l = \delta(x - x_l)$$,where $\delta(x-x_l)$ is the Dirac delta function

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$,\mathbf{f} = [\,f_l\,], \mathbf{K} = [K_{lm}], \mathbf{a}^T = [a_1 \;\, a_2 \;\, \cdots \;\, a_M\,]$$

$$f_l = \int_\Omega \delta(x - x_l)(\phi - \psi)d\Omega$$

$$= (\phi - \psi)\Big|_{x=x_l}$$

$$K_{lm} = \int_\Omega \delta(x - x_l)\,N_m d\Omega$$

$$= N_m\Big|_{x=x_l}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [\,f_l\,],$$

$$f_l = \int_\Omega W_l\,(\phi - \psi)d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l\,N_m d\Omega$$

$$\mathbf{a}^T = [a_1 \;\, a_2 \;\, \cdots \;\, a_M\,]$$

properties of the Dirac delta function

$$\delta(x - x_l) = 0 \;\;,\; x \neq x_l$$

$$\delta(x - x_l) = \infty \;,\; x = x_l$$

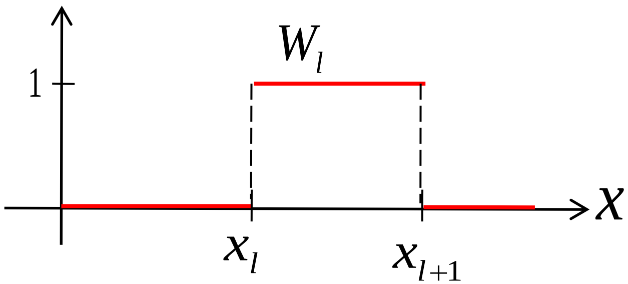
$$\int_{x_2}^{x_1} G(x)\,\delta(x - x_l)dx = G(x_l)$$

$$,\text{where } x_1 < x_l < x_2$$

Weighted Residual Method
- Subdomain Collocation

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$W_l = \begin{cases} 1 & , x_l < x < x_{l+1} \\ 0 & , x < x_l, \ x_{l+1} < x \end{cases}$$



$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$,\mathbf{f} = [f_l], \mathbf{K} = [K_{lm}], \mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$$

$$f_l = \int_{\Omega_{x<x_l}} 0 \cdot (\phi - \psi) dx + \int_{x_l}^{x_{l+1}} 1 \cdot (\phi - \psi) dx + \int_{\Omega_{x_{l+1}<x}} 0 \cdot (\phi - \psi) dx$$

$$= \int_{x_l}^{x_{l+1}} (\phi - \psi) dx$$

$$K_{lm} = \int_{\Omega_{x<x_l}} 0 \cdot N_m dx + \int_{x_l}^{x_{l+1}} 1 \cdot N_m dx + \int_{\Omega_{x_{l+1}<x}} 0 \cdot N_m dx$$

$$= \int_{x_l}^{x_{l+1}} N_m \, dx$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [f_l],$$

$$f_l = \int_\Omega W_l (\phi - \psi) d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l N_m d\Omega$$

$$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$$

Weighted Residual Method
- The Galerkin Method

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$W_l = N_l$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{K} = [K_{lm}]$$

$$K_{lm} = \int_\Omega N_l \, N_m dx$$

$$l, m = 1, 2, 3, \ldots, M$$

$$\mathbf{K} = \begin{bmatrix} \int_\Omega N_1 N_1 dx & \int_\Omega N_1 N_2 dx & \int_\Omega N_1 N_3 dx & \cdots & \int_\Omega N_1 N_M dx \\ \int_\Omega N_2 N_1 dx & \int_\Omega N_2 N_2 dx & \int_\Omega N_2 N_3 dx & \cdots & \int_\Omega N_2 N_M dx \\ \int_\Omega N_3 N_1 dx & \int_\Omega N_3 N_2 dx & \int_\Omega N_3 N_3 dx & \cdots & \int_\Omega N_3 N_M dx \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \int_\Omega N_M N_1 dx & \int_\Omega N_M N_2 dx & \int_\Omega N_M N_3 dx & \cdots & \int_\Omega N_M N_M dx \end{bmatrix}$$

$$\mathbf{f} = [f_l],$$

$$f_l = \int_\Omega W_l (\phi - \psi) d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l \, N_m d\Omega$$

$$\mathbf{a}^T = [a_1 \; a_2 \; \cdots \; a_M]$$

we notice the computational advantages of the method in that the matrix $\mathbf{K}$ is *symmetric*

Summary : Weighted Residual Method

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

Various forms of weighting functions sets lead to a different *weighted residual approximation methods*

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [\, f_l \,],$$

$$f_l = \int_\Omega W_l\,(\phi - \psi)\,d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l\, N_m\, d\Omega$$

$$\mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M\,]$$

Point Collocation

$$W_l = \delta(x - x_l)$$,where $\delta(x - x_l)$ is the Dirac delta function

Subdomain Collocation

$$W_l = \begin{cases} 1 & , x_l < x < x_{l+1} \\ 0 & , x < x_l,\ \ x_{l+1} < x \end{cases}$$

The Galerkin Method

$$W_l = N_l$$

Weighted Residual Method
- The Galerkin Method : Case 1

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

- Given function $\phi(x)$ over the range $0 \le x \le L$
- The trial function set for the approximation of the given function is $\{N_m = \sin(m\pi x / L); m = 1, 2, 3, \ldots\}$
- This trial function leads to typical coefficients

$$K_{lm} = \int_0^L \sin\frac{l\pi x}{L} \sin\frac{m\pi x}{L} dx$$

$$f_l = \int_0^L (\phi - \psi) \sin\frac{l\pi x}{L} dx$$

$$\mathbf{K}\mathbf{a} = \mathbf{f}$$

$\mathbf{f} = [f_l]$,

$f_l = \int_\Omega W_l (\phi - \psi) d\Omega$,

$\mathbf{K} = [K_{lm}]$,

$K_{lm} = \int_\Omega W_l N_m d\Omega$

$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$

↓

The Galerkin Method

$W_l = N_l$

$$f_l = \int_\Omega N_l (\phi - \psi) dx$$

$$K_{lm} = \int_\Omega N_l N_m dx$$

$$K_{lm} = \int_0^L \sin\frac{l\pi x}{L} \sin\frac{m\pi x}{L} dx$$

if $l = m$

$$K_{ll} = \int_0^L \sin^2\frac{l\pi x}{L} dx$$

$$= \frac{1}{2}\int_0^L \left(1-\cos\frac{2l\pi x}{L}\right) dx$$

$$= \frac{1}{2}[x]_0^L - \frac{1}{2}\left[\frac{L}{2l\pi}\sin\frac{2l\pi x}{L}\right]_0^L$$

$$= \frac{1}{2}L - \frac{1}{2}\cdot 0 - \frac{1}{2}\left(\frac{L}{2l\pi}\sin\frac{2l\pi L}{L} - \frac{L}{2l\pi}\sin\frac{2l\pi 0}{L}\right)$$

$$= \frac{1}{2}L$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$\mathbf{f} = [f_l],$

$f_l = \int_\Omega W_l\,(\phi - \psi) d\Omega,$

$\mathbf{K} = [K_{lm}],$

$K_{lm} = \int_\Omega W_l\, N_m d\Omega$

$\mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M]$

$\downarrow$

The Galerkin Method

$W_l = N_l$

$$\boxed{\begin{aligned} f_l &= \int_\Omega N_l\,(\phi - \psi) dx \\ K_{lm} &= \int_\Omega N_l\, N_m dx \end{aligned}}$$

$$K_{lm} = \int_0^L \sin\frac{l\pi x}{L}\sin\frac{m\pi x}{L}\,dx$$

,else if $l \neq m$

$$K_{lm} = \int_0^L \sin\frac{l\pi x}{L}\sin\frac{m\pi x}{L}\,dx$$

$$= -\frac{1}{2}\int_0^L \left(\cos\left(\frac{l\pi}{L}+\frac{m\pi}{L}\right)x - \cos\left(\frac{l\pi}{L}-\frac{m\pi}{L}\right)x\right)dx$$

$$= -\frac{1}{2}\left[\frac{\sin\left(\frac{l\pi}{L}+\frac{m\pi}{L}\right)x}{\frac{l\pi}{L}+\frac{m\pi}{L}} - \frac{\sin\left(\frac{l\pi}{L}-\frac{m\pi}{L}\right)x}{\frac{l\pi}{L}-\frac{m\pi}{L}}\right]_0^L$$

$$= 0$$

$$\mathbf{K}\mathbf{a} = \mathbf{f}$$

$\mathbf{f} = [f_l]$,

$f_l = \int_\Omega W_l(\phi-\psi)d\Omega$,

$\mathbf{K} = [K_{lm}]$,

$K_{lm} = \int_\Omega W_l N_m d\Omega$

$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$

$\downarrow$

The Galerkin Method

$W_l = N_l$

$$\boxed{f_l = \int_\Omega N_l(\phi-\psi)dx \qquad K_{lm} = \int_\Omega N_l N_m dx}$$

$$K_{lm} = \begin{cases} \frac{1}{2}L, & l = m \\ 0, & l \neq m \end{cases}$$

$$f_l = \int_0^L (\phi - \psi) \sin \frac{l\pi x}{L} dx$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

Diagonal form!!!

$$\begin{bmatrix} L/2 & 0 & \cdots & 0 & 0 \\ 0 & L/2 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & L/2 & 0 \\ 0 & 0 & \cdots & 0 & L/2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{M-1} \\ a_M \end{bmatrix} = \begin{bmatrix} \int_0^L (\phi - \psi) \sin \frac{\pi x}{L} dx \\ \int_0^L (\phi - \psi) \sin \frac{2\pi x}{L} dx \\ \vdots \\ \int_0^L (\phi - \psi) \sin \frac{(M-1)\pi x}{L} dx \\ \int_0^L (\phi - \psi) \sin \frac{M\pi x}{L} dx \end{bmatrix}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$\mathbf{f} = [f_l]$,

$f_l = \int_\Omega W_l (\phi - \psi) d\Omega$,

$\mathbf{K} = [K_{lm}]$,

$K_{lm} = \int_\Omega W_l N_m d\Omega$

$\mathbf{a}^T = [a_1 \; a_2 \; \cdots \; a_M]$

↓

The Galerkin Method

$W_l = N_l$

$$\boxed{\begin{aligned} f_l &= \int_\Omega N_l (\phi - \psi) dx \\ K_{lm} &= \int_\Omega N_l N_m dx \end{aligned}}$$

$$\begin{bmatrix} L/2 & 0 & \cdots & 0 & 0 \\ 0 & L/2 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & L/2 & 0 \\ 0 & 0 & \cdots & 0 & L/2 \end{bmatrix} \boxed{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{M-1} \\ a_M \end{bmatrix}} = \begin{bmatrix} \int_0^L (\phi-\psi)\sin\frac{\pi x}{L}dx \\ \int_0^L (\phi-\psi)\sin\frac{2\pi x}{L}dx \\ \vdots \\ \int_0^L (\phi-\psi)\sin\frac{(M-1)\pi x}{L}dx \\ \int_0^L (\phi-\psi)\sin\frac{M\pi x}{L}dx \end{bmatrix}$$

Find

$$\downarrow$$

$$\frac{L}{2}a_1 = \int_0^L (\phi-\psi)\sin\frac{\pi x}{L}dx$$

$$\frac{L}{2}a_2 = \int_0^L (\phi-\psi)\sin\frac{2\pi x}{L}dx \quad \rightarrow \quad \boxed{a_m} = \frac{2}{L}\int_0^L (\phi-\psi)\sin\frac{m\pi x}{L}dx$$

$$\vdots$$

$$\frac{L}{2}a_M = \int_0^L (\phi-\psi)\sin\frac{M\pi x}{L}dx$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

Using Galerkin method with the trial function set $\{N_m = \sin(m\pi x / L); m = 1, 2, 3, \ldots\}$ to approximate a function $\phi(x)$ over the range $0 \le x \le L$ leads to typical coefficients

$$a_m = \frac{2}{L}\int_0^L (\phi - \psi)\sin\frac{m\pi x}{L}\,dx$$

The truncated Fourier sine series representation of a function can be regarded as a Galerkin weighted residual approximation.

The particular simplicity of the equations produced by the Galerkin approximation on this case was due to the orthogonality property of the trial function.

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [f_l],$$

$$f_l = \int_\Omega W_l(\phi - \psi)d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l N_m d\Omega$$

$$\mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M]$$

$\downarrow$

The Galerkin Method

$$W_l = N_l$$

$$\boxed{\begin{aligned} f_l &= \int_\Omega N_l(\phi - \psi)dx \\ K_{lm} &= \int_\Omega N_l N_m dx \end{aligned}}$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

- Given function $\phi(x)$ over the range $0 \le x \le L$
- The trial function set for the approximation of the given function is $\left\{ N_m = x^{m-1}; m = 1, 2, 3, \ldots \right\}$
- This trial function leads to typical coefficients

$$l, m = 1, 2, 3, \ldots, M$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [\,f_l\,],$$

$$f_l = \int_\Omega W_l\,(\phi - \psi)\,d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l\,N_m\,d\Omega$$

$$\mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M]$$

$$\downarrow$$

The Galerkin Method

$$W_l = N_l$$

$$f_l = \int_\Omega N_l\,(\phi - \psi)\,dx$$

$$\boxed{K_{lm} = \int_\Omega N_l\,N_m\,dx}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

The trial function set $\{N_m = x^{m-1}; m = 1, 2, 3, \ldots\}$

$$l, m = 1, 2, 3, \ldots, M$$

$$\mathbf{K} = \begin{bmatrix} \int_0^L x^0 x^0 dx & \int_0^L x^0 x^1 dx & \int_0^L x^0 x^2 dx & \cdots & \int_0^L x^0 x^{M-1} dx \\ \int_0^L x^1 x^0 dx & \int_0^L x^1 x^1 dx & \int_0^L x^1 x^2 dx & \cdots & \int_0^L x^1 x^{M-1} dx \\ \int_0^L x^2 x^0 dx & \int_0^L x^2 x^1 dx & \int_0^L x^2 x^2 dx & \cdots & \int_0^L x^2 x^{M-1} dx \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \int_0^L x^{M-1} x^0 dx & \int_0^L x^{M-1} x^1 dx & \int_0^L x^{M-1} x^2 dx & \cdots & \int_0^L x^{M-1} x^{M-1} dx \end{bmatrix}$$

we notice the computational advantages of the method in that the matrix $\mathbf{K}$ is *symmetric*

$\mathbf{f} = [\,f_l\,]$,

$f_l = \int_\Omega W_l\,(\phi - \psi)\,d\Omega$,

$\mathbf{K} = [K_{lm}]$,

$K_{lm} = \int_\Omega W_l\, N_m\, d\Omega$

$\mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M]$

↓

The Galerkin Method

$W_l = N_l$

$f_l = \int_\Omega N_l\,(\phi - \psi)\,dx$

$$\boxed{K_{lm} = \int_\Omega N_l\, N_m\, dx}$$

Weighted Residual Method
- Least square method

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

The method of least squares can be shown to belong to a weighted residual method

The sum of the squares of the residual, (or error), at each point in the domain Ω

$$I(a_1, a_2, \dots, a_M) = \int_\Omega (\phi - \hat{\phi})^2 \, d\Omega$$

The standard least-squares approach is to attempt to minimize I

$$\frac{\partial I}{\partial a_l} = 0, \qquad l = 1, 2, \dots, M$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{f} = [f_l],$$

$$f_l = \int_\Omega W_l (\phi - \psi) d\Omega,$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l N_m d\Omega$$

$$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$$

$$\frac{\partial I}{\partial a_l} = 0, \quad l = 1, 2, \ldots, M$$

$$\frac{\partial I}{\partial a_l} = \frac{\partial}{\partial a_l}\left(\int_\Omega (\phi - \hat{\phi})^2 \, d\Omega\right) = \int_\Omega \left[\frac{\partial}{\partial a_l}(\phi - \hat{\phi})^2\right] d\Omega$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{l=1}^{M} a_l N_l$$

$$= \int_\Omega \left[2(\phi - \hat{\phi})\frac{\partial}{\partial a_l}(\phi - \hat{\phi})\right] d\Omega \longleftarrow \frac{\partial}{\partial a_l}(\phi - \hat{\phi}) = \frac{\partial}{\partial a_l}\left(\phi - \psi - \sum_{l=1}^{M} a_l N_l\right)$$

$$= \left[\frac{\partial}{\partial a_l}(\phi - \psi) - \frac{\partial}{\partial a_l}\sum_{l=1}^{M} a_l N_l\right]$$

$$= -\frac{\partial}{\partial a_l}\sum_{l=1}^{M} a_l N_l$$

$$= -N_l$$

$$= -2\int_\Omega (\phi - \hat{\phi}) N_l \, d\Omega$$

$$\therefore \int_\Omega (\phi - \hat{\phi}) N_l \, d\Omega = 0$$

The standard least-squares leads to

$$\int_\Omega (\phi - \hat{\phi}) N_l \, d\Omega = 0 \quad , l = 1, 2, \ldots, M$$

The standard weighted residual statement for function approximation by trial functions

$$\int_\Omega W_l (\phi - \hat{\phi}) \, d\Omega = 0 \quad , l = 1, 2, \ldots, M$$

Using Galerkin method ($W_l = N_l$)

$$\int_\Omega N_l (\phi - \hat{\phi}) \, d\Omega = 0 \quad , l = 1, 2, \ldots, M$$

These are exactly the same form

[Zienkiewicz 1983] Ch. 2.3

Approximation to the Solutions of Differential Equations and the Use of Trial Function

Comparison

Function Approximation by Trial Function

a function

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$R_\Omega \equiv \phi - \hat{\phi}$$

Approximation to the Solutions of Differential Equations and the Use of Trial Function

D.E. $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$

B.C. $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$

a function which is the solution of the D.E with B.C.

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$R_\Omega \equiv A(\hat{\phi}) - A(\phi)$$

Approximation to the Solutions of Differential Equations and the Use of Trial Function

possibilities in which we choose trial functions such that..

	Differential Equation	Boundary Condition	
Case 1	Not Satisfied	Satisfied	ex) Rayleigh–Ritz method
Case 2	Not Satisfied	Not Satisfied	
Case 3	Satisfied	Not Satisfied	ex) Linearized hydrodynamics using Rankin source

the coefficients of the trial functions will be determined to satisfy them by "weighted residual" process

linear differential operator

$$\phi \simeq \hat{\phi} = \psi + \sum_{l=1}^{M} a_l N_l$$

General expression with an appropriate linear differential operator

Example)

A steady-state problem of heat flow in a two-dimensional domain Ω ⇨ we will now write quite generally with an appropriate linear differential operators $\mathscr{L}$ as

differential equation

$$\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right) + Q = 0 \longrightarrow \boxed{A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega}$$

, where $\mathscr{L} = \frac{\partial}{\partial x}\left(k\frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial}{\partial y}\right), \; p = Q$

$$\downarrow$$

$$\phi = \frac{\partial}{\partial x}\left(k\frac{\partial}{\partial x}\right)\phi + \frac{\partial}{\partial y}\left(k\frac{\partial}{\partial y}\right)\phi$$

$$= \frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial \phi}{\partial y}\right)$$

we will now write quite generally with an appropriate linear differential operators as $\mathcal{M}$ **boundary condition**

Dirichlet B/C

$$\phi - \bar{\phi} = 0 \quad on \quad \Gamma_\phi$$

Neumann B/C

$$-k\frac{d\phi}{dn} - \bar{q} = 0 \quad on \quad \Gamma_q$$

$$\rightarrow \boxed{B(\phi) = \mathcal{M}\phi + r = 0 \quad on \quad \Gamma}$$

for Dirichlet B/C

$$\mathcal{M} = 1, \quad r = -\bar{\phi} \quad on \quad \Gamma_\phi$$

for Neumann B/C

$$\mathcal{M} = -k\frac{d}{dn}, \quad r = -\bar{q} \quad on \quad \Gamma_q$$

General expression with an appropriate *linear differential operator*

we will now write quite generally with an appropriate linear differential operators $\mathscr{L}$, $\mathscr{M}$ to a differential equation and boundary conditions as

differential equation

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

boundary conditions

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

Case 1 : Boundary conditions are satisfied by choice of trial function while differential equations are not satisfied

$$A(\phi) = 0 \quad in \ \Omega$$

① Original Differential Equation & B/C

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$m, l = 1, 2, \ldots, M$$

② Approximation by Trial Functions

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

the function ψ and the trial functions N_m are chosen such that

$$\mathscr{M}\left[\psi + \sum_{m=1}^{M} a_m N_m\right] + r = 0 \quad on \ \Gamma$$

$$\boxed{\mathscr{M}\psi + r} + \sum_{m=1}^{M} a_m \boxed{\mathscr{M}N_m} = 0$$

,or

$$\left.\begin{aligned} \mathscr{M}\psi + r &= 0 \\ \mathscr{M}N_m &= 0 \end{aligned}\right\} on \ \Gamma$$

and then, $\hat{\phi}$ automatically satisfies the boundary condition

$$\mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

③ Weighted Residual Method

We need only ensure that $\hat{\phi}$ approximately satisfies the differential equation

We shall now attempt to develop a general method for determining the parameters a_m

Residual $R_\Omega \equiv A(\hat{\phi}) - \cancel{A(\phi)} = \mathscr{L}\hat{\phi} + p \quad in \ \Omega$

$$\int_\Omega W_l \boxed{R_\Omega} \, d\Omega = 0$$

It can be represented as matrix form.

$$\int_\Omega W_l\left[\mathscr{L}\hat{\phi}+p\right]d\Omega=0$$

$$\int_\Omega W_l\left[\mathscr{L}\left(\psi+\sum_{m=1}^{M}a_m N_m\right)+p\right]d\Omega=0$$

$$\int_\Omega W_l\left[\mathscr{L}\psi+\sum_{m=1}^{M}a_m\,\mathscr{L}N_m+p\right]d\Omega=0$$

$$\int_\Omega W_l\,\mathscr{L}\psi\,d\Omega+\int_\Omega\sum_{m=1}^{M}a_m W_l\,\mathscr{L}N_m d\Omega+\int_\Omega W_l\,p\,d\Omega=0$$

$$\sum_{m=1}^{M}a_m\left[\int_\Omega W_l\mathscr{L}N_m d\Omega\right]=-\int_\Omega W_l\,p\,d\Omega-\int_\Omega W_l\,\mathscr{L}\psi\,d\Omega$$

$$A(\phi)=\mathscr{L}\phi+p=0\quad in\ \ \Omega$$

$$B(\phi)=\mathscr{M}\phi+r=0\quad on\ \ \Gamma$$

$$\phi\ \simeq\ \hat{\phi}=\psi+\sum_{m=1}^{M}a_m N_m$$

$$R_\Omega\equiv A(\hat{\phi})-A(\phi)=\mathscr{L}\hat{\phi}+p\ \ in\ \ \Omega$$

$$\boxed{\int_\Omega W_l\left[\mathscr{L}\hat{\phi}+p\right]d\Omega=0}$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$R_\Omega \equiv A(\hat{\phi}) - A(\phi) = \mathscr{L}\hat{\phi} + p \ \ in \ \ \Omega$$

$$\boxed{\int_\Omega W_l \left[\mathscr{L}\hat{\phi} + p \right] d\Omega = 0}$$

$$\sum_{m=1}^{M} a_m \left[\int_\Omega W_l \mathscr{L} N_m d\Omega \right] = -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \, \mathscr{L}\psi d\Omega$$

$$m, \ l = 1, 2, \ldots, M$$

$$\Downarrow$$

$$a_1 \left[\int_\Omega W_l \mathscr{L} N_1 d\Omega \right] + a_2 \left[\int_\Omega W_l \mathscr{L} N_2 d\Omega \right] + \cdots + a_M \left[\int_\Omega W_l \mathscr{L} N_M d\Omega \right] = -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \, \mathscr{L}\psi d\Omega$$

$$\Downarrow$$

$$a_1 \left[\int_\Omega W_1 \mathscr{L} N_1 d\Omega \right] + a_2 \left[\int_\Omega W_1 \mathscr{L} N_2 d\Omega \right] + \cdots + a_M \left[\int_\Omega W_1 \mathscr{L} N_M d\Omega \right] = -\int_\Omega W_1 \, p \, d\Omega - \int_\Omega W_1 \, \mathscr{L}\psi d\Omega$$

$$a_1 \left[\int_\Omega W_2 \mathscr{L} N_1 d\Omega \right] + a_2 \left[\int_\Omega W_2 \mathscr{L} N_2 d\Omega \right] + \cdots + a_M \left[\int_\Omega W_2 \mathscr{L} N_M d\Omega \right] = -\int_\Omega W_2 \, p \, d\Omega - \int_\Omega W_2 \, \mathscr{L}\psi d\Omega$$

$$\vdots$$

$$a_1 \left[\int_\Omega W_M \mathscr{L} N_1 d\Omega \right] + a_2 \left[\int_\Omega W_M \mathscr{L} N_2 d\Omega \right] + \cdots + a_M \left[\int_\Omega W_M \mathscr{L} N_M d\Omega \right] = -\int_\Omega W_M \, p \, d\Omega - \int_\Omega W_M \, \mathscr{L}\psi d\Omega$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$R_\Omega \equiv A(\hat{\phi}) - A(\phi) = \mathscr{L}\hat{\phi} + p \ \ in \ \ \Omega$$

$$\boxed{\int_\Omega W_l \left[\mathscr{L}\hat{\phi} + p \right] d\Omega = 0}$$

$$a_1\left[\int_\Omega W_1 \mathscr{L} N_1 d\Omega\right] + a_2\left[\int_\Omega W_1 \mathscr{L} N_2 d\Omega\right] + \cdots + a_M\left[\int_\Omega W_1 \mathscr{L} N_M d\Omega\right] = -\int_\Omega W_1 \, p \, d\Omega - \int_\Omega W_1 \, \mathscr{L}\psi d\Omega$$

$$a_1\left[\int_\Omega W_2 \mathscr{L} N_1 d\Omega\right] + a_2\left[\int_\Omega W_2 \mathscr{L} N_2 d\Omega\right] + \cdots + a_M\left[\int_\Omega W_2 \mathscr{L} N_M d\Omega\right] = -\int_\Omega W_2 \, p \, d\Omega - \int_\Omega W_2 \, \mathscr{L}\psi d\Omega$$

$$\vdots$$

$$a_1\left[\int_\Omega W_M \mathscr{L} N_1 d\Omega\right] + a_2\left[\int_\Omega W_M \mathscr{L} N_2 d\Omega\right] + \cdots + a_M\left[\int_\Omega W_M \mathscr{L} N_M d\Omega\right] = -\int_\Omega W_M \, p \, d\Omega - \int_\Omega W_M \, \mathscr{L}\psi d\Omega$$

in matrix form

$$\begin{bmatrix} \int_\Omega W_1 \mathscr{L} N_1 d\Omega & \int_\Omega W_1 \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_1 \mathscr{L} N_M d\Omega \\ \int_\Omega W_2 \mathscr{L} N_1 d\Omega & \int_\Omega W_2 \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_2 \mathscr{L} N_M d\Omega \\ \vdots & & & \vdots \\ \int_\Omega W_M \mathscr{L} N_1 d\Omega & \int_\Omega W_M \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_M \mathscr{L} N_M d\Omega \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} -\int_\Omega W_1 \, p \, d\Omega - \int_\Omega W_1 \mathscr{L}\psi d\Omega \\ -\int_\Omega W_2 \, p \, d\Omega - \int_\Omega W_2 \mathscr{L}\psi d\Omega \\ \vdots \\ -\int_\Omega W_M \, p \, d\Omega - \int_\Omega W_M \mathscr{L}\psi d\Omega \end{bmatrix}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \;\simeq\; \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$R_\Omega \equiv A(\hat{\phi}) - A(\phi) = \mathscr{L}\hat{\phi} + p \;\; in \;\; \Omega$$

$$\boxed{\int_\Omega W_l \left[\mathscr{L}\hat{\phi} + p \right] d\Omega = 0}$$

$$\begin{bmatrix} \int_\Omega W_1 \mathscr{L} N_1 d\Omega & \int_\Omega W_1 \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_1 \mathscr{L} N_M d\Omega \\ \int_\Omega W_2 \mathscr{L} N_1 d\Omega & \int_\Omega W_2 \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_2 \mathscr{L} N_M d\Omega \\ & \vdots & & \\ \int_\Omega W_M \mathscr{L} N_1 d\Omega & \int_\Omega W_M \mathscr{L} N_2 d\Omega & \cdots & \int_\Omega W_M \mathscr{L} N_M d\Omega \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} = \begin{bmatrix} -\int_\Omega W_1 \, p \, d\Omega - \int_\Omega W_1 \mathscr{L}\psi d\Omega \\ -\int_\Omega W_2 \, p \, d\Omega - \int_\Omega W_2 \mathscr{L}\psi d\Omega \\ \vdots \\ -\int_\Omega W_M \, p \, d\Omega - \int_\Omega W_M \mathscr{L}\psi d\Omega \end{bmatrix}$$

$$\Downarrow$$

$$\boxed{\mathbf{K}\mathbf{a} = \mathbf{f}}$$

$$\mathbf{K} = [K_{lm}], \;\; K_{lm} = \int_\Omega W_l \mathscr{L} N_m d\Omega,$$

$$\mathbf{f} = [f_l], \;\; f_l = -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \mathscr{L}\psi d\Omega,$$

$$\mathbf{a}^T = [a_1 \;\; a_2 \;\; \cdots \;\; a_M]$$

Example 2.2

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

From Boundary Condition: $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$

Approximation by Trial Functions $\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

Example 2.2

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

governing equation $\dfrac{d^2\phi}{dx^2} = \phi$

$$\left.\begin{aligned} \mathscr{M}\psi + r &= 0 \\ \mathscr{M}N_m &= 0 \end{aligned}\right\} on \ \Gamma$$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

Boundary Condition $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$

$\phi = 0$ **at** $x = 0$ $\rightarrow$ $\phi + 0 = 0$ **at** $x = 0$ $\rightarrow$ $\mathscr{M} = 1 \rightarrow \mathscr{M}\phi = \phi$

$\phi = 1$ **at** $x = 1$ $\quad$ $\phi - 1 = 0$ **at** $x = 1$

$\boxed{r = 0 \quad \textbf{at} \quad x = 0}$

$r = -1$ **at** $x = 1$

substituting

$$\boxed{\left.\begin{aligned} \mathscr{M}\psi + r &= 0 \\ \mathscr{M}N_m &= 0 \end{aligned}\right\} on \ \Gamma} \longrightarrow \boxed{\left.\begin{aligned} \psi &= 0 \\ N_m &= 0 \end{aligned}\right\} at \ \ x = 0}$$

$$\left.\begin{aligned} \psi &= 1 \\ N_m &= 0 \end{aligned}\right\} at \ \ x = 1$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\left.\begin{aligned} \mathscr{M}\psi + r &= 0 \\ \mathscr{M}N_m &= 0 \end{aligned}\right\} on \ \Gamma$$

$$\left.\begin{aligned} \psi &= 0 \\ N_m &= 0 \end{aligned}\right\} at \ \ x = 0 \qquad \left.\begin{aligned} \psi &= 1 \\ N_m &= 0 \end{aligned}\right\} at \ \ x = 1$$

The function $\boxed{\psi = x}$ satisfies the required conditions on ψ, and as trial functions, vanishing at $x = 0$ and at $x = 1$,
we can take the set $\boxed{\{N_m = \sin(m\pi x); m = 1, 2, 3, \ldots\}}$

Governing equation $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$

$$\frac{d^2\phi}{dx^2} = \phi \rightarrow -\frac{d^2\phi}{dx^2} + \phi = 0 \xrightarrow{\mathscr{L} = -\frac{d^2}{dx^2} + 1} \boxed{\mathscr{L}\phi = -\frac{d^2\phi}{dx^2} + \phi}, \ \boxed{p = 0}$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$
$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\boxed{\mathscr{L} = -\frac{d^2}{dx^2} + 1, \;\; p = 0 \quad , \psi = x \;\; , N_m = \sin(m\pi x)}$$

substituting

W_l **is not defined**

$$\begin{aligned}
K_{lm} &= \int_{\Omega} W_l \, \mathscr{L} N_m d\Omega, \\
&= \int_0^1 W_l \left(-\frac{d^2 \sin(m\pi x)}{dx^2} + \sin(m\pi x) \right) dx \\
&= \int_0^1 W_l \left(m^2\pi^2 \sin(m\pi x) + \sin(m\pi x) \right) dx \\
&= \int_0^1 W_l \left(\left(1 + m^2\pi^2 \right) \sin(m\pi x) \right) dx
\end{aligned}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\boxed{\begin{aligned} \mathbf{K} &= [K_{lm}], \\ K_{lm} &= \int_{\Omega} W_l \, \mathscr{L} N_m d\Omega, \end{aligned}}$$

$$\mathbf{f} = [\, f_l \,],$$

$$f_l = -\int_{\Omega} W_l \, p \, d\Omega - \int_{\Omega} W_l \, \mathscr{L} \psi \, d\Omega,$$

$$\mathbf{a}^T = [a_1 \;\; a_2 \;\; \cdots \;\; a_M \,]$$

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\boxed{\mathscr{L} = -\frac{d^2}{dx^2} + 1, \ \ p = 0 \quad , \psi = x \ , N_m = \sin(m\pi x)}$$

substituting

$$K_{lm} = \int_0^1 W_l \left(\left(1 + m^2\pi^2\right) \sin(m\pi x) \right) dx$$

W_l **is not defined**

$$\begin{aligned} f_l &= -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \, \mathscr{L}\psi \, d\Omega \\ &= -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \left(-\frac{d^2 x}{dx^2} + x \right) d\Omega \\ &= -\int_0^1 W_l x dx \end{aligned}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l \, \mathscr{L} N_m d\Omega,$$

$$\boxed{\begin{aligned} &\mathbf{f} = [\, f_l \,], \\ &f_l = -\int_\Omega W_l \, p \, d\Omega - \int_\Omega W_l \, \mathscr{L}\psi \, d\Omega, \end{aligned}}$$

$$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

We shall take $M = 2$, so the two unknown parameters a_1 and a_2 are involved.

$$K_{lm} = \int_0^1 W_l \left(\left(1 + m^2\pi^2\right) \sin(m\pi x) \right) dx$$

$$f_l = -\int_0^1 W_l x dx$$

$$, where \quad l, m = 1, 2$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l \, \mathscr{L} N_m d\Omega,$$

$$\mathbf{f} = [\, f_l \,],$$

$$f_l = -\int_\Omega W_l \; p \, d\Omega - \int_\Omega W_l \, \mathscr{L}\psi \, d\Omega,$$

$$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M \,]$$

$$A(\phi)=\mathscr{L}\phi+p=0 \quad in \ \Omega$$

$$B(\phi)=\mathscr{M}\phi+r=0 \quad on \ \Gamma$$

$$N_m=\sin(m\pi x)$$

$$K_{lm}=\int_0^1 W_l\left(\left(1+m^2\pi^2\right)\sin(m\pi x)\right)dx, \qquad f_l=-\int_0^1 W_l x dx \quad , where \ l,m=1,2$$

For Galerkin method

$$W_l=N_l$$

$$K_{11}=\left(1+\pi^2\right)\int_0^1 \sin\pi x\cdot\sin\pi x dx$$

$\begin{pmatrix} l=1 \\ m=1 \end{pmatrix}$

$$=\frac{1}{2}\left(1+\pi^2\right)$$

$$K_{12}=\left(1+4\pi^2\right)\int_0^1 \sin\pi x\cdot\sin 2\pi x dx$$

$\begin{pmatrix} l=1 \\ m=2 \end{pmatrix}$

$$=0$$

$$K_{21}=\left(1+4\pi^2\right)\int_0^1 \sin 2\pi x\cdot\sin\pi x dx$$

$\begin{pmatrix} l=2 \\ m=1 \end{pmatrix}$

$$=0$$

$$K_{22}=\left(1+4\pi^2\right)\int_0^1 \sin 2\pi x\cdot\sin 2\pi x dx$$

$\begin{pmatrix} l=2 \\ m=2 \end{pmatrix}$

$$=\frac{1}{2}\left(1+4\pi^2\right)$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$N_m = \sin(m\pi x)$$

$$K_{lm} = \int_0^1 W_l \left(\left(1 + m^2\pi^2\right) \sin(m\pi x) \right) dx, \quad f_l = -\int_0^1 W_l x dx \quad , where \ l, m = 1, 2$$

For Galerkin method

$$W_l = N_l$$

$K_{11} = \frac{1}{2}\left(1+\pi^2\right)$	$K_{12} = 0$
$K_{21} = 0$	$K_{22} = \frac{1}{2}\left(1+4\pi^2\right)$
$f_1 = -\int_0^1 \sin \pi x \cdot x dx = -\frac{1}{\pi}$	$f_2 = -\int_0^1 \sin 2\pi x \cdot x dx = \frac{1}{2\pi}$

It should be noted that again the Galerkin process used in conjunction with suitable trigonometric functions has resulted in a diagonal form of the matrix K due to the orthogonality of the shape functions.

$$\rightarrow \begin{bmatrix} \frac{1}{2}\left(1+\pi^2\right) & 0 \\ 0 & \frac{1}{2}\left(1+4\pi^2\right) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\pi} \\ -\frac{1}{2\pi} \end{bmatrix}$$

$$\downarrow$$

$$\boxed{a_1 = -0.05857, \ a_2 = 0.007864}$$

For Galerkin method

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$N_m = \sin(m\pi x)$$

$$K_{lm} = \int_0^1 W_l \left(\left(1 + m^2\pi^2\right)\sin(m\pi x) \right) dx, \qquad f_l = -\int_0^1 W_l x dx \quad , where \ l, m = 1, 2$$

For Galerkin method

$$W_l = N_l$$

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi = x \quad , N_m = \sin(m\pi x)$$

$$a_1 = -0.05857, \ a_2 = 0.007864$$

$$\boxed{\hat{\phi} = x - 0.05857\sin(\pi x) + 0.007864\sin(2\pi x)}$$

For Galerkin method

$$K_{lm} = \int_0^1 W_l \left(\left(1 + m^2\pi^2\right) \sin(m\pi x) \right) dx, \qquad f_l = -\int_0^1 W_l x dx \quad , where\ l, m = 1, 2$$

For point collocation process, with $\mathbf{R}_\Omega$ made equal to zero at $x_1 = \dfrac{1}{3}$ at $l = 1$ and $x_2 = \dfrac{2}{3}$ at $l = 2$.

$$W_l = \delta\left(x - x_l\right)$$

$K_{11} = \int_0^1 \delta\left(x - x_1\right)\left(\left(1 + \pi^2\right)\sin \pi x\right) dx$ $\begin{pmatrix} l = 1 \\ m = 1 \end{pmatrix}$ $= \left(1 + \pi^2\right)\sin\dfrac{\pi}{3}$	$K_{12} = \int_0^1 \delta\left(x - x_1\right)\left(\left(1 + 4\pi^2\right)\sin 2\pi x\right) dx$ $\begin{pmatrix} l = 1 \\ m = 2 \end{pmatrix}$ $= \left(1 + 4\pi^2\right)\sin\dfrac{2\pi}{3}$
$K_{21} = \int_0^1 \delta\left(x - x_2\right)\left(\left(1 + \pi^2\right)\sin \pi x\right) dx$ $\begin{pmatrix} l = 2 \\ m = 1 \end{pmatrix}$ $= \left(1 + \pi^2\right)\sin\dfrac{2\pi}{3}$	$K_{22} = \int_0^1 \delta\left(x - x_2\right)\left(\left(1 + 4\pi^2\right)\sin 2\pi x\right) dx$ $\begin{pmatrix} l = 2 \\ m = 2 \end{pmatrix}$ $= \left(1 + 4\pi^2\right)\sin\dfrac{4\pi}{3}$

$K_{11} = \left(1+\pi^2\right)\sin\frac{\pi}{3}$	$K_{12} = \left(1+4\pi^2\right)\sin\frac{2\pi}{3}$
$K_{21} = \left(1+\pi^2\right)\sin\frac{2\pi}{3}$	$K_{22} = \left(1+4\pi^2\right)\sin\frac{4\pi}{3}$
$f_1 = -\int_0^1 \delta\left(x-x_1\right)xdx = -\frac{1}{3}$ $(l=1)$	$f_2 = -\int_0^1 \delta\left(x-x_2\right)xdx = -\frac{2}{3}$ $(l=2)$

$$\rightarrow \begin{bmatrix} \left(1+\pi^2\right)\sin\frac{\pi}{3} & \left(1+4\pi^2\right)\sin\frac{2\pi}{3} \\ \left(1+\pi^2\right)\sin\frac{2\pi}{3} & \left(1+4\pi^2\right)\sin\frac{4\pi}{3} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

$$\downarrow$$

$$\boxed{a_1 = -0.05312, \ a_2 = 0.004754}$$

For point collocation

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi = x \quad , N_m = \sin(m\pi x)$$

$$a_1 = -0.05312, \; a_2 = 0.004754$$

$$\boxed{\hat{\phi} = x - 0.05312\sin(\pi x) + 0.004754\sin(2\pi x)}$$

For point collocation

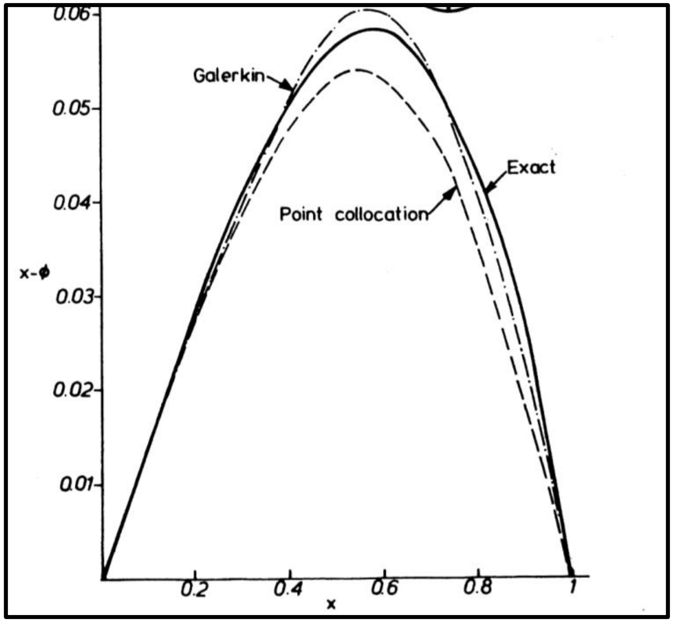
$$\boxed{\hat{\phi} = x - 0.05857\sin(\pi x) + 0.007864\sin(2\pi x)}$$

cf) For Galerkin method

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$N_m = \sin(m\pi x)$$



The approximate values and the exact values at the finite difference mesh point $x = 1/3$ and $x = 2/3$.

x	Finite Difference	Point Collocation	Galerkin method		Exact
			$\hat{\phi} = x - 0.05857\sin(\pi x) + 0.007864\sin(2\pi x)$	$\hat{\phi} = 0.068 + 0.632x + 0.226x^2$	
1/3	0.2893	0.2941	0.2894	0.3038	0.2889
2/3	0.6107	0.6165	0.6091	0.5898	0.6102

Case 2

	Differential Equation	Boundary Condition
Case 1	Not Satisfied	Satisfied
Case 2	Not Satisfied	Not Satisfied
Case 3	Satisfied	Not Satisfied

the coefficients of the trial functions will be determine
to satisfy them by "weighted residual" process

Case 2: Simultaneous approximation to the solutions of differential equations and to the boundary conditions

① Original Differential Equation & B/C

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

② Approximation by Trial Functions

If now we postulate that an expansion

$$\phi \ \simeq \ \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

cf) in previous section $\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$

③ Weighted Residual Method

$\hat{\phi}$ **does not satisfy a priori some or all of the problem boundary conditions**

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \mathscr{L}\hat{\phi} + p \ \ in \ \ \Omega$$

The boundary residual

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \mathscr{M}\hat{\phi} + r \ \ on \ \ \Gamma$$

The weighted sum of the residual

$$\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Where, in general, W_l **and** $\overline{W}_l$ **can be chosen independently**

↑

It can be represented as matrix form.

Approximation by Trial Functions

$$\hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

The residual in domain

$$\mathbf{R}_\Omega = \mathscr{L}\hat{\phi} + p \quad in \quad \Omega$$

The boundary residual

$$\mathbf{R}_\Gamma = \mathscr{M}\hat{\phi} + r \quad on \quad \Gamma$$

Substituting $\hat{\phi}$ **in** $\mathbf{R}_\Omega$ **and** $\mathbf{R}_\Gamma$

$$\mathbf{R}_\Omega = \mathscr{L}\left(\sum_{m=1}^{M} a_m N_m\right) + p = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\mathbf{R}_\Gamma = \mathscr{M}\left(\sum_{m=1}^{M} a_m N_m\right) + r = \sum_{m=1}^{M} a_m \mathscr{M} N_m + r$$

The weighted sum of the residual

$$\boxed{\int_{\Omega} W_l \mathbf{R}_{\Omega} d\Omega} + \int_{\Gamma} \bar{W}_l \mathbf{R}_{\Gamma} d\Gamma = 0$$

$$\mathbf{R}_{\Omega} = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\mathbf{R}_{\Gamma} = \sum_{m=1}^{M} a_m \mathscr{M} N_m + r$$

Substituting $\mathbf{R}_{\Omega}$ in the first term of the weighted sum of the residual

$$\int_{\Omega} W_l \left(\sum_{m=1}^{M} a_m \mathscr{L} N_m + p \right) d\Omega$$

$$= \sum_{m=1}^{M} \left(a_m \int_{\Omega} W_l \mathscr{L} N_m d\Omega \right) + \int_{\Omega} W_l p d\Omega$$

$$m = 1, 2, 3, \ldots, M$$

$$= \boxed{a_1 \int_{\Omega} W_l \mathscr{L} N_1 d\Omega + a_2 \int_{\Omega} W_l \mathscr{L} N_2 d\Omega + \ldots + a_M \int_{\Omega} W_l \mathscr{L} N_M d\Omega + \int_{\Omega} W_l p d\Omega}$$

The weighted sum of the residual

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega} + \boxed{\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma} = 0$$

$$\mathbf{R}_\Omega = \mathscr{L}\left(\sum_{m=1}^{M} a_m N_m\right) + p = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\mathbf{R}_\Gamma = \mathscr{M}\left(\sum_{m=1}^{M} a_m N_m\right) + r = \sum_{m=1}^{M} a_m \mathscr{M} N_m + r$$

Substituting $\mathbf{R}_\Omega$ in the first term of the weighted sum of the residual

$$\boxed{a_1\int_\Omega W_l \mathscr{L} N_1 d\Omega + a_2\int_\Omega W_l \mathscr{L} N_2 d\Omega + \ldots + a_M\int_\Omega W_l \mathscr{L} N_M d\Omega + \int_\Omega W_l p d\Omega}$$

Substituting $\mathbf{R}_\Gamma$ in the second term of the weighted sum of the residual

$$\int_\Gamma \bar{W}_l \left(\sum_{m=1}^{M} a_m \mathscr{M} N_m + r\right) d\Omega$$

$$= \sum_{m=1}^{M}\left(a_m \int_\Gamma \bar{W}_l \mathscr{M} N_m d\Omega\right) + \int_\Gamma \bar{W}_l r d\Omega \qquad m = 1, 2, 3, \ldots, M$$

$$= \boxed{a_1\int_\Gamma \bar{W}_l \mathscr{M} N_1 d\Omega + a_2\int_\Gamma \bar{W}_l \mathscr{M} N_2 d\Omega + \ldots + a_M\int_\Gamma \bar{W}_l \mathscr{M} N_M d\Omega + \int_\Gamma \bar{W}_l r d\Omega}$$

$$\hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Omega = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\mathbf{R}_\Gamma = \sum_{m=1}^{M} a_m \mathcal{M} N_m + r$$

$$\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

The weighted sum of the residual

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega} + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

$$m = 1, 2, 3, \dots, M$$

$$\downarrow$$

$$a_1 \int_\Omega W_l \mathscr{L} N_1 d\Omega + a_2 \int_\Omega W_l \mathscr{L} N_2 d\Omega + \dots + a_M \int_\Omega W_l \mathscr{L} N_M d\Omega + \boxed{\int_\Omega W_l p d\Omega}$$

$$+ a_1 \int_\Gamma \bar{W}_l \mathcal{M} N_1 d\Gamma + a_2 \int_\Gamma \bar{W}_l \mathcal{M} N_2 d\Gamma + \dots + a_M \int_\Gamma \bar{W}_l \mathcal{M} N_M d\Gamma + \boxed{\int_\Gamma \bar{W}_l r d\Gamma} = 0$$

$$\downarrow$$

$$a_1 \left(\int_\Omega W_l \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_1 d\Gamma \right) + a_2 \left(\int_\Omega W_l \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_2 d\Gamma \right)$$

$$+ \dots + a_M \left(\int_\Omega W_l \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_M d\Gamma \right) + \boxed{\int_\Omega W_l p d\Omega} + \boxed{\int_\Gamma \bar{W}_l r d\Gamma} = 0$$

$$\downarrow$$

$$a_1 \left(\int_\Omega W_l \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_1 d\Gamma \right) + a_2 \left(\int_\Omega W_l \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_2 d\Gamma \right)$$

$$+ \dots + a_M \left(\int_\Omega W_l \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_l \mathcal{M} N_M d\Gamma \right) = -\int_\Omega W_l p d\Omega - \int_\Gamma \bar{W}_l r d\Gamma$$

$$\hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Omega = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\mathbf{R}_\Gamma = \sum_{m=1}^{M} a_m \mathscr{M} N_m + r$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

The weighted sum of the residual

$$a_1\left(\int_\Omega W_l \mathscr{L} N_1 d\Omega + \int_\Gamma \overline{W}_l \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_l \mathscr{L} N_2 d\Omega + \int_\Gamma \overline{W}_l \mathscr{M} N_2 d\Gamma\right)$$
$$+ \ldots + a_M\left(\int_\Omega W_l \mathscr{L} N_M d\Omega + \int_\Gamma \overline{W}_l \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_l p d\Omega - \int_\Gamma \overline{W}_l r d\Gamma$$

$$\downarrow l = 1, 2, 3, \ldots, M$$

$$a_1\left(\int_\Omega W_1 \mathscr{L} N_1 d\Omega + \int_\Gamma \overline{W}_1 \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_1 \mathscr{L} N_2 d\Omega + \int_\Gamma \overline{W}_1 \mathscr{M} N_2 d\Gamma\right)$$
$$+ \ldots + a_M\left(\int_\Omega W_1 \mathscr{L} N_M d\Omega + \int_\Gamma \overline{W}_1 \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_1 p d\Omega - \int_\Gamma \overline{W}_1 r d\Gamma$$

$$a_1\left(\int_\Omega W_2 \mathscr{L} N_1 d\Omega + \int_\Gamma \overline{W}_2 \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_2 \mathscr{L} N_2 d\Omega + \int_\Gamma \overline{W}_2 \mathscr{M} N_2 d\Gamma\right)$$
$$+ \ldots + a_M\left(\int_\Omega W_2 \mathscr{L} N_M d\Omega + \int_\Gamma \overline{W}_2 \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_2 p d\Omega - \int_\Gamma \overline{W}_2 r d\Gamma$$

$$\vdots$$

$$a_1\left(\int_\Omega W_M \mathscr{L} N_1 d\Omega + \int_\Gamma \overline{W}_M \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_M \mathscr{L} N_2 d\Omega + \int_\Gamma \overline{W}_M \mathscr{M} N_2 d\Gamma\right)$$
$$+ \ldots + a_M\left(\int_\Omega W_M \mathscr{L} N_M d\Omega + \int_\Gamma \overline{W}_M \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_M p d\Omega - \int_\Gamma \overline{W}_M r d\Gamma$$

$$\hat{\phi} = \sum_{m=1}^{M} a_m N_m \qquad \mathbf{R}_\Omega = \sum_{m=1}^{M} a_m \mathscr{L} N_m + p$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0} \qquad \mathbf{R}_\Gamma = \sum_{m=1}^{M} a_m \mathscr{M} N_m + r$$

The weighted sum of the residual

$$a_1\left(\int_\Omega W_1 \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_1 \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_2 d\Gamma\right) + \ldots + a_M\left(\int_\Omega W_1 \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_1 p d\Omega - \int_\Gamma \bar{W}_1 r d\Gamma$$

$$a_1\left(\int_\Omega W_2 \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_2 \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_2 d\Gamma\right) + \ldots + a_M\left(\int_\Omega W_2 \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_2 p d\Omega - \int_\Gamma \bar{W}_2 r d\Gamma$$

$$\vdots$$

$$a_1\left(\int_\Omega W_M \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_1 d\Gamma\right) + a_2\left(\int_\Omega W_M \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_2 d\Gamma\right) + \ldots + a_M\left(\int_\Omega W_M \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_M d\Gamma\right) = -\int_\Omega W_M p d\Omega - \int_\Gamma \bar{W}_M r d\Gamma$$

↓ **Matrix representation**

$$\underbrace{\begin{bmatrix} \int_\Omega W_1 \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_1 d\Gamma & \int_\Omega W_1 \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_2 d\Gamma & \cdots & \int_\Omega W_1 \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_1 \mathscr{M} N_M d\Gamma \\ \int_\Omega W_2 \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_1 d\Gamma & \int_\Omega W_2 \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_2 d\Gamma & \cdots & \int_\Omega W_2 \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_2 \mathscr{M} N_M d\Gamma \\ & & \vdots & \\ \int_\Omega W_M \mathscr{L} N_1 d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_1 d\Gamma & \int_\Omega W_M \mathscr{L} N_2 d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_2 d\Gamma & \cdots & \int_\Omega W_M \mathscr{L} N_M d\Omega + \int_\Gamma \bar{W}_M \mathscr{M} N_M d\Gamma \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} -\int_\Omega W_1 p d\Omega - \int_\Gamma \bar{W}_1 r d\Gamma \\ -\int_\Omega W_2 p d\Omega - \int_\Gamma \bar{W}_2 r d\Gamma \\ \vdots \\ -\int_\Omega W_M p d\Omega - \int_\Gamma \bar{W}_M r d\Gamma \end{bmatrix}}_{\mathbf{f}}$$

$$\mathbf{K} = [K_{lm}],\ K_{lm} = \int_\Omega W_l \mathscr{L} N_m d\Omega + \int_\Gamma \bar{W}_l \mathscr{M} N_m d\Gamma,\ \mathbf{f} = [f_l],\quad f_l = -\int_\Omega W_l p d\Omega - \int_\Gamma \bar{W}_l r d\Gamma,\ \mathbf{a}^T = [a_1\ a_2\ \cdots\ a_M]$$

Summary of case 2

① Original Differential Equation & B/C

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

② Approximation by Trial Functions

If now we postulate that an expansion

$$\phi \ \simeq \ \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

cf) in previous section $\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$

③ Weighted Residual Method

$\hat{\phi}$ **does not satisfy a priori some or all of the problem boundary conditions**

$$\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

↓ **matrix representation**

$$\mathbf{Ka} = \mathbf{f}$$

, where

$$\mathbf{K} = [K_{lm}],$$

$$K_{lm} = \int_\Omega W_l \mathscr{L} N_m d\Omega + \int_\Gamma \bar{W}_l \mathscr{M} N_m d\Gamma,$$

$$\mathbf{f} = [f_l], \ f_l = -\int_\Omega W_l p d\Omega - \int_\Gamma \bar{W}_l r d\Gamma,$$

$$\mathbf{a}^T = [a_1 \ a_2 \ \cdots \ a_M]$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

Example 2.4

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ **in** $0 < x < 1$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

Governing equation $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$

$$\frac{d^2\phi}{dx^2} = \phi \rightarrow \frac{d^2\phi}{dx^2} - \phi = 0 \rightarrow \boxed{A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \ \Omega}$$

Boundary Conditions $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$

$\phi = 0$ **at** $x = 0$ → $\phi - 0 = 0$ **at** $x = 0$ → $\boxed{B(\phi) = \phi = 0 \ at \ x = 0}$

$\phi = 1$ **at** $x = 1$ → $\phi - 1 = 0$ **at** $x = 1$ → $\boxed{B(\phi) = \phi - 1 = 0 \ at \ x = 1}$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$$

$$\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Example 2.4

$$A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \ \ 0 < x < 1$$

$$B(\phi) = \phi = 0 \ \ at \ \ x = 0$$

$$B(\phi) = \phi - 1 = 0 \ \ at \ \ x = 1$$

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = A(\hat{\phi})$$

$$= \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \ \ in \ \ 0 < x < 1$$

The boundary residual

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \hat{\phi} \ \ at \ \ x = 0$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \hat{\phi} - 1 \ \ at \ \ x = 1$$

cf) Using the trial function set which satisfy the boundary conditions:

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = A(\hat{\phi})$$

$$= \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \ \ in \ \ 0 < x < 1$$

The boundary residual

$$\mathbf{R}_\Gamma = \cancel{B(\hat{\phi})} - \cancel{B(\phi)} = 0$$

,since $B(\hat{\phi}) = 0$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = \hat{\phi} \;\; at \;\; x = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \;\; at \;\; x = 1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \;\; , \bar{W}_l = -N_l\big|_\Gamma$$

The residual in domain

$$\boxed{\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi}} \;\; in \;\; 0 < x < 1$$

The boundary residual

$$\boxed{\begin{aligned} \mathbf{R}_\Gamma &= \hat{\phi} \;\; at \;\; x = 0 \\ \mathbf{R}_\Gamma &= \hat{\phi} - 1 \;\; at \;\; x = 1 \end{aligned}}$$

$$\int_\Omega W_l \boxed{\mathbf{R}_\Omega} d\Omega + \boxed{\int_\Gamma} \bar{W}_l \boxed{\mathbf{R}_\Gamma} d\Gamma = 0$$

In this case the boundary curve Γ consists of the two points $x = 0$ and $x = 1$, so that the integration over the boundary reduces to two discrete residuals

$$\int_0^1 W_l \left(\boxed{\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi}} \right) dx + \left[\bar{W}_l \boxed{\hat{\phi}} \right]_{\boxed{x=0}} + \left[\bar{W}_l \boxed{(\hat{\phi} - 1)} \right]_{\boxed{x=1}} = 0$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = \hat{\phi} \quad at \;\; x = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \quad at \;\; x = 1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \quad in \;\; 0 < x < 1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \;\; , \bar{W}_l = -N_l\big|_\Gamma$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\bar{W}_l \hat{\phi} \right]_{x=0} + \left[\bar{W}_l (\hat{\phi} - 1) \right]_{x=1} = 0$$

The weighting functions will be defined by $W_l = \boxed{N_l}$ and $\bar{W}_l = \boxed{-N_l\big|_\Gamma}$

In general, W_l and $\bar{W}_l$ can be chosen independently.

$$\int_0^1 \boxed{N_l} \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\boxed{-N_l} \hat{\phi} \right]_{x=0} + \left[\boxed{-N_l} (\hat{\phi} - 1) \right]_{x=1} = 0$$

$$\int_0^1 \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) N_l dx - \left[N_l \hat{\phi} \right]_{x=0} - \left[N_l (\hat{\phi} - 1) \right]_{x=1} = 0$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = \hat{\phi} \;\; at \;\; x = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \;\; at \;\; x = 1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \;\; , \bar{W}_l = -N_l\big|_\Gamma$$

$$\int_0^1 \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) N_l dx - \left[N_l \, \hat{\phi} \right]_{x=0} - \left[N_l \, (\hat{\phi} - 1) \right]_{x=1} = 0$$

A possible trial function set is taken now simply as $\{N_m = x^{m-1}; m = 1, 2, 3 \ldots\}$

And **using a three-term expansion** $\hat{\phi} = a_1 + a_2 x + a_3 x^2 \longrightarrow l, m = 1, 2, 3$

$$\int_0^1 \left(\frac{d^2(a_1 + a_2 x + a_3 x^2)}{dx^2} - a_1 - a_2 x - a_3 x^2 \right) x^{l-1} dx$$

$$- \left[x^{l-1}(a_1 + a_2 x + a_3 x^2) \right]_{x=0} - \left[x^{l-1}(a_1 + a_2 x + a_3 x^2 - 1) \right]_{x=1} = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} \ \ at \ \ x=0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \ \ at \ \ x=1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \ \ in \ \ 0 < x < 1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \ \ , \bar{W}_l = -N_l\big|_\Gamma$$

$$\int_0^1 \left(\frac{d^2(a_1 + a_2 x + a_3 x^2)}{dx^2} - a_1 - a_2 x - a_3 x^2 \right) x^{l-1} dx$$
$$- \left[x^{l-1}(a_1 + a_2 x + a_3 x^2) \right]_{x=0} - \left[x^{l-1}(a_1 + a_2 x + a_3 x^2 - 1) \right]_{x=1} = 0$$

$$\downarrow$$

$$\int_0^1 \left(2a_3 - a_1 - a_2 x - a_3 x^2 \right) x^{l-1} dx$$
$$- \left[x^{l-1}(a_1 + a_2 x + a_3 x^2) \right]_{x=0} - \left[x^{l-1}(a_1 + a_2 x + a_3 x^2 - 1) \right]_{x=1} = 0$$

$$\downarrow \ \ x^{l-1}\big|_{x=1} = 1^{l-1} = 1$$

$$\int_0^1 \left(2a_3 - a_1 - a_2 x - a_3 x^2 \right) x^{l-1} dx - \left[x^{l-1}\big|_{x=0} \cdot \left(a_1 + a_2 \cdot 0 + a_3 \cdot 0 \right) \right] - \left[1 \cdot (a_1 + a_2 \cdot 1 + a_3 \cdot 1 - 1) \right] = 0$$

$$\downarrow$$

$$\int_0^1 \left(2a_3 - a_1 - a_2 x - a_3 x^2 \right) x^{l-1} dx - x^{l-1}\big|_{x=0} \cdot a_1 - a_1 - a_2 - a_3 + 1 = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} \ \ at \ \ x=0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \ \ at \ \ x=1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \ \ in \ \ 0<x<1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \ \ , \bar{W}_l = -N_l\big|_\Gamma$$

$$\int_0^1 \left(2a_3 - a_1 - a_2 x - a_3 x^2\right) x^{l-1} dx - x^{l-1}\big|_{x=0} \cdot a_1 - a_1 - a_2 - a_3 + 1 = 0$$

$$\downarrow$$

$$2\int_0^1 x^{l-1} dx\, a_3 - \int_0^1 x^{l-1} dx\, a_1 - \int_0^1 x^{l-1} x dx\, a_2 - \int_0^1 x^{l-1} x^2 dx a_3 - x^{l-1}\big|_{x=0} \cdot a_1 - a_1 - a_2 - a_3 + 1 = 0$$

$$\downarrow$$

$$2\int_0^1 x^{l-1} dx\, a_3 - \int_0^1 x^{l-1} dx\, a_1 - \int_0^1 x^l dx\, a_2 - \int_0^1 x^{l+1} dx\, a_3 - \left(x^{l-1}\big|_{x=0} + 1\right) \cdot a_1 - a_2 - a_3 + 1 = 0$$

$$\downarrow$$

$$-\left(\int_0^1 x^{l-1} dx + x^{l-1}\big|_{x=0} + 1\right) a_1 - \left(\int_0^1 x^l dx + 1\right) a_2 + \left(2\int_0^1 x^{l-1} dx - \int_0^1 x^{l+1} dx - 1\right) a_3 = -1$$

$$\mathbf{R}_\Gamma = \hat{\phi} \ \ at \ \ x=0$$

$$\mathbf{R}_\Gamma = \hat{\phi}-1 \ \ at \ \ x=1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \ \ in \ \ 0<x<1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \ \ , \bar{W}_l = -N_l\Big|_\Gamma$$

$$-\left(\int_0^1 x^{l-1}dx + x^{l-1}\Big|_{x=0} + 1\right)a_1 - \left(\int_0^1 x^l dx + 1\right)a_2 + \left(2\int_0^1 x^{l-1}dx - \int_0^1 x^{l+1}dx - 1\right)a_3 = -1$$

$$l = 1, 2, 3$$

$$x^{l-1}\Big|_{l=1,x=0} = x^0\Big|_{x=0} = 1\Big|_{x=0} = 1$$

$$, x^{l-1}\Big|_{l=2,x=0} = x^1\Big|_{x=0} = 0 \quad , x^{l-1}\Big|_{l=3,x=0} = x^2\Big|_{x=0} = 0$$

$$-\left(\int_0^1 1dx + 2\right)a_1 - \left(\int_0^1 xdx + 1\right)a_2 + \left(2\int_0^1 1dx - \int_0^1 x^2dx - 1\right)a_3 = -1$$

$$-\left(\int_0^1 x^1dx + 1\right)a_1 - \left(\int_0^1 x^2dx + 1\right)a_2 + \left(2\int_0^1 x^1dx - \int_0^1 x^3dx - 1\right)a_3 = -1$$

$$-\left(\int_0^1 x^2dx + 1\right)a_1 - \left(\int_0^1 x^3dx + 1\right)a_2 + \left(2\int_0^1 x^2dx - \int_0^1 x^4dx - 1\right)a_3 = -1$$

$$\mathbf{R}_\Gamma = \hat{\phi} \;\; at \;\; x=0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \;\; at \;\; x=1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0<x<1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \;\; , \bar{W}_l = -N_l\big|_\Gamma$$

$$-\left(\int_0^1 1dx + 2\right)a_1 - \left(\int_0^1 xdx + 1\right)a_2 + \left(2\int_0^1 1dx - \int_0^1 x^2 dx - 1\right)a_3 = -1$$

$$-\left(\int_0^1 x^1 dx + 1\right)a_1 - \left(\int_0^1 x^2 dx + 1\right)a_2 + \left(2\int_0^1 x^1 dx - \int_0^1 x^3 dx - 1\right)a_3 = -1$$

$$-\left(\int_0^1 x^2 dx + 1\right)a_1 - \left(\int_0^1 x^3 dx + 1\right)a_2 + \left(2\int_0^1 x^2 dx - \int_0^1 x^4 dx - 1\right)a_3 = -1$$

↓ **Matrix representation**

$$\begin{bmatrix} -\left(\int_0^1 1dx + 2\right) & -\left(\int_0^1 xdx + 1\right) & 2\int_0^1 1dx - \int_0^1 x^2 dx - 1 \\ -\left(\int_0^1 x^1 dx + 1\right) & -\left(\int_0^1 x^2 dx + 1\right) & 2\int_0^1 x^1 dx - \int_0^1 x^3 dx - 1 \\ -\left(\int_0^1 x^2 dx + 1\right) & -\left(\int_0^1 x^3 dx + 1\right) & 2\int_0^1 x^2 dx - \int_0^1 x^4 dx - 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\mathbf{R}_\Gamma = \hat{\phi} \;\; at \;\; x = 0$$

$$\mathbf{R}_\Gamma = \hat{\phi} - 1 \;\; at \;\; x = 1$$

$$\mathbf{R}_\Omega = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0}$$

$$W_l = N_l \;\; , \bar{W}_l = -N_l\big|_\Gamma$$

$$\begin{bmatrix} -\left(\int_0^1 1dx + 2\right) & -\left(\int_0^1 xdx + 1\right) & 2\int_0^1 1dx - \int_0^1 x^2 dx - 1 \\ -\left(\int_0^1 x^1 dx + 1\right) & -\left(\int_0^1 x^2 dx + 1\right) & 2\int_0^1 x^1 dx - \int_0^1 x^3 dx - 1 \\ -\left(\int_0^1 x^2 dx + 1\right) & -\left(\int_0^1 x^3 dx + 1\right) & 2\int_0^1 x^2 dx - \int_0^1 x^4 dx - 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\downarrow$$

$$\mathbf{Ka} = \mathbf{f} \;\; \text{, where } \mathbf{K} = \begin{bmatrix} -3 & -\frac{3}{2} & \frac{2}{3} \\ -\frac{3}{2} & -\frac{4}{3} & -\frac{1}{4} \\ -\frac{4}{3} & -\frac{5}{4} & -\frac{8}{15} \end{bmatrix}, \mathbf{f} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \quad \rightarrow \quad \mathbf{K} = \begin{bmatrix} 3 & \frac{3}{2} & -\frac{2}{3} \\ \frac{3}{2} & \frac{4}{3} & \frac{1}{4} \\ \frac{4}{3} & \frac{5}{4} & \frac{8}{15} \end{bmatrix}, \; \mathbf{f} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Behavior of the one-, two-, and three-term approximations

Example 2.4

$$\mathbf{Ka}=\mathbf{f} \quad \text{, where} \quad \mathbf{K}=\begin{bmatrix} 3 & \frac{3}{2} & -\frac{2}{3} \\ \frac{3}{2} & \frac{4}{3} & \frac{1}{4} \\ \frac{4}{3} & \frac{5}{4} & \frac{8}{15} \end{bmatrix} \quad \mathbf{f}=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \longrightarrow \mathbf{a}=\begin{bmatrix} 0.068 \\ 0.632 \\ 0.226 \end{bmatrix}$$

$$\hat{\phi}=a_1+a_2x+a_3x^2=0.068+0.632x+0.226x^2$$

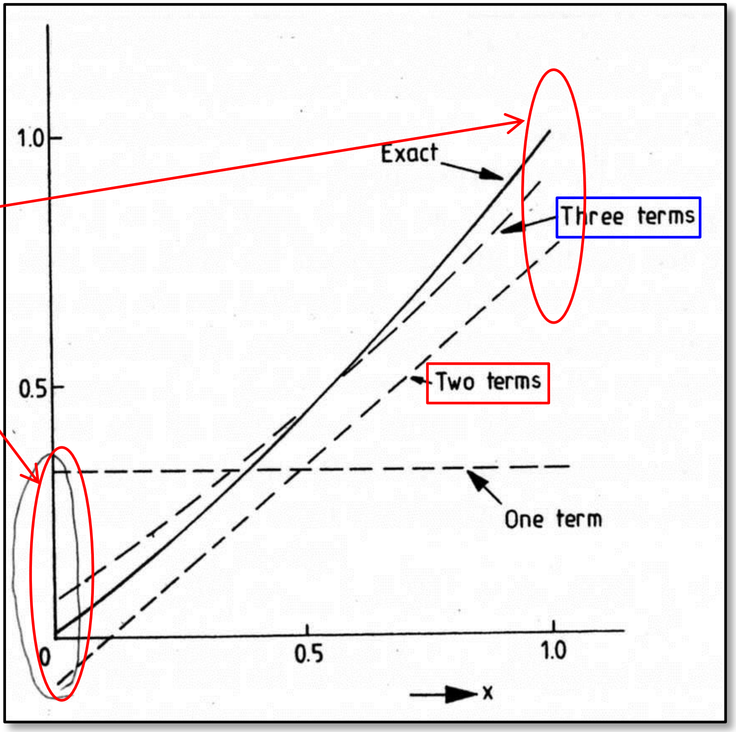
$$\downarrow$$

x	One term	Two terms	Three terms	Exact
0	1/3	-0.095	0.068	0
1	1/3	0.762	0.925	1

The convergence of the approximation to the prescribed conditions at $x=0$ and at $x=1$ is shown in the above table, which compares the behavior of the one-, two-, and three-term approximations at these two points.

Example 2.4

The approximation functions do not satisfy the boundary condition.



x	One term	Two terms	Three terms	Exact
0	1/3	-0.095	0.068	0
1	1/3	0.762	0.925	1

Approximation to the Solutions of Differential Equations and the Use of Trial Function

Case 3

	Differential Equation	Boundary Condition
Case 1	Not Satisfied	Satisfied
Case 2	Not Satisfied	Not Satisfied
Case 3	Satisfied	Not Satisfied

the coefficients of the trial functions will be determine
to satisfy them by "weighted residual" process

Case 3 : Differential equations are satisfied by choice of trial function while boundary conditions are not satisfied

① Original Differential Equation & B/C

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

② Approximation by Trial Functions

We choose trial function such that the approximation $\hat{\phi}$ automatically satisfies the differential equation, but does not satisfy the B/Cs

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

③ Weighted Residual Method

Since $\hat{\phi}$ satisfy differential equations

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \boxed{\mathscr{L}\hat{\phi} + p = 0} \quad in \ \Omega$$

The boundary residual

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

The weighted sum of the boundary residual

$$\int_\Gamma \overline{W}_l \mathbf{R}_\Gamma \, d\Gamma = 0$$

① Original Differential Equation & B/C

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

② Approximation by Trial Functions

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

③ Weighted Residual Method

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

$$\int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Given from the boundary condition.

$$\int_\Gamma \left(\boxed{\overline{W}_l} \boxed{\mathscr{M}} \left(\sum_{m=1}^{M} \boxed{a_m} \boxed{N_m} \right) + \boxed{r} \right) d\Gamma = 0$$

Weighting functions are chosen

Find

Trial function set is chosen

However, this set of trial function is more difficult to choose.

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Choice of the trial function set which satisfy the differential equation

Considering the example of the Laplace differential equation in which the choice of the trial function set is particularly easy.

$$\nabla^2 f = 0$$

What is trial function which satisfy the Laplace equation?

If $f(z)$ is an analytic function of the complex variable $z = x + iy$, then $f(z)$ automatically satisfy the Laplace equation

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Choice of the trial function set which satisfy the differential equation

The Laplace equation: $\nabla^2 f = 0$

Analytic function of the complex variable $f(z)$

,where $z = x + iy$

$$\frac{\partial^2 f(z)}{\partial x^2} = \frac{\partial\left(\frac{\partial f(z)}{\partial x}\right)}{\partial x} = \frac{\partial\left(\frac{df(z)}{dz}\frac{\partial z}{\partial x}\right)}{\partial x} = \frac{\partial\left(\frac{df(z)}{dz}\right)}{\partial x} = \frac{d\left(\frac{df(z)}{dz}\right)}{dz}\frac{\partial z}{\partial x} = \frac{d\left(\frac{df(z)}{dz}\right)}{dz} = \frac{d^2 f(z)}{dz^2}$$

$$\frac{\partial^2 f(z)}{\partial y^2} = \frac{\partial\left(\frac{\partial f(z)}{\partial y}\right)}{\partial y} = \frac{\partial\left(\frac{df(z)}{dz}\frac{\partial z}{\partial y}\right)}{\partial y} = \frac{\partial\left(\frac{df(z)}{dz}i\right)}{\partial y} = \frac{d\left(\frac{df(z)}{dz}i\right)}{dz}\frac{\partial z}{\partial y} = \frac{d\left(\frac{df(z)}{dz}i\right)}{dz}i = -\frac{d^2 f(z)}{dz^2}$$

$$\nabla^2 f(z) = \frac{\partial^2 f(z)}{\partial x^2} + \frac{\partial^2 f(z)}{\partial y^2} = \frac{d^2 f(z)}{dz^2} - \frac{d^2 f(z)}{dz^2} = 0 \;\longrightarrow$$

$f(z)$ **satisfies the Laplace equation**

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

Choice of the trial function set which satisfy the differential equation

The Laplace equation: $\nabla^2 f = 0$

$f(z)$ **satisfies the Laplace equation, where** $z = x + iy$

We can immediately use an analytic function such as

$f(z) = z^n = u + iv$ **,where** u **and** v **are real**

This leads to the follow set:

$n = 1,$	$u = x,$	$v = y$
$n = 2,$	$u = x^2 - y^2,$	$v = 2xy$
$n = 3,$	$u = x^3 - 3xy^2,$	$v = 3x^2 y - y^3$
$n = 4,$	$u = x^4 - 6x^2 y^2 + y^4,$	$v = 4x^3 y - 4xy^3$

Example 2.8

It is required to obtain the function $\phi(x)$

which satisfies the differential equation $\dfrac{d^2\phi}{dx^2}+\dfrac{d^2\phi}{dy^2}=-2$ in $-3\le x\le 3 \quad -2\le y\le 2$

$\phi=0$ on the boundary

To enable us to use the trial set of functions which satisfy the Laplace equation, we introduce a new variable θ

$$\phi=\theta-\frac{1}{2}\left(x^2+y^2\right)$$

Example 2.8

the differential equation $\frac{d^2\phi}{dx^2}+\frac{d^2\phi}{dy^2}=-2$ **in** $-3\le x\le 3 \quad -2\le y\le 2$

$\phi=0$ **on the boundary**

Substituting $\phi=\theta-\frac{1}{2}\left(x^2+y^2\right)$ **into the differential equation**

$$\frac{d^2\left(\theta-\frac{1}{2}\left(x^2+y^2\right)\right)}{dx^2}+\frac{d^2\left(\theta-\frac{1}{2}\left(x^2+y^2\right)\right)}{dy^2}=-2$$

$$\frac{d^2\theta}{dx^2}-\frac{d^2\left(\frac{1}{2}\left(x^2+y^2\right)\right)}{dx^2}+\frac{d^2\theta}{dy^2}-\frac{d^2\left(\frac{1}{2}\left(x^2+y^2\right)\right)}{dy^2}=-2$$

$$\frac{d^2\theta}{dx^2}-1+\frac{d^2\theta}{dy^2}-1=-2$$

$$\frac{d^2\theta}{dx^2}+\frac{d^2\theta}{dy^2}=0$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \;\; \Gamma$$

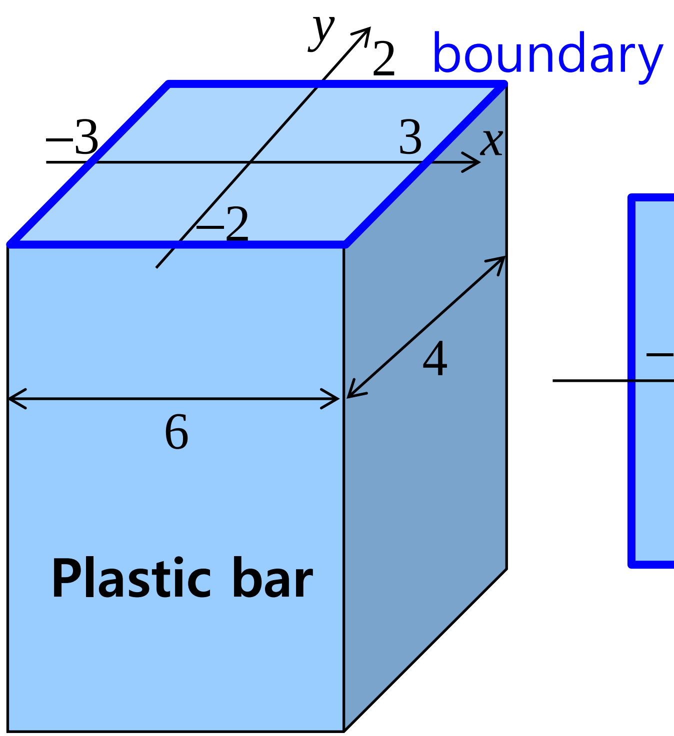
$$\int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

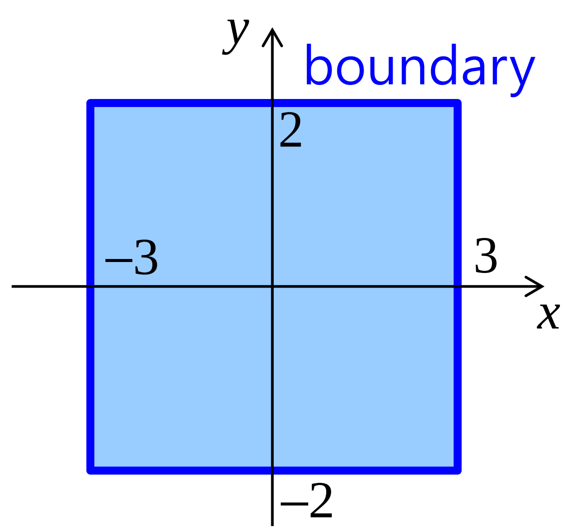
the differential equation $\dfrac{d^2\phi}{dx^2} + \dfrac{d^2\phi}{dy^2} = -2$ **in** $-3 \le x \le 3,\ -2 \le y \le 2$

$\phi = 0$ **on the boundary**

$$\phi = \theta - \frac{1}{2}\left(x^2 + y^2\right) \downarrow$$

the differential equation $\dfrac{d^2\theta}{dx^2} + \dfrac{d^2\theta}{dy^2} = 0$ **in** $-3 \le x \le 3,\ -2 \le y \le 2$





The required solution will be symmetric in x and y, and so we can use as trial function the set

$$f = u + iv$$

$n = 1,\ u = x,\quad v = y$

$\boxed{n = 2,\ u = x^2 - y^2,}\quad v = 2xy$

$n = 3,\ u = x^3 - 3xy^2,\quad v = 3x^2y - y^3$

$\boxed{n = 4,\ u = x^4 - 6x^2y^2 + y^4,}\quad v = 4x^3y - 4xy^3$

symmetric trial function in x and y

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \ \Gamma$$

$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

$$f = u + iv$$

$n = 1,\ u = x,\quad v = y$

$\boxed{n = 2,\ u = x^2 - y^2,}\quad v = 2xy$

$n = 3,\ u = x^3 - 3xy^2,\quad v = 3x^2y - y^3$

$\boxed{n = 4,\ u = x^4 - 6x^2y^2 + y^4,}\quad v = 4x^3y - 4xy^3$

$$N_1 = 1,\ N_2 = x^2 - y^2,\ N_3 = x^4 - 6x^2y^2 + y^4$$

A three-term approximation would be

$$\hat{\theta} = a_1 + a_2\left(x^2 - y^2\right) + a_3\left(x^4 - 6x^2y^2 + y^4\right)$$

which satisfy the differential equation in terms of θ exactly

$$A(\phi)=\mathscr{L}\phi+p=0 \quad in \ \Omega$$
$$B(\phi)=\mathscr{M}\phi+r=0 \quad on \ \Gamma$$
$$\phi \simeq \hat{\phi}=\sum_{m=1}^{M} a_m N_m$$
$$\mathbf{R}_\Gamma = B(\hat{\phi})=\mathscr{M}\hat{\phi}+r \quad on \ \Gamma$$
$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma=0$$

the differential equation $\frac{d^2\phi}{dx^2}+\frac{d^2\phi}{dy^2}=-2$ **in** $-3\le x\le 3,\ -2\le y\le 2$

$$\phi=\theta-\frac{1}{2}\left(x^2+y^2\right)\downarrow$$

the differential equation $\frac{d^2\theta}{dx^2}+\frac{d^2\theta}{dy^2}=0$ **in** $-3\le x\le 3,\ -2\le y\le 2$

$$\hat{\theta}=a_1+a_2\left(x^2-y^2\right)+a_3\left(x^4-6x^2y^2+y^4\right)$$

$\phi=0$ **on the boundary**

$\hat{\phi}=\hat{\theta}-\frac{1}{2}\left(x^2+y^2\right)$ **should satisfy** $\hat{\phi}=0$ **on the boundary**

The weighted residual statement is thus

$$\int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma=0$$

$$\int_\Gamma \left(\hat{\phi}\right)\bar{W}_l d\Gamma=\int_\Gamma \left(\hat{\theta}-\frac{1}{2}\left(x^2+y^2\right)\right)\bar{W}_l d\Gamma=0$$

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

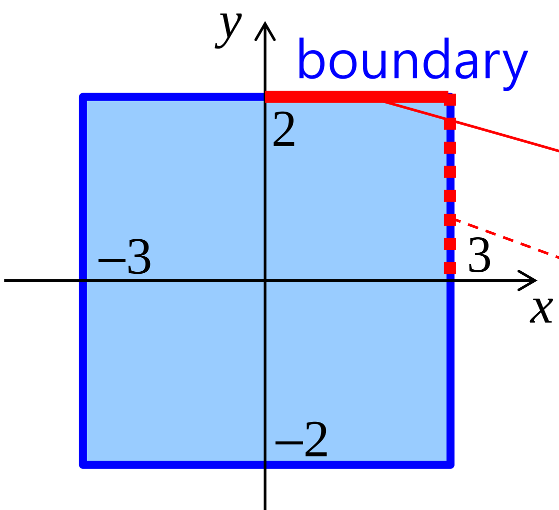
$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \;\simeq\; \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \;\; \Gamma$$

$$\int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

$$\int_\Gamma \left(\hat{\theta} - \frac{1}{2}\left(x^2 + y^2\right) \right) \overline{W}_l d\Gamma = 0$$



Since the trial function set is symmetric in x and y, boundary to be satisfied will be chosen as follows:

$$0 \le x \le 3, y = 2$$

$$x = 3, 0 \le y \le 2$$

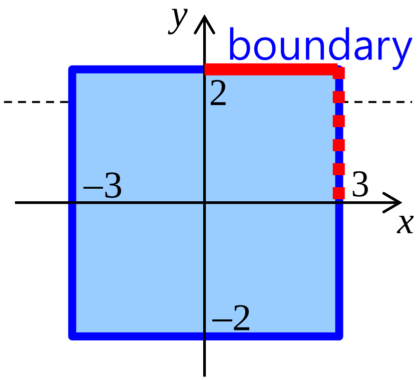
$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{M} a_m N_m$$

$$\mathbf{R}_\Gamma = B(\hat{\phi}) = \mathscr{M}\hat{\phi} + r \quad on \;\; \Gamma$$

$$\int_\Gamma \overline{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$



$$\int_\Gamma \left(\hat{\theta} - \frac{1}{2}\left(x^2 + y^2\right) \right) \overline{W}_l d\Gamma = 0$$

If we choose weighting functions $\overline{W}_l$ **defined by** $\overline{W}_l = N_l\big|_\Gamma$,

$$\int_0^2 \left(\hat{\theta}\Big|_{x=3} - \frac{1}{2}\left(9 + y^2\right) \right) N_l\Big|_{x=3} dy + \int_0^3 \left(\hat{\theta}\Big|_{y=2} - \frac{1}{2}\left(x^2 + 4\right) \right) N_l\Big|_{y=2} dx = 0$$

$$x = 3, 0 \le y \le 2 \qquad\qquad 0 \le x \le 3, y = 2$$

$$\int_0^2 \left(\hat{\theta}\Big|_{x=3} - \frac{1}{2}\left(9+y^2\right) \right) N_l\Big|_{x=3} dy + \int_0^3 \left(\hat{\theta}\Big|_{y=2} - \frac{1}{2}\left(x^2+4\right) \right) N_l\Big|_{y=2} dx = 0$$

$$\downarrow \quad \hat{\theta} = a_1 + a_2\left(x^2 - y^2\right) + a_3\left(x^4 - 6x^2y^2 + y^4\right)$$

$$\int_0^2 \left(a_1 + a_2\left(3^2 - y^2\right) + a_3\left(3^4 - 6\cdot 3^2 y^2 + y^4\right) - \frac{1}{2}\left(9+y^2\right) \right) N_l\Big|_{x=3} dy$$

$$+ \int_0^3 \left(a_1 + a_2\left(x^2 - 2^2\right) + a_3\left(x^4 - 6\cdot x^2 2^2 + 2^4\right) - \frac{1}{2}\left(x^2+4\right) \right) N_l\Big|_{y=2} dx = 0$$

$$\downarrow$$

$$a_1\left(\int_0^2 N_l\Big|_{x=3} dy + \int_0^3 N_l\Big|_{y=2} dx \right) + a_2\left(\int_0^2 \left(3^2 - y^2\right) N_l\Big|_{x=3} dy + \int_0^3 \left(x^2 - 2^2\right) N_l\Big|_{y=2} dx \right)$$

$$+ a_3\left(\int_0^2 \left(3^4 - 6\cdot 3^2 y^2 + y^4\right) N_l\Big|_{x=3} dy + \int_0^3 \left(x^4 - 6\cdot x^2 2^2 + 2^4\right) N_l\Big|_{y=2} dx \right)$$

$$+ \int_0^2 \left(-\frac{1}{2}\left(9+y^2\right) \right) N_l\Big|_{x=3} dy + \int_0^3 \left(-\frac{1}{2}\left(x^2+4\right) \right) N_l\Big|_{y=2} dx = 0$$

$$a_1\left(\underbrace{\int_0^2 N_l\Big|_{x=3}dy+\int_0^3 N_l\Big|_{y=2}dx}_{K_{l1}}\right)+a_2\left(\underbrace{\int_0^2\left(3^2-y^2\right)N_l\Big|_{x=3}dy+\int_0^3\left(x^2-2^2\right)N_l\Big|_{y=2}dx}_{K_{l2}}\right)$$

$$+a_3\left(\underbrace{\int_0^2\left(3^4-6\cdot 3^2y^2+y^4\right)N_l\Big|_{x=3}dy+\int_0^3\left(x^4-6\cdot x^2 2^2+2^4\right)N_l\Big|_{y=2}dx}_{K_{l3}}\right)$$

$$+\underbrace{\int_0^2\left(-\frac{1}{2}\left(9+y^2\right)\right)N_l\Big|_{x=3}dy+\int_0^3\left(-\frac{1}{2}\left(x^2+4\right)\right)N_l\Big|_{y=2}dx}_{f_l}=0$$

$$\downarrow$$

$$a_1K_{l1}+a_2K_{l2}+a_3K_{l3}+f_l=0$$

, where

$$K_{l1}=\int_0^2 N_l\Big|_{x=3}dy+\int_0^3 N_l\Big|_{y=2}dx$$

$$K_{l2}=\int_0^2\left(3^2-y^2\right)N_l\Big|_{x=3}dy+\int_0^3\left(x^2-2^2\right)N_l\Big|_{y=2}dx$$

$$K_{l3}=\int_0^2\left(3^4-6\cdot 3^2y^2+y^4\right)N_l\Big|_{x=3}dy+\int_0^3\left(x^4-6\cdot x^2 2^2+2^4\right)N_l\Big|_{y=2}dx$$

$$f_l=\int_0^2\left(-\frac{1}{2}\left(9+y^2\right)\right)N_l\Big|_{x=3}dy+\int_0^3\left(-\frac{1}{2}\left(x^2+4\right)\right)N_l\Big|_{y=2}dx$$

$$a_1K_{l1} + a_2K_{l2} + a_3K_{l3} + f_l = 0$$

, **where**

$$K_{l1} = \int_0^2 N_l\big|_{x=3}\, dy + \int_0^3 N_l\big|_{y=2}\, dx$$

$$K_{l2} = \int_0^2 \left(3^2 - y^2\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(x^2 - 2^2\right) N_l\big|_{y=2}\, dx$$

$$K_{l3} = \int_0^2 \left(3^4 - 6\cdot 3^2 y^2 + y^4\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(x^4 - 6\cdot x^2 2^2 + 2^4\right) N_l\big|_{y=2}\, dx$$

$$f_l = \int_0^2 \left(-\frac{1}{2}\left(9 + y^2\right)\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(-\frac{1}{2}\left(x^2 + 4\right)\right) N_l\big|_{y=2}\, dx$$

$$\downarrow$$

$$a_1K_{l1} + a_2K_{l2} + a_3K_{l3} = -f_l$$

$$N_1 = 1,\ N_2 = x^2 - y^2,\ N_3 = x^4 - 6x^2y^2 + y^4$$

$$\downarrow l = 1, 2, 3$$

$$\begin{aligned} a_1K_{11} + a_2K_{12} + a_3K_{13} &= -f_1 \\ a_1K_{21} + a_2K_{22} + a_3K_{23} &= -f_2 \\ a_1K_{31} + a_2K_{32} + a_3K_{33} &= -f_3 \end{aligned} \longrightarrow \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -f_1 \\ -f_2 \\ -f_3 \end{bmatrix}$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

$$\begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -f_1 \\ -f_2 \\ -f_3 \end{bmatrix}$$

, where

$$K_{l1} = \int_0^2 N_l\big|_{x=3}\, dy + \int_0^3 N_l\big|_{y=2}\, dx$$

$$K_{l2} = \int_0^2 \left(3^2 - y^2\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(x^2 - 2^2\right) N_l\big|_{y=2}\, dx$$

$$K_{l3} = \int_0^2 \left(3^4 - 6\cdot 3^2 y^2 + y^4\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(x^4 - 6\cdot x^2 2^2 + 2^4\right) N_l\big|_{y=2}\, dx$$

$$f_l = \int_0^2 \left(-\frac{1}{2}\left(9 + y^2\right)\right) N_l\big|_{x=3}\, dy + \int_0^3 \left(-\frac{1}{2}\left(x^2 + 4\right)\right) N_l\big|_{y=2}\, dx$$

$$N_1 = 1,\ N_2 = x^2 - y^2,\ N_3 = x^4 - 6x^2y^2 + y^4$$

$$\downarrow$$

$$\mathbf{K}\,\mathbf{a} = \mathbf{f}$$

, where $\mathbf{K} = \begin{bmatrix} 5 & 12.3333 & -95 \\ 12.3333 & 145 & 98.543 \\ -95 & 98.543 & 18{,}170.4 \end{bmatrix}$, $\mathbf{f} = \begin{bmatrix} 20.8333 \\ 78.1 \\ -539.643 \end{bmatrix}$

$$\downarrow$$

$$\mathbf{a} = \begin{bmatrix} 3.2154 \\ -0.2749 \\ -0.01438 \end{bmatrix} \longrightarrow$$

$$\hat{\theta} = a_1 + a_2\left(x^2 - y^2\right) + a_3\left(x^4 - 6x^2y^2 + y^4\right)$$

$$\phi = \theta - \frac{1}{2}\left(x^2 + y^2\right)$$

$$\hat{\phi} = 3.2154 - 0.2749\left(x^2 - y^2\right) - 0.01438\left(x^4 - 6x^2y^2 + y^4\right) - \frac{1}{2}\left(x^2 + y^2\right)$$

NATURAL BOUNDARY CONDITIONS

Weak form

The weighted sum of the residual

$$\int_\Omega W_l \mathbf{R}_\Omega d\Omega + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \mathscr{L}\hat{\phi} + p \ \ in \ \ \Omega$$

The boundary residual

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \mathscr{M}\hat{\phi} + r \ \ on \ \ \Gamma$$

The weighted residual form could require the evaluation of integrals involving derivatives of $\hat{\phi}$ along the boundaries which may present difficulties if these boundaries are of curved or complicated.

In this section we show, for certain equations and boundary conditions how such boundary derivative evaluations can be made unnecessary.

Weak form

The residual in domain

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \mathscr{L}\hat{\phi} + p \quad in \quad \Omega$$

The weighted sum of the residual

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega} + \int_\Gamma \bar{W}_l \mathbf{R}_\Gamma d\Gamma = 0$$

The boundary residual

$$\mathbf{R}_\Gamma = B(\hat{\phi}) - \cancel{B(\phi)} = \mathscr{M}\hat{\phi} + r \quad on \quad \Gamma$$

The first term of the residual statement

$$\boxed{\int_\Omega W_l \mathbf{R}_\Omega d\Omega} \equiv \int_\Omega W_l \left(\mathscr{L}\hat{\phi} + p \right) d\Omega$$

can be frequently be rearranged to yield an expression of the form

$$\int_\Omega W_l \left(\mathscr{L}\hat{\phi} \right) d\Omega \equiv \int_\Omega \left(\mathscr{C} W_l \right) \left(\mathscr{D}\hat{\phi} \right) d\Omega + \int_\Gamma W_l \mathscr{E}\hat{\phi} d\Gamma$$

ex) integration by part

Where $\mathscr{C}$, $\mathscr{D}$ and $\mathscr{E}$ are linear differential operators involving an order of differentiation lower than that of the original operator $\mathscr{L}$.

The resulting expression is often termed the weak form of the weighted residual statement, which relaxes the requirement on the trial functions.

Natural Boundary Condition

The weighted sum of the residual

$$\boxed{\int_{\Omega} W_l \mathbf{R}_{\Omega} d\Omega} + \int_{\Gamma} \bar{W}_l \mathbf{R}_{\Gamma} d\Gamma = 0$$

$$\boxed{\int_{\Omega} W_l \mathbf{R}_{\Omega} d\Omega} \equiv \boxed{\int_{\Omega} W_l \left(\mathscr{L}\hat{\phi}\right.}\left. + p \right) d\Omega$$

$$\boxed{\int_{\Omega} W_l \left(\mathscr{L}\hat{\phi} \right)} d\Omega \equiv \int_{\Omega} \left(\mathscr{C} W_l \right) \left(\mathscr{D}\hat{\phi} \right) d\Omega + \int_{\Gamma} W_l \mathscr{E}\hat{\phi} d\Gamma$$

$$\downarrow$$

$$\int_{\Omega} \left(\mathscr{C} W_l \right) \left(\mathscr{D}\hat{\phi} \right) d\Omega + \int_{\Gamma} W_l \mathscr{E}\hat{\phi} d\Gamma + \int_{\Omega} W_l \left(p \right) d\Omega + \int_{\Gamma} \bar{W}_l \mathbf{R}_{\Gamma} d\Gamma = 0$$

$$\downarrow$$

$$\int_{\Omega} \left(\mathscr{C} W_l \right) \left(\mathscr{D}\hat{\phi} \right) d\Omega + \int_{\Omega} W_l \left(p \right) d\Omega + \int_{\Gamma} W_l \mathscr{E}\hat{\phi} d\Gamma + \int_{\Gamma} \bar{W}_l \mathbf{R}_{\Gamma} d\Gamma = 0$$

$$\int_{\Omega}(\mathscr{C}W_l)(\mathscr{D}\hat{\phi})\,d\Omega+\int_{\Omega}W_l(p)\,d\Omega+\underbrace{\int_{\Gamma}W_l\mathscr{E}\hat{\phi}\,d\Gamma}_{(2)}+\underbrace{\int_{\Gamma}\bar{W}_l\mathbf{R}_{\Gamma}\,d\Gamma}_{(1)}=0$$

It may be possible to arrange for the last term (1) to cancel with the term (2) by a suitable choice of the boundary weighting function $\bar{W}_l$

,thus eliminating the integral involving $\hat{\phi}$ or its derivatives along the boundary.

This will only be possible for certain boundary conditions that we term *natural.*

In general, boundary conditions involving prescribed values of the function itself will not benefit from this treatment, while certain boundary conditions on derivatives will.

Natural Boundary Condition

Example 2.6

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ in $0 \le x \le 1$

Boundary Condition $\phi = 0$ at $x = 0$ and $d\phi / dx = 20$ at $x = 1$

Let us assume that we choose an approximation

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

where ψ and the set N_m is such that the condition at $x = 0$ is automatically satisfied, for example $\psi = 0$; $\left\{ N_m = x^m; m = 1, 2, 3, \ldots \right\}$ could be a suitable choice here

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$\psi = 0 \quad \{N_m = x^m; m = 1, 2, 3, \ldots$

Governing equation $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \ \Omega$

$$\frac{d^2\phi}{dx^2} = \phi \rightarrow \frac{d^2\phi}{dx^2} - \phi = 0 \rightarrow \boxed{A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \ \ \Omega}$$

Boundary Conditions $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \ \Gamma$

$\phi = 0$ **at** $x = 0$ → $\phi - 0 = 0$ **at** $x = 0$ → $\boxed{B(\phi) = \phi = 0 \ at \ \ x = 0}$

$d\phi / dx = 1$ **at** $x = 1$ → $d\phi / dx - 20 = 0$ **at** $x = 1$ → $\boxed{B(\phi) = d\phi / dx - 20 = 0 \ at \ \ x = 1}$

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

Example 2.6

$$\psi = 0 \quad \left\{ N_m = x^m ; m = 1, 2, 3, \ldots \right\}$$

$$\boxed{A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \;\; 0 < x < 1}$$

$$B(\phi) = \phi = 0 \;\; at \;\; x = 0$$

$$B(\phi) = d\phi / dx - 20 = 0 \;\; at \;\; x = 1$$

The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

The boundary residual:

$$\mathbf{R}_{\Gamma,0} = B(\hat{\phi}) - \cancel{B(\phi)} = \phi = 0 \;\; at \;\; x = 0$$

$$\mathbf{R}_{\Gamma,1} = B(\hat{\phi}) - \cancel{B(\phi)} = d\phi / dx - 20 = 0 \;\; at \;\; x = 1$$

The residual at x=0 being omitted, as the trial function satisfy the boundary condition at $x = 0$

The weighted residual form:

$$\int_0^1 W_l \mathbf{R}_\Omega dx + \overline{W}_l \mathbf{R}_{\Gamma,1}\Big|_{x=1}$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right]\Bigg|_{x=1} = 0$$

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\psi = 0 \quad \left\{N_m = x^m; m = 1, 2, 3, \ldots\right\}$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right] \Bigg|_{x=1} = 0$$

$$\downarrow$$

$$\int_0^1 W_l \frac{d^2\hat{\phi}}{dx^2} dx - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right] \Bigg|_{x=1} = 0$$

$\downarrow$ **Carrying out integration by parts gives**

$$-\int_0^1 \frac{dW_l}{dx} \frac{d\hat{\phi}}{dx} dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right] \Bigg|_{x=1} = 0$$

The resulting expression is often termed the weak form of the weighted residual statement, which relaxes the requirement on the trial functions.

$$\psi = 0 \quad \left\{ N_m = x^m ; m = 1, 2, 3, \ldots \right\}$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right] \Bigg|_{x=1} = 0$$

$$\downarrow \quad \overline{W}_l \Big|_{x=1} = -W_l \Big|_{x=1}$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx - \left[W_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right] \Bigg|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx - W_l \frac{d\hat{\phi}}{dx} \Bigg|_{x=1} + 20 W_l \Big|_{x=1} = 0$$

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$$\int_0^1 W_l \frac{d^2\hat{\phi}}{dx^2}dx - \int_0^1 W_l \hat{\phi}dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 20 \right) \right]_{x=1} = 0 \qquad \psi = 0 \quad \left\{ N_m = x^m ; m = 1, 2, 3, \ldots \right\}$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi}dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1} + 20 W_l\big|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \cancel{W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1}} - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} - \int_0^1 W_l \hat{\phi}dx - \cancel{W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1}} + 20 W_l\big|_{x=1} = 0$$

$\downarrow$ **eliminating the integral involving $\hat{\phi}$ or its derivatives along the boundary.**

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx - \int_0^1 W_l \hat{\phi}dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} + 20 W_l\big|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \left(\frac{dW_l}{dx}\frac{d\hat{\phi}}{dx} + W_l \hat{\phi} \right) dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} + 20 W_l\big|_{x=1} = 0$$

$$\hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

$\psi = 0 \quad \{N_m = x^m; m = 1, 2, 3, \ldots\}$

$$-\int_0^1 \left(\frac{dW_l}{dx} \frac{d\hat{\phi}}{dx} + W_l \hat{\phi} \right) dx - W_l \left. \frac{d\hat{\phi}}{dx} \right|_{x=0} + 20 W_l \big|_{x=1} = 0$$

$\overline{W}_l \big|_{x=1} = -W_l \big|_{x=1}$

$$\downarrow \; W_l \big|_{x=0} = 0$$

$$-\int_0^1 \left(\frac{dW_l}{dx} \frac{d\hat{\phi}}{dx} + W_l \hat{\phi} \right) dx + 20 W_l \big|_{x=1} = 0$$

$$\downarrow$$

$$\int_0^1 \left(\frac{dW_l}{dx} \frac{d\hat{\phi}}{dx} + W_l \hat{\phi} \right) dx = 20 W_l \big|_{x=1}$$

Thus in this formulation there is no need to evaluate the derivative of $\hat{\phi}$ at $x = 1$, and the boundary condition to be applied at this point is a natural condition

Piecewise Defined Trial Functions and the Finite Element Method

E : total no. of elements

Introduction

Function Approximation by Trial Functions

We assumed implicitly that the trial functions were, defined by a single expression, valid throughout the whole domain Ω

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$$

and the integral of the approximating equations were evaluated in one operation over the domain

$$\int_{\Omega} W_l \, R_{\Omega} \, d\Omega = 0, \quad R_{\Omega} = \phi - \hat{\phi}$$

$$\sum_{e=1}^{E} \Omega^e, \sum_{e=1}^{E} \Gamma^e$$

divide the domain Ω and the boundary Γ into a number of nonoverlapping elements Ω^e and Γ^e

The trial functions were can be also defined in a piecewise manner by using various expressions in the various subdomains.

$$\phi \simeq \hat{\phi} = \psi + \sum_{m=1}^{E} \phi_m N_m$$

The definite integral can be obtained simply by summing the contributions from each subdomain as

$$\int_{\Omega} W_l \, R_{\Omega} \, d\Omega = \sum_{e=1}^{E} \int_{\Omega^e} W_l \, R_{\Omega} \, d\Omega$$

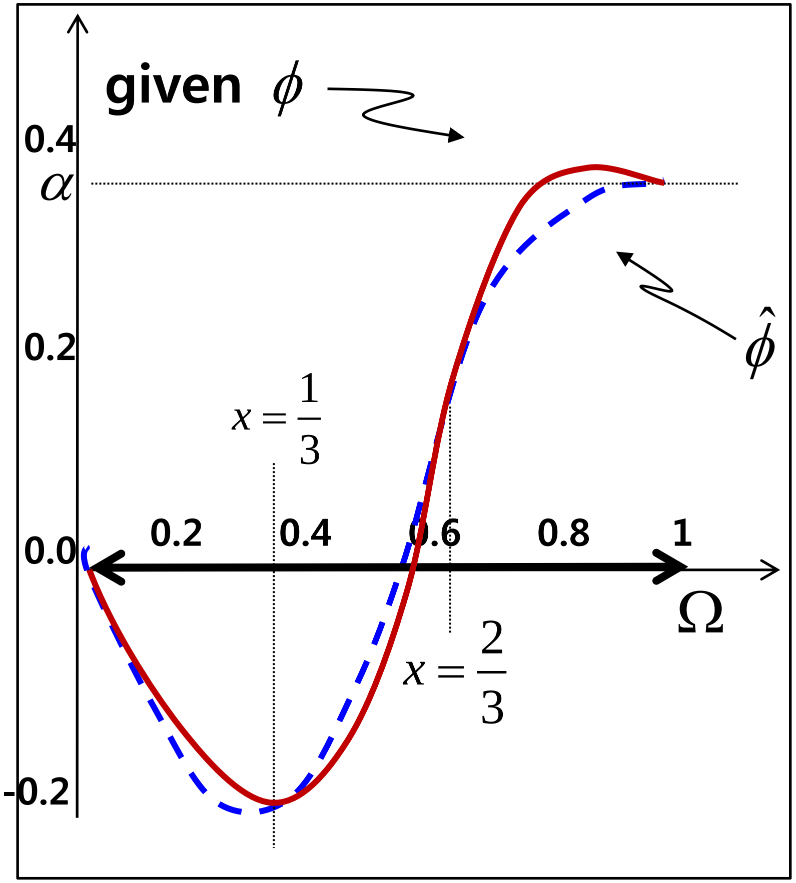
simplified domain

banded matrix

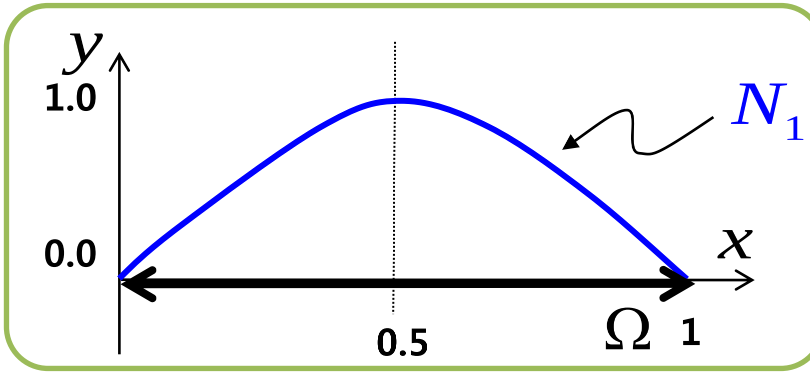
if the subdomains are of a relaticvely simple Trial and if the definition of the trial functions over therse subdomains can be made in a repeatable manner, it is possible to deal in this fashion with assembled regions of complex Trials quite readily.

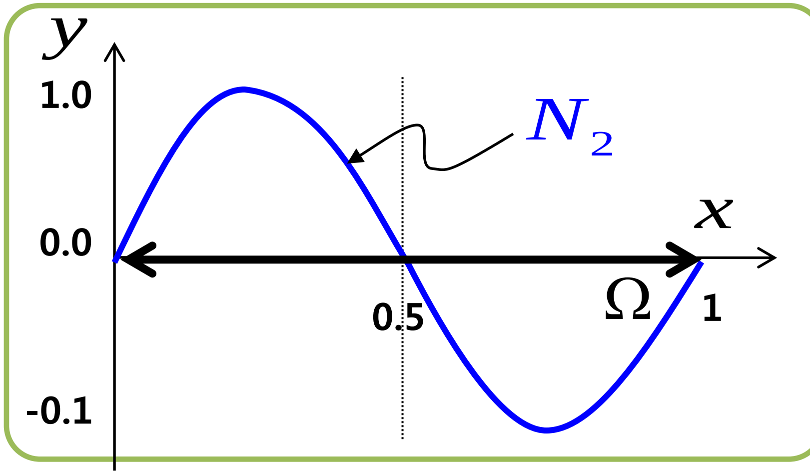
If the trial functions are to be defined in a piecewise manner, it is advantageous to assign to them a narros "base" and make their value zero everywhere except in the element in question and in the subdomains immediately adjacent to this element. This, as we shall see later, will give banded matrices.

Function Approximation by Trial Functions defined by "a single expression, valid throughout the whole domain"



two trial functions are chosen such as

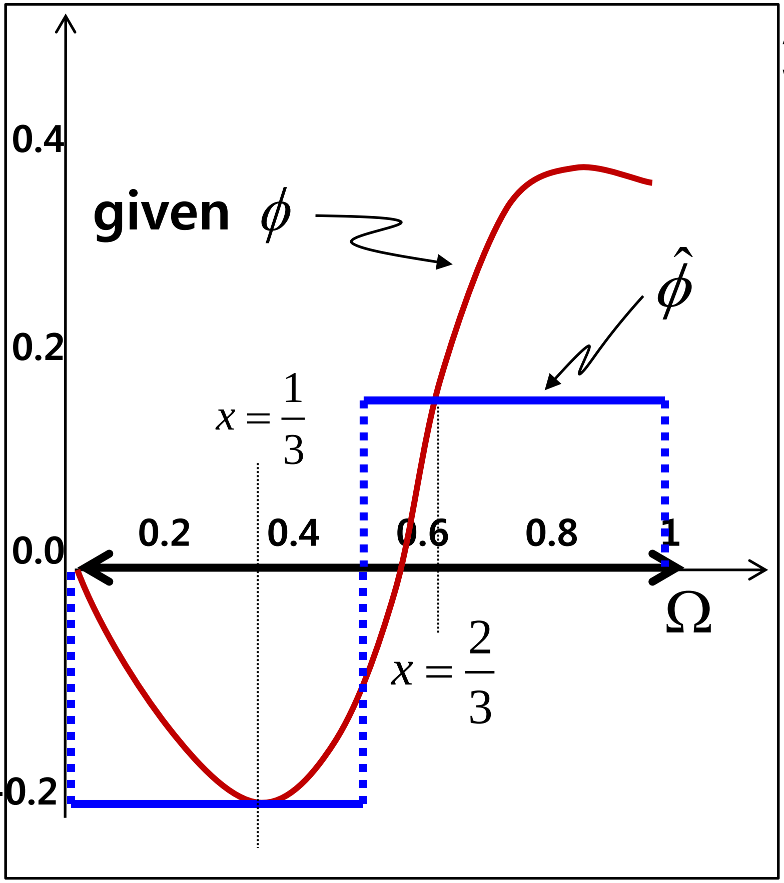




function approximation : $\hat{\phi}=\psi+a_1N_1+a_2N_2 \quad ,\psi=\alpha x$

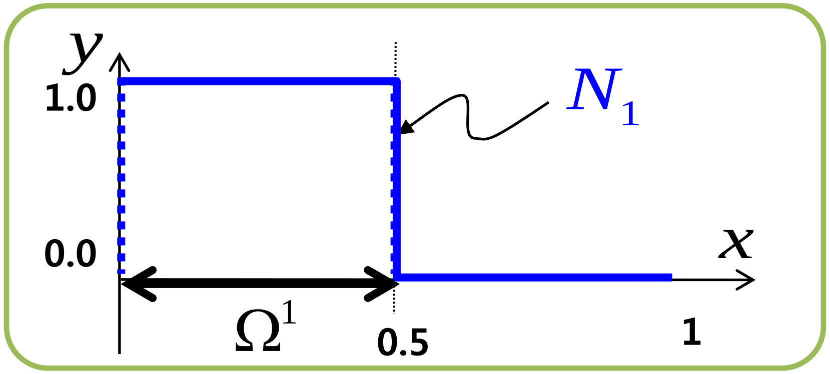
in this case a_1, a_2 **are determined to satisfy** $\phi=\hat{\phi}$ **at** $x=\frac{1}{3}, x=\frac{2}{3}$

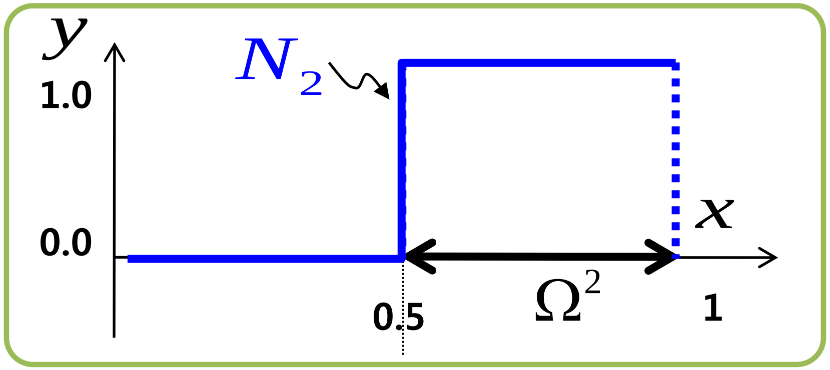
Function Approximation by Trial Functions defined by "in a piecewise manner in the various subdomains"



Approximation of a given function ϕ by means of a function which takes a constant value on each element

when two trial functions are chosen such as

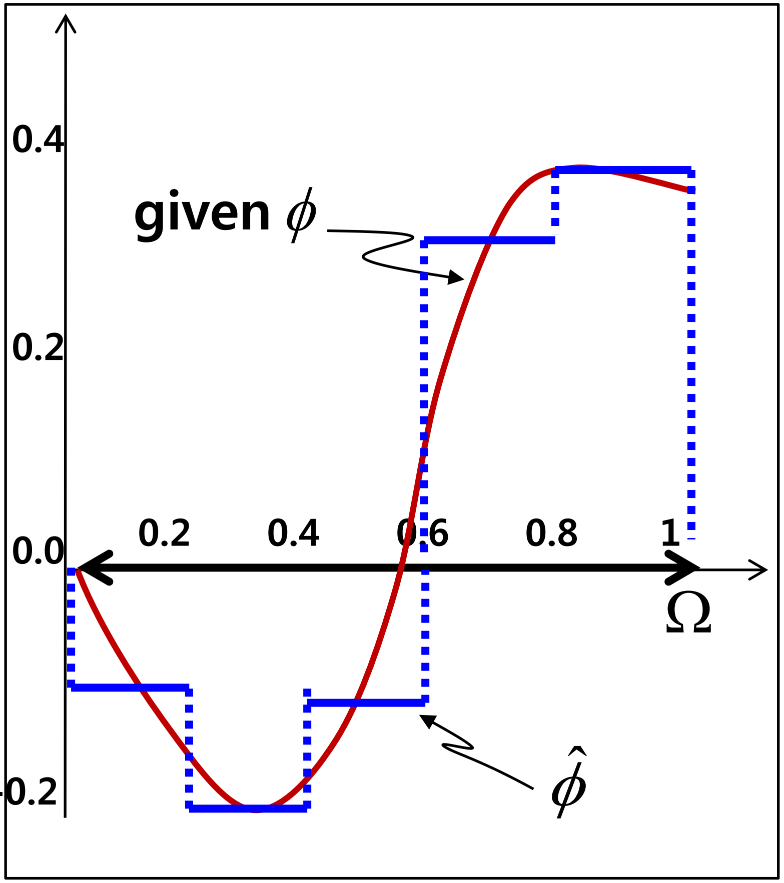




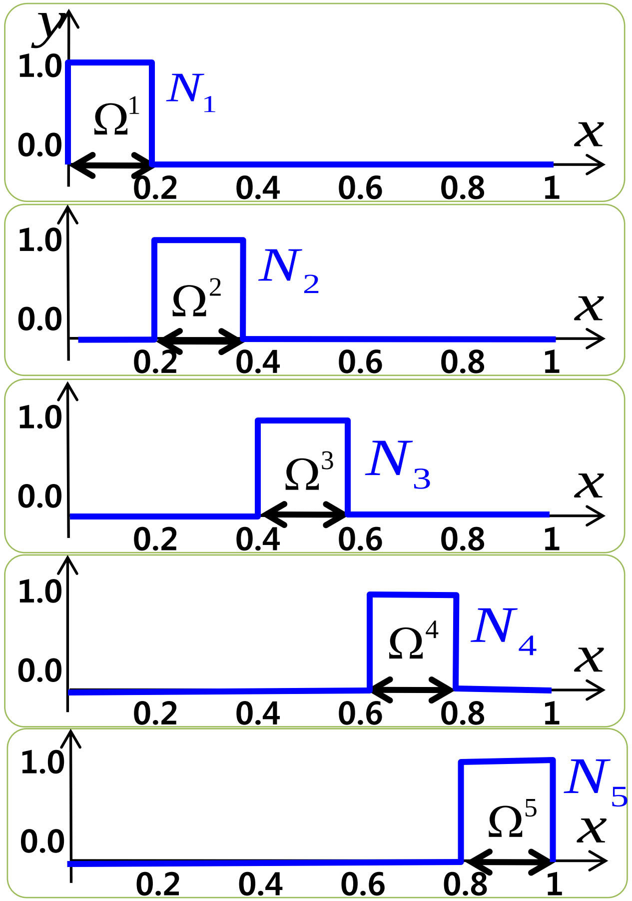
function approximation : $\hat{\phi}=a_1N_1+a_2N_2 \quad ,\psi=\alpha x$

in this case a_1, a_2 are determined to satisfy $\phi=\hat{\phi}$ at $x=\frac{1}{3},\ x=\frac{2}{3}$

Note) the arbitrary function ψ has been omitted. The end values , however, can be satisfied as closely as required by suitable reduction in the length of the elements at the end points ([Zienkiewicz 1983] pp.97-98



if five trial functions are chosen such as



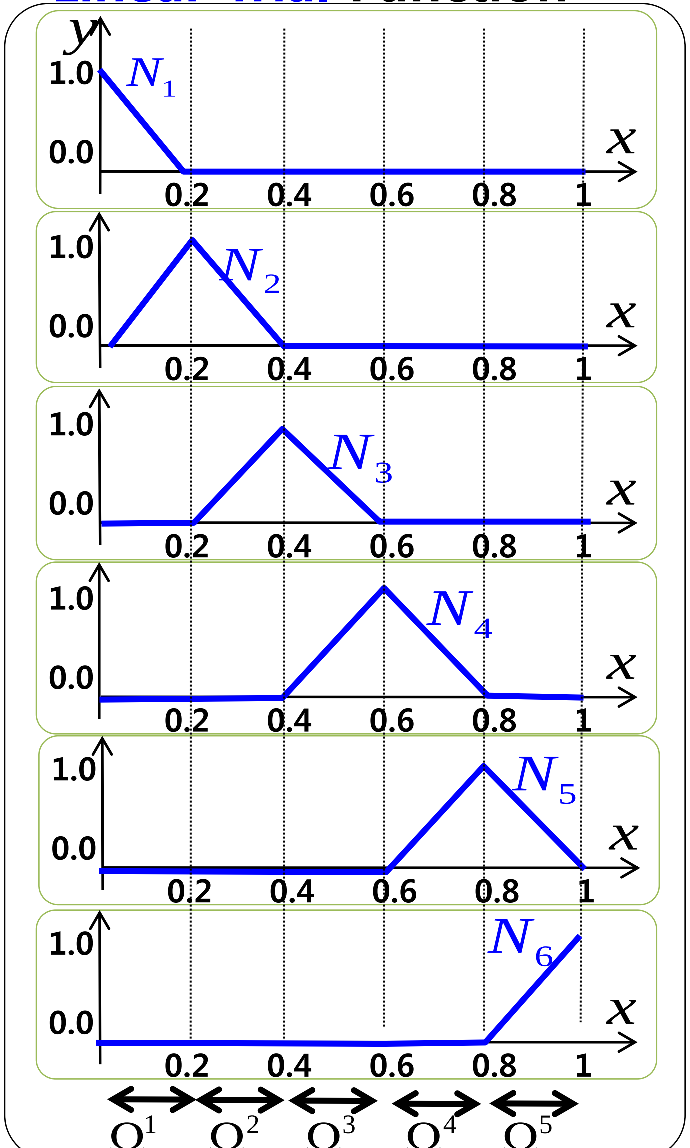
function approximation : $\hat{\phi} = a_1 N_1 + a_2 N_2 + a_3 N_3 + a_4 N_4 + a_5 N_5$

in this case a_m **are determined to satisfy** $\phi = \hat{\phi}$ **at midpoint of each** Ω^m

$\therefore a_m = \phi_m \quad , m = 1, \cdots, 5$

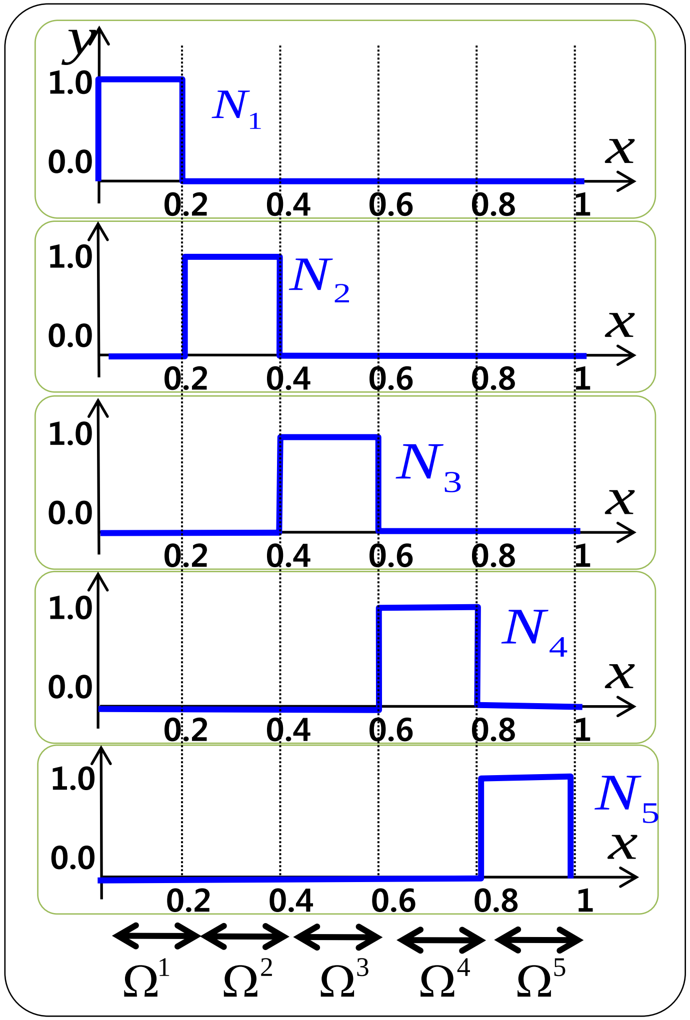
"Point collocation method"

Linear Trial Function



An improved approximation has been produced by using an approximation function that **varies linearly** with over each subdomain called '**element**'

Constant Trial Function



Weighted Residual Method for the Function Approximation by Piecewise Linear Trial Functions

The Weighted Residual Statement

$$\int_\Omega W_l \, R_\Omega \, d\Omega = 0, \quad R_\Omega \equiv \phi - \hat{\phi}$$

$$\phi \simeq \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m \ in\ \Omega$$

Piecewise Linear Trial Functions

,where E is number of the elements.

$$\int_\Omega W_l \left(\phi - \sum_{m=1}^{E+1} \phi_m N_m \right) dx = 0$$

The Galerkin Method

$$\int_\Omega N_l \left(\phi - \sum_{m=1}^{E+1} \phi_m N_m \right) dx = 0$$

(derivation)

$$\int_\Omega N_l \left(\phi - \sum_{m=1}^{E+1} \phi_m N_m \right) dx = 0$$

$$\downarrow$$

$$\int_\Omega N_l \phi \, dx - \int_\Omega N_l \left(\sum_{m=1}^{E+1} \phi_m N_m \right) dx = 0$$

$$\downarrow$$

$$\int_\Omega N_l \left(\sum_{m=1}^{E+1} \phi_m N_m \right) dx = \int_\Omega N_l \phi \, dx$$

$$\int_\Omega N_l \left(\sum_{m=1}^{E+1} \phi_m N_m \right) dx = \int_\Omega N_l \phi \, dx$$

for instance,

$$\Omega = \sum_{e=1}^{E} \Omega^e \quad ,$$

$$\Omega : 0 < x < 1$$

$$\Omega^1 : x_1 < x < x_2$$

$$\Omega^2 : x_2 < x < x_3$$

$$\Omega^3 : x_3 < x < x_4$$

$$E = 3$$

$$l, m = 1, 2, 3, 4$$

Trial Function N_1

Trial Function N_2

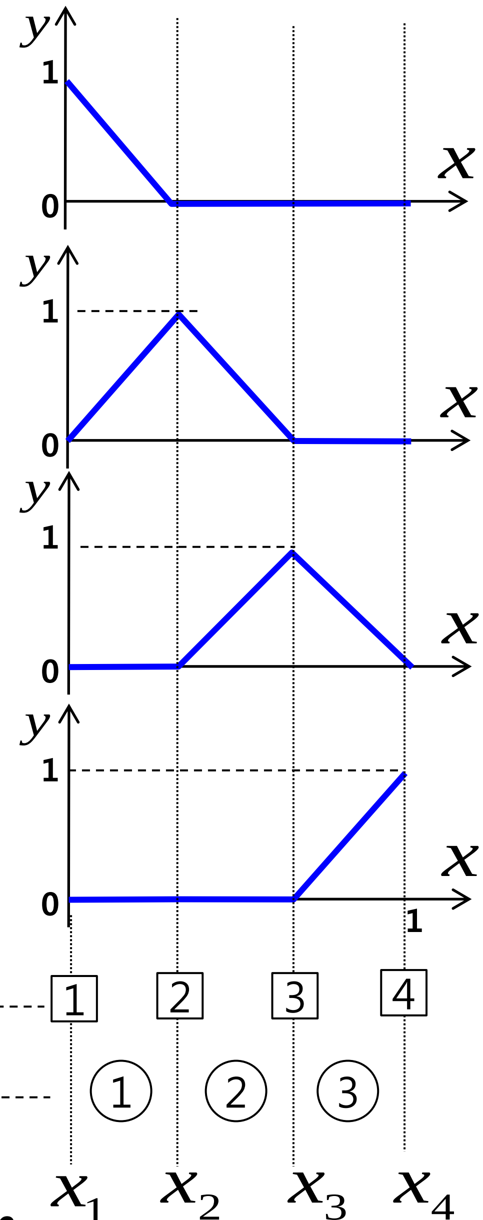
Trial Function N_3

Trial Function N_4

No. of Node

No. of Element

Coordinates in 1-Dimensional Domain



(derivation)

$$\int_\Omega N_l \left(\sum_{m=1}^{E+1} \phi_m N_m \right) dx = \int_\Omega N_l \phi \, dx \, , E = 3, \quad l, m = 1,2,3,4$$

$$\sum_{m=1}^{E+1} \phi_m \left[\int_\Omega N_l N_m dx \right] dx = \int_\Omega N_l \phi \, dx$$

$$\downarrow$$

$$\begin{array}{cccc} m=1 & m=2 & m=3 & m=4 \end{array}$$

$$\phi_1 \int_{x_1}^{x_4} N_l N_1 dx + \phi_2 \int_{x_1}^{x_4} N_l N_2 dx + \phi_3 \int_{x_1}^{x_4} N_l N_3 dx + \phi_4 \int_{x_1}^{x_4} N_l N_4 dx = \int_{x_1}^{x_4} N_l \phi \, dx$$

$$\downarrow$$

$$\begin{array}{cccc} m=1 & m=2 & m=3 & m=4 \end{array}$$

$$l=1 \quad \phi_1 \int_{x_1}^{x_4} N_1 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_1 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_1 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_1 N_4 dx = \int_{x_1}^{x_4} N_1 \phi \, dx$$

$$l=2 \quad \phi_1 \int_{x_1}^{x_4} N_2 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_2 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_2 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_2 N_4 dx = \int_{x_1}^{x_4} N_2 \phi \, dx$$

$$l=3 \quad \phi_1 \int_{x_1}^{x_4} N_3 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_3 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_3 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_3 N_4 dx = \int_{x_1}^{x_4} N_3 \phi \, dx$$

$$l=4 \quad \phi_1 \int_{x_1}^{x_4} N_4 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_4 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_4 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_4 N_4 dx = \int_{x_1}^{x_4} N_4 \phi \, dx$$

(derivation)

$$\boxed{\phi_1\int_{x_1}^{x_4} N_1N_1dx}+\phi_2\int_{x_1}^{x_4} N_1N_2dx+\phi_3\int_{x_1}^{x_4} N_1N_3dx+\phi_4\int_{x_1}^{x_4} N_1N_4dx=\int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx+\phi_2\int_{x_1}^{x_4} N_2N_2dx+\phi_3\int_{x_1}^{x_4} N_2N_3dx+\phi_4\int_{x_1}^{x_4} N_2N_4dx=\int_{x_1}^{x_4} N_2\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_3N_1dx+\phi_2\int_{x_1}^{x_4} N_3N_2dx+\phi_3\int_{x_1}^{x_4} N_3N_3dx+\phi_4\int_{x_1}^{x_4} N_3N_4dx=\int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx+\phi_2\int_{x_1}^{x_4} N_4N_2dx+\phi_3\int_{x_1}^{x_4} N_4N_3dx+\phi_4\int_{x_1}^{x_4} N_4N_4dx=\int_{x_1}^{x_4} N_4\phi\, dx$$

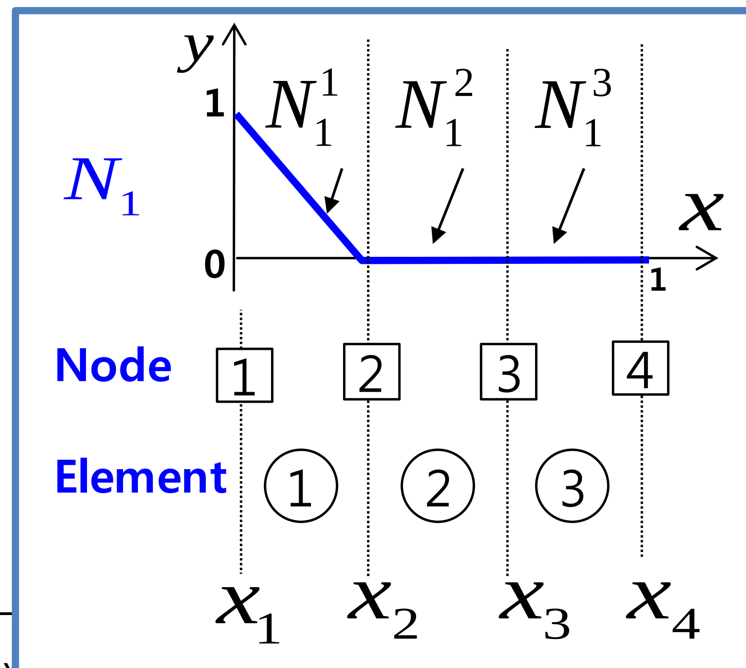
$$\boxed{\phi_1\int_{x_1}^{x_4} N_1N_1dx} => \phi_1\int_{x_1}^{x_2} N_1^1N_1^1dx+\phi_1\int_{x_2}^{x_3} \cancel{N_1^2}\cancel{N_1^2}dx+\phi_1\int_{x_3}^{x_4} \cancel{N_1^3}\cancel{N_1^3}dx$$

$$\Omega: x_1<x<x_4 \;=>\; \Omega=\Omega^1+\Omega^2+\Omega^3$$

$$where,\Omega^1: x_1<x<x_2,\ \Omega^2: x_2<x<x_3,\ \Omega^3: x_3<x<x_4$$

notation

N_1^1 — upper index: **No. of Element**; lower index: **No. Trial Function**



$$N_1^2=0,\ N_1^3=0$$

$$\boxed{\therefore \phi_1\int_{x_1}^{x_4} N_1N_1dx=\phi_1\int_{x_1}^{x_2} N_1^1N_1^1dx}$$

(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx+\boxed{\phi_2\int_{x_1}^{x_4} N_1N_2dx}+\phi_3\int_{x_1}^{x_4} N_1N_3dx+\phi_4\int_{x_1}^{x_4} N_1N_4dx=\int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx+\phi_2\int_{x_1}^{x_4} N_2N_2dx+\phi_3\int_{x_1}^{x_4} N_2N_3dx+\phi_4\int_{x_1}^{x_4} N_2N_4dx=\int_{x_1}^{x_4} N_2\phi\, dx$$

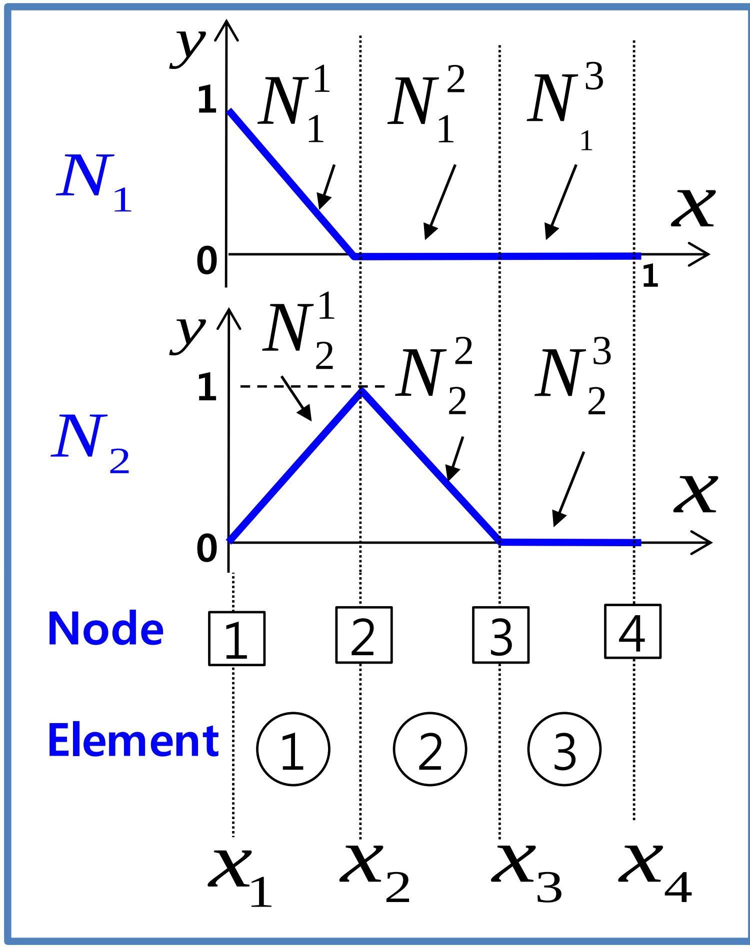
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx+\phi_2\int_{x_1}^{x_4} N_3N_2dx+\phi_3\int_{x_1}^{x_4} N_3N_3dx+\phi_4\int_{x_1}^{x_4} N_3N_4dx=\int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx+\phi_2\int_{x_1}^{x_4} N_4N_2dx+\phi_3\int_{x_1}^{x_4} N_4N_3dx+\phi_4\int_{x_1}^{x_4} N_4N_4dx=\int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_2\int_{x_1}^{x_4} N_1N_2dx}$$

$$=\phi_2\int_{x_1}^{x_2} N_1^1N_2^1dx+\phi_2\int_{x_2}^{x_3} \cancel{N_1^2N_2^2}dx+\phi_2\int_{x_3}^{x_4} \cancel{N_1^3N_2^3}dx$$

$$\boxed{\therefore \phi_2\int_{x_1}^{x_4} N_1N_2dx=\phi_2\int_{x_1}^{x_2} N_1^1N_2^1dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \boxed{\phi_3\int_{x_1}^{x_4} N_1N_3dx} + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

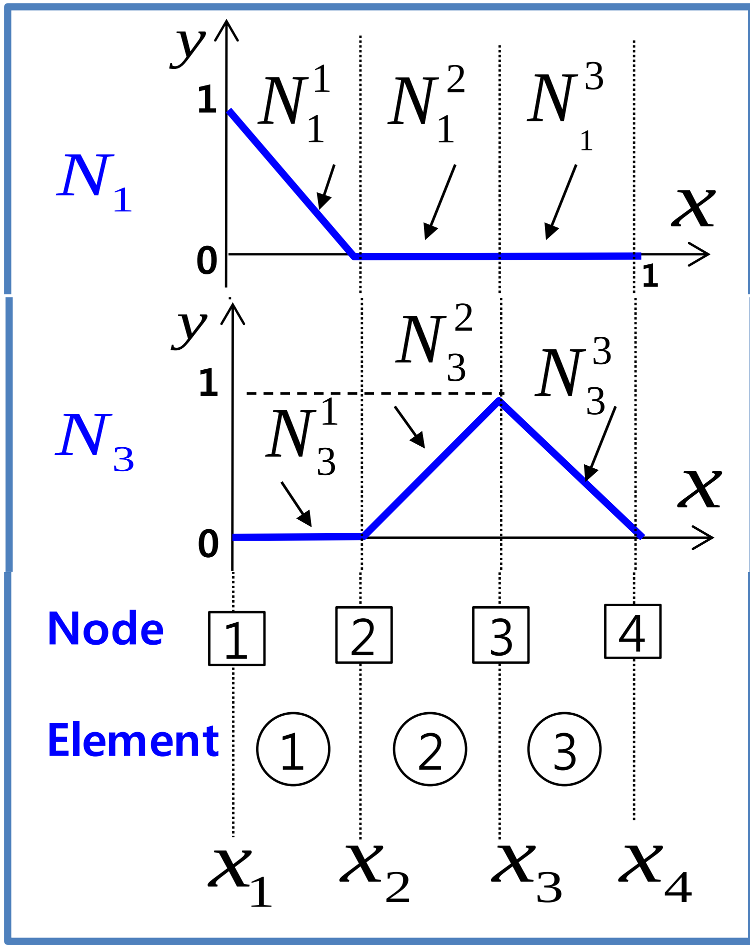
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_3\int_{x_1}^{x_4} N_1N_3dx}$$

$$= \phi_3\int_{x_1}^{x_2} N_1^1N_3^1dx + \phi_3\int_{x_2}^{x_3} N_1^2N_3^2dx + \phi_3\int_{x_3}^{x_4} N_1^3N_3^3dx$$

$$\boxed{\therefore \phi_3\int_{x_1}^{x_4} N_1N_3dx = 0}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \boxed{\phi_4\int_{x_1}^{x_4} N_1N_4dx} = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

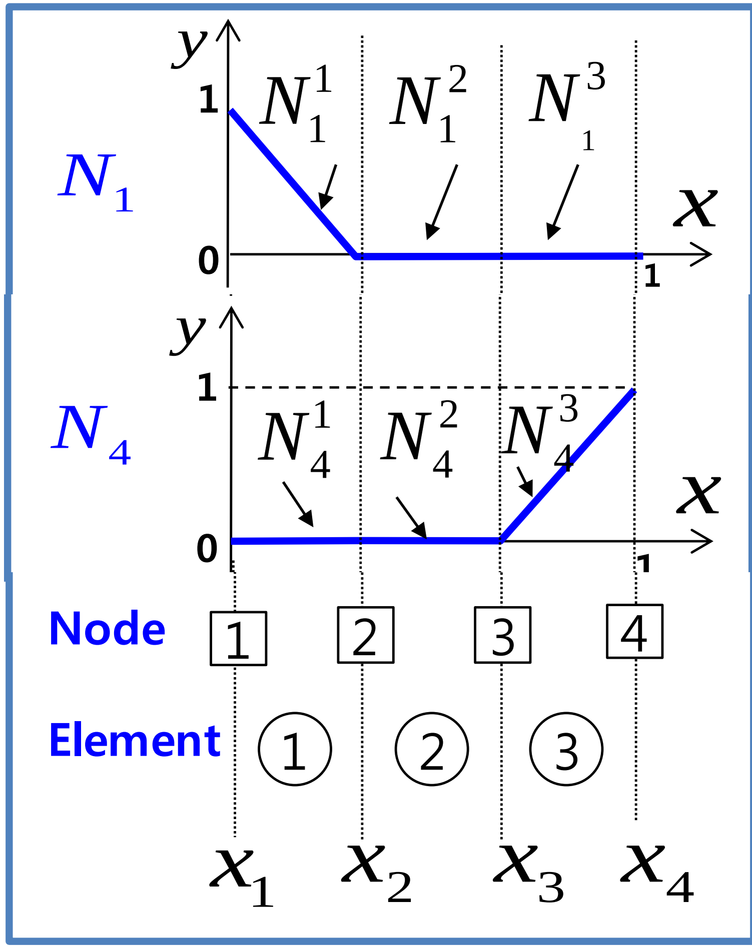
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_4\int_{x_1}^{x_4} N_1N_4dx}$$

$$= \phi_4\int_{x_1}^{x_2} N_1^1N_4^1dx + \phi_4\int_{x_2}^{x_3} N_1^2N_4^2dx + \phi_4\int_{x_3}^{x_4} N_1^3N_4^3dx$$

$$\boxed{\therefore \phi_4\int_{x_1}^{x_4} N_1N_4dx = 0}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\boxed{\phi_1\int_{x_1}^{x_4} N_2N_1dx} + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

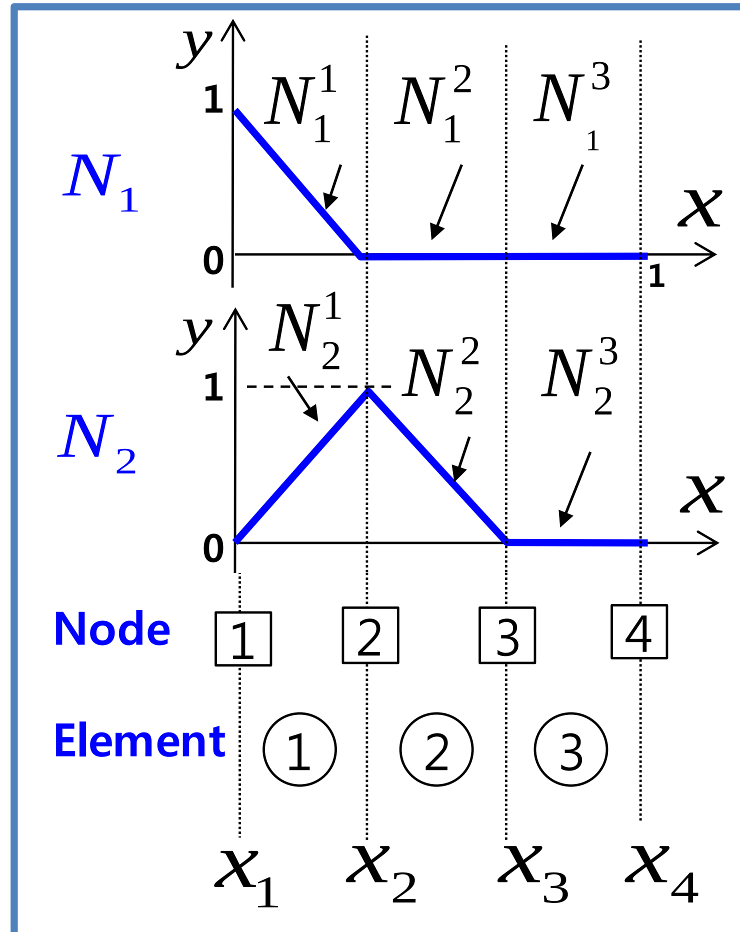
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_1\int_{x_1}^{x_4} N_2N_1dx}$$

$$= \phi_1\int_{x_1}^{x_2} N_2^1N_1^1dx + \phi_1\int_{x_2}^{x_3} N_2^2\cancel{N_1^2}dx + \phi_1\int_{x_3}^{x_4} \cancel{N_2^3}\,\cancel{N_1^3}dx$$

$$\boxed{\therefore \phi_1\int_{x_1}^{x_4} N_2N_1dx = \phi_1\int_{x_1}^{x_2} N_2^1N_1^1dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \boxed{\phi_2\int_{x_1}^{x_4} N_2N_2dx} + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

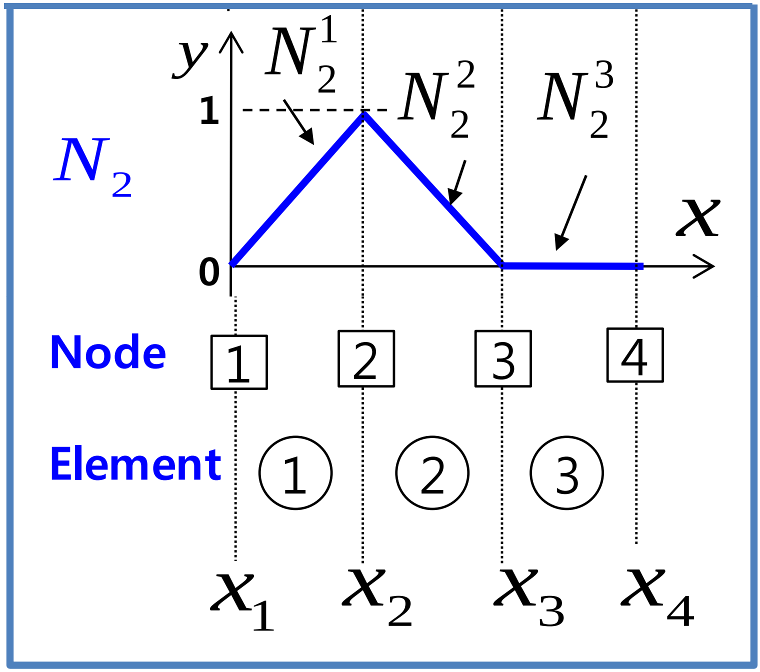
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_2\int_{x_1}^{x_4} N_2N_2dx}$$

$$= \phi_2\int_{x_1}^{x_2} N_2^1N_2^1dx + \phi_2\int_{x_2}^{x_3} N_2^2N_2^2dx + \phi_2\int_{x_3}^{x_4} N_2^3N_2^3dx$$

$$\boxed{\therefore \phi_2\int_{x_1}^{x_4} N_2N_2dx = \phi_2\int_{x_1}^{x_2} N_2^1N_2^1dx + \phi_2\int_{x_2}^{x_3} N_2^2N_2^2dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx+\phi_2\int_{x_1}^{x_4} N_1N_2dx+\phi_3\int_{x_1}^{x_4} N_1N_3dx+\phi_4\int_{x_1}^{x_4} N_1N_4dx=\int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx+\phi_2\int_{x_1}^{x_4} N_2N_2dx+\boxed{\phi_3\int_{x_1}^{x_4} N_2N_3dx}+\phi_4\int_{x_1}^{x_4} N_2N_4dx=\int_{x_1}^{x_4} N_2\phi\, dx$$

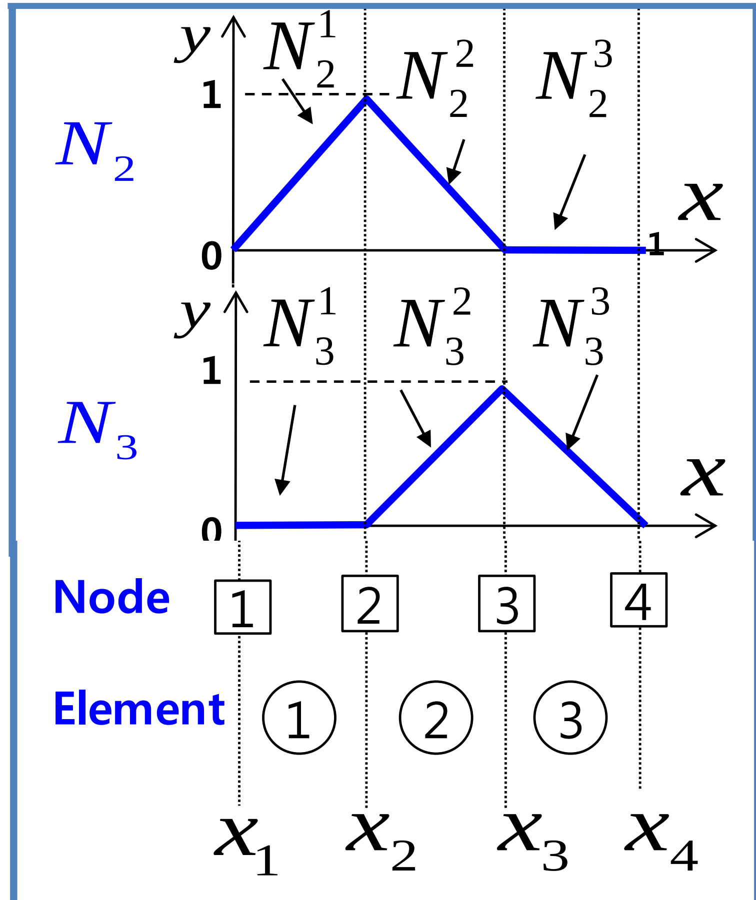
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx+\phi_2\int_{x_1}^{x_4} N_3N_2dx+\phi_3\int_{x_1}^{x_4} N_3N_3dx+\phi_4\int_{x_1}^{x_4} N_3N_4dx=\int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx+\phi_2\int_{x_1}^{x_4} N_4N_2dx+\phi_3\int_{x_1}^{x_4} N_4N_3dx+\phi_4\int_{x_1}^{x_4} N_4N_4dx=\int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_3\int_{x_1}^{x_4} N_2N_3dx}$$

$$=\phi_3\int_{x_1}^{x_2} N_2^1N_3^1dx+\phi_3\int_{x_2}^{x_3} N_2^2N_3^2dx+\phi_3\int_{x_3}^{x_4} N_2^3N_3^3dx$$

$$\boxed{\therefore \phi_3\int_{x_1}^{x_4} N_2N_3dx=\phi_3\int_{x_2}^{x_3} N_2^2N_3^2dx}$$



(derivation)

$$\phi_1 \int_{x_1}^{x_4} N_1 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_1 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_1 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_1 N_4 dx = \int_{x_1}^{x_4} N_1 \phi \, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_2 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_2 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_2 N_3 dx + \boxed{\phi_4 \int_{x_1}^{x_4} N_2 N_4 dx} = \int_{x_1}^{x_4} N_2 \phi \, dx$$

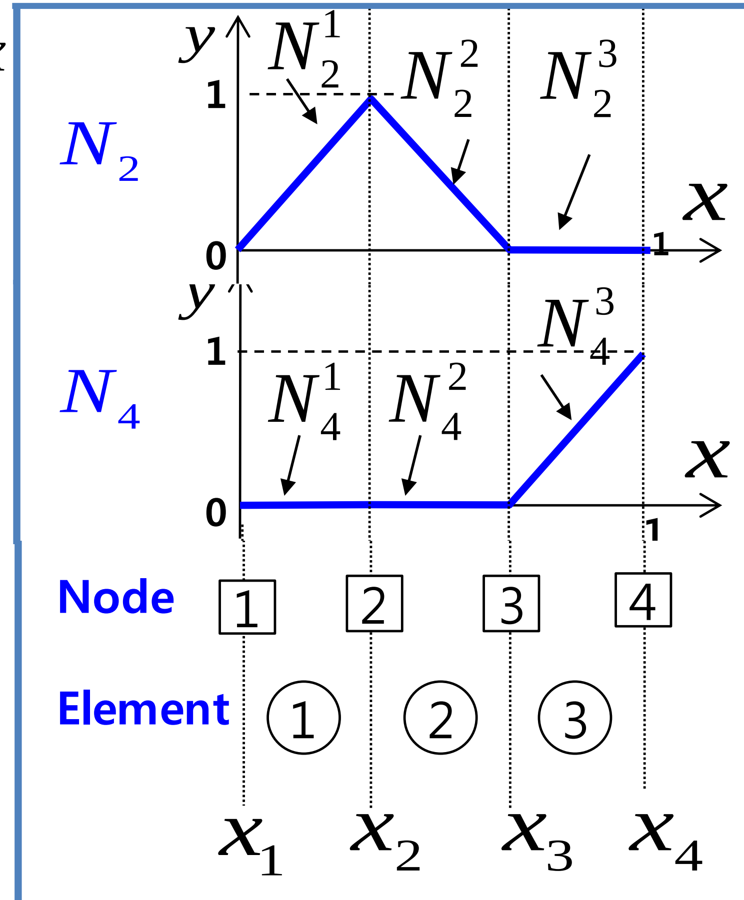
$$\phi_1 \int_{x_1}^{x_4} N_3 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_3 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_3 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_3 N_4 dx = \int_{x_1}^{x_4} N_3 \phi \, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_4 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_4 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_4 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_4 N_4 dx = \int_{x_1}^{x_4} N_4 \phi \, dx$$

$$\boxed{\phi_4 \int_{x_1}^{x_4} N_2 N_4 dx}$$

$$= \phi_4 \int_{x_1}^{x_2} N_2^1 N_4^1 dx + \phi_4 \int_{x_2}^{x_3} N_2^2 N_4^2 dx + \phi_4 \int_{x_3}^{x_4} N_2^3 N_4^3 dx$$

$$\boxed{\therefore \phi_4 \int_{x_1}^{x_4} N_2 N_4 dx = 0}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

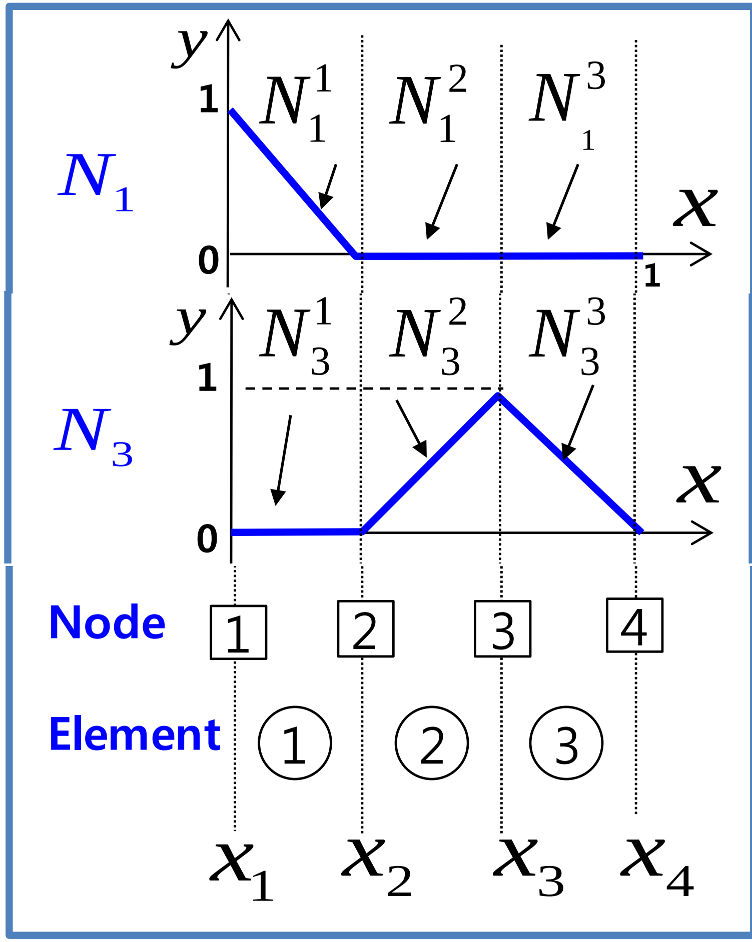
$$\boxed{\phi_1\int_{x_1}^{x_4} N_3N_1dx} + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_1\int_{x_1}^{x_4} N_3N_1dx}$$

$$= \phi_1\int_{x_1}^{x_2} N_3^1N_1^1dx + \phi_1\int_{x_2}^{x_3} N_3^2N_1^2dx + \phi_1\int_{x_3}^{x_4} N_3^3N_1^3dx$$

$$\boxed{\therefore \phi_1\int_{x_1}^{x_4} N_3N_1dx = 0}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

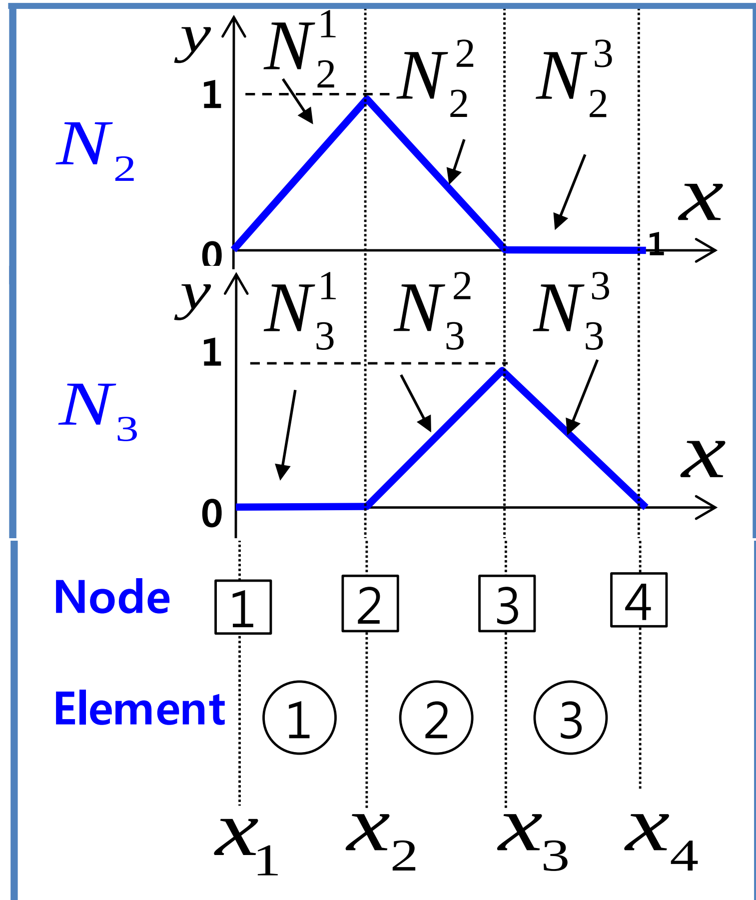
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \boxed{\phi_2\int_{x_1}^{x_4} N_3N_2dx} + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_2\int_{x_1}^{x_4} N_3N_2dx}$$

$$= \phi_2\int_{x_1}^{x_2} \cancel{N_3^1N_2^1dx} + \phi_2\int_{x_2}^{x_3} N_3^2N_2^2dx + \phi_2\int_{x_3}^{x_4} \cancel{N_3^3N_2^3dx}$$

$$\boxed{\therefore \phi_2\int_{x_1}^{x_4} N_3N_2dx = \phi_2\int_{x_2}^{x_3} N_3^2N_2^2dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

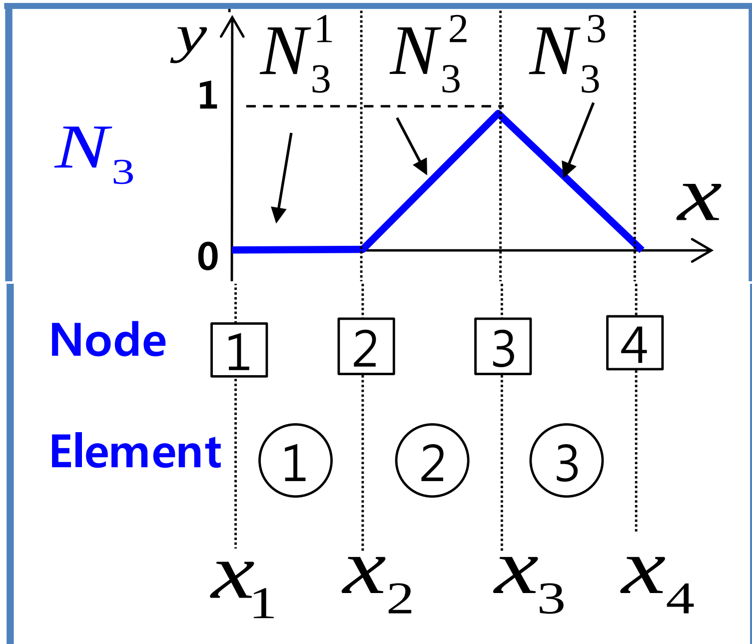
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \boxed{\phi_3\int_{x_1}^{x_4} N_3N_3dx} + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_3\int_{x_1}^{x_4} N_3N_3dx}$$

$$= \phi_3\int_{x_1}^{x_2} N_3^1N_3^1dx + \phi_3\int_{x_2}^{x_3} N_3^2N_3^2dx + \phi_3\int_{x_3}^{x_4} N_3^3N_3^3dx$$

$$\boxed{\therefore\ \phi_3\int_{x_1}^{x_4} N_3N_3dx = \phi_3\int_{x_2}^{x_3} N_3^2N_3^2dx + \phi_3\int_{x_3}^{x_4} N_3^3N_3^3dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx+\phi_2\int_{x_1}^{x_4} N_1N_2dx+\phi_3\int_{x_1}^{x_4} N_1N_3dx+\phi_4\int_{x_1}^{x_4} N_1N_4dx=\int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx+\phi_2\int_{x_1}^{x_4} N_2N_2dx+\phi_3\int_{x_1}^{x_4} N_2N_3dx+\phi_4\int_{x_1}^{x_4} N_2N_4dx=\int_{x_1}^{x_4} N_2\phi\, dx$$

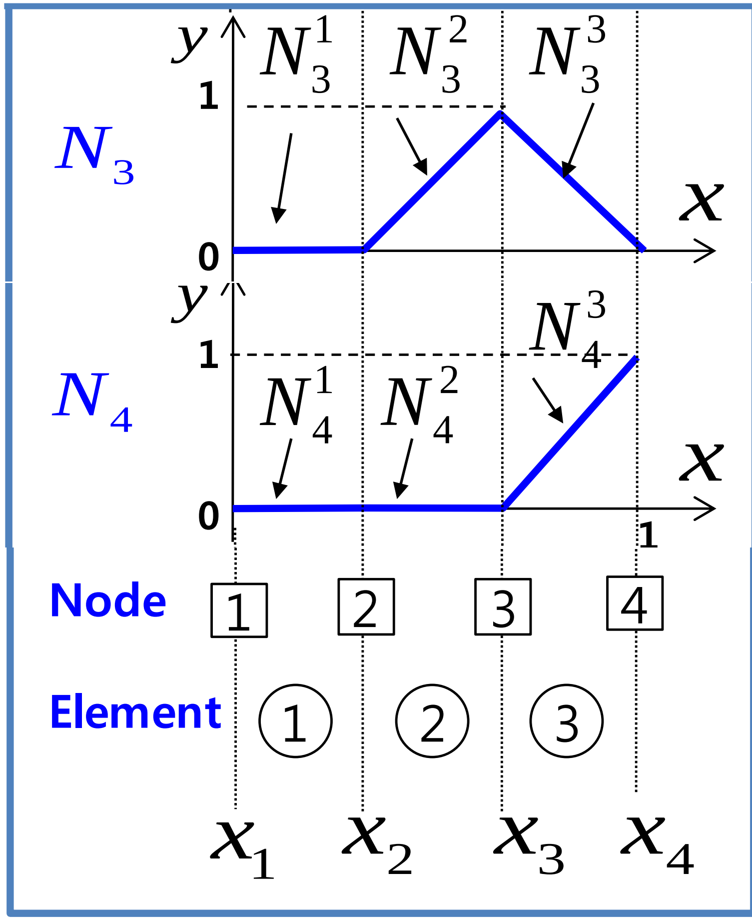
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx+\phi_2\int_{x_1}^{x_4} N_3N_2dx+\phi_3\int_{x_1}^{x_4} N_3N_3dx+\boxed{\phi_4\int_{x_1}^{x_4} N_3N_4dx}=\int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx+\phi_2\int_{x_1}^{x_4} N_4N_2dx+\phi_3\int_{x_1}^{x_4} N_4N_3dx+\phi_4\int_{x_1}^{x_4} N_4N_4dx=\int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_4\int_{x_1}^{x_4} N_3N_4dx}$$

$$=\phi_4\int_{x_1}^{x_2} N_3^1N_4^1dx+\phi_4\int_{x_2}^{x_3} N_3^2N_4^2dx+\phi_4\int_{x_3}^{x_4} N_3^3N_4^3dx$$

$$\boxed{\therefore \phi_4\int_{x_1}^{x_4} N_3N_4dx=\phi_4\int_{x_3}^{x_4} N_3^3N_4^3dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

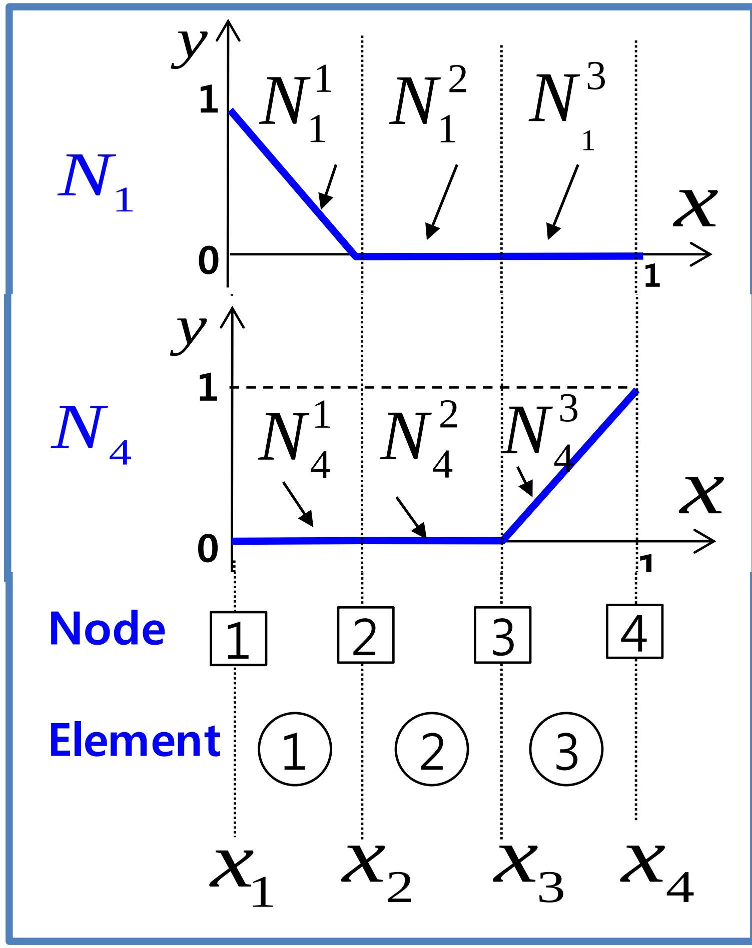
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\boxed{\phi_1\int_{x_1}^{x_4} N_4N_1dx} + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_1\int_{x_1}^{x_4} N_4N_1dx}$$

$$= \phi_1\int_{x_1}^{x_2} N_4^1N_1^1dx + \phi_1\int_{x_2}^{x_3} N_4^2N_1^2dx + \phi_1\int_{x_3}^{x_4} N_4^3N_1^3dx$$

$$\boxed{\therefore \phi_1\int_{x_1}^{x_4} N_4N_1dx = 0}$$



(derivation)

$$\phi_1 \int_{x_1}^{x_4} N_1 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_1 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_1 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_1 N_4 dx = \int_{x_1}^{x_4} N_1 \phi\, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_2 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_2 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_2 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_2 N_4 dx = \int_{x_1}^{x_4} N_2 \phi\, dx$$

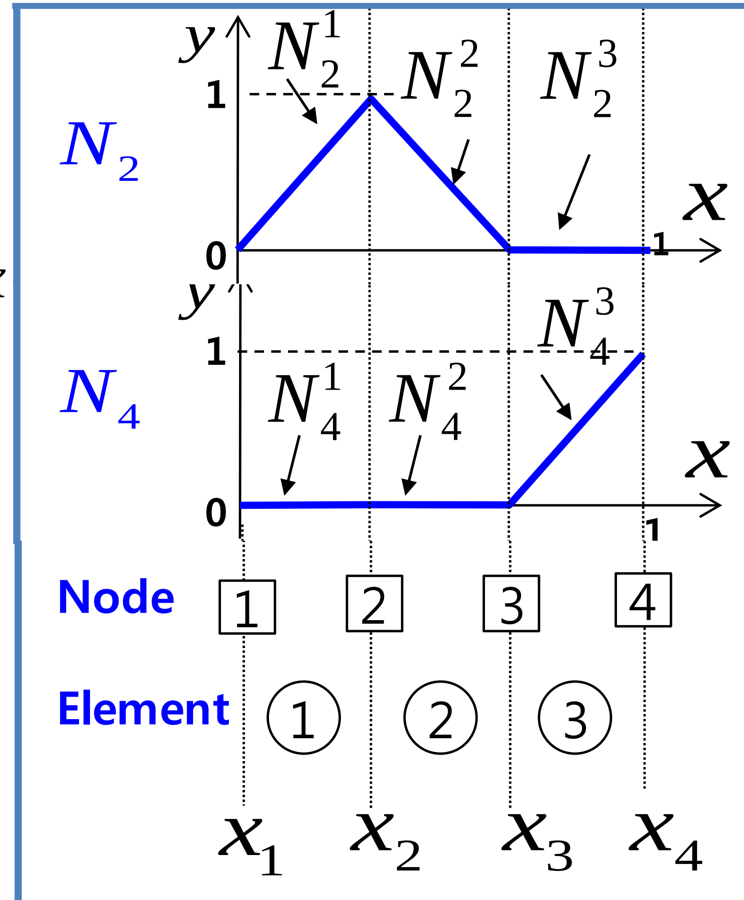
$$\phi_1 \int_{x_1}^{x_4} N_3 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_3 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_3 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_3 N_4 dx = \int_{x_1}^{x_4} N_3 \phi\, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_4 N_1 dx + \boxed{\phi_2 \int_{x_1}^{x_4} N_4 N_2 dx} + \phi_3 \int_{x_1}^{x_4} N_4 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_4 N_4 dx = \int_{x_1}^{x_4} N_4 \phi\, dx$$

$$\boxed{\phi_2 \int_{x_1}^{x_4} N_4 N_2 dx}$$

$$= \phi_2 \int_{x_1}^{x_2} N_4^1 N_2^1 dx + \phi_2 \int_{x_2}^{x_3} N_4^2 N_2^2 dx + \phi_2 \int_{x_3}^{x_4} N_4^3 N_2^3 dx$$

$$\boxed{\therefore \phi_2 \int_{x_1}^{x_4} N_4 N_2 dx = 0}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

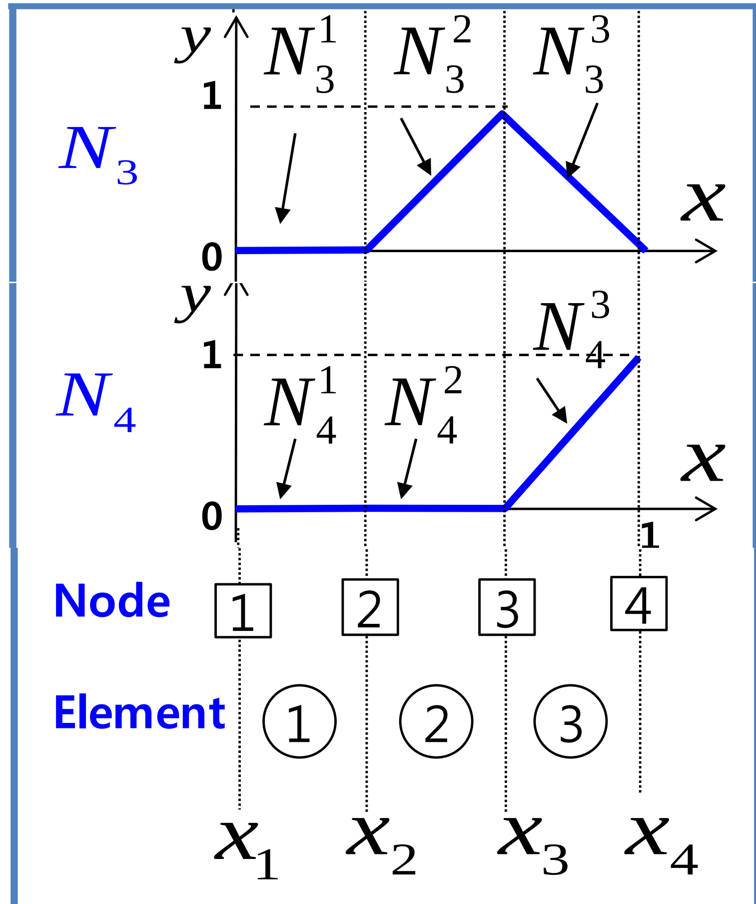
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \boxed{\phi_3\int_{x_1}^{x_4} N_4N_3dx} + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_3\int_{x_1}^{x_4} N_4N_3dx}$$

$$= \phi_3\int_{x_1}^{x_2} N_4^1N_3^1dx + \phi_3\int_{x_2}^{x_3} N_4^2N_3^2dx + \phi_3\int_{x_3}^{x_4} N_4^3N_3^3dx$$

$$\boxed{\therefore \phi_3\int_{x_1}^{x_4} N_4N_3dx = \phi_3\int_{x_3}^{x_4} N_4^3N_3^3dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

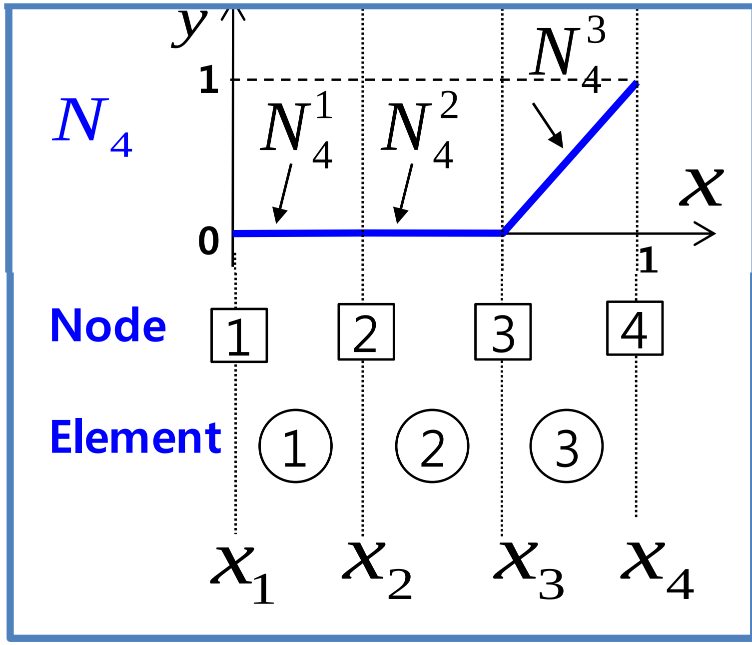
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \boxed{\phi_4\int_{x_1}^{x_4} N_4N_4dx} = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\phi_4\int_{x_1}^{x_4} N_4N_4dx}$$

$$= \phi_4\int_{x_1}^{x_2} N_4^1N_4^1dx + \phi_4\int_{x_2}^{x_3} N_4^2N_4^2dx + \phi_4\int_{x_3}^{x_4} N_4^3N_4^3dx$$

$$\boxed{\therefore \phi_4\int_{x_1}^{x_4} N_4N_4dx = \phi_4\int_{x_3}^{x_4} N_4^3N_4^3dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \boxed{\int_{x_1}^{x_4} N_1\phi\, dx}$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

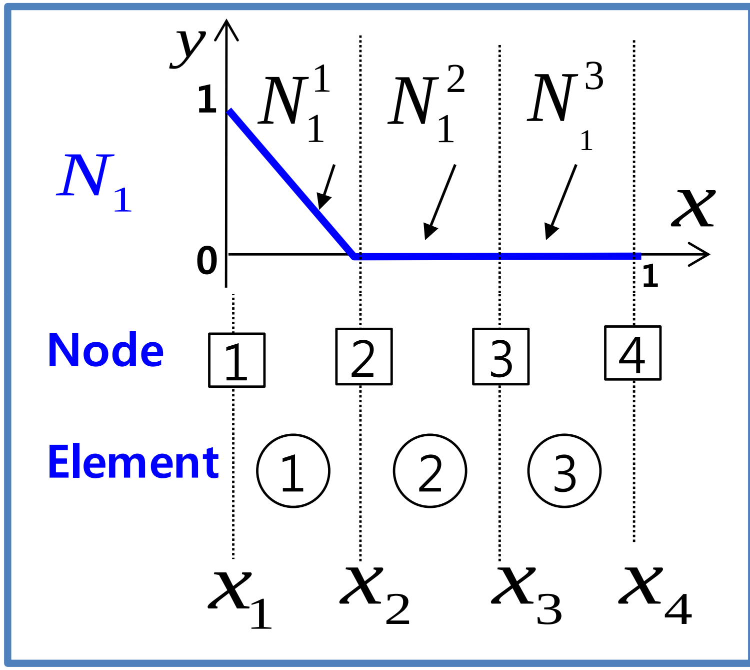
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

$$\boxed{\int_{x_1}^{x_4} N_1\,\phi\, dx}$$

$$= \int_{x_1}^{x_2} N_1^1\,\phi dx + \int_{x_2}^{x_3} \cancel{N_1^2}\,\phi dx + \int_{x_3}^{x_4} \cancel{N_1^3}\,\phi dx$$

$$\boxed{\therefore \int_{x_1}^{x_4} N_1\,\phi\, dx = \int_{x_1}^{x_2} N_1^1\,\phi dx}$$



(derivation)

$$\phi_1 \int_{x_1}^{x_4} N_1 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_1 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_1 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_1 N_4 dx = \int_{x_1}^{x_4} N_1 \phi \, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_2 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_2 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_2 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_2 N_4 dx = \boxed{\int_{x_1}^{x_4} N_2 \phi \, dx}$$

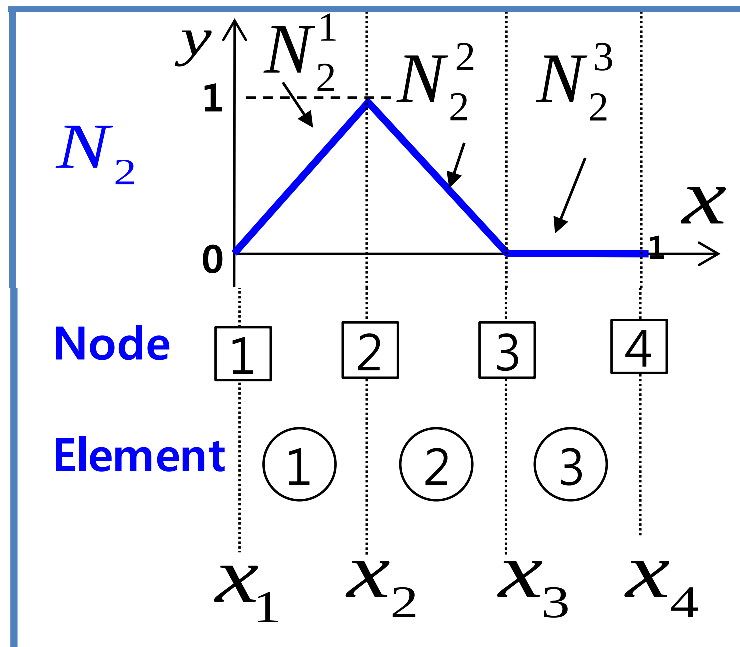
$$\phi_1 \int_{x_1}^{x_4} N_3 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_3 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_3 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_3 N_4 dx = \int_{x_1}^{x_4} N_3 \phi \, dx$$

$$\phi_1 \int_{x_1}^{x_4} N_4 N_1 dx + \phi_2 \int_{x_1}^{x_4} N_4 N_2 dx + \phi_3 \int_{x_1}^{x_4} N_4 N_3 dx + \phi_4 \int_{x_1}^{x_4} N_4 N_4 dx = \int_{x_1}^{x_4} N_4 \phi \, dx$$

$$\boxed{\int_{x_1}^{x_4} N_2 \, \phi \, dx}$$

$$= \int_{x_1}^{x_2} N_2^1 \, \phi dx + \int_{x_2}^{x_3} N_2^2 \, \phi dx + \int_{x_3}^{x_4} N_2^3 \, \phi dx$$

$$\boxed{\therefore \int_{x_1}^{x_4} N_1 \, \phi \, dx = \int_{x_1}^{x_2} N_2^1 \, \phi dx + \int_{x_2}^{x_3} N_2^2 \, \phi dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx+\phi_2\int_{x_1}^{x_4} N_1N_2dx+\phi_3\int_{x_1}^{x_4} N_1N_3dx+\phi_4\int_{x_1}^{x_4} N_1N_4dx=\int_{x_1}^{x_4} N_1\phi\,dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx+\phi_2\int_{x_1}^{x_4} N_2N_2dx+\phi_3\int_{x_1}^{x_4} N_2N_3dx+\phi_4\int_{x_1}^{x_4} N_2N_4dx=\int_{x_1}^{x_4} N_2\phi\,dx$$

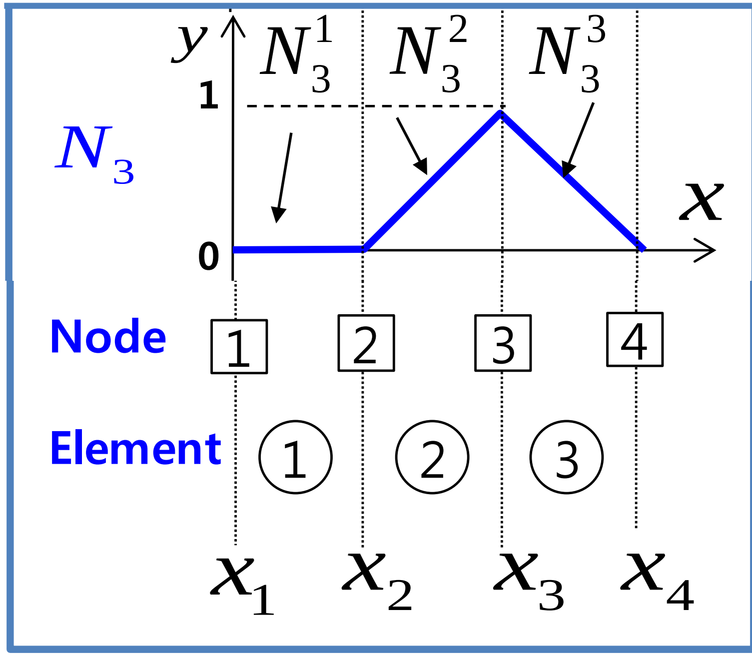
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx+\phi_2\int_{x_1}^{x_4} N_3N_2dx+\phi_3\int_{x_1}^{x_4} N_3N_3dx+\phi_4\int_{x_1}^{x_4} N_3N_4dx=\boxed{\int_{x_1}^{x_4} N_3\phi\,dx}$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx+\phi_2\int_{x_1}^{x_4} N_4N_2dx+\phi_3\int_{x_1}^{x_4} N_4N_3dx+\phi_4\int_{x_1}^{x_4} N_4N_4dx=\int_{x_1}^{x_4} N_4\phi\,dx$$

$$\boxed{\int_{x_1}^{x_4} N_3\,\phi\,dx}$$

$$=\int_{x_1}^{x_2} N_3^1\,\phi dx+\int_{x_2}^{x_3} N_3^2\,\phi dx+\int_{x_3}^{x_4} N_3^3\,\phi dx$$

$$\boxed{\therefore\int_{x_1}^{x_4} N_3\,\phi\,dx=\int_{x_2}^{x_3} N_3^2\,\phi dx+\int_{x_3}^{x_4} N_3^3\,\phi dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

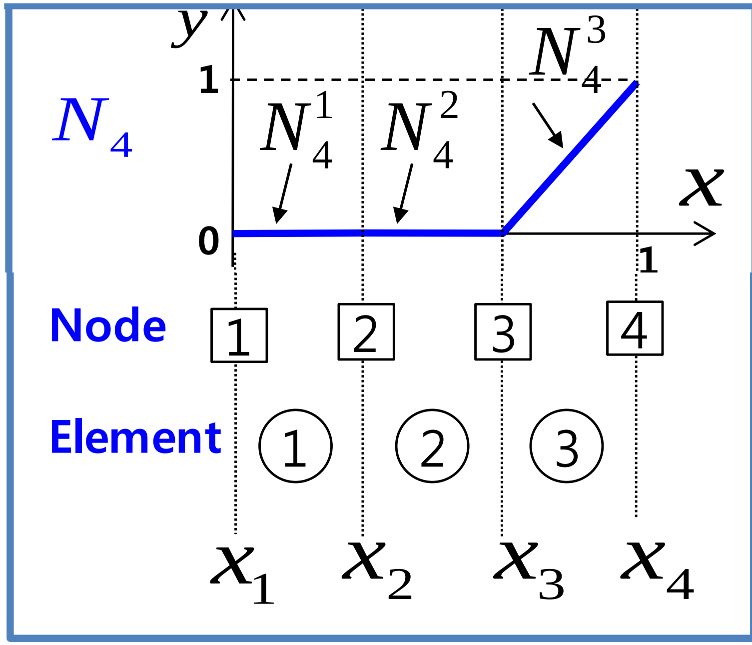
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \boxed{\int_{x_1}^{x_4} N_4\phi\, dx}$$

$$\boxed{\int_{x_1}^{x_4} N_4\,\phi\, dx}$$

$$= \int_{x_1}^{x_2} N_4^1\,\phi dx + \int_{x_2}^{x_3} N_4^2\,\phi dx + \int_{x_3}^{x_4} N_4^3\,\phi dx$$

$$\boxed{\therefore \int_{x_1}^{x_4} N_4\,\phi\, dx = \int_{x_3}^{x_4} N_4^3\,\phi dx}$$



(derivation)

$$\phi_1\int_{x_1}^{x_4} N_1N_1dx + \phi_2\int_{x_1}^{x_4} N_1N_2dx + \phi_3\int_{x_1}^{x_4} N_1N_3dx + \phi_4\int_{x_1}^{x_4} N_1N_4dx = \int_{x_1}^{x_4} N_1\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_2N_1dx + \phi_2\int_{x_1}^{x_4} N_2N_2dx + \phi_3\int_{x_1}^{x_4} N_2N_3dx + \phi_4\int_{x_1}^{x_4} N_2N_4dx = \int_{x_1}^{x_4} N_2\phi\, dx$$

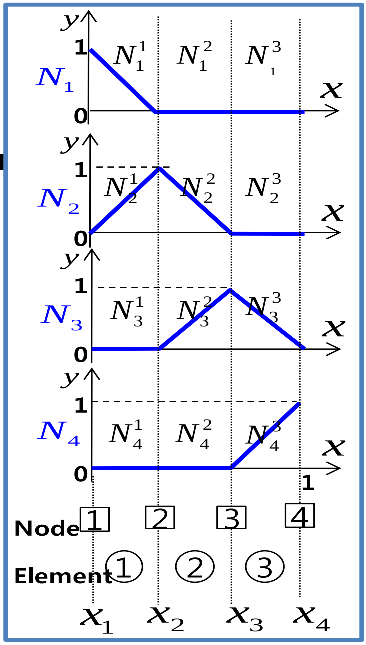
$$\phi_1\int_{x_1}^{x_4} N_3N_1dx + \phi_2\int_{x_1}^{x_4} N_3N_2dx + \phi_3\int_{x_1}^{x_4} N_3N_3dx + \phi_4\int_{x_1}^{x_4} N_3N_4dx = \int_{x_1}^{x_4} N_3\phi\, dx$$

$$\phi_1\int_{x_1}^{x_4} N_4N_1dx + \phi_2\int_{x_1}^{x_4} N_4N_2dx + \phi_3\int_{x_1}^{x_4} N_4N_3dx + \phi_4\int_{x_1}^{x_4} N_4N_4dx = \int_{x_1}^{x_4} N_4\phi\, dx$$

in matrix form

$$\begin{bmatrix} \int_{x_1}^{x_4} N_1N_1dx & \int_{x_1}^{x_4} N_1N_2dx & \int_{x_1}^{x_4} N_1N_3dx & \int_{x_1}^{x_4} N_1N_4dx \\ \int_{x_1}^{x_4} N_2N_1dx & \int_{x_1}^{x_4} N_2N_2dx & \int_{x_1}^{x_4} N_2N_3dx & \int_{x_1}^{x_4} N_2N_4dx \\ \int_{x_1}^{x_4} N_3N_1dx & \int_{x_1}^{x_4} N_3N_2dx & \int_{x_1}^{x_4} N_3N_3dx & \int_{x_1}^{x_4} N_3N_4dx \\ \int_{x_1}^{x_4} N_4N_1dx & \int_{x_1}^{x_4} N_4N_2dx & \int_{x_1}^{x_4} N_4N_3dx & \int_{x_1}^{x_4} N_4N_4dx \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} \int_{x_1}^{x_4} N_1\phi\, dx \\ \int_{x_1}^{x_4} N_2\phi\, dx \\ \int_{x_1}^{x_4} N_3\phi\, dx \\ \int_{x_1}^{x_4} N_4\phi\, dx \end{bmatrix}$$

(derivation)



$$\begin{bmatrix} \int_{x_1}^{x_4} N_1N_1dx & \int_{x_1}^{x_4} N_1N_2dx & \int_{x_1}^{x_4} N_1N_3dx & \int_{x_1}^{x_4} N_1N_4dx \\ \int_{x_1}^{x_4} N_2N_1dx & \int_{x_1}^{x_4} N_2N_2dx & \int_{x_1}^{x_4} N_2N_3dx & \int_{x_1}^{x_4} N_2N_4dx \\ \int_{x_1}^{x_4} N_3N_1dx & \int_{x_1}^{x_4} N_3N_2dx & \int_{x_1}^{x_4} N_3N_3dx & \int_{x_1}^{x_4} N_3N_4dx \\ \int_{x_1}^{x_4} N_4N_1dx & \int_{x_1}^{x_4} N_4N_2dx & \int_{x_1}^{x_4} N_4N_3dx & \int_{x_1}^{x_4} N_4N_4dx \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} \int_{x_1}^{x_4} N_1\phi\, dx \\ \int_{x_1}^{x_4} N_2\phi\, dx \\ \int_{x_1}^{x_4} N_3\phi\, dx \\ \int_{x_1}^{x_4} N_4\phi\, dx \end{bmatrix}$$

by using the integration on each element

"narrow banded (symmetric)"

$$\begin{bmatrix} \int_{x_1}^{x_2} N_1^1N_1^1dx & \int_{x_1}^{x_2} N_1^1N_2^1dx & 0 & 0 \\ \int_{x_1}^{x_2} N_2^1N_1^1dx & \int_{x_1}^{x_2} N_2^1N_2^1dx + \int_{x_2}^{x_3} N_2^2N_2^2dx & \int_{x_2}^{x_3} N_2^2N_3^2dx & 0 \\ 0 & \int_{x_2}^{x_3} N_3^2N_2^2dx & \int_{x_2}^{x_3} N_3^2N_3^2dx + \int_{x_3}^{x_4} N_3^3N_3^3dx & \int_{x_3}^{x_4} N_3^3N_4^3dx \\ 0 & 0 & \int_{x_3}^{x_4} N_4^3N_3^3dx & \int_{x_3}^{x_4} N_4^3N_4^3dx \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} \int_{x_1}^{x_2} N_1^1\phi dx \\ \int_{x_1}^{x_2} N_2^1\phi dx + \int_{x_2}^{x_3} N_2^2\phi dx \\ \int_{x_2}^{x_3} N_3^2\phi dx + \int_{x_3}^{x_4} N_3^3\phi dx \\ \int_{x_3}^{x_4} N_4^3\phi dx \end{bmatrix}$$

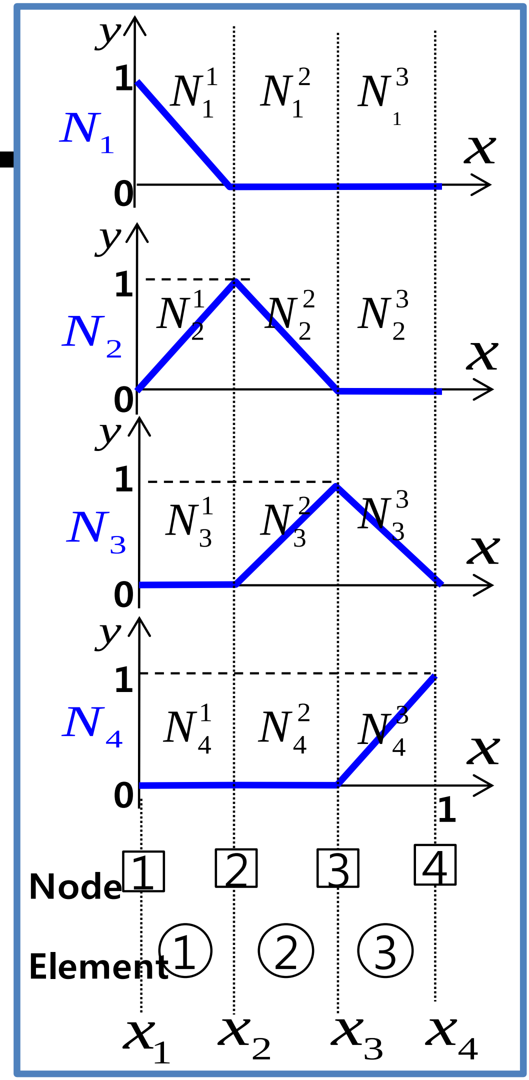
$$\mathbf{K}\boldsymbol{\phi} = \mathbf{f}$$

(derivation)

$$\mathbf{K}\boldsymbol{\phi}=\mathbf{f}$$

$$\begin{cases} \mathbf{K}=[K_{lm}], \quad K_{lm}=\int_{x_1}^{x_4} N_l N_m dx=\sum_{e=1}^{3}\left[\int_{x_e}^{x_{e+1}} N_l^e N_m^e dx\right] \\ \boldsymbol{\phi}=[\phi_m]^T \\ \mathbf{f}=[f_l], \quad f_l=\int_{x_1}^{x_4} N_l \phi\, dx=\sum_{e=1}^{3}\left[\int_{x_e}^{x_{e+1}} N_l^e \phi\, dx\right] \end{cases}$$

$l, m=1,2,3,4$



notation of the indices

N_l^e: e — No. of Element; l — No. Trial Function

(derivation)

notation

N_l^{e} — e: No. of Element; l: No. Trial Function

in general,

$$\mathbf{K}\boldsymbol{\phi} = \mathbf{f}$$

$$\mathbf{K} = [K_{lm}],$$

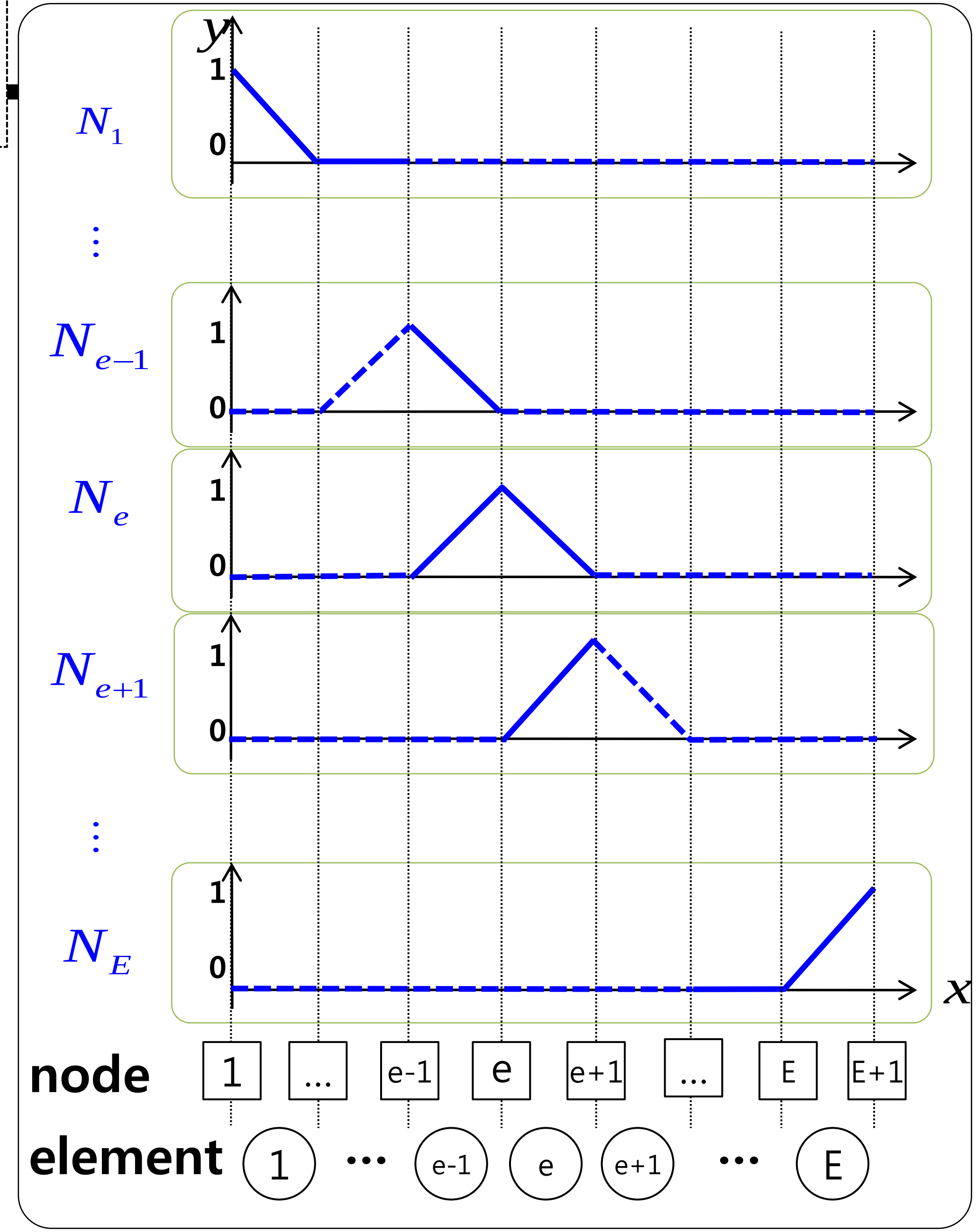
$$K_{lm} = \int_{x_1}^{x_{E+1}} N_l N_m dx = \sum_{e=1}^{E}\left[\int_{x_e}^{x_{e+1}} N_l^e N_m^e dx\right]$$

$$\boldsymbol{\phi} = [\phi_m]^T$$

$$\mathbf{f} = [f_l],$$

$$f_l = \int_{x_1}^{x_{E+1}} N_l \phi\, dx = \sum_{e=1}^{E}\left[\int_{x_e}^{x_{e+1}} N_l^e \phi\, dx\right]$$

$$l, m = 1, \ldots, E+1$$



C.f. B-Spline Basis Functions

Examples

Ex2.1 Let $U = \{u_0 = 0, u_1 = 0, u_2 = 0, u_3 = 1, u_4 = 1, u_5 = 1\}$ and $p = 2$. We now compute the B-spline basis functions of degrees 0, 1, and 2

$$N_{0,0} = N_{1,0} = 0 \qquad -\infty < u < \infty$$

$$N_{2,0} = \begin{cases} 1 & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$N_{3,0} = N_{4,0} = 0 \qquad -\infty < u < \infty$$

$$N_{0,1} = \frac{u-0}{0-0}N_{0,0} + \frac{0-u}{0-0}N_{1,0} = 0 \qquad -\infty < u < \infty$$

$$N_{1,1} = \frac{u-0}{0-0}N_{1,0} + \frac{1-u}{1-0}N_{2,0} = \begin{cases} 1-u & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$N_{2,1} = \frac{u-0}{1-0}N_{2,0} + \frac{1-u}{1-1}N_{3,0} = \begin{cases} u & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$N_{3,1} = \frac{u-1}{1-1}N_{3,0} + \frac{1-u}{1-1}N_{4,0} = 0 \qquad -\infty < u < \infty$$

$$N_{0,2} = \frac{u-0}{0-0}N_{0,1} + \frac{1-u}{1-0}N_{1,1} = \begin{cases} (1-u)^2 & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$N_{1,2} = \frac{u-0}{1-0}N_{1,1} + \frac{1-u}{1-0}N_{2,1} = \begin{cases} 2u(1-u) & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$N_{2,2} = \frac{u-0}{1-0}N_{2,1} + \frac{1-u}{1-1}N_{3,1} = \begin{cases} u^2 & 0 \le u < 1 \\ 0 & \text{otherwise} \end{cases}$$

Note that the $N_{i,2}$, restricted to the interval $u \in [0,1]$, are the quadratic Bernstein polynomials (Section 1.3 and Figure 1.13b). For this reason, the B-spline representation with a knot vector of the form

$$U = \{\underbrace{0,\ldots,0}_{p+1},\underbrace{1,\ldots,1}_{p+1}\}$$

is a generalization of the Bézier representation.

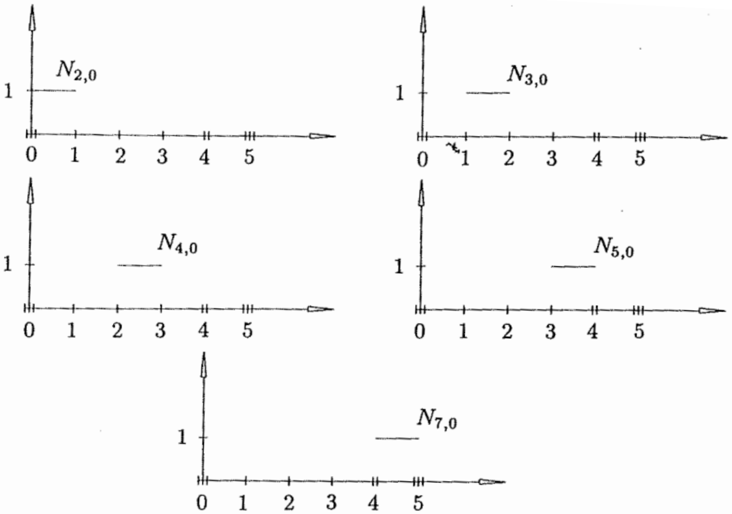


Figure 2.4. The nonzero zeroth-degree basis functions, $U = \{0,0,0,1,2,3,4,4,5,5,5\}$.

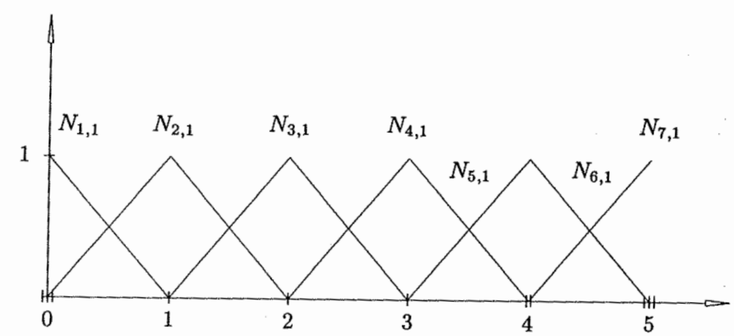


Figure 2.5. The nonzero first-degree basis functions, $U = \{0,0,0,1,2,3,4,4,5,5,5\}$.

Approximation to Solutions of Differential Equations and Continuity Requirements

Approximation to Solutions of Differential Equations

Recall,

Approximation to the Solutions of Differential Equations and the Use of Trial Function

Differential Equation

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

Boundary Conditions

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

	Differential Equation	Boundary Condition
Case 1	Not Satisfied	Satisfied
Case 2	Not Satisfied	Not Satisfied
Case 3	Satisfied	Not Satisfied

the coefficients of the trial functions will be determine to satisfy them by "weighted residual" process

we shall obtain out(?) discrete approximation equations in weighted residual form as

$$\int_{\Omega} W_l \, R_{\Omega} \, d\Omega + \int_{\Gamma} \bar{W}_l \, R_{\Gamma} \, d\Gamma = 0$$

with $R_{\Omega} = \mathscr{L}\hat{\phi} + p, \;\; R_{\Gamma} = \mathscr{M}\hat{\phi} + r$

Now if integrals of the weighted residual type are evaluated, it is desirable to avoid infinite value.

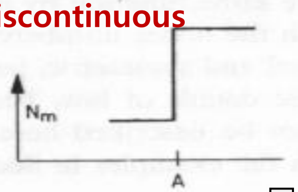
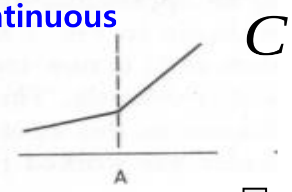
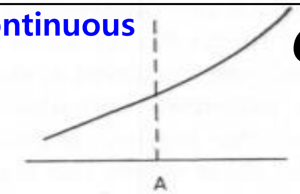
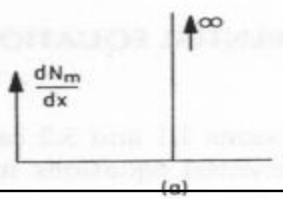
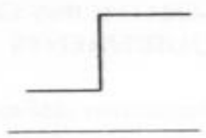
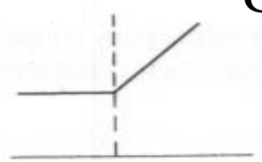
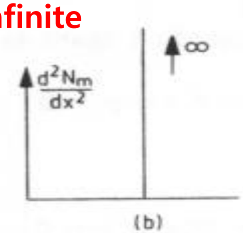
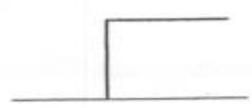
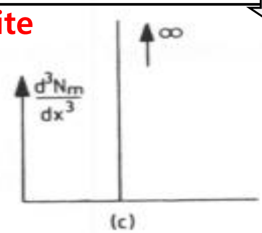
Continuity requirements for the trial functions

If the integrals contain derivatives of order S (i.e., the operators $\mathscr{L}$ or $\mathscr{M}$ contain such derivatives) , we must ensure that derivatives of the order $S-1$ are continuous in the trial functions N_m used in the approximation.

"C^{s-1} continuity"

Continuity requirements

We consider the behavior of three types of one dimensional Trial functions N_m near a junction A of two elements

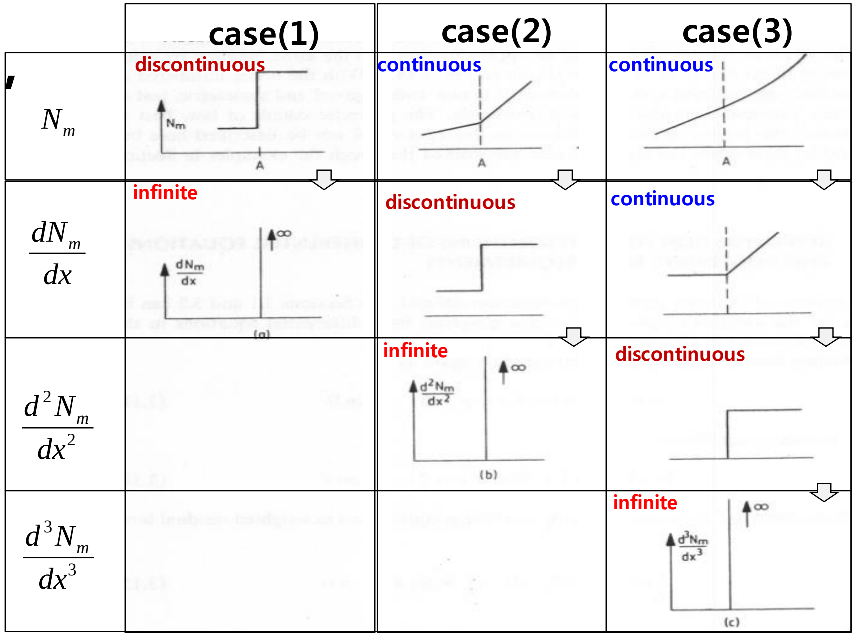
	case(1)	case(2)	case(3)
N_m	discontinuous 	continuous C^0 	continuous C^1 
$\dfrac{dN_m}{dx}$	infinite 	discontinuous 	continuous C^0 
$\dfrac{d^2N_m}{dx^2}$		infinite 	discontinuous 
$\dfrac{d^3N_m}{dx^3}$	If integrals of the weighted residual type are evaluated, it is desirable to avoid infinite value		infinite 

Continuity Requirements

"C^{s-1} continuity"

If the first derivatives occur in $\mathscr{L}$ or $\mathscr{M}$, that is $S=1$

then, C^0 continuity is necessary
→trial function such as case (2) is required



If the second derivatives occur in $\mathscr{L}$ or $\mathscr{M}$, that is $S=2$

then, C^1 continuity is necessary
→trial function such as case (3) is required

The continuity requirements are also applicable to the weighting function W_l

Weak Formulation and the Galerkin Method

Example 3.1

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ in $0 \le x \le 1$

Boundary Condition $\phi = 0$ at $x = 0$ and $\phi = 1$ at $x = 1$

Governing equation $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$

$$\frac{d^2\phi}{dx^2} = \phi \rightarrow \frac{d^2\phi}{dx^2} - \phi = 0 \rightarrow \boxed{A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \;\; \Omega}$$

We shall now attempt to solve this problem by the finite element method. And associate a piecewise linear global shape function N_m .

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \;\; 0 \le x \le 1$$,where E is the number of the elements

Weak Formulation and the Galerkin Method

Example 3.1

$$A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \;\; 0 < x < 1$$

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \quad 0 \le x \le 1$$

,where E is the number of the elements

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

The weighted residual form:

$$\int_0^1 W_l \mathbf{R}_\Omega dx = 0, \; l = 1, 2, \ldots, E+1$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx = 0, \; l = 1, 2, \ldots, E+1$$

The weighted residual form:

$$\int_0^1 W_l \left(\boxed{\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi}} \right) dx = 0,\ l = 1, 2, \ldots, E+1$$

Derivatives of order two → C^1 continuity is necessary for trial function

In its present form, this statement requires continuity of first derivatives of the trial functions if infinite values are to be avoided.

Integration by parts relaxes this requirement on the trial functions and leads to a weak form of the weighted residual statement.

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \quad 0 \le x \le 1$$

The weighted residual form:

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx = 0, \; l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$\int_0^1 W_l \frac{d^2\hat{\phi}}{dx^2} dx - \int_0^1 W_l \hat{\phi} dx = 0$$

$\downarrow$ Integration by parts

$$-\int_0^1 \boxed{\frac{dW_l}{dx} \frac{d\hat{\phi}}{dx}} dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx = 0$$

Derivatives of order one

Boundary Condition

$\phi = 0$ at $x = 0$ and $\phi = 1$ at $x = 1$

Now it is apparent that only C^0 continuity of $\hat{\phi}$ (and hence of N_m) and W_l is demanded

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \quad 0 \le x \le 1$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx} \frac{d\sum_{m=1}^{E+1}\phi_m N_m}{dx} dx - \int_0^1 W_l \sum_{m=1}^{E+1} \phi_m N_m dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

Boundary Condition

$\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \quad 0 \le x \le 1$$

$$-\int_0^1 \frac{dW_l}{dx} \frac{d\sum_{m=1}^{E+1}\phi_m N_m}{dx} dx - \int_0^1 W_l \sum_{m=1}^{E+1} \phi_m N_m dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

Boundary Condition

$\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

$\downarrow$ **Galerkin methods** $W_l = N_l$

$$-\int_0^1 \frac{dN_l}{dx} \frac{d\sum_{m=1}^{E+1}\phi_m N_m}{dx} dx - \int_0^1 N_l \sum_{m=1}^{E+1} \phi_m N_m dx + \left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

$\downarrow$

$$-\sum_{m=1}^{E+1} \phi_m \int_0^1 \frac{dN_l}{dx} \frac{dN_m}{dx} dx - \sum_{m=1}^{E+1} \phi_m \int_0^1 N_l N_m dx + \left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

$$\phi \approx \hat{\phi} = \sum_{m=1}^{E+1} \phi_m N_m, \quad 0 \le x \le 1$$

$$-\sum_{m=1}^{E+1} \phi_m \int_0^1 \frac{dN_l}{dx}\frac{dN_m}{dx}dx - \sum_{m=1}^{E+1} \phi_m \int_0^1 N_l N_m dx + \left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

Boundary Condition

$\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

$$\downarrow \ m = 1, 2, 3, \ldots, E+1$$

$$-\phi_1 \int_0^1 \frac{dN_l}{dx}\frac{dN_1}{dx}dx - \phi_2 \int_0^1 \frac{dN_l}{dx}\frac{dN_2}{dx}dx - \phi_3 \int_0^1 \frac{dN_l}{dx}\frac{dN_3}{dt}dx \ldots - \phi_{E+1} \int_0^1 \frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx$$

$$-\phi_1 \int_0^1 N_l N_1 dx - \phi_2 \int_0^1 N_l N_2 dx - \phi_3 \int_0^1 N_l N_3 dx \ldots - \phi_{E+1} \int_0^1 N_l N_{E+1} dx + \left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1 = 0$$

$$\downarrow$$

$$-\phi_1 \int_0^1 \left(\frac{dN_l}{dx}\frac{dN_1}{dx} + N_l N_1 \right) dx - \phi_2 \int_0^1 \left(\frac{dN_l}{dx}\frac{dN_2}{dx} + N_l N_2 \right) dx - \phi_3 \int_0^1 \left(\frac{dN_l}{dx}\frac{dN_3}{dx} + N_l N_3 \right) dx$$

$$\ldots - \phi_{E+1} \int_0^1 \left(\frac{dN_l}{dx}\frac{dN_{E+1}}{dx} + N_l N_{E+1} \right) dx = -\left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1$$

$$-\phi_1\int_0^1\left(\frac{dN_l}{dx}\frac{dN_1}{dx}+N_lN_1\right)dx-\phi_2\int_0^1\left(\frac{dN_l}{dx}\frac{dN_2}{dx}+N_lN_2\right)dx-\phi_3\int_0^1\left(\frac{dN_l}{dx}\frac{dN_3}{dx}+N_lN_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_l}{dx}\frac{dN_{E+1}}{dx}+N_lN_{E+1}\right)dx=-\left[N_l\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$\downarrow\ l=1,2,3,\ldots,E+1$$

$$-\phi_1\int_0^1\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_1}{dx}\frac{dN_2}{dt}+N_1N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_1}{dx}\frac{dN_{E+1}}{dx}+N_1N_{E+1}\right)dx=-\left[N_1\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$-\phi_1\int_0^1\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_2}{dx}\frac{dN_{E+1}}{dx}+N_2N_{E+1}\right)dx=-\left[N_2\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$-\phi_1\int_0^1\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_3}{dx}\frac{dN_{E+1}}{dx}+N_3N_{E+1}\right)dx=-\left[N_3\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$\vdots$$

$$-\phi_1\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_1}{dx}+N_{E+1}N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_2}{dx}+N_{E+1}N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_3}{dx}+N_{E+1}N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}+N_{E+1}N_{E+1}\right)dx=-\left[N_{E+1}\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$-\phi_1\int_0^1\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_2}{dx}\frac{dN_{E+1}}{dx}+N_2N_{E+1}\right)dx=-\left[N_2\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$-\phi_1\int_0^1\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_3}{dx}\frac{dN_{E+1}}{dx}+N_3N_{E+1}\right)dx=-\left[N_3\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$\vdots$$

$$-\phi_1\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_1}{dx}+N_{E+1}N_1\right)dx-\phi_2\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_2}{dx}+N_{E+1}N_2\right)dx-\phi_3\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_3}{dx}+N_{E+1}N_3\right)dx\ldots-\phi_{E+1}\int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}+N_{E+1}N_{E+1}\right)dx=-\left[N_{E+1}\frac{d\hat{\phi}}{dx}\right]_0^1$$

$$\downarrow$$

$$\boxed{\begin{bmatrix}\int_0^1\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_0^1\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_0^1\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \ldots & \int_0^1\left(\frac{dN_1}{dx}\frac{dN_{E+1}}{dx}+N_1N_{E+1}\right)dx \\ \int_0^1\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_0^1\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_0^1\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \ldots & \int_0^1\left(\frac{dN_2}{dx}\frac{dN_{E+1}}{dx}+N_3N_{E+1}\right)dx \\ \int_0^1\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_0^1\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_0^1\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \ldots & \int_0^1\left(\frac{dN_2}{dx}\frac{dN_{E+1}}{dx}+N_3N_{E+1}\right)dx \\ \vdots & & & & \\ \int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_1}{dx}+N_{E+1}N_1\right)dx & \int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_2}{dx}+N_{E+1}N_2\right)dx & \int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_3}{dx}+N_{E+1}N_3\right)dx & \ldots & \int_0^1\left(\frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}+N_{E+1}N_{E+1}\right)dx\end{bmatrix}}\ \boxed{\begin{bmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \vdots\\ \phi_{E+1}\end{bmatrix}}=\boxed{\begin{bmatrix}\left[N_1\frac{d\hat{\phi}}{dx}\right]_0^1\\ \left[N_2\frac{d\hat{\phi}}{dx}\right]_0^1\\ \left[N_3\frac{d\hat{\phi}}{dx}\right]_0^1\\ \vdots\\ \left[N_{E+1}\frac{d\hat{\phi}}{dx}\right]_0^1\end{bmatrix}}$$

$$\mathbf{K}\,\Phi=\mathbf{f}$$

$$\mathbf{K}\,\Phi = \mathbf{f}$$

, where $K_{lm} = \int_0^1 \left(\frac{dN_l}{dx}\frac{dN_m}{dx} + N_l N_m \right) dx \qquad 1 \le l, m \le E+1$

$$f_l = \left[N_l \frac{d\hat{\phi}}{dx} \right]_0^1$$

The definite integrals occurring in the approximating equations can be obtained simply by summing the contributions from each elements

$$\mathbf{K}\,\Phi = \mathbf{f}$$

, where $K_{lm} = \sum_{e=1}^{M} K_{lm}^e \qquad 1 \le l, m \le E+1$

$$= \sum_{e=1}^{M} \int_e^{e+1} \left(\frac{dN_l^e}{dx}\frac{dN_m^e}{dx} + N_l^e N_m^e \right) dx$$

$$f_l = \sum_{e=1}^{M} f_l^e = \sum_{e=1}^{M} \left[N_l^e \frac{d\hat{\phi}}{dx} \right]_0^1$$

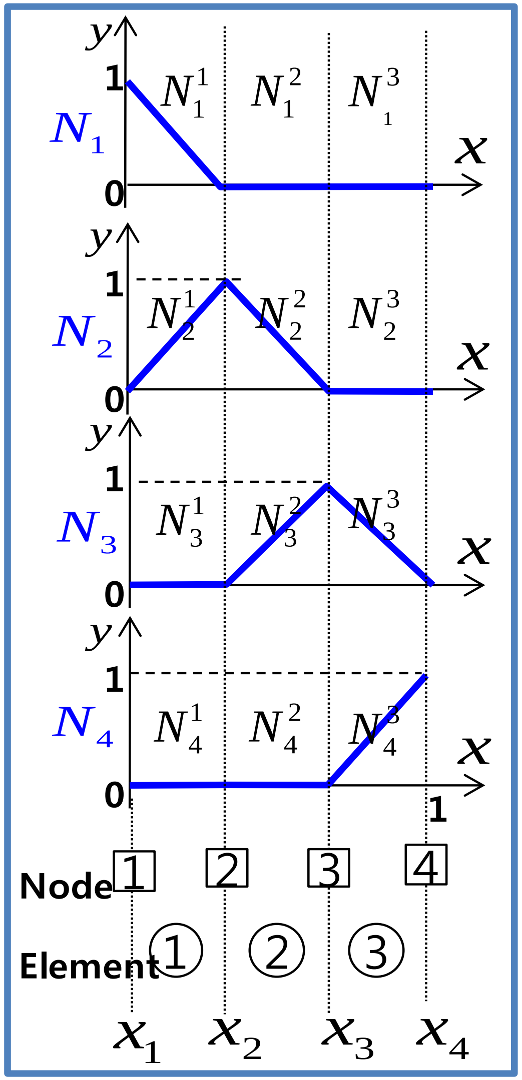
Suppose that the number of the elements "E" is 3.

$$\mathbf{K\Phi} = \mathbf{f}$$

$$\mathbf{K} = \begin{bmatrix}
\int_0^1 \left(\frac{dN_1}{dx}\frac{dN_1}{dx} + N_1N_1 \right) dx & \int_0^1 \left(\frac{dN_1}{dx}\frac{dN_2}{dx} + N_1N_2 \right) dx & \int_0^1 \left(\frac{dN_1}{dx}\frac{dN_3}{dx} + N_1N_3 \right) dx & \int_0^1 \left(\frac{dN_1}{dx}\frac{dN_4}{dx} + N_1N_4 \right) dx \\
\int_0^1 \left(\frac{dN_2}{dx}\frac{dN_1}{dx} + N_2N_1 \right) dx & \int_0^1 \left(\frac{dN_2}{dx}\frac{dN_2}{dx} + N_2N_2 \right) dx & \int_0^1 \left(\frac{dN_2}{dx}\frac{dN_3}{dx} + N_2N_3 \right) dx & \int_0^1 \left(\frac{dN_2}{dx}\frac{dN_4}{dx} + N_2N_4 \right) dx \\
\int_0^1 \left(\frac{dN_3}{dx}\frac{dN_1}{dx} + N_3N_1 \right) dx & \int_0^1 \left(\frac{dN_3}{dx}\frac{dN_2}{dx} + N_3N_2 \right) dx & \int_0^1 \left(\frac{dN_3}{dx}\frac{dN_3}{dx} + N_3N_3 \right) dx & \int_0^1 \left(\frac{dN_2}{dx}\frac{dN_4}{dx} + N_3N_4 \right) dx \\
\int_0^1 \left(\frac{dN_4}{dx}\frac{dN_1}{dx} + N_4N_1 \right) dx & \int_0^1 \left(\frac{dN_4}{dx}\frac{dN_2}{dx} + N_4N_2 \right) dx & \int_0^1 \left(\frac{dN_4}{dx}\frac{dN_3}{dx} + N_4N_3 \right) dx & \int_0^1 \left(\frac{dN_4}{dx}\frac{dN_4}{dx} + N_4N_4 \right) dx
\end{bmatrix}$$

$$\Phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix}$$

$$\mathbf{f} = \left[\left[N_1 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_2 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_3 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_4 \frac{d\hat{\phi}}{dx} \right]_0^1 \right]^T$$



$$\mathbf{K}=\begin{bmatrix}
\int_{x_1}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_1}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_1}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_1}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_1}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \boxed{\int_{x_1}^{x_4}}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix} \longrightarrow \boxed{\int_{x_1}^{x_4}}$$

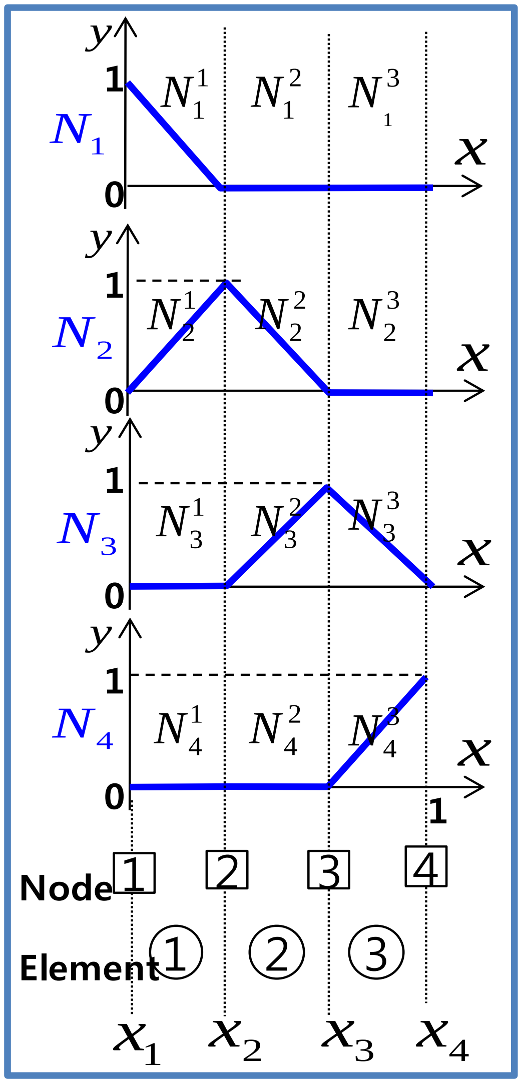
The definite integrals occurring in the approximating equations can be obtained simply by summing the contributions from each elements

$$\mathbf{K}=\mathbf{K}^1+\mathbf{K}^2+\mathbf{K}^3$$

$$\mathbf{K}^1=\begin{bmatrix}
\int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \boxed{\int_{x_1}^{x_2}}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix} \longrightarrow \boxed{\int_{x_1}^{x_2}}$$

$$\mathbf{K}^2=\begin{bmatrix}
\int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \boxed{\int_{x_2}^{x_3}}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix} \longrightarrow \boxed{\int_{x_2}^{x_3}}$$

$$\mathbf{K}^3=\begin{bmatrix}
\int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \boxed{\int_{x_3}^{x_4}}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix} \longrightarrow \boxed{\int_{x_3}^{x_4}}$$



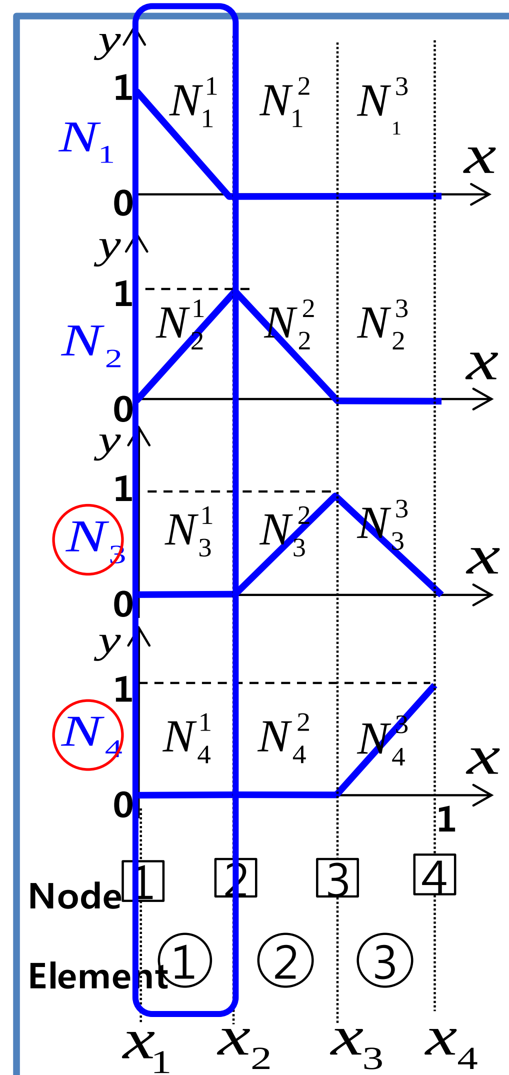
$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$\mathbf{K}^1 =$$

$$\begin{bmatrix}
\int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix}$$

The value of the trial function N_3, N_4 are zero, in element 1

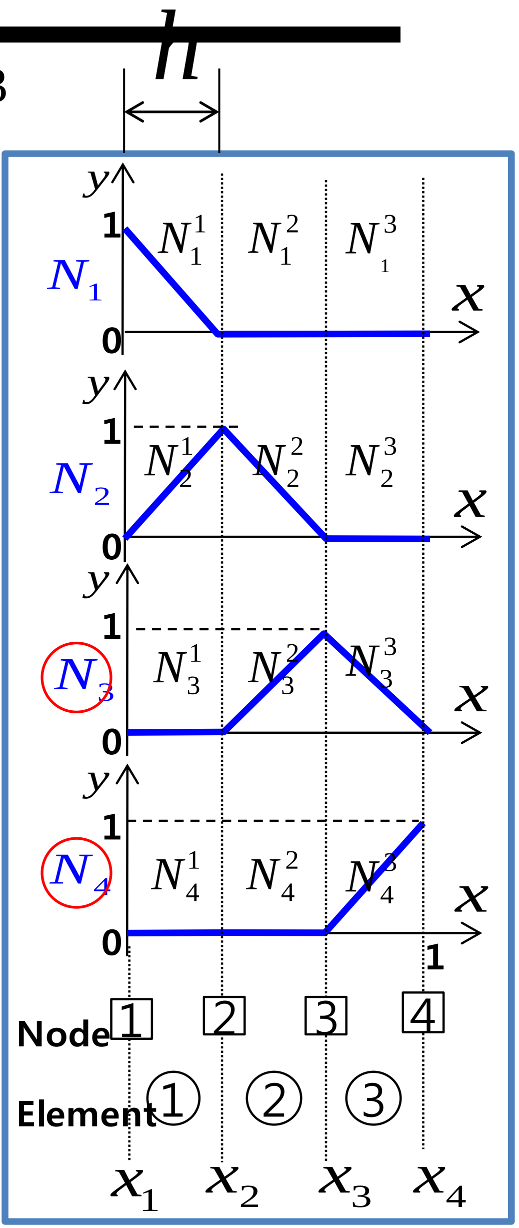
$$\mathbf{K}^1 = \begin{bmatrix}
\int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & 0 & 0 \\
\int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}$$



$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$\mathbf{K}^1 = \begin{bmatrix} \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_1}{dx} + N_1N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_1}{dx}\frac{dN_2}{dx} + N_1N_2\right)dx & 0 & 0 \\ \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_1}{dx} + N_2N_1\right)dx & \int_{x_1}^{x_2}\left(\frac{dN_2}{dx}\frac{dN_2}{dx} + N_2N_2\right)dx & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{h} + \frac{h}{3} & -\frac{1}{h} + \frac{h}{6} & 0 & 0 \\ -\frac{1}{h} + \frac{h}{6} & \frac{1}{h} + \frac{h}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

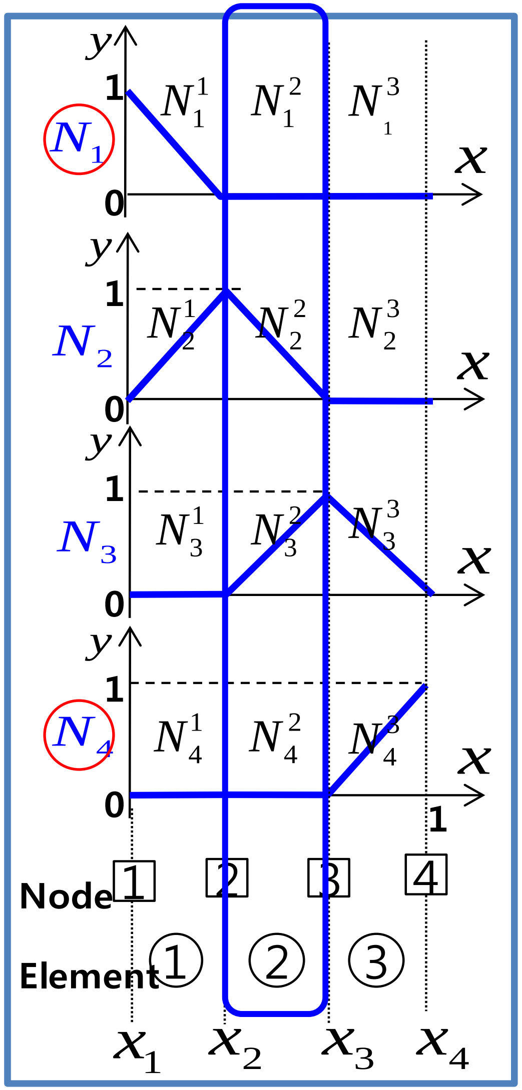


$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$\mathbf{K}^2 = \begin{bmatrix}
\int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\
\int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx
\end{bmatrix}$$

The value of the trial function N_1, N_4 are zero, in element 2

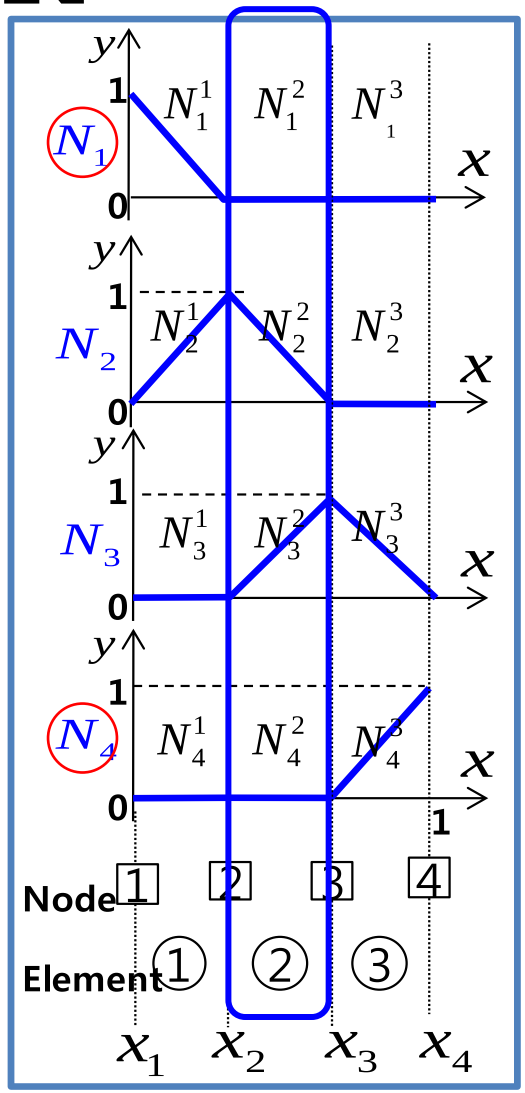
$$\mathbf{K}^2 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & 0 \\
0 & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}$$



$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$\mathbf{K}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_2}{dx} + N_2N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_2}{dx}\frac{dN_3}{dx} + N_2N_3\right)dx & 0 \\ 0 & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_2}{dx} + N_3N_2\right)dx & \int_{x_2}^{x_3}\left(\frac{dN_3}{dx}\frac{dN_3}{dx} + N_3N_3\right)dx & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{h} + \frac{h}{3} & -\frac{1}{h} + \frac{h}{6} & 0 \\ 0 & -\frac{1}{h} + \frac{h}{6} & \frac{1}{h} + \frac{h}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

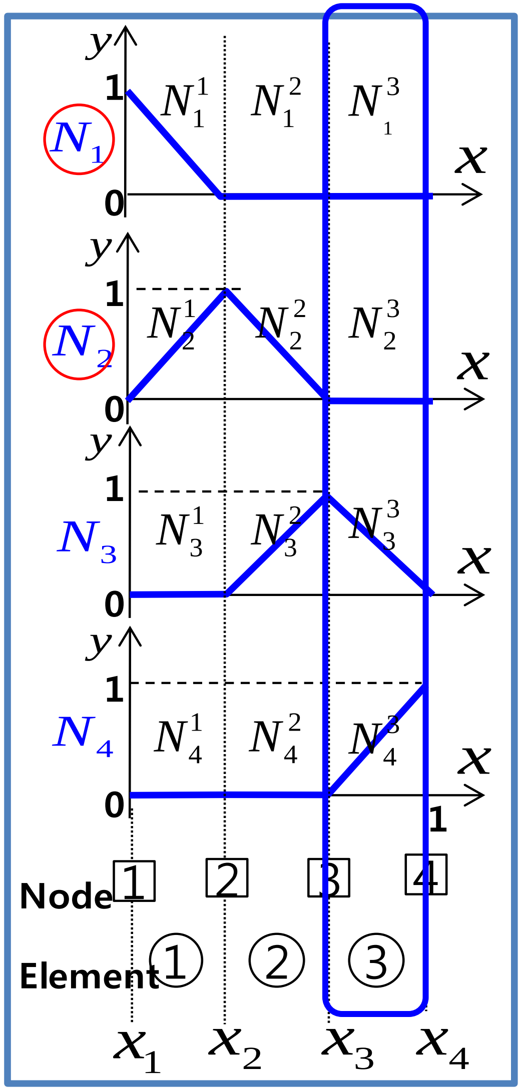


$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$\mathbf{K}^3 = \begin{bmatrix} \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_1}{dx}+N_1N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_2}{dx}+N_1N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_3}{dx}+N_1N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_1}{dx}\frac{dN_4}{dx}+N_1N_4\right)dx \\ \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_1}{dx}+N_2N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_2}{dx}+N_2N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_3}{dx}+N_2N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_2N_4\right)dx \\ \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_1}{dx}+N_3N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_2}{dx}+N_3N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\ \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_1}{dx}+N_4N_1\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_2}{dx}+N_4N_2\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx \end{bmatrix}$$

The value of the trial function N_1, N_2 are zero, in element 3

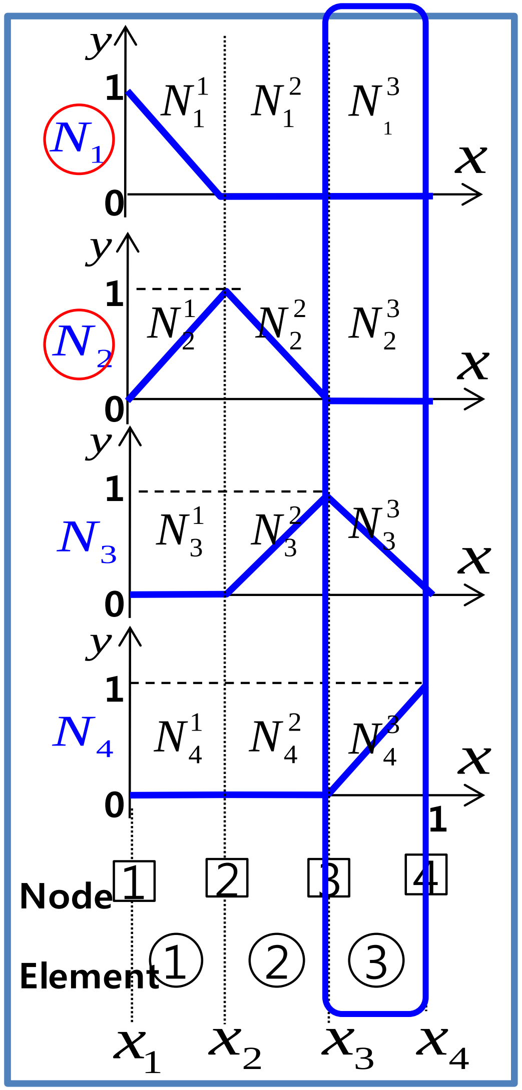
$$\mathbf{K}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_3}{dx}+N_3N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx}+N_3N_4\right)dx \\ 0 & 0 & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_3}{dx}+N_4N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_4}{dx}+N_4N_4\right)dx \end{bmatrix}$$



$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

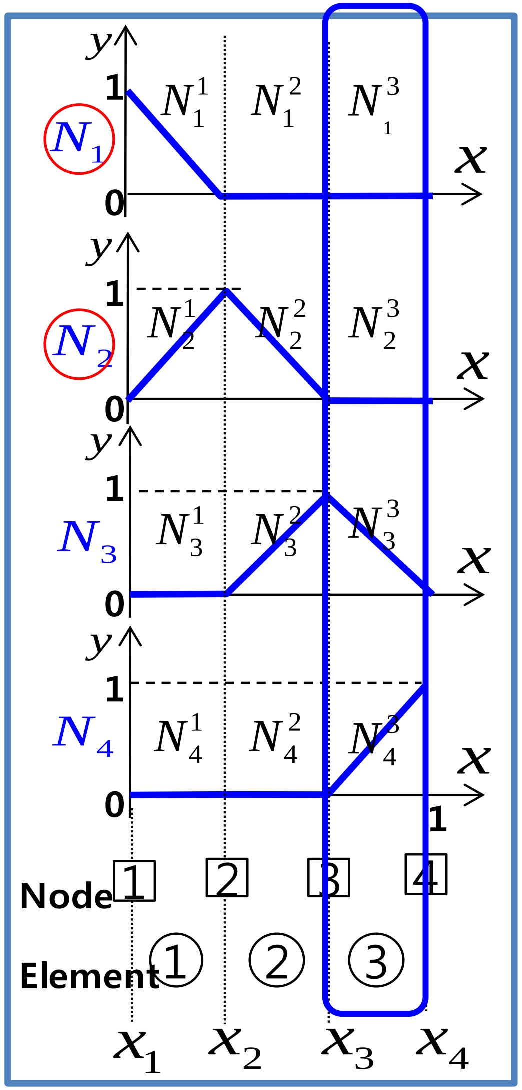
$$\mathbf{K}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \int_{x_3}^{x_4}\left(\frac{dN_3}{dx}\frac{dN_3}{dx} + N_3N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_2}{dx}\frac{dN_4}{dx} + N_3N_4\right)dx \\ 0 & 0 & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_3}{dx} + N_4N_3\right)dx & \int_{x_3}^{x_4}\left(\frac{dN_4}{dx}\frac{dN_4}{dx} + N_4N_4\right)dx \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{h} + \frac{h}{3} & -\frac{1}{h} + \frac{h}{6} \\ 0 & 0 & -\frac{1}{h} + \frac{h}{6} & \frac{1}{h} + \frac{h}{3} \end{bmatrix}$$



$$\mathbf{K} = \mathbf{K}^1 + \mathbf{K}^2 + \mathbf{K}^3$$

$$= \begin{bmatrix} \frac{1}{h} + \frac{h}{3} & -\frac{1}{h} + \frac{h}{6} & 0 & 0 \\ -\frac{1}{h} + \frac{h}{6} & 2\left(\frac{1}{h} + \frac{h}{3}\right) & -\frac{1}{h} + \frac{h}{6} & 0 \\ 0 & -\frac{1}{h} + \frac{h}{6} & 2\left(\frac{1}{h} + \frac{h}{3}\right) & -\frac{1}{h} + \frac{h}{6} \\ 0 & 0 & -\frac{1}{h} + \frac{h}{6} & \frac{1}{h} + \frac{h}{3} \end{bmatrix}$$



$$\mathbf{K\Phi = f}$$

$$\mathbf{f} = \left[\left[N_1 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_2 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_3 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_4 \frac{d\hat{\phi}}{dx} \right]_0^1 \right]^T$$

(1) **(2)** **(3)** **(4)**

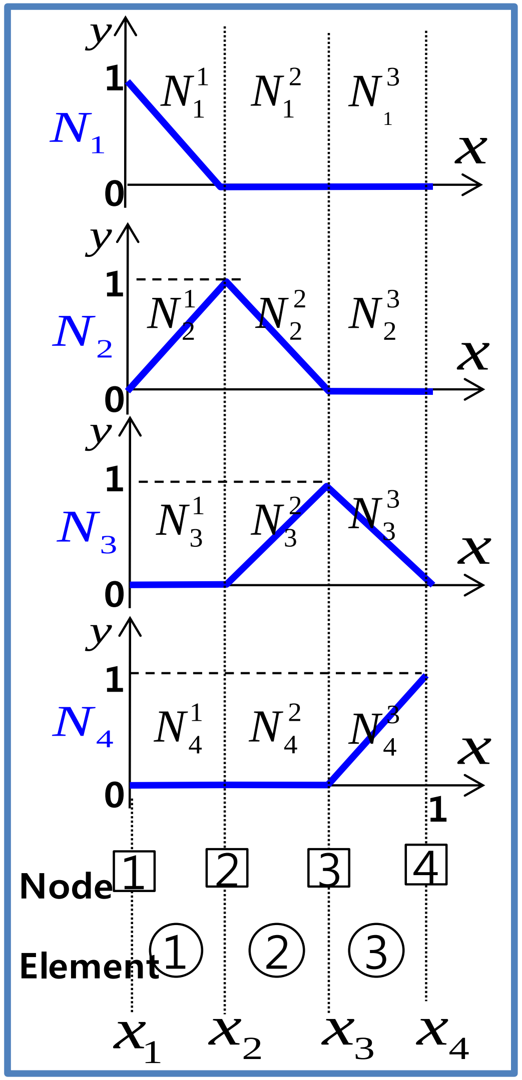
$\downarrow$

(1) $\left[N_1 \frac{d\hat{\phi}}{dx} \right]_0^1 = N_1 \frac{d\hat{\phi}}{dx}\bigg|_{x=1} - N_1 \frac{d\hat{\phi}}{dx}\bigg|_{x=0} = -\frac{d\hat{\phi}}{dx}\bigg|_{x=0}$ $\quad$ *where* $N_1(0)=1, N_1(1)=0$

(2) $\left[N_2 \frac{d\hat{\phi}}{dx} \right]_0^1 = N_2 \frac{d\hat{\phi}}{dx}\bigg|_{x=1} - N_2 \frac{d\hat{\phi}}{dx}\bigg|_{x=0} = 0$ $\quad$ *where* $N_2(0)=0, N_2(1)=0$

(3) $\left[N_3 \frac{d\hat{\phi}}{dx} \right]_0^1 = N_3 \frac{d\hat{\phi}}{dx}\bigg|_{x=1} - N_3 \frac{d\hat{\phi}}{dx}\bigg|_{x=0} = 0$ $\quad$ *where* $N_3(0)=0, N_3(1)=0$

(4) $\left[N_4 \frac{d\hat{\phi}}{dx} \right]_0^1 = N_4 \frac{d\hat{\phi}}{dx}\bigg|_{x=1} - N_4 \frac{d\hat{\phi}}{dx}\bigg|_{x=0} = \frac{d\hat{\phi}}{dx}\bigg|_{x=1}$ $\quad$ *where* $N_4(0)=0, N_4(1)=1$



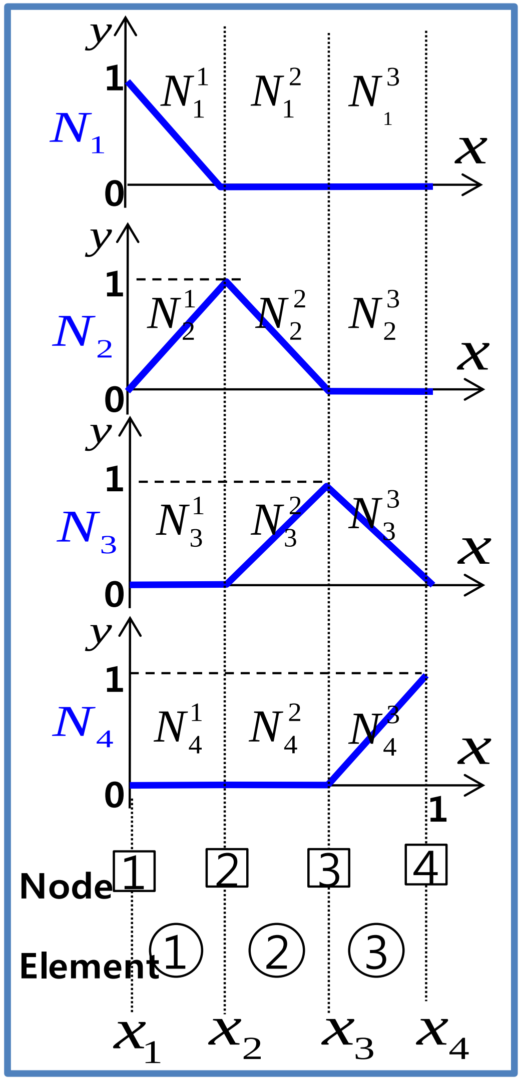
$$\mathbf{K}\boldsymbol{\Phi} = \mathbf{f}$$

$$\mathbf{K} = \begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix}$$

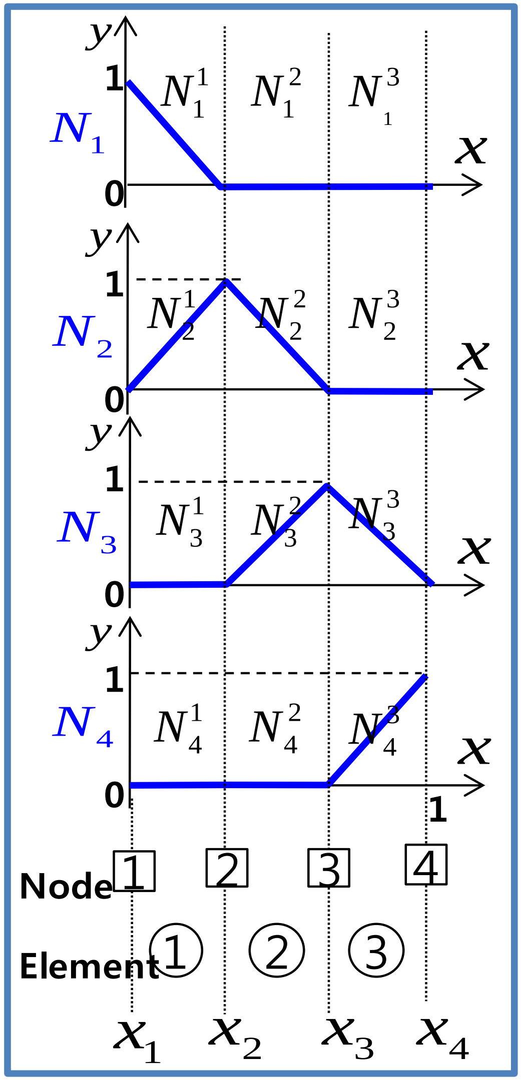
$$\mathbf{f} = \left[\left[N_1 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_2 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_3 \frac{d\hat{\phi}}{dx} \right]_0^1 \quad \left[N_4 \frac{d\hat{\phi}}{dx} \right]_0^1 \right]^T$$

$$\left[-\left.\frac{d\hat{\phi}}{dx}\right|_{x=0} \quad 0 \quad 0 \quad \left.\frac{d\hat{\phi}}{dx}\right|_{x=1} \right]^T$$



$$\mathbf{K\Phi = f}$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} -\left.\frac{d\hat{\phi}}{dx}\right|_{x=0} \\ 0 \\ 0 \\ \left.\frac{d\hat{\phi}}{dx}\right|_{x=1} \end{bmatrix}$$



By using Boundary Condition

$\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

$$\rightarrow \phi_1 = 0, \phi_4 = 1$$

$$\mathbf{K\Phi} = \mathbf{f}$$

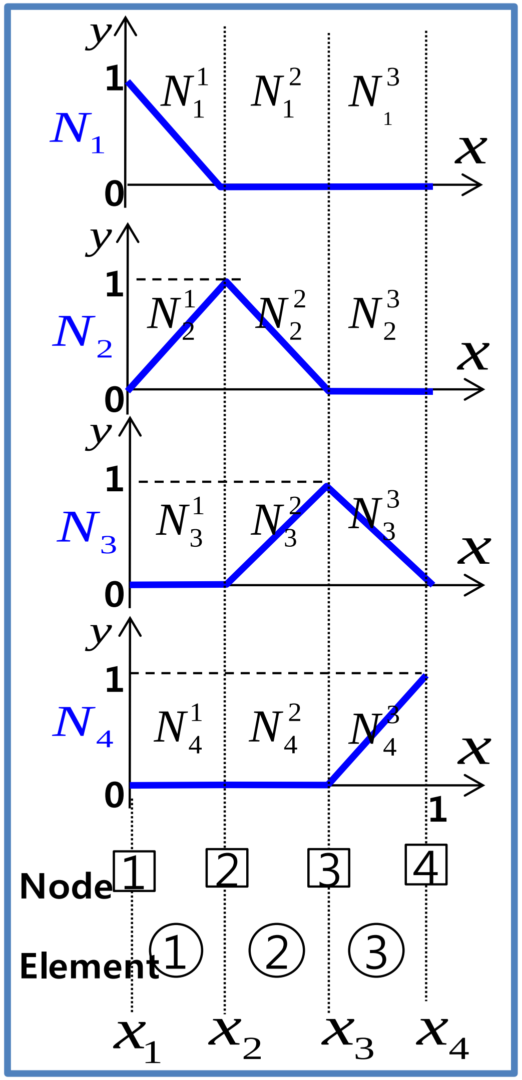
$$\boxed{\phi_1 = 0, \phi_4 = 1}$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix} \begin{bmatrix} 0 \\ \phi_2 \\ \phi_3 \\ 1 \end{bmatrix} = \begin{bmatrix} -\left.\frac{d\hat{\phi}}{dx}\right|_{x=0} \\ 0 \\ 0 \\ \left.\frac{d\hat{\phi}}{dx}\right|_{x=1} \end{bmatrix}$$

$$2\left(\frac{1}{h}+\frac{h}{3}\right)\phi_2 + \left(-\frac{1}{h}+\frac{h}{6}\right)\phi_3 = 0$$

$$\left(-\frac{1}{h}+\frac{h}{6}\right)\phi_2 + 2\left(\frac{1}{h}+\frac{h}{3}\right)\phi_3 = -\left(-\frac{1}{h}+\frac{h}{6}\right)$$

$$h = \frac{1}{3} \longrightarrow \boxed{\phi_2 = 0.2855, \phi_3 = 0.6098}$$



[Recall]

Example 2.2

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \ \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$$

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $\phi = 1$ **at** $x = 1$

From Boundary Condition: $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \ \Gamma$

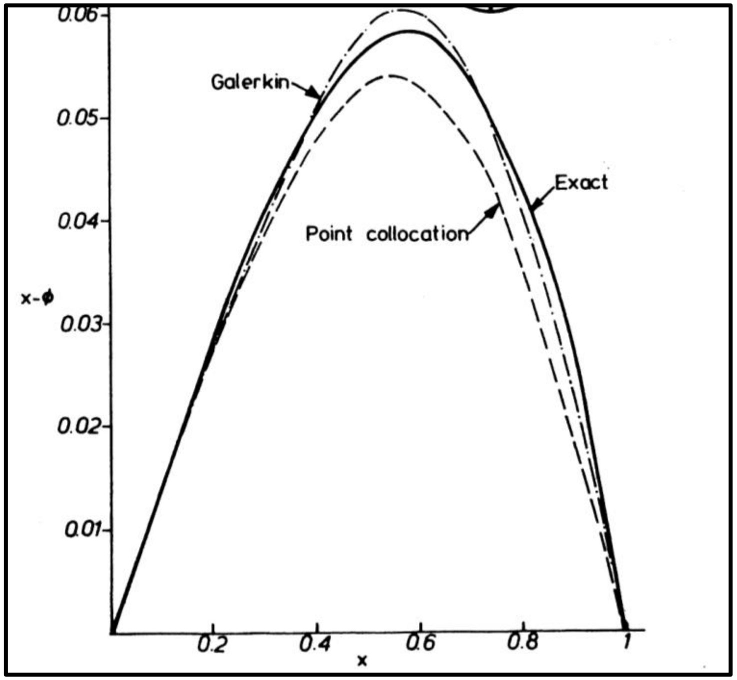
Approximation by Trial Functions $\phi \approx \hat{\phi} = \psi + \sum_{m=1}^{M} a_m N_m$

[Recall]

$$A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$$

$$B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$$

$$N_m = \sin(m\pi x)$$



The approximate values and the exact values at the finite difference mesh point $x = 1/3$ and $x = 2/3$.

x	Finite Difference	Point Collocation	Galerkin method	Exact
1/3	0.2893	0.2941	0.2894	0.2889
2/3	0.6107	0.6165	0.6091	0.6102

Example 3.2

It is required to obtain the function $\phi(x)$

which satisfies the governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ **in** $0 \le x \le 1$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

Governing equation $A(\phi) = \mathscr{L}\phi + p = 0 \quad in \;\; \Omega$

$$\frac{d^2\phi}{dx^2} = \phi \rightarrow \frac{d^2\phi}{dx^2} - \phi = 0 \rightarrow \boxed{A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \;\; \Omega}$$

Boundary Conditions $B(\phi) = \mathscr{M}\phi + r = 0 \quad on \;\; \Gamma$

$\phi = 0$ **at** $x = 0$ → $\phi - 0 = 0$ **at** $x = 0$ → $B(\phi) = \phi = 0 \; at \; x = 0$

$d\phi / dx = 1$ **at** $x = 1$ → $d\phi / dx - 1 = 0$ **at** $x = 1$ → $B(\phi) = d\phi / dx - 1 = 0 \; at \; x = 1$

Example 3.2

$$A(\phi) = \frac{d^2\phi}{dx^2} - \phi = 0 \quad in \;\; 0 < x < 1$$

$$B(\phi) = \phi = 0 \;\; at \;\; x = 0$$

$$B(\phi) = d\phi / dx - 1 = 0 \;\; at \;\; x = 1$$

The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{\phi}) - \cancel{A(\phi)} = \frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \;\; in \;\; 0 < x < 1$$

The boundary residual:

$$\mathbf{R}_{\Gamma,0} = B(\hat{\phi}) - \cancel{B(\phi)} = \hat{\phi} = 0 \;\; at \;\; x = 0$$

$$\mathbf{R}_{\Gamma,1} = B(\hat{\phi}) - \cancel{B(\phi)} = d\hat{\phi} / dx - 1 = 0 \;\; at \;\; x = 1$$

The residual at x=0 being omitted, as this will be made identically zero later, as in the example 3.1

The weighted residual form:

$$\int_0^1 W_l \mathbf{R}_\Omega dx + \overline{W}_l \mathbf{R}_{\Gamma,1}\Big|_{x=1}, \; l = 1, 2, \ldots, E+1$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 1 \right) \right]\Bigg|_{x=1} = 0, \;\; l = 1, 2, \ldots, E+1$$

$$\int_0^1 W_l \left(\frac{d^2\hat{\phi}}{dx^2} - \hat{\phi} \right) dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 1 \right) \right] \Bigg|_{x=1} = 0, \quad l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$\int_0^1 W_l \frac{d^2\hat{\phi}}{dx^2} dx - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 1 \right) \right] \Bigg|_{x=1} = 0$$

$\downarrow$ **Carrying out integration by parts gives**

$$-\int_0^1 \frac{dW_l}{dx} \frac{d\hat{\phi}}{dx} dx + \left[W_l \frac{d\hat{\phi}}{dx} \right]_0^1 - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 1 \right) \right] \Bigg|_{x=1} = 0$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx}\right]_0^1 - \int_0^1 W_l \hat{\phi} dx + \left[\overline{W}_l \left(\frac{d\hat{\phi}}{dx} - 1\right)\right]\Bigg|_{x=1} = 0$$

$$\downarrow \quad \overline{W}_l\Big|_{x=1} = -W_l\Big|_{x=1}$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx}\right]_0^1 - \int_0^1 W_l \hat{\phi} dx - \left[W_l \left(\frac{d\hat{\phi}}{dx} - 1\right)\right]\Bigg|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx}\right]_0^1 - \int_0^1 W_l \hat{\phi} dx - W_l \frac{d\hat{\phi}}{dx}\Bigg|_{x=1} + W_l\Big|_{x=1} = 0$$

Boundary condition $d\phi / dx = 1$ at $x = 1$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \left[W_l \frac{d\hat{\phi}}{dx}\right]_0^1 - \int_0^1 W_l\hat{\phi}dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1} + W_l\Big|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx + \cancel{W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1}} - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} - \int_0^1 W_l\hat{\phi}dx - \cancel{W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=1}} + W_l\Big|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}dx - \int_0^1 W_l\hat{\phi}dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} + W_l\Big|_{x=1} = 0$$

$$\downarrow$$

$$-\int_0^1 \left(\frac{dW_l}{dx}\frac{d\hat{\phi}}{dx} + W_l\hat{\phi}\right)dx - W_l \frac{d\hat{\phi}}{dx}\bigg|_{x=0} + W_l\Big|_{x=1} = 0$$

The boundary condition to be imposed at x=1 can be seen to be a natural condition for this problem.

cf) Example 3.1

$$-\int_0^1\left(\frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}+W_l\hat{\phi}\right)dx+\left[W_l\frac{d\hat{\phi}}{dx}\right]_0^1=0$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \phi_4\end{bmatrix}=\begin{bmatrix}\left[N_1\frac{d\hat{\phi}}{dx}\right]_0^1\\ \left[N_2\frac{d\hat{\phi}}{dx}\right]_0^1\\ \left[N_3\frac{d\hat{\phi}}{dx}\right]_0^1\\ \left[N_4\frac{d\hat{\phi}}{dx}\right]_0^1\end{bmatrix}$$

$$\int_0^1\left(\frac{dW_l}{dx}\frac{d\hat{\phi}}{dx}+W_l\hat{\phi}\right)dx+W_l\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}-W_l\big|_{x=1}=0$$

When the weighting functions are defined by $W_l = N_l$, and with the three equal elements, just as in example 3.1.

$$\mathbf{K\Phi}=\mathbf{f}$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \phi_4\end{bmatrix}=\begin{bmatrix} -N_1\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}+N_1\big|_{x=1} \\ -N_2\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}+N_2\big|_{x=1} \\ -N_3\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}+N_3\big|_{x=1} \\ -N_4\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}+N_4\big|_{x=1} \end{bmatrix}=\begin{bmatrix}-\left.\frac{d\hat{\phi}}{dx}\right|_{x=0}\\ 0\\ 0\\ 1\end{bmatrix}$$

$$\mathbf{K\Phi} = \mathbf{f}$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} -\left.\frac{d\hat{\phi}}{dx}\right|_{x=0} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The boundary condition at x=0 now imposed by deleting the first equation from this set and setting $\phi_1 = 0$

$$2\left(\frac{1}{h}+\frac{h}{3}\right)\phi_2 + \left(-\frac{1}{h}+\frac{h}{6}\right)\phi_3 = 0$$

$$\left(-\frac{1}{h}+\frac{h}{6}\right)\phi_2 + 2\left(\frac{1}{h}+\frac{h}{3}\right)\phi_3 + \left(-\frac{1}{h}+\frac{h}{6}\right)\phi_4 = 0 \longrightarrow$$

$$\left(-\frac{1}{h}+\frac{h}{6}\right)\phi_3 + \left(-\frac{1}{h}+\frac{h}{6}\right)\phi_4 = 1$$

$$\phi_2 = 0.2193$$
$$\phi_3 = 0.4634$$
$$\phi_4 = 0.7600$$

$$\mathbf{K\Phi} = \mathbf{f}$$

$$\begin{bmatrix} \frac{1}{h}+\frac{h}{3} & -\frac{1}{h}+\frac{h}{6} & 0 & 0 \\ -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} & 0 \\ 0 & -\frac{1}{h}+\frac{h}{6} & 2\left(\frac{1}{h}+\frac{h}{3}\right) & -\frac{1}{h}+\frac{h}{6} \\ 0 & 0 & -\frac{1}{h}+\frac{h}{6} & \frac{1}{h}+\frac{h}{3} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{bmatrix} = \begin{bmatrix} -\left.\frac{d\hat{\phi}}{dx}\right|_{x=0} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Then $-\frac{d\widehat{\phi}}{dx}\Big|_{x=0}$ **can be determined**

[Recall]

Example 1.3

governing equation $\dfrac{d^2\phi}{dx^2} = \phi$ **A mesh spacing:** $\Delta x = \dfrac{1}{3}$

Boundary Condition $\phi = 0$ **at** $x = 0$ **and** $d\phi / dx = 1$ **at** $x = 1$

Solution)

$d\phi / dx\big|_3 = 1$

$\phi_0 = 0$ $\quad \boxed{\phi_1}$ $\quad \boxed{\phi_2}$ $\quad \boxed{\phi_3}$ $\quad \boxed{\phi_4}$ **4 unknowns**

$\Delta x = \dfrac{1}{3}$

$x_0 = 0$ $\quad x_1 = \dfrac{1}{3}$ $\quad x_2 = \dfrac{2}{3}$ $\quad x_3 = 1$ $\quad x_4 = \dfrac{4}{3}$

The fictitious mesh point

$$\frac{\phi_{l+1} - 2\phi_l + \phi_{l-1}}{\Delta x^2} = \phi_l \;\Rightarrow\; \phi_{l+1} - 2\phi_l + \phi_{l-1} = \Delta x^2 \phi_l$$

when $l = 1,\quad \phi_2 - 2\phi_1 + \phi_0 = \Delta x^2 \phi_1$

$l = 2,\quad \phi_3 - 2\phi_2 + \phi_1 = \Delta x^2 \phi_2$

$l = 3,\quad \phi_4 - 2\phi_3 + \phi_2 = \Delta x^2 \phi_3$

3 equations

[Recall]

Example 1.2, 1.3

governing equation $\frac{d^2\phi}{dx^2} = \phi$ A mesh spacing: $\Delta x = \frac{1}{3}$

Boundary Condition $\phi = 0$ at $x = 0$ and $d\phi / dx = 1$ at $x = 1$

Solution using backward difference representation of the derivative boundary condition

$$\phi_1 = 0.2477, \phi_2 = 0.5229, \phi_3 = 0.8563$$

Solution using central differencing of the derivative boundary condition

$$\phi_1 = 0.2168, \phi_2 = 0.4576, \phi_3 = 0.7493$$

Exact Solution

$$\phi_1 = 0.2200, \phi_2 = 0.4648, \phi_3 = 0.7616$$

Solution using central differencing can be seen to be considerably more accurate than the solution calculated using the backward difference representation of the derivative B.C

Overview (move to the end)

$$\int_0^1(-u''v)dx=[-u'v]_0^1-\int_0^1(-u'v')dx$$

whenever a smooth 'classical(strong)' solution to a (D.E.) problem exists, it is also the solution of the weak problem[3]

Differential Equation (ODE)

$$ma_y=\sum F$$

$$\sum F_y=T\frac{y(x+\Delta x)-y(x)}{\Delta x}$$

$$ma_y=-\rho\cdot\Delta x\cdot y\cdot\omega^2$$

$$\therefore -\rho\cdot y\cdot\omega^2=\frac{d}{\Delta x}\left(T\frac{y(x+\Delta x)-y(x)}{\Delta x}\right)$$

$$\frac{d}{dx}\left(T\frac{dy}{dx}\right)+\rho\cdot\omega^2\cdot y=0$$

with. external force

$$\frac{d}{dx}\left(T\frac{dy}{dx}\right)+\rho\cdot\omega^2\cdot y=-p$$

ex.) $-u''+u=x,\ 0<x<1,$

$u(0)=0,\ u(1)=0$

$y(0)=0,\ y(l)=0$

$$\mathbf{L}y+\Phi=0$$

Leibnitz formula[1]

$$\frac{d}{dx}\int_{A(x)}^{B(x)}F(x,\xi)d\xi=\frac{d}{dx}\int_{A(x)}^{B(x)}\frac{\partial F(x,\xi)}{\partial x}d\xi+F[x,B(x)]\frac{dB}{dx}-F[x,A(x)]\frac{dA}{dx}$$

A power series defines a function

$$y(x)=\sum_k c_k x^k$$

Series Solution

Approximated Solution of I.E

$$y(x)\approx\sum_k^n c_k\phi_k(x)$$

system of equations

$$\int_0^1(-u''+u-x)v\,dx=0$$

$$u(0)=0,\ u(1)=0$$

multiply a weight function, or test function, v and integration

Integral form

$$\sum_{i=1}^N\beta_i\left(\sum_{j=1}^N\alpha_j\underbrace{\left\{\int_0^1\left[\phi_i'(x)\phi_j'(x)+\phi_i(x)\phi_j(x)\right]dx\right\}}_{K_{ij}}-\underbrace{\int_0^1 x\phi_i(x)dx}_{F_i}\right)=0$$

Weighted Residual[2]

integration by part and demand the test function v vanish at the endpoints

$$\int_0^1(u'v'+uv-xv)\,dx=0$$

Approximate Method[4]

Galerkin

$$u(x)=\sum_{i=1}^n\alpha_i\phi_i(x)$$

$$v(x)=\sum_{i=1}^n\beta_i\phi_i(x)$$

$$\sum_{j=1}^n\beta_i\left(K_{ij}\alpha_j-F_i\right)$$

$$K_{ij}\alpha_j=F_i$$

discretization

→ Algebraic Equation

assume:

$$y(x)=\sum_{k=1}^n c_k\phi_k(x)$$

Work and Energy Principle

Calculus of Variation

multiply virtual displacement δy

integration by part and B/C

$$\left(\frac{d}{dx}\left(T\frac{dy}{dx}\right)+\rho\omega^2y+p\right)\delta y$$

virtual work δW

$$\delta\int_0^l\left[\frac{1}{2}\rho\omega^2y^2+py-\frac{T}{2}\left(\frac{dy}{dx}\right)^2\right]dx=0$$

interpretation : minimize I

$$L=T-V$$

$$\frac{\partial}{\partial x}\left(\frac{\partial L}{\partial y'}\right)-\frac{\partial L}{\partial y}=0$$

Euler-Lagrange Eqn.

$$I=\int_0^l\left[\frac{1}{2}\rho\omega^2y^2-\left(\frac{T}{2}\left(\frac{dy}{dx}\right)^2-py\right)\right]dx$$

kinetic energy T potential Energy V

in static equilibrium

$$I\equiv\Pi=\Pi_{in}-\Pi_{virtual}=\int_0^l\left[\frac{T}{2}\left(\frac{dy}{dx}\right)^2-py\right]dx$$

total potential

internal energy

potential due to external work

Approximate Method

Approximate Method

Rayleigh-Ritz

assume:

$$y(x)=\sum_{k=1}^n c_k\phi_k(x)$$

- Variation and integration
- Integration and variation

Finite Element Method

$$\sum_{k=1}^n c_k s_k(x)\approx F(x)$$

Collocation

$$\sum_{k=1}^n c_k s_k(x_i)=F(x_i)$$

Least Square

$$\min\int_a^b\left[\sum_{k=1}^n c_k s_k(x)-F(x)\right]^2dx$$

Galerkin

$$\sum_{k=1}^n c_k\int_a^b\psi_i(x)s_k(x)dx=\int_a^b\psi_i(x)F(x)dx$$

$$,\psi(x)=\sum_{k=1}^n a_k\phi_i(x)$$

Approximate Method

problem of a "hereditary" nature[5]

Integral Equations

'kernel'

Volterra $\alpha(x)y(x)=F(x)+\lambda\int_a^x K(x,\xi)y(\xi)d\xi$

Fredholm $\alpha(x)y(x)=F(x)+\lambda\int_a^b K(x,\xi)y(\xi)d\xi$

$$y(x)=\int_a^b G(x,\xi)\Phi(\xi)d\xi$$

'Green's function'

$$\sum_{k=1}^n c_k s_k(x)\approx F(x)$$

$$,\ s_k(x)=\phi_k(x)-\lambda\int_a^b K(x,\xi)y(\xi)d\xi$$

1) Jerry, A.j., Introduction to Integral Equations with Applications, Marcel Dekker Inc., 1985, p19~25
2) 'variational statement of the problem' -Becker, E.B., et al, Finite Elements An Introduction, Volume 1, Prentice-Hall, 1981, p4
3) Becker, E.B., et al, Finite Elements An Introduction, Volume 1, Prentice-Hall, 1981, p2 . See also Betounes, Partial Differential Equations for Computational Science, Springer, 1988, p408 "...the weak solution is actually a strong (or classical) solution..."
4) some books refer as 'Method of Weighted Residue' from the Finite Element Equation point of view and they have different type depending on how to choose the weight functions. See also Fletcher,C.A.J., "Computational Galerkin Methods", Springer, 1984
5) Jerry, A.j., Introduction to Integral Equations with Applications, Marcel Dekker Inc., 1985, p1 "Problems of a 'hereditary' nature fall under the first category, since the state of the system u(t) at any time t depends by the definition on all the previous states u(t-τ) at the previous time t-τ ,which means that we must sum over them, hence involve them under the integral sign in an integral equation.

Supplementary Slide

Galerkin's Residual Method

Seoul National Univ.
@SDAL Advanced Ship Design Automation Lab. http://asdal.snu.ac.kr

$$\int_0^l \left[EI \frac{d^4 v(x)}{dx^4} + f \right] N_i \, dx = 0 \quad ,(i = 1,2,3,4)$$

integration by parts

$$\left[N_i EI \frac{d^3 v}{dx^3} \right]_0^l - \int_0^l EI \boxed{\frac{d^3 v}{dx^3}} \frac{dN_i}{dx} dx + \int_0^l f \, N_i \, dx = 0 \quad ,(i = 1,2,3,4)$$

the order of derivative: 3

This equation involves an order of differentiation lower than other equations

integration by parts 2 times

$$EI \int_0^l \boxed{\frac{d^2 N_i}{dx^2} \frac{d^2 v}{dx^2}} dx + EI \left[N_i \frac{d^3 v}{dx^3} - \frac{dN_i}{dx} \frac{d^2 v}{dx^2} \right]_0^l + \int_0^l f \, N_i \, dx = 0 \quad ,(i = 1,2,3,4)$$

the order of derivative: 2

integration by parts 3 times

$$-EI \int_0^l \boxed{\frac{d^3 N_i}{dx^3}} \frac{dv}{dx} dx + EI \left[\frac{d^2 N_i}{dx^2} \frac{dv}{dx} \right] + EI \left[N_i \frac{d^3 v}{dx^3} - \frac{dN_i}{dx} \frac{d^2 v}{dx^2} \right]_0^l + \int_0^l f \, N_i \, dx = 0 \quad ,(i = 1,2,3,4)$$

the order of derivative: 3

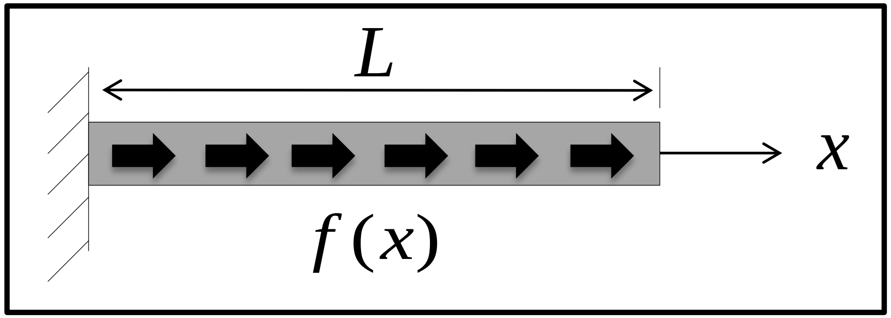
Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

Differential Equation

$$EA\frac{d^2u(x)}{dx^2}+f(x)=0 \qquad 0<x<L$$

Boundary Condition

$$u\big|_{x=0}=0 \quad , \quad EA\frac{du}{dx}\bigg|_{x=L}=0$$



$f(x)$: **External force**

$$f(x)=f : const$$

Governing equation $A(u)=\mathscr{L}u+p=0 \quad in \ \ \Omega$

$$EA\frac{d^2u}{dx^2}+f=0 \quad \longrightarrow \quad \boxed{A(u)=EA\frac{d^2u}{dx^2}+f=0 \quad in \ \ \Omega}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

$$\boxed{A(u) = EA\frac{d^2u}{dx^2} + f = 0 \quad in \;\; 0 < x < L}$$

$$u \approx \hat{u} = \sum_{m=1}^{E+1} u_m N_m, \;\; 0 < x < L$$

,where E is the number of the elements

The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{u}) - \cancel{A(u)} = EA\frac{d^2\hat{u}}{dx^2} + f \quad in \;\; 0 < x < L$$

The weighted residual form:

$$\int_0^L W_l \mathbf{R}_\Omega dx = 0, \; l = 1, 2, \ldots, E+1$$

$$\int_0^L W_l \left(EA\frac{d^2\hat{u}}{dx^2} + f \right) dx = 0, \; l = 1, 2, \ldots, E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$u \approx \hat{u} = \sum_{m=1}^{E+1} u_m N_m, 0 < x < L$$

,where E is the number of the elements

$$\int_0^L W_l \left(EA \frac{d^2\hat{u}}{dx^2} + f \right) dx = 0,\ l = 1, 2, \dots, E+1$$

$$\downarrow$$

$$\int_0^L W_l EA \frac{d^2\hat{u}}{dx^2} dx - \int_0^L W_l f\, dx = 0,\ l = 1, 2, \dots, E+1$$

$$\downarrow$$

$$EA \int_0^L W_l \frac{d^2\hat{u}}{dx^2} dx - \int_0^L W_l f\, dx = 0,\ l = 1, 2, \dots, E+1$$

$\downarrow$ **Integration by parts**

$$-EA \int_0^L \frac{dW_l}{dx} \frac{d\hat{u}}{dx} dx + EA \left[W_l \frac{d\hat{u}}{dx} \right]_0^L - \int_0^L W_l f dx = 0,\ l = 1, 2, \dots, E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$-EA\int_0^L \frac{dW_l}{dx}\frac{d\hat{u}}{dx}dx + EA\left[W_l\frac{d\hat{u}}{dx}\right]_0^L - \int_0^L W_l f dx = 0,\ l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$-EA\int_0^L \frac{dW_l}{dx}\frac{d\sum_{m=1}^{E+1} u_m N_m}{dx}dx - \int_0^L W_l f dx + EA\left[W_l\frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$EA\int_0^L \frac{dW_l}{dx}\frac{d\sum_{m=1}^{E+1} u_m N_m}{dx}dx + \int_0^L W_l f dx - EA\left[W_l\frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, \ldots, E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\int_0^L \frac{dW_l}{dx}\frac{d\sum_{m=1}^{E+1} u_m N_m}{dx}dx + \int_0^L W_l\, f dx - EA\left[W_l \frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, ..., E+1$$

$\downarrow$ **Galerkin methods** $W_l = N_l$

$$EA\int_0^L \frac{dN_l}{dx}\frac{d\sum_{m=1}^{E+1} u_m N_m}{dx}dx + \int_0^L N_l\, f dx - EA\left[N_l \frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, ..., E+1$$

$\downarrow$

$$EA\sum_{m=1}^{E+1}\int_0^L u_m \frac{dN_l}{dx}\frac{dN_m}{dx}dx + \int_0^L N_l\, f dx - \left[N_l EA \frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, ..., E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\sum_{m=1}^{E+1}\int_0^L u_m \frac{dN_l}{dx}\frac{dN_m}{dx}dx+\int_0^L N_l f dx-\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L=0,\ l=1,2,...,E+1$$

$$\downarrow \quad m=1,2,...,E+1$$

$$EA\left(\int_0^L u_1 \frac{dN_l}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_l}{dx}\frac{dN_2}{dx}dx+...+\int_0^L u_{E+1} \frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right)$$

$$+\int_0^L N_l f dx-\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L=0,\quad l=1,2,...,E+1$$

$$\downarrow$$

$$EA\left(\int_0^L u_1 \frac{dN_l}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_l}{dx}\frac{dN_2}{dx}dx+...+\int_0^L u_{E+1} \frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right)=$$

$$-\int_0^L N_l f dx+\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L,\quad l=1,2,...,E+1$$

Element : Bar (2 elements , 3 nodes) - Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^L u_1 \frac{dN_l}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_l}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right) =$$

$$-\int_0^L N_l f dx + \left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\downarrow \; l = 1, 2, \ldots, E+1$$

$$EA\left(\int_0^L u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_1 f dx + \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$EA\left(\int_0^L u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_2 f dx + \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\vdots$$

$$EA\left(\int_0^L u_1 \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_{E+1} f dx + \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L$$

Element : Bar (2 elements , 3 nodes) - Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^L u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_1 f dx + \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$EA\left(\int_0^L u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_2 f dx + \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\vdots$$

$$EA\left(\int_0^L u_1 \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx\right) = -\int_0^L N_{E+1} f dx + \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\downarrow$$

$$\underbrace{EA\begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{E+1} \end{bmatrix}}_{\mathbf{d}} = \underbrace{\begin{bmatrix} -\int_0^L N_1 f dx + \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L \\ -\int_0^L N_2 f dx + \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L \\ \vdots \\ -\int_0^L N_{E+1} f dx + \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L \end{bmatrix}}_{\mathbf{F}}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

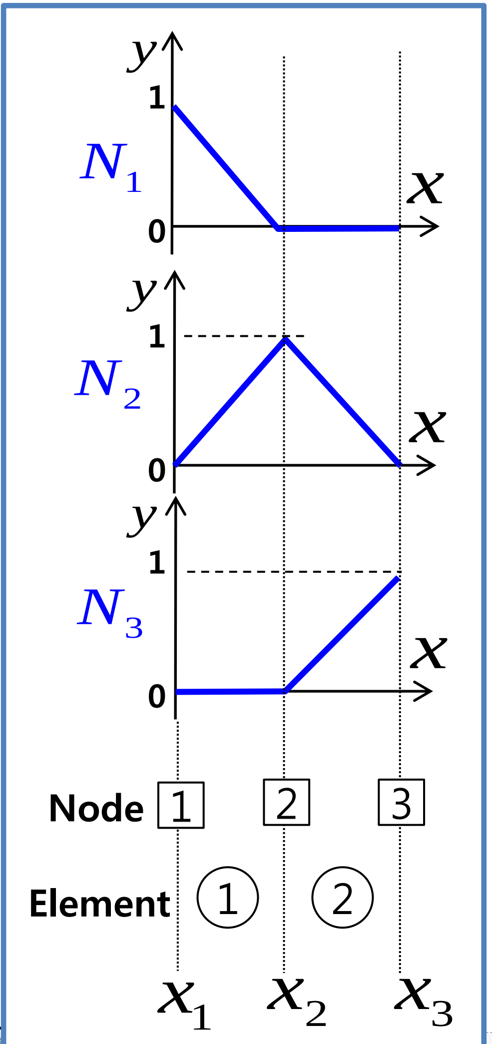
The weighted residual form:

The number of the elements E is 2

$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = EA\begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{E+1} \end{bmatrix} \qquad \mathbf{F} = \begin{bmatrix} -\int_0^L N_1 f dx + \left[N_1 EA \frac{d\hat{u}}{dx} \right]_0^L \\ -\int_0^L N_2 f dx + \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^L \\ \vdots \\ -\int_0^L N_{E+1} f dx + \left[N_{E+1} EA \frac{d\hat{u}}{dx} \right]_0^L \end{bmatrix}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $\mathbf{Kd} = \mathbf{F}$

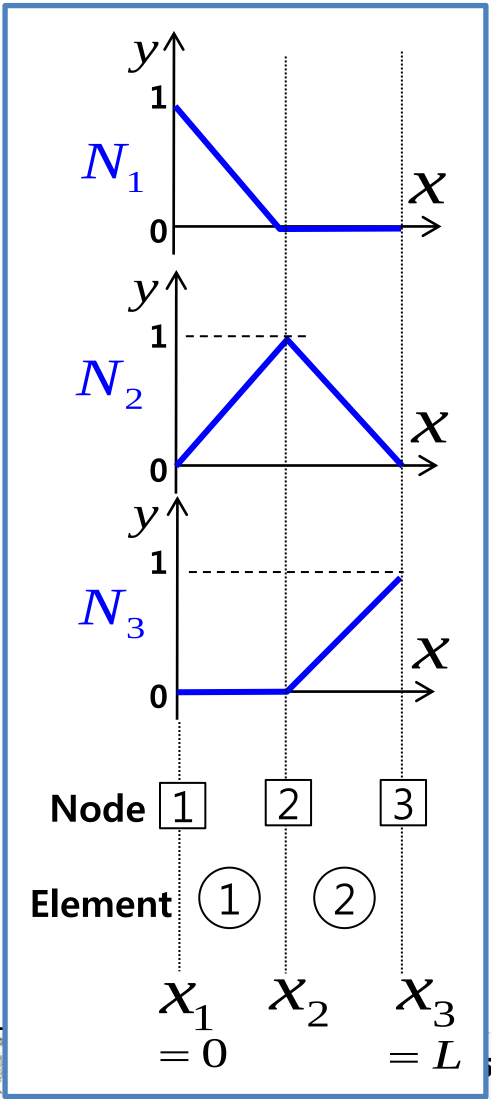
The number of the elements E is 2

$$\mathbf{K} = EA \begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA \begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_{x_1}^{x_2} \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \boxed{\int_{x_1}^{x_2}} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_1}^{x_2}}$$

$$\mathbf{K}_2 = EA \begin{bmatrix} \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_{x_2}^{x_3} \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \boxed{\int_{x_2}^{x_3}} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_2}^{x_3}}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

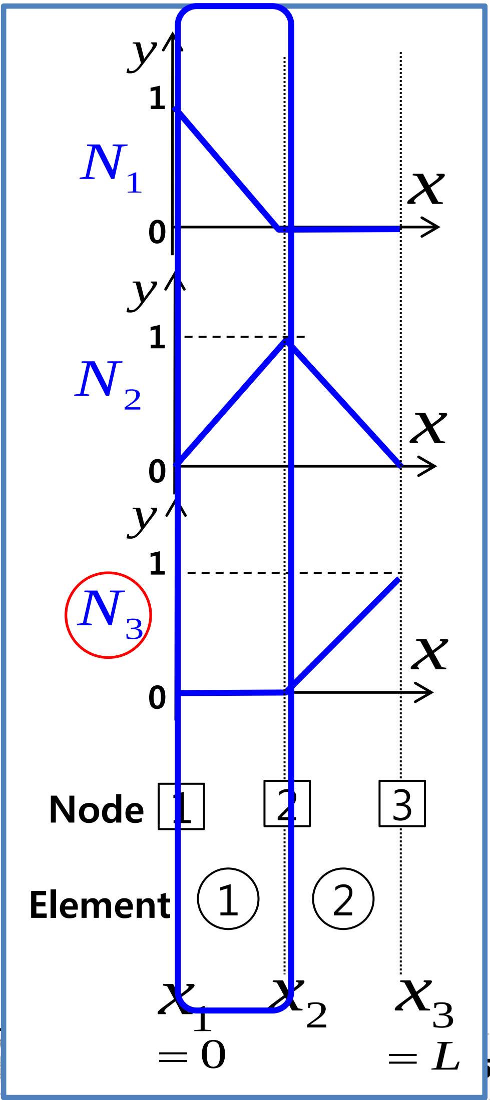
$$\mathbf{Kd} = \mathbf{F}$$

The number of the elements E is 2

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA\begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_3}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_2}{dx}dx & \int_{x_1}^{x_2} \frac{dN_3}{dx}\frac{dN_3}{dx}dx \end{bmatrix}$$

$$= EA\begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & 0 \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

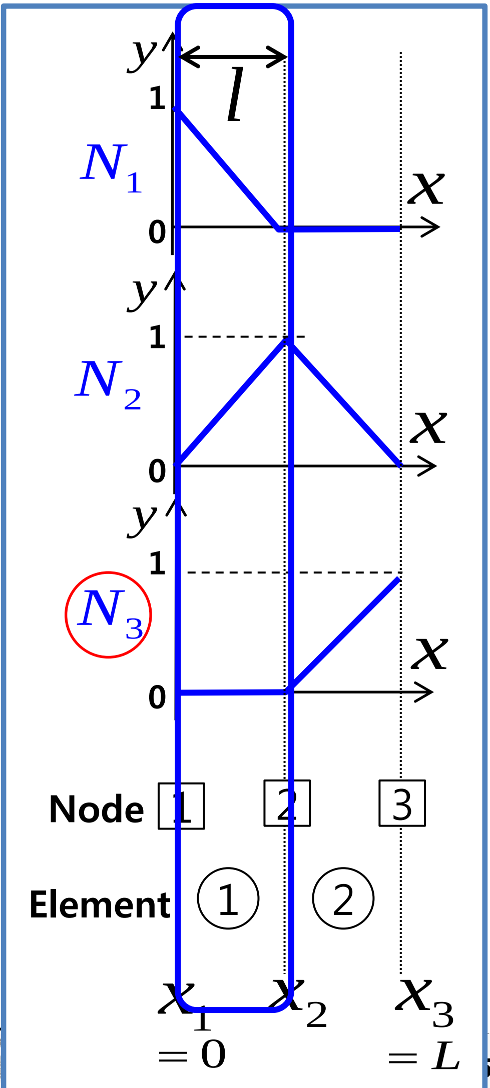
The weighted residual form:

$$\mathbf{Kd} = \mathbf{F}$$

The number of the elements E is 2

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA \begin{bmatrix} \boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx} & \boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx} & 0 \\ \boxed{\int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx} & \boxed{\int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



$$\boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx} = \int_0^l \left(-\frac{1}{l}\right)\left(-\frac{1}{l}\right)dx$$

$$= \int_0^l \frac{1}{l^2}dx = \left[\frac{1}{l^2}x\right]_0^l = \frac{1}{l}$$

$$\frac{dN_1}{dx} = -\frac{1}{l}, \quad \frac{dN_2}{dx} = \frac{1}{l}$$

$$\boxed{\int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx} = \int_0^l \left(-\frac{1}{l}\right)\left(\frac{1}{l}\right)dx$$

$$= -\int_0^l \frac{1}{l^2}dx = -\left[\frac{1}{l^2}x\right]_0^l = -\frac{1}{l}$$

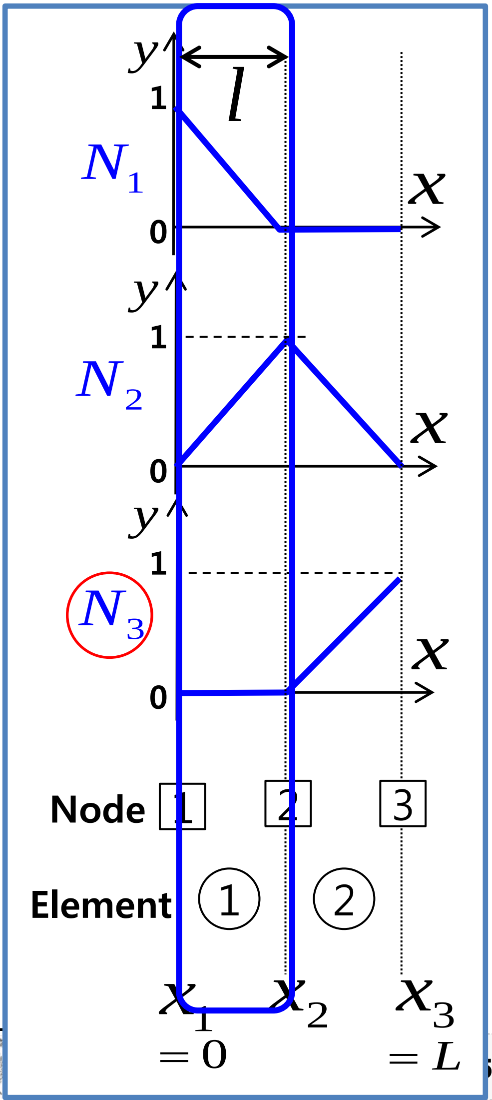
Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $$\mathbf{Kd} = \mathbf{F}$$

The number of the elements E is 2

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA\begin{bmatrix} \frac{1}{l} & -\frac{1}{l} & 0 \\ -\frac{1}{l} & \frac{1}{l} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



In a manner similar to calculate K_1,

$$\mathbf{K}_2 = EA\begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{l} & -\frac{1}{l} \\ 0 & -\frac{1}{l} & \frac{1}{l} \end{bmatrix} = \frac{EA}{l}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

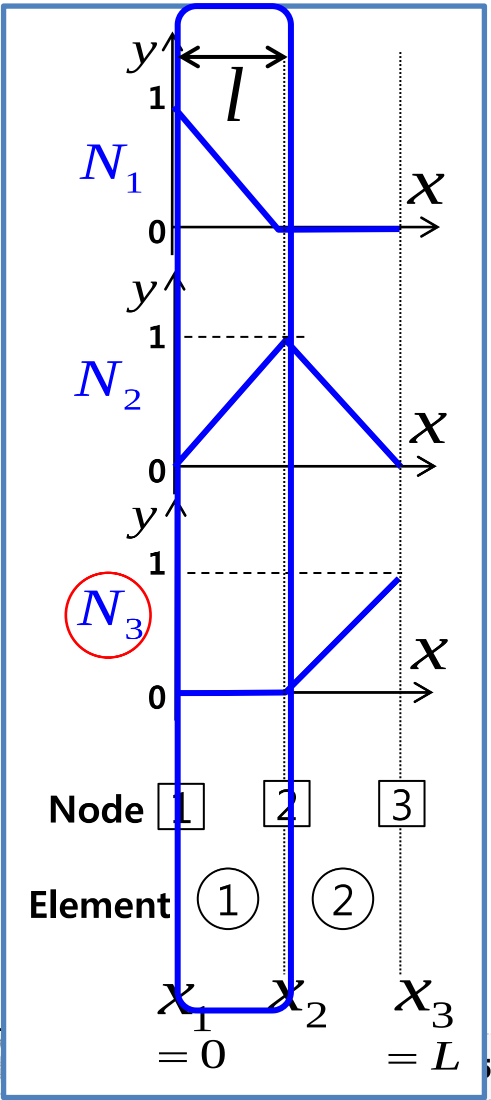
Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $\mathbf{Kd} = \mathbf{F}$

The number of the elements E is 2

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2 = \frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{K}_1 = \frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{K}_2 = \frac{EA}{l}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

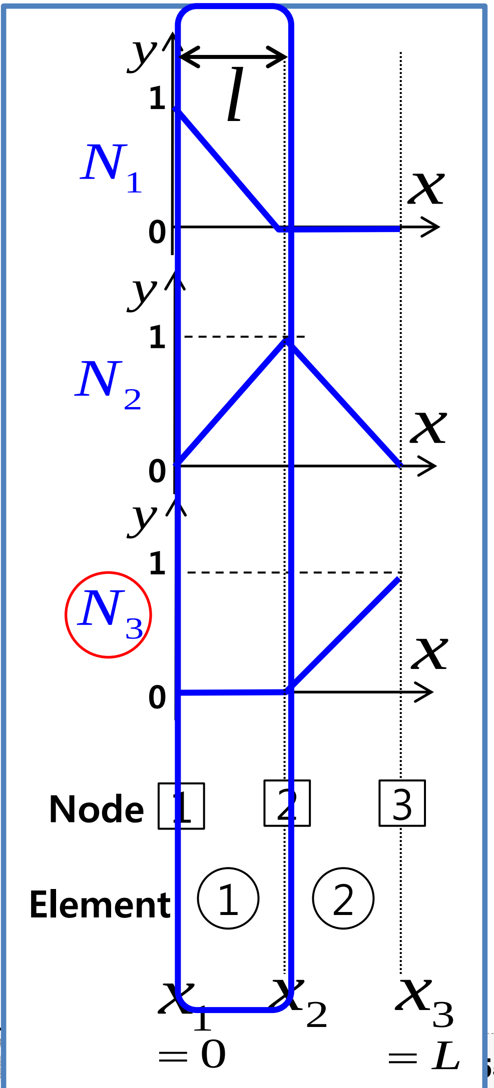
$$\mathbf{F} = \begin{bmatrix} -\int_0^L N_1 f dx + \left[N_1 EA \frac{d\hat{u}}{dx} \right]_0^L \\ -\int_0^L N_2 f dx + \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^L \\ -\int_0^L N_3 f dx + \left[N_3 EA \frac{d\hat{u}}{dx} \right]_0^L \end{bmatrix}$$

Boundary Condition

$$u\big|_{x=0} = 0 \quad , \quad EA\frac{du}{dx}\bigg|_{x=L} = 0$$

$$f(x) = f : const$$

The number of the elements E is 2



$$-\int_0^L N_1 f dx + \left[N_1 EA \frac{d\hat{u}}{dx} \right]_0^L = -f\int_0^{\frac{L}{2}} \left(-\frac{1}{L/2}x + 1 \right) dx + N_1 EA \frac{d\hat{u}}{dx}\bigg|_{x=0} + \cancel{N_1} EA \cancel{\frac{d\hat{u}}{dx}}\bigg|_{x=L}$$

$$= f\left[\frac{1}{L/2}\frac{1}{2}x^2 - x \right]_0^{\frac{L}{2}} + EA\frac{d\hat{u}}{dx}\bigg|_{x=0}$$

$$= f\left(\frac{1}{4}L - \frac{1}{2}L\right) + EA\frac{d\hat{u}}{dx}\bigg|_{x=0}$$

$$= -\frac{1}{4} f \cdot L + EA\frac{d\hat{u}}{dx}\bigg|_{x=0}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

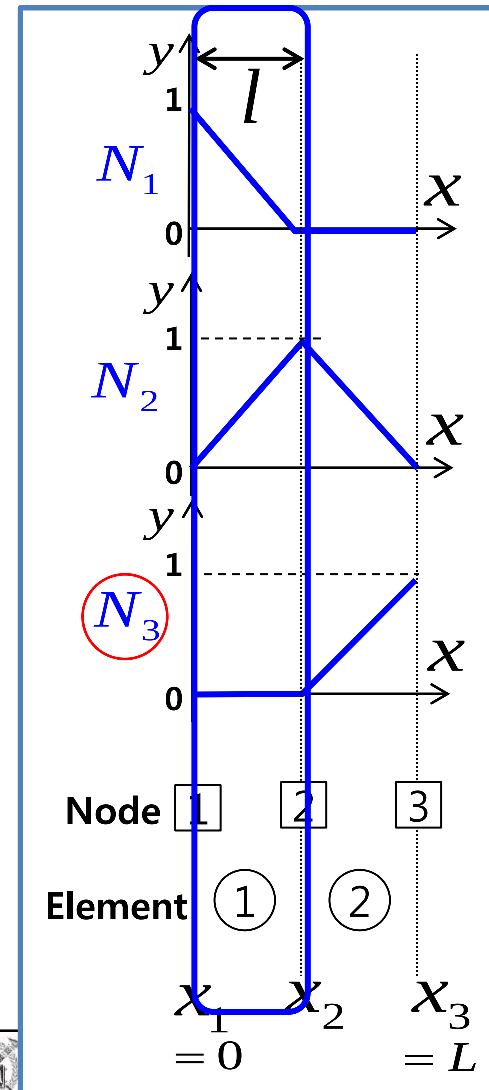
$$\mathbf{F} = \begin{bmatrix} -\int_0^L N_1 f dx + \left[N_1 EA \frac{d\hat{u}}{dx} \right]_0^L \\ -\int_0^L N_2 f dx + \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^L \\ -\int_0^L N_3 f dx + \left[N_3 EA \frac{d\hat{u}}{dx} \right]_0^L \end{bmatrix}$$

Boundary Condition

$$u\big|_{x=0} = 0 \quad , \quad EA \frac{du}{dx}\bigg|_{x=L} = 0$$

$$f(x) = f : const$$

The number of the elements E is 2



$$-\int_0^L N_2 f dx + \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^L = -\int_0^L N_2 f dx + \cancel{N_2} EA \frac{d\hat{u}}{dx}\bigg|_{x=0} + \cancel{N_2} EA \cancel{\frac{d\hat{u}}{dx}}\bigg|_{x=L}$$

$$= -\frac{1}{2} f \cdot L$$

$$-\int_0^L N_2 f dx + \left[N_2 EA \frac{d\hat{u}}{dx} \right]_0^L = -\int_0^L N_3 f dx + \cancel{N_3} EA \frac{d\hat{u}}{dx}\bigg|_{x=0} + N_3 EA \cancel{\frac{d\hat{u}}{dx}}\bigg|_{x=L}$$

$$= -\frac{1}{4} f \cdot L$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$\mathbf{K}\mathbf{d} = \mathbf{F}$$

$$\mathbf{K} = \frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

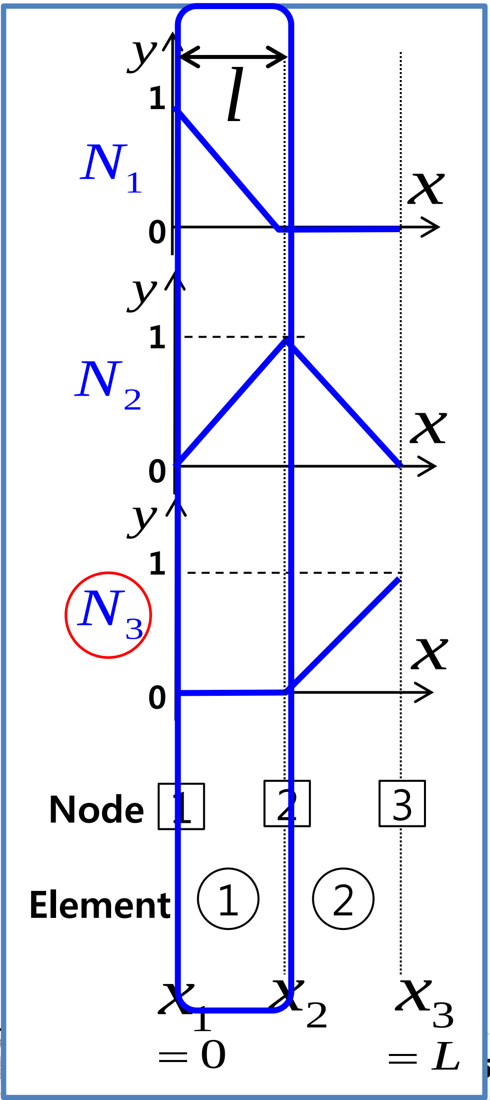
Boundary Condition

$$u\big|_{x=0} = 0 \;,\qquad EA\frac{du}{dx}\bigg|_{x=L} = 0$$

$$u_1 = 0$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad \mathbf{F} = \begin{bmatrix} -\frac{1}{4} f \cdot L + EA\frac{d\hat{u}}{dx}\Big|_{x=0} \\ -\frac{1}{2} f \cdot L \\ -\frac{1}{4} f \cdot L \end{bmatrix}$$

The number of the elements E is 2



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

The number of the elements E is 2

$$\mathbf{Kd} = \mathbf{F}$$

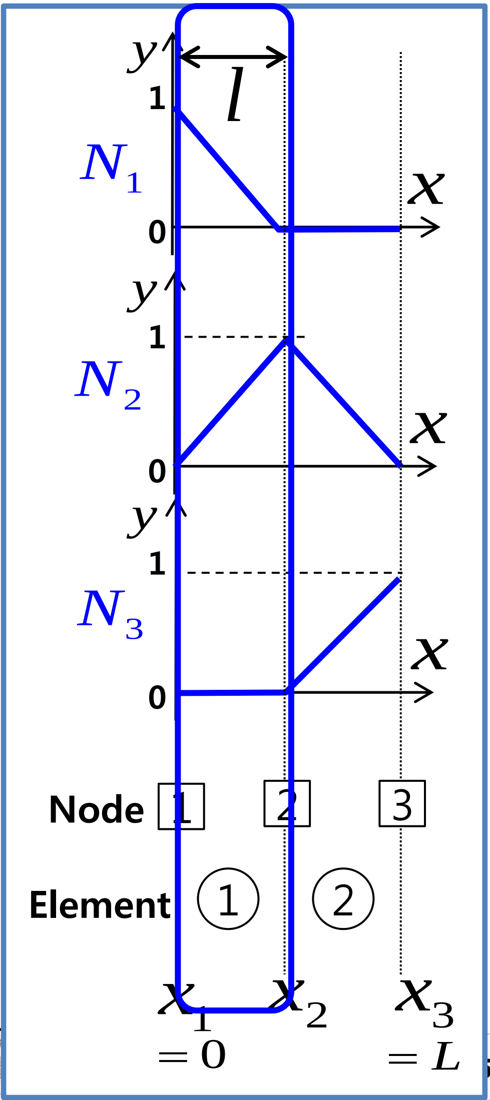
$$\frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} f \cdot L + EA \left.\frac{d\hat{u}}{dx}\right|_{x=0} \\ -\frac{1}{2} f \cdot L \\ -\frac{1}{4} f \cdot L \end{bmatrix}$$

$$2u_2 - u_3 = -\frac{1}{2} f \cdot L \frac{l}{EA}$$

$$where\ l = L/2$$

$$-u_2 + u_3 = -\frac{1}{4} f \cdot L \frac{l}{EA}$$

$$\Rightarrow \quad u_2 = -\frac{3}{8} f \cdot \frac{L^2}{EA}, \quad u_3 = -\frac{1}{2} f \cdot \frac{L^2}{EA}$$



Element : Bar (2 elements , 3 nodes) - Solving D/E using Galerkin's Residual Method

$u_2 = -\frac{3}{8} f \cdot \frac{L^2}{EA}, \quad u_3 = -\frac{1}{2} f \cdot \frac{L^2}{EA}$

The weighted residual form:

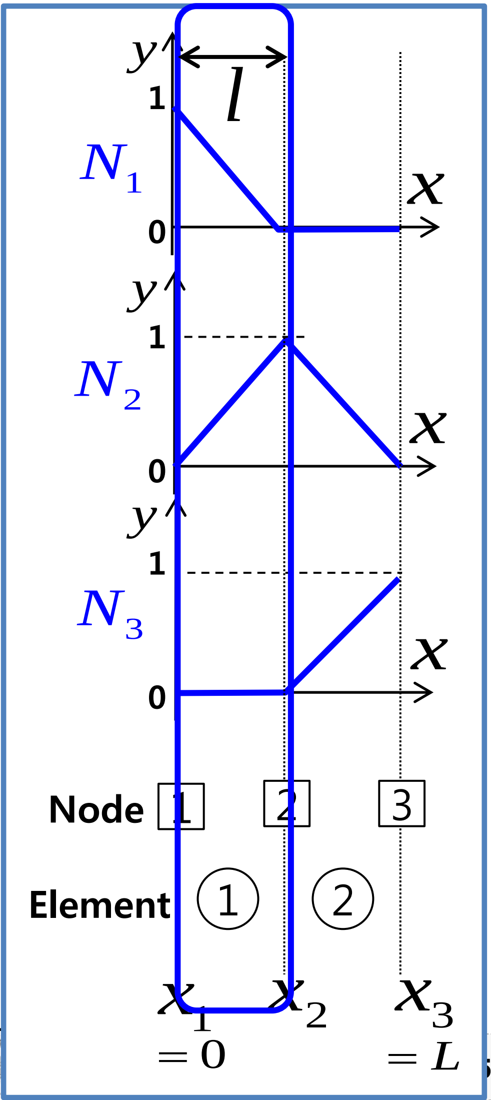
The number of the elements E is 2

$$\mathbf{Kd} = \mathbf{F}$$

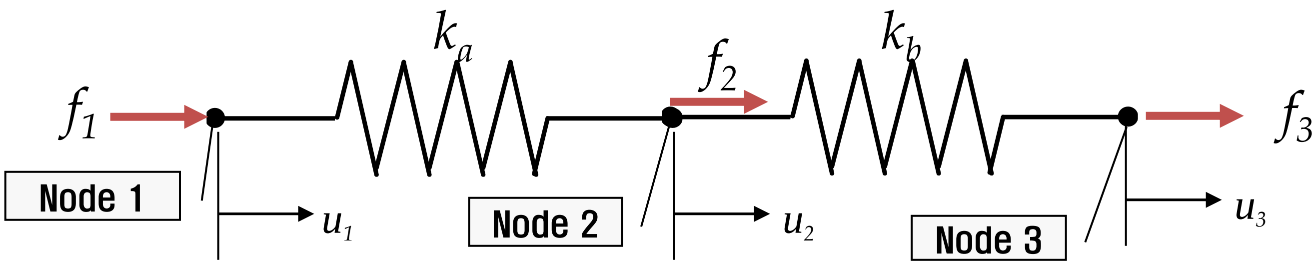
$$\frac{EA}{l}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \boxed{-\frac{1}{4} f \cdot L + EA \left.\frac{d\hat{u}}{dx}\right|_{x=0}} \\ -\frac{1}{2} f \cdot L \\ -\frac{1}{4} f \cdot L \end{bmatrix}$$

$$\boxed{} = \left(-\frac{EA}{l}\right) u_2 = \left(-\frac{EA}{l}\right)\left(-\frac{3}{8}\right) f \cdot \frac{L^2}{EA}$$

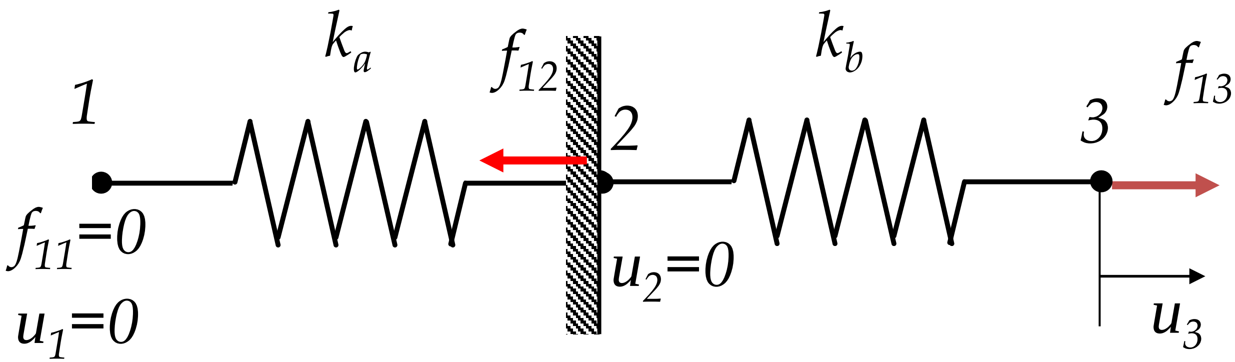
$$= \left(-\frac{EA}{L/2}\right)\left(-\frac{3}{8}\right) f \cdot \frac{L^2}{EA} = \frac{3}{4} f \cdot L$$



Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach

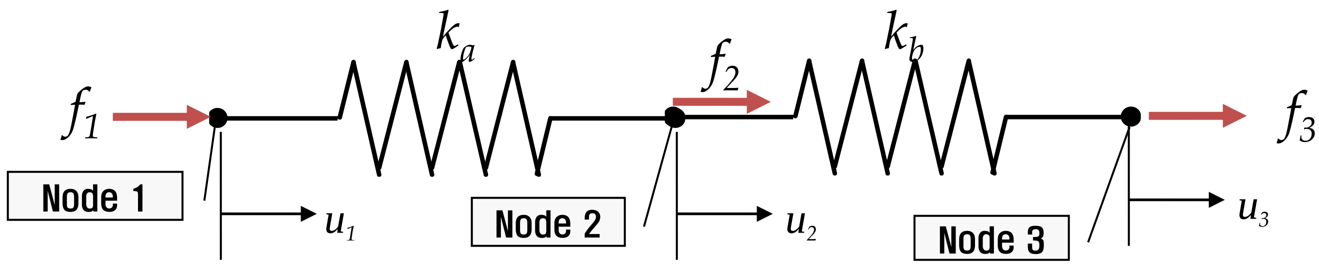


① **Case #1: The node 1 and 2 are fixed** $(u_1 = u_2 = 0)$

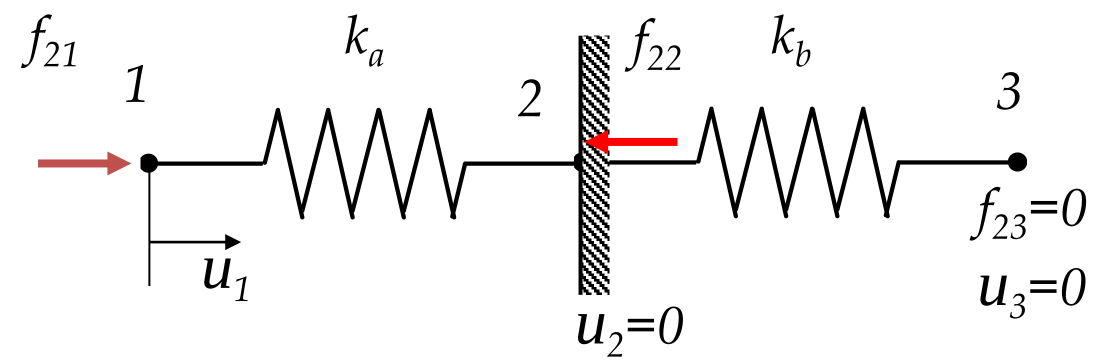


$$f_{11} = 0 \qquad f_{12} = -k_b u_3 \qquad f_{13} = k_b u_3$$

Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach

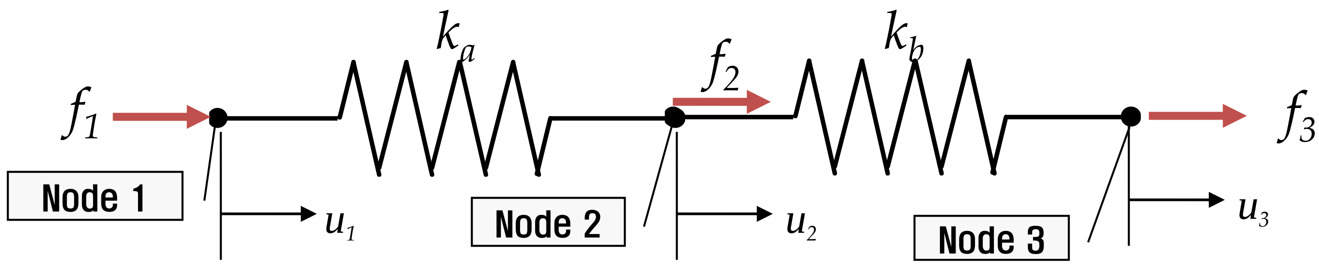


② **Case #2: The node 2 and 3 are fixed.** ($u_2 = u_3 = 0$)

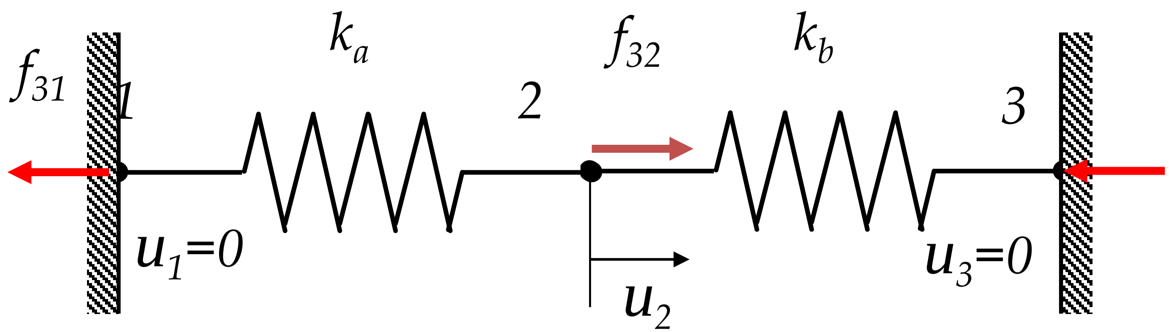


$$f_{21} = k_a u_1 \qquad f_{22} = -k_a u_1 \qquad f_{23} = 0$$

Element : Bar (2 elements , 3 nodes)
- Direct equilibrium approach



③ **Case #3: The node 1 and 3 are fixed.** ($u_1 = u_3 = 0$)

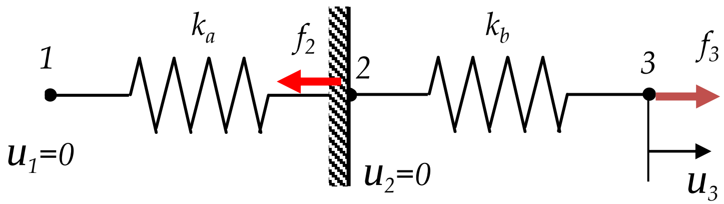


$$f_{31} = -k_a u_2 \qquad f_{32} = k_a u_2 + k_b u_2 \qquad f_{33} = -k_b u_2$$

$$= (k_a + k_b) u_2$$

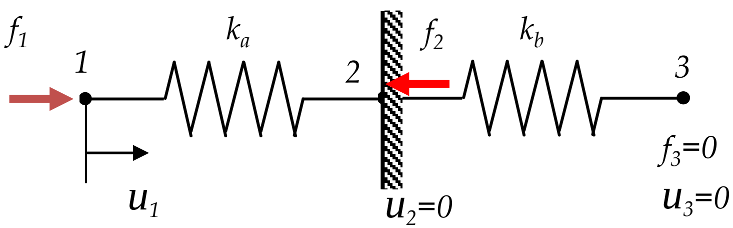
Element : Bar (2 elements , 3 nodes)
- Structural analysis using direct stiffness method

① **The node 1 and 2 are fixed.** ($u_1 = u_2 = 0$)



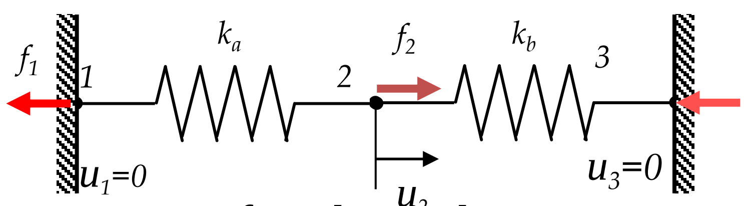
$$f_{11} = 0 \qquad f_{12} = -k_b u_3 \qquad f_{13} = k_b u_3$$

② **The node 2 and 3 are fixed.** ($u_2 = u_3 = 0$)



$$f_{21} = k_a u_1 \qquad f_{22} = -k_a u_1 \qquad f_{23} = 0$$

③ **The node 1 and 3 are fixed.** ($u_1 = u_3 = 0$)



$$f_{31} = -k_a u_2 \qquad f_{32} = k_a u_2 + k_b u_2 = (k_a + k_b) u_2 \qquad f_{33} = -k_b u_2$$

Case #4: The nodes are not fixed.

$$\begin{cases} f_1 = k_a u_1 - k_a u_2 \\ f_2 = -k_a u_1 + (k_a + k_b) u_2 - k_b u_3 \\ f_3 = -k_b u_2 + k_b u_3 \end{cases}$$

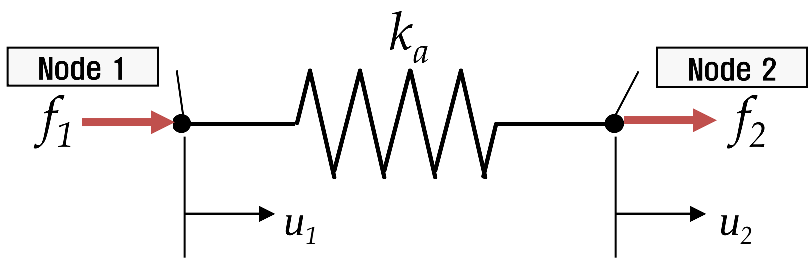
Matrix Form

Stiffness Matrix

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \boxed{\begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix}} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

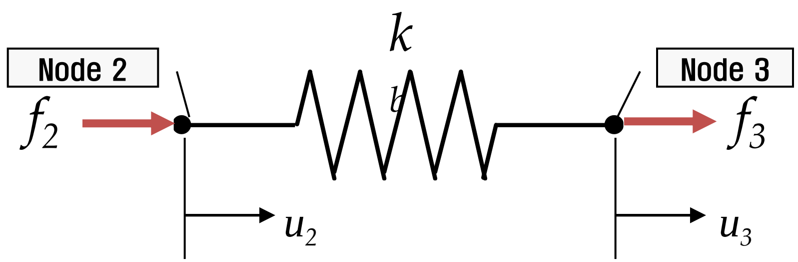
$$[f] = [K][u]$$

Element : Bar (2 elements , 3 nodes)
- Structural analysis using direct stiffness method



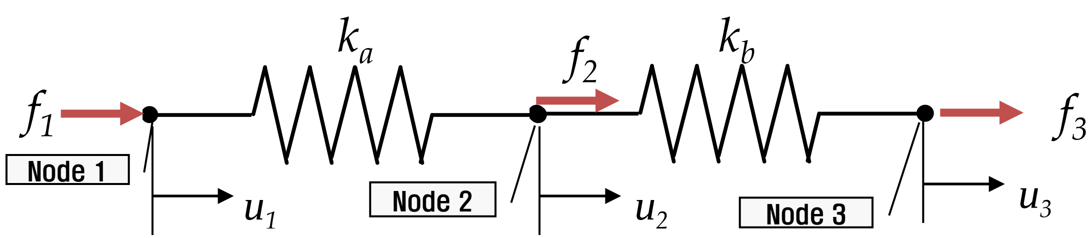
$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

+



$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

=

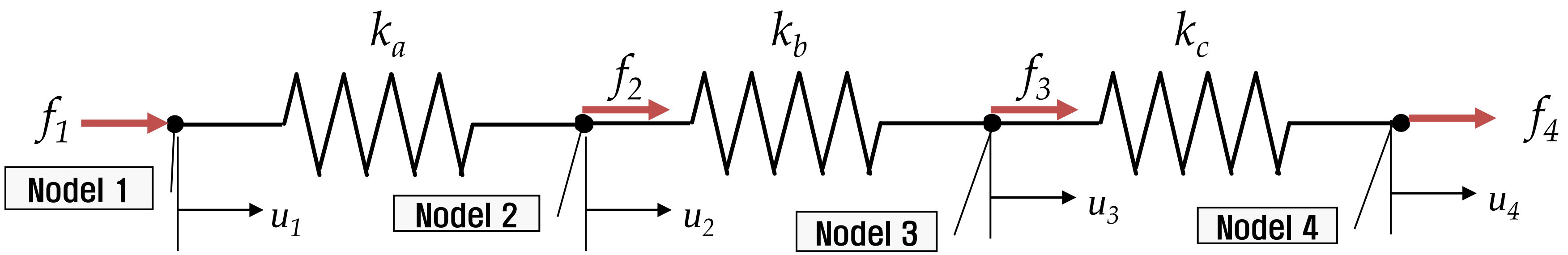


$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

※ superposition of stiffness matrix

Element : Bar (3 elements , 4 nodes)
- Structural analysis using direct stiffness method

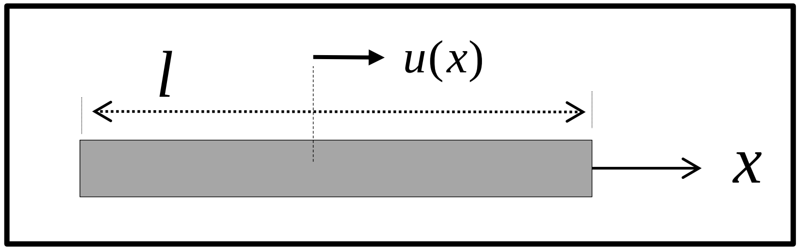
ex) 3 elements



$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 & 0 \\ -k_a & k_a & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & k_b & -k_b & 0 \\ 0 & -k_b & k_b & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & k_c & -k_c \\ 0 & 0 & -k_c & k_c \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}$$

$$= \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 & 0 \\ -k_a & k_a + k_b & -k_b & 0 \\ 0 & -k_b & k_b + k_c & -k_c \\ 0 & 0 & -k_c & k_c \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}$$

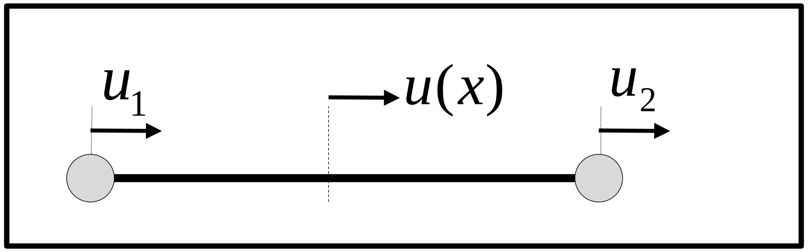
Element : Bar - Finite Element Method



discretization

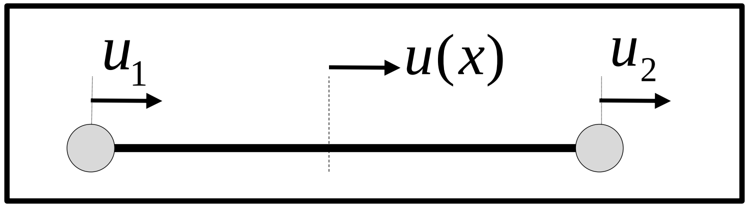
finite element method

1 element , 2 nodes



assume: $u(x) = c_1 + c_2 x \quad , u(0) = u_1, \;\; u(l) = u_2$

Element : Bar - Finite Element Method



assume: $u(x) = c_1 + c_2 x \ , u(0) = u_1, \ u(l) = u_2$

$$u(0) = c_1 \Rightarrow c_1 = u_1$$

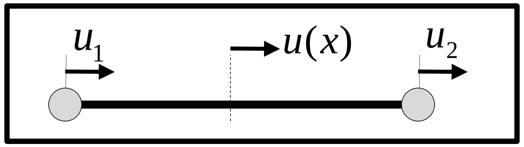
$$u(l) = c_1 + c_2 l \Rightarrow c_2 = \frac{u_2 - u_1}{l}$$

$$\therefore u(x) = u_1 + \left(\frac{u_2 - u_1}{l} \right) x$$

$$or, \ u(x) = \left(1 - \frac{x}{l} \right) u_1 + \frac{x}{l} u_2$$

Element : Bar - Finite Element Method

$$u(x) = \left(1 - \frac{x}{l}\right) u_1 + \frac{x}{l} u_2$$



$$u(x) = \begin{bmatrix} 1 - \frac{x}{l} & \frac{x}{l} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

differentiation with respect to x

$$\frac{du(x)}{dx} = \begin{bmatrix} -\frac{1}{l} & \frac{1}{l} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

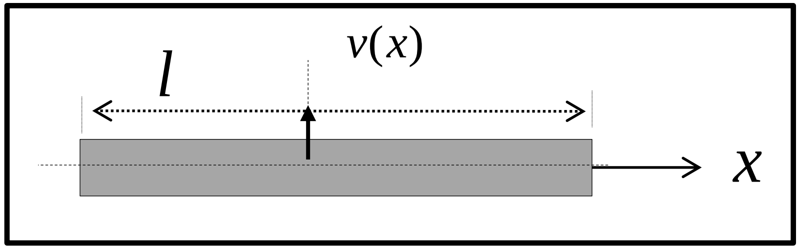
$$\therefore u(x) = \mathbf{N}\,\mathbf{d}, \quad \frac{du(x)}{dx} = \mathbf{B}\,\mathbf{d}$$

$$where\ \mathbf{N} = \begin{bmatrix} 1 - \frac{x}{l} & \frac{x}{l} \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} -\frac{1}{l} & \frac{1}{l} \end{bmatrix}, \ \mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Element : Beam - Finite Element Method

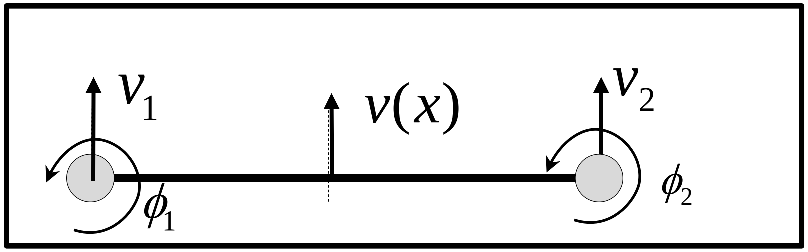
Variational Method

$$\delta\int_0^l \left[\frac{EI}{2}\left(\frac{d^2v}{dx^2}\right)^2 + (fv) \right] dx$$



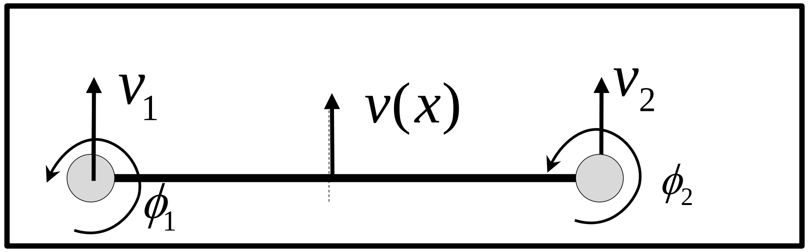
finite element method ↓ **discretization**

1 element , 2 nodes



assume: $v(x) = c_0 + c_1x + c_2x^2 + c_3x^3$, $v(0) = v_1,\ v(l) = v_2$, $\frac{dv}{dx}(0) = \phi_1,\ \frac{dv}{dx}(l) = \phi_2$

Element : Beam - Finite Element Method



assume: $v(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$, $v(0) = v_1,\ v(l) = v_2$, $\dfrac{dv}{dx}(0) = \phi_1,\ \dfrac{dv}{dx}(l) = \phi_2$

$$v(0) = c_0 \quad \Rightarrow \quad c_0 = v_1$$

$$v(l) = c_0 + c_1 l + c_2 l^2 + c_3 l^3 = v_2$$

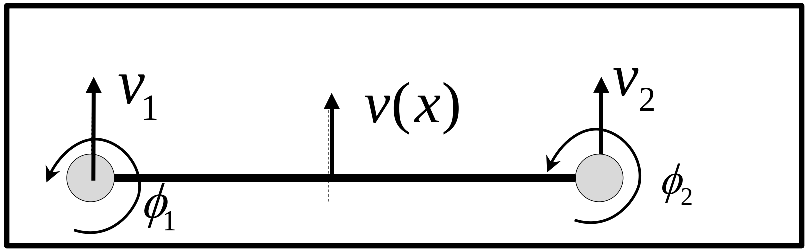
$$\frac{dv}{dx}(0) = c_1 \quad \Rightarrow \quad c_1 = \phi_1$$

$$\frac{dv}{dx}(l) = c_1 + 2c_2 l + 3c_3 l^2 = \phi_2$$

$$\Rightarrow \quad c_2 = -\frac{3}{l^2}(v_1 - v_2) - \frac{1}{l}(2\phi_1 + \phi_2)$$

$$c_3 = \frac{2}{l^3}(v_1 - v_2) + \frac{1}{l^2}(\phi_1 + \phi_2)$$

Element : Beam - Finite Element Method



assume: $v(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$, $v(0) = v_1$, $v(l) = v_2$

, $\dfrac{dv}{dx}(0) = \phi_1$, $\dfrac{dv}{dx}(l) = \phi_2$

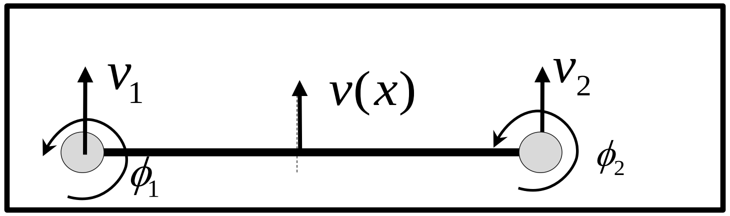
$$c_0 = v_1 \quad , c_1 = \phi_1 \quad , c_2 = -\frac{3}{l^2}(v_1 - v_2) - \frac{1}{l}(2\phi_1 + \phi_2)$$

$$c_3 = \frac{2}{l^3}(v_1 - v_2) + \frac{1}{l^2}(\phi_1 + \phi_2)$$

$$v(x) = v_1 + \phi_1 x + \left[-\frac{3}{l^2}(v_1 - v_2) - \frac{1}{l}(2\phi_1 + \phi_2)\right]x^2 + \left[\frac{2}{l^3}(v_1 - v_2) + \frac{1}{l^2}(\phi_1 + \phi_2)\right]x^3$$

$$or\ v(x) = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)v_1 + \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)\phi_1 + \frac{1}{l^3}(-2x^3 + 3x^2 l)v_2 + \frac{1}{l^3}(x^3 l - x^2 l^2)\phi_2$$

Element : Beam - Finite Element Method



$$v(x) = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)v_1 + \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)\phi_1 + \frac{1}{l^3}(-2x^3 + 3x^2 l)v_2 + \frac{1}{l^3}(x^3 l - x^2 l^2)\phi_2$$

$\Downarrow$

$$v(x) = \begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix} \begin{bmatrix} v_1 \\ \phi_1 \\ v_2 \\ \phi_2 \end{bmatrix}$$

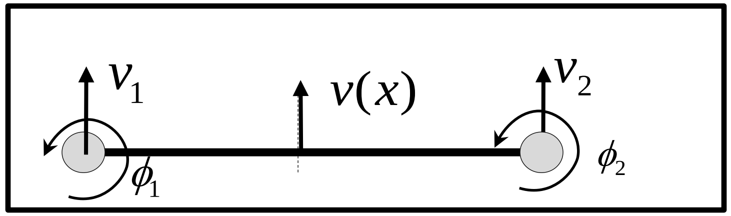
$$N_1 = \frac{1}{l^3}(2x^3 - 3x^2 l + l^3)$$

$$N_2 = \frac{1}{l^3}(x^3 l - 2x^2 l^2 + x l^3)$$

$$N_3 = \frac{1}{l^3}(-2x^3 + 3x^2 l)$$

$$N_4 = \frac{1}{l^3}(x^3 l - x^2 l^2)$$

Element : Beam - Finite Element Method



$$v(x)=\frac{1}{l^3}(2x^3-3x^2l+l^3)v_1+\frac{1}{l^3}(x^3l-2x^2l^2+xl^3)\phi_1+\frac{1}{l^3}(-2x^3+3x^2l)v_2+\frac{1}{l^3}(x^3l-x^2l^2)\phi_2$$

⇩

$$v(x)=\begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix}\begin{bmatrix} v_1 \\ \phi_1 \\ v_2 \\ \phi_2 \end{bmatrix}$$

$$N_1=\frac{1}{l^3}(2x^3-3x^2l+l^3) \qquad N_2=\frac{1}{l^3}(x^3l-2x^2l^2+xl^3)$$

$$N_3=\frac{1}{l^3}(-2x^3+3x^2l) \qquad N_4=\frac{1}{l^3}(x^3l-x^2l^2)$$

⇩ **differentiation with respect to** x **twice**

$$\frac{d^2v(x)}{dx^2}=\begin{bmatrix} B_1 & B_2 & B_3 & B_4 \end{bmatrix}\begin{bmatrix} v_1 \\ \phi_1 \\ v_2 \\ \phi_2 \end{bmatrix}$$

$$B_1=\frac{1}{l^3}(12x-6l) \qquad B_2=\frac{1}{l^3}(6xl-4l^2)$$

$$B_3=\frac{1}{l^3}(-12x+6l) \qquad B_4=\frac{1}{l^3}(6xl-2l^2)$$

$$\therefore v(x)=\mathbf{N}\mathbf{d}, \quad \frac{d^2v(x)}{dx^2}=\mathbf{B}\mathbf{d}$$

Element : Bar

E : Young's Modulus, A : sectional area

Element: Differential Equation → Variational Method → •Discretization •Approximation → Finite Element Method

$\mathbf{M}\ddot{\mathbf{x}} = \sum \mathbf{F}$, *where* $\ddot{\mathbf{x}} = 0$ $\qquad\qquad \mathbf{Kd} = \mathbf{F}$

Bar

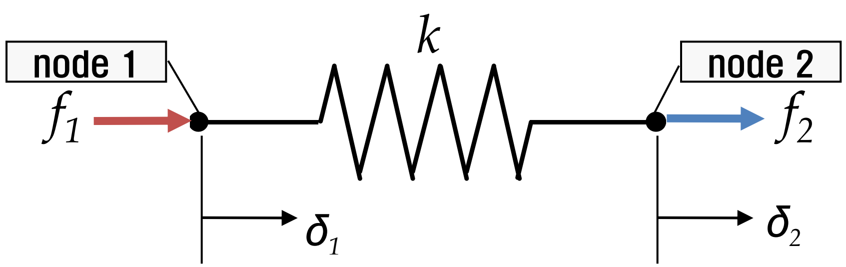
$$EA\frac{d^2u(x)}{dx^2} + f(x) = 0 \Rightarrow \delta\int_0^l \left[\frac{EA}{2}\left(\frac{du}{dx}\right)^2 - (f\,u)\right]dx = 0 \Rightarrow \frac{EA}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

$u(x) = a_0 + a_1 x$

! Notation

$$k = \frac{EA}{l}$$

bar element



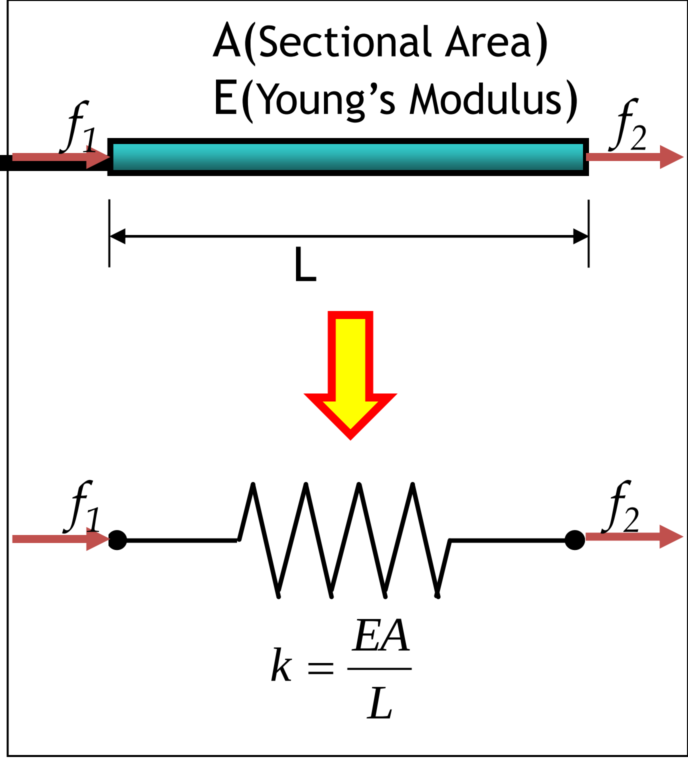
$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} k & -k \\ -k & k \end{bmatrix}\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$$

stiffness matrix

$$[f] = [K][\delta]$$

stiffness equation

Element : Bar - Linearity

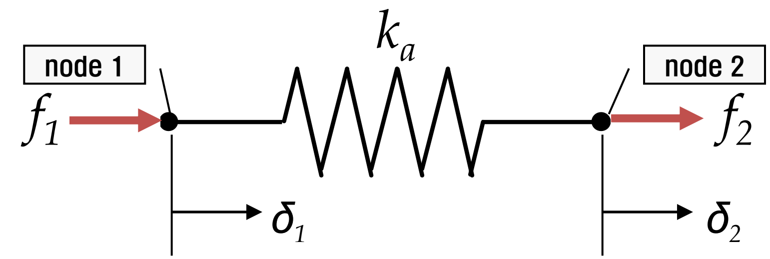


Linearity (α : Scalar)
$L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2)$
$L(\alpha\,\mathbf{v}_1) = \alpha \cdot L(\mathbf{v}_1)$
⇒ Definition of Linearity

▪ Bar - Linearity (α : Scalar)

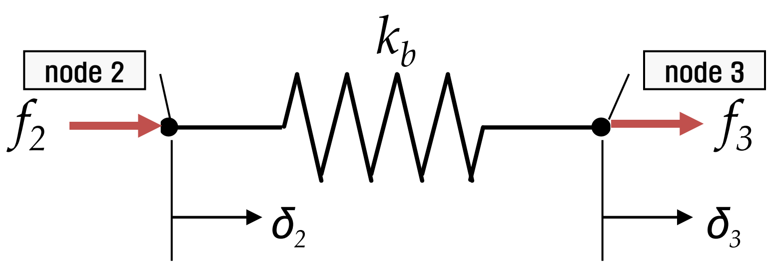
$f(\delta_1) = k\delta_1 \quad , \quad f(\delta_2) = k\delta_2$	$f(\delta_1) = k\delta_1$
$f(\delta_1) + f(\delta_2) = k\delta_1 + k\delta_2 = k(\delta_1 + \delta_2)$	$f(\alpha \cdot \delta_1) = k(\alpha\delta_1) = k\alpha\delta_1$
$f(\delta_1 + \delta_2) = k(\delta_1 + \delta_2)$	$= \alpha(k\delta_1) = \alpha \cdot f(\delta_1)$
$\therefore f(\delta_1 + \delta_2) = f(\delta_1) + f(\delta_2)$	$\therefore f(\alpha \cdot \delta_1) = \alpha \cdot f(\delta_1)$

Element : Bar - Superposition



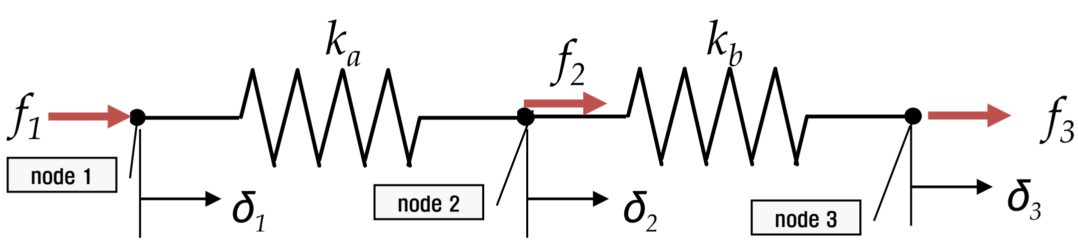
$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

+



$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

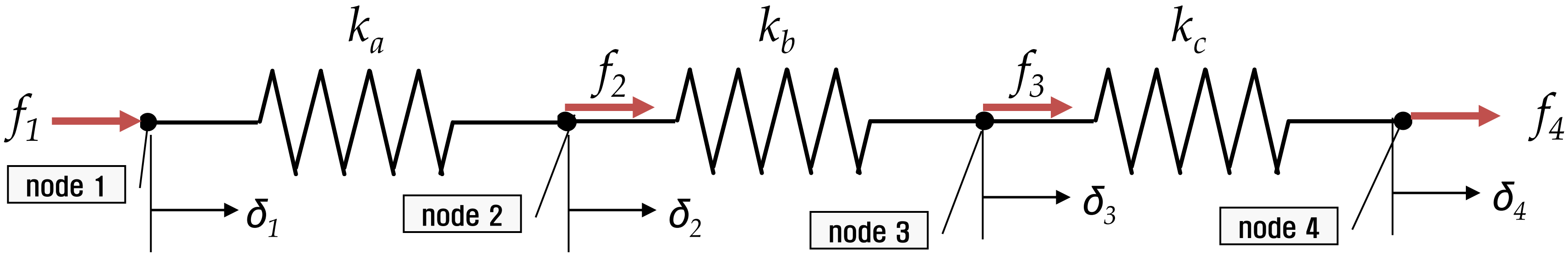
=



$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

Element : Bar - Superposition

ex.) Find a stiffness equation of the following system:



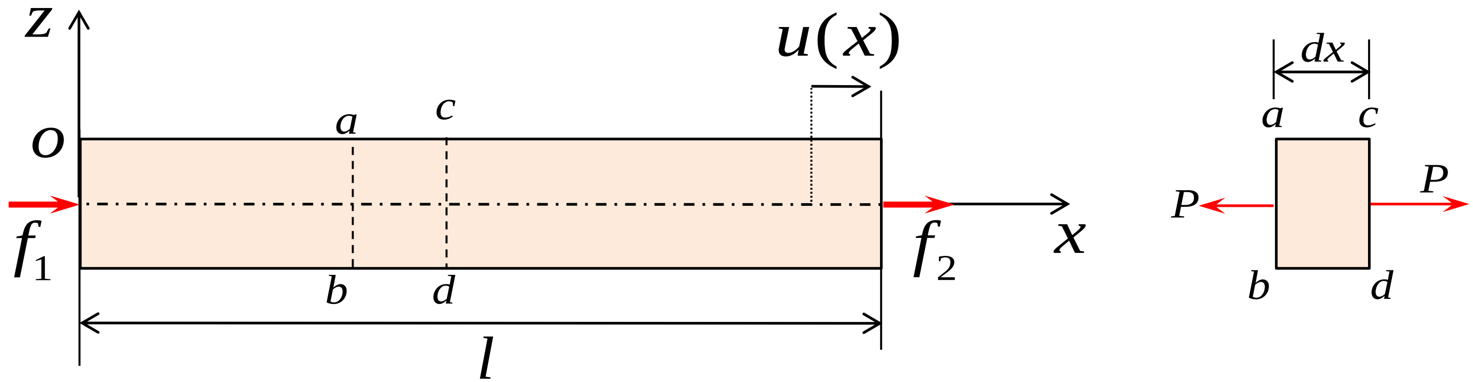
$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 & 0 \\ -k_a & k_a & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & k_b & -k_b & 0 \\ 0 & -k_b & k_b & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & k_c & -k_c \\ 0 & 0 & -k_c & k_c \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

$$= \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 & 0 \\ -k_a & k_a + k_b & -k_b & 0 \\ 0 & -k_b & k_b + k_c & -k_c \\ 0 & 0 & -k_c & k_c \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix}$$

Reference

EXPLANATION ABOUT BAR ELEMENT IN KOREAN

Element : Bar - Differential Eq.



길이가 l인 bar의 양 끝에 힘 f1, f2가 작용하고, distributed force가 작용하지 않을 때, 미분 방정식은 아래와 같이 유도 됨

$$P = A(x)\sigma$$

$$\sigma = E\varepsilon$$

$$P = EA(x)\varepsilon$$

$$\varepsilon = \frac{du(x)}{dx}$$

$$P = EA(x)\frac{du(x)}{dx}$$

From the force equilibrium, "P" dose not change along the x-axis

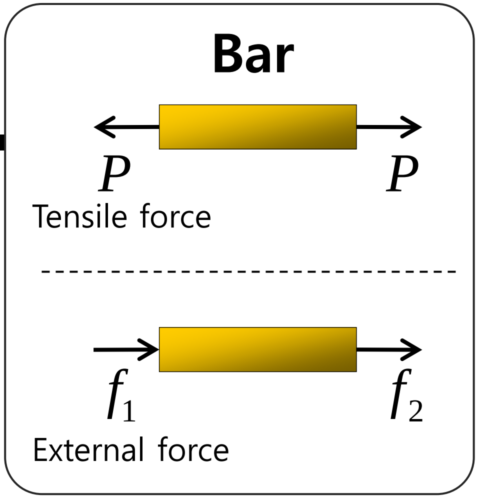
$$\frac{dP}{dx} = \frac{d}{dx}\left(EA(x)\frac{du(x)}{dx}\right) = 0$$

If A(x) is constant "A"

$$EA\frac{d^2u(x)}{dx^2} = 0$$

주의: 여기에서 "P"는 외력이 아닌 외력 f1, f2에 의해서 발생되는 **stress resultants**임

Element : Bar (1 element , 2 nodes)
- Galerkin's Residual Method



이전 장에서 유도한 미분 방정식은 **Bar**의 모든 미소 **element**에 대하여 성립해야 하므로 **Galerkin's residual method**를 적용하여 아래와 같이 식을 유도할 수 있다.

$$\int_0^l AE\frac{d^2u(x)}{dx^2}N_i\,dx=0 \quad ,(i=1,2)$$

$$where,\ u(x)=\mathbf{N}\mathbf{d}=N_1u_1+N_2u_2 \quad ,N_1=1-\frac{x}{l},\ N_2=\frac{x}{l}$$

integration by parts

$$\left[N_iAE\frac{du}{dx}\right]_0^l-\int_0^l AE\frac{du}{dx}\frac{dN_i}{dx}dx=0 \qquad \textbf{since} \quad \frac{du}{dx}=\frac{dN_1}{dx}u_1+\frac{dN_2}{dx}u_2=\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}\begin{bmatrix}u_1\\u_2\end{bmatrix}$$

$$AE\int_0^l\frac{dN_i}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\left[N_iAE\frac{du}{dx}\right]_0^l \quad ,(i=1,2)$$

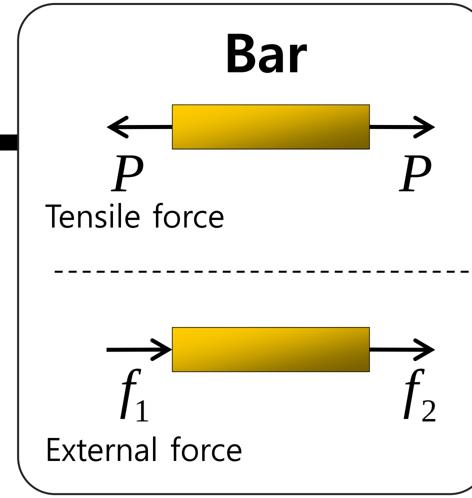
$$\Rightarrow\begin{cases} i=1: \quad AE\int_0^l\frac{dN_1}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\boxed{\left[N_1AE\frac{du}{dx}\right]_0^l} \Rightarrow \cancel{N_1}AE\left.\frac{du}{dx}\right|_{x=l}-N_1AE\left.\frac{du}{dx}\right|_{x=0} \Rightarrow -N_1P\big|_{x=0} \Rightarrow -P \\ i=2: \quad AE\int_0^l\frac{dN_2}{dx}\begin{bmatrix}-\frac{1}{l} & \frac{1}{l}\end{bmatrix}dx\begin{bmatrix}u_1\\u_2\end{bmatrix}=\boxed{\left[N_2AE\frac{du}{dx}\right]_0^l} \Rightarrow N_2AE\left.\frac{du}{dx}\right|_{x=l}-\cancel{N_2}AE\left.\frac{du}{dx}\right|_{x=0} \Rightarrow N_2P\big|_{x=l} \Rightarrow P \end{cases}$$

since $N_1(0)=1,\ N_1(l)=0$, $N_2(0)=0,\ N_2(l)=1$ **since** $AE\frac{du}{dx}=AE\varepsilon=A\sigma=P$

여기서 주의할 점은 $AE\frac{du}{dx}=P$ 라는 것이다.

즉, $AE\left.\frac{du}{dx}\right|_{x=0}$, $AE\left.\frac{du}{dx}\right|_{x=l}$ 는 각각 **x=0**과 **x=l** 에서의 **tensile force**를 의미하는 것이다. <u>외력 **f_1, f_2**가 아님!!!</u>

Element : Bar (1 element , 2 nodes) - Galerkin's Residual Method



Bar - Galerkin's Residual Method

$$\int_0^l AE \frac{d^2u(x)}{dx^2} N_i \, dx = 0 \quad ,(i = 1, 2)$$

$$where, \; u(x) = \mathbf{N}\mathbf{d} = N_1 u_1 + N_2 u_2 \quad , N_1 = 1 - \frac{x}{l}, \; N_2 = \frac{x}{l}$$

Differential Equation

$$EA\frac{d^2u(x)}{dx^2} = 0$$

integration by parts

$$\begin{cases} AE\int_0^l \frac{dN_1}{dx}\begin{bmatrix} -\frac{1}{l} & \frac{1}{l}\end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = -P \\ AE\int_0^l \frac{dN_2}{dx}\begin{bmatrix} -\frac{1}{l} & \frac{1}{l}\end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = P \end{cases}$$

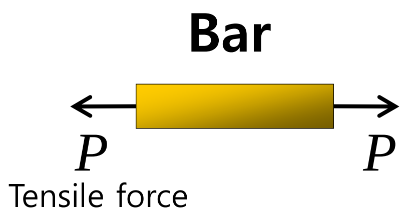
$$\Rightarrow \begin{cases} AE\int_0^l \begin{bmatrix} -\frac{1}{l} \end{bmatrix}\begin{bmatrix} -\frac{1}{l} & \frac{1}{l}\end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = AE\frac{1}{l^2}\int_0^l \begin{bmatrix} 1 & -1 \end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = AE\frac{1}{l^2}\begin{bmatrix} l & -l \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{AE}{l}\begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ AE\int_0^l \begin{bmatrix} \frac{1}{l} \end{bmatrix}\begin{bmatrix} -\frac{1}{l} & \frac{1}{l}\end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = AE\frac{1}{l^2}\int_0^l \begin{bmatrix} -1 & 1 \end{bmatrix} dx \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = AE\frac{1}{l^2}\begin{bmatrix} -l & l \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{AE}{l}\begin{bmatrix} -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \end{cases}$$

$$\therefore \begin{cases} \frac{AE}{l}\begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = -P \\ \frac{AE}{l}\begin{bmatrix} -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = P \end{cases}$$

$$\Rightarrow \boxed{\therefore \mathbf{Kd} = \mathbf{F} \quad where \; , \mathbf{K} = \frac{EA}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} -P \\ P \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}}$$

문제 정의에서 처음에 주어진 것이 외력 **f1**과 **f2**이다. 따라서 **tensile force**인 **-P, P**를 **f1**과 **f2**를 이용하여 계산하고, 치환하여 표현한 것이 위 식의 파란색 Column matrix이다.

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method



Differential Equation

$$EA\frac{d^2u(x)}{dx^2}=0 \qquad 0<x<L$$

Boundary Condition

$$EA\left.\frac{du}{dx}\right|_{x=0}=P \quad , \quad EA\left.\frac{du}{dx}\right|_{x=L}=P$$

distributed load가 작용하지 않는 bar에 대하여 미분 방정식을 세우고(미분 방정식에 f(x)가 포함되어 있지 않음), Galerkin's residual method를 적용한다. 여기서 element를 2개로 정의할 예정인데, 그렇다면 두 element 사이의 node에 작용하는 외력이 어떻게 표현 되는지 살펴보자

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

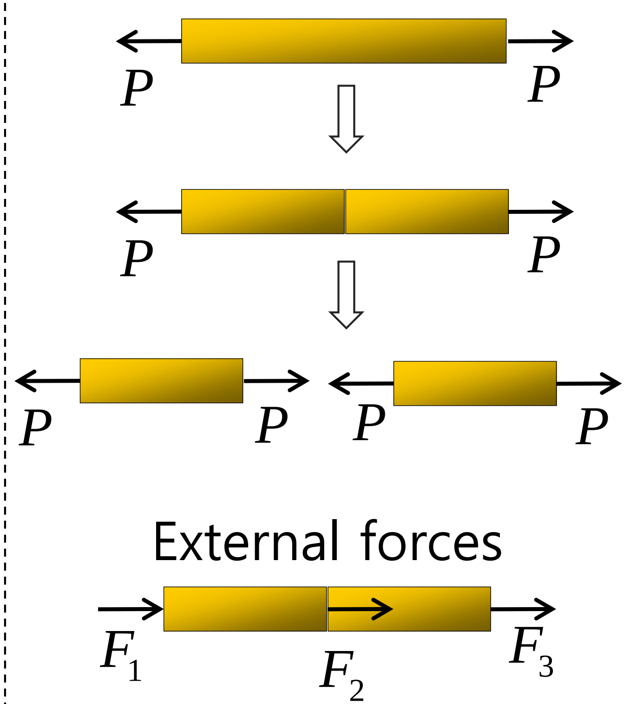
The weighted residual form:

$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = \frac{EA}{L/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} \left[N_1 EA \dfrac{d\hat{u}}{dx} \right]_0^l \\ \left[N_2 EA \dfrac{d\hat{u}}{dx} \right]_0^l \\ \left[N_3 EA \dfrac{d\hat{u}}{dx} \right]_0^l \end{bmatrix} = \begin{bmatrix} -P \\ 0 \\ P \end{bmatrix}$$

최종 유도한 식은 위와 같다. 여기에서 두 element 사이에 작용하는 외력 F_2, 즉 중간 node에 작용하는 외력은 0으로 유도 되었다. 이는 bar의 양 끝 단에 작용하는 힘 F_1, F_3 이외의 외력(분포 하중)은 없다고 가정하고 미분 방정식을 유도한 결과이다.

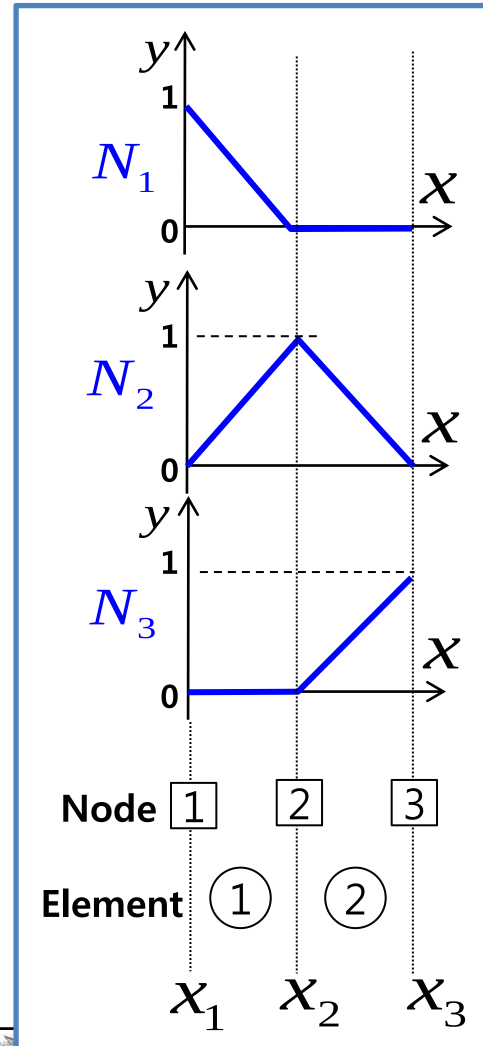


$$\therefore F_1 = -P$$

$$F_2 = -P + P = 0$$

$$F_3 = P$$

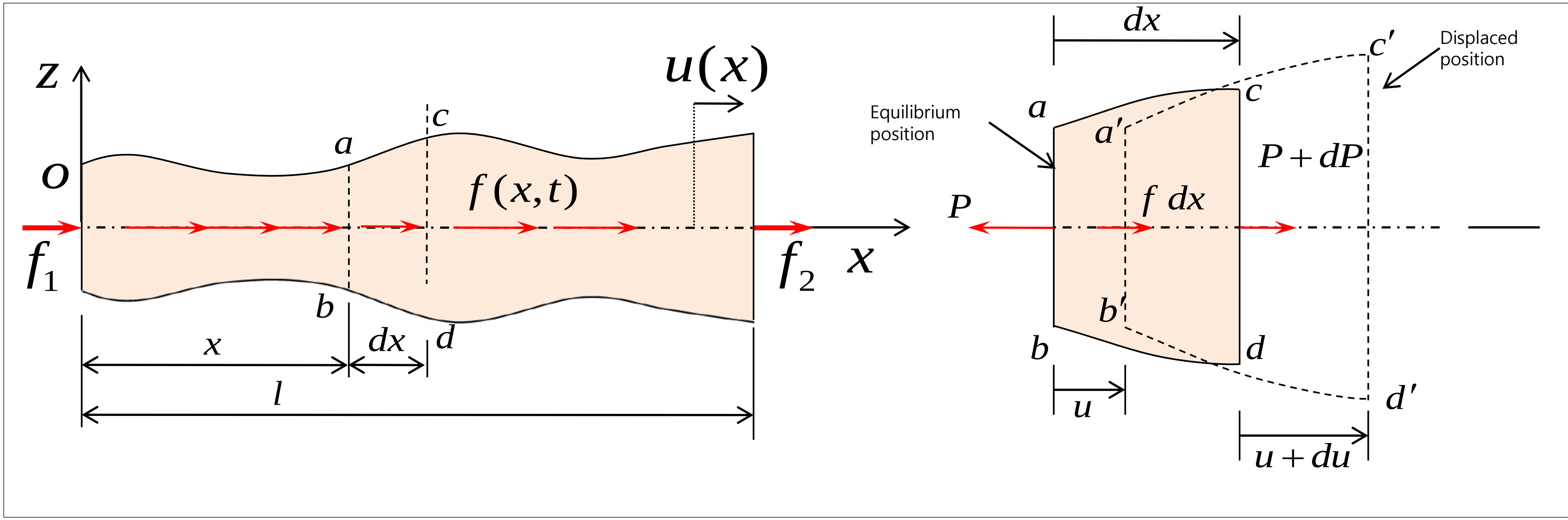
The number of the elements E is 2



Chapter 1. Element : Bar

Element : Bar - Differential Eq.

참고자료



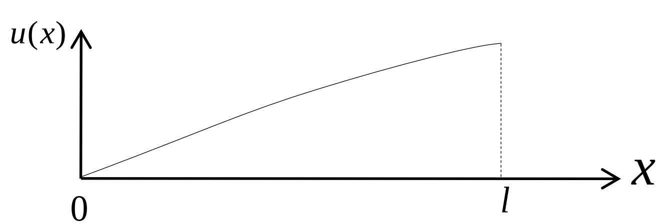
P: the (internal) forces acting on the cross sections of a small element of the bar of length dx

$f(x,t)$: external force per unit length, distributed force

f_1, f_2: concentrated forces exerted on the ends of the bar

ρ: density, E: Young's Modulus, A: sectional area

an example of displacement in x-direction at x



"Longitudinal Vibration of a Bar or Rod : Rao,S.S., Mechanical Vibrations, Fourth Edition, Prentice Hall, 2004, pp597-600

Element : Bar - Differential Eq.

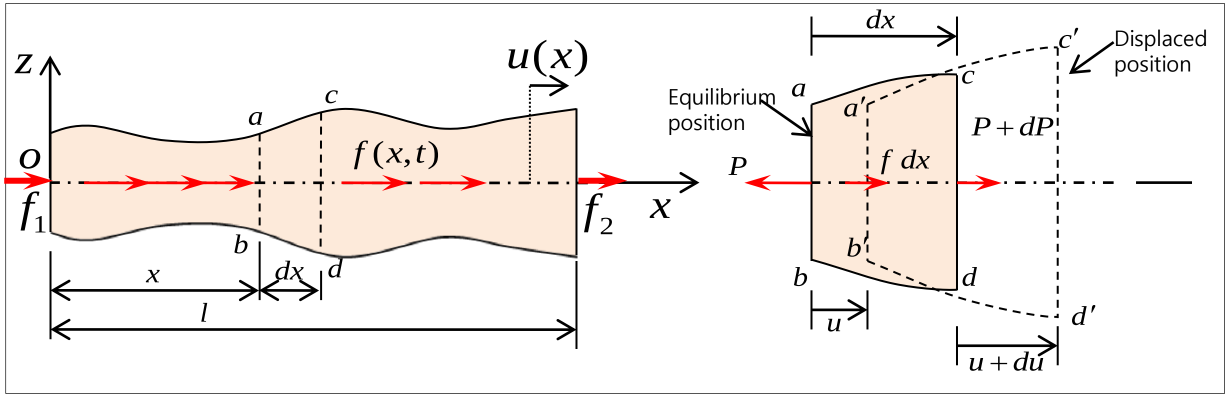
참고자료

A Bar in Axial Vibration

$$\sum F = ma$$

$$(P+dP) + f\,dx - P = \rho A(x)\,dx \frac{\partial^2 u(x,t)}{\partial t^2}$$

$$dP + f\,dx = \rho A(x)\,dx \frac{\partial^2 u(x,t)}{\partial t^2}$$



$f(x,t)$: external force per unit length, distributed force

f_1, f_2: concentrated forces exerted on the ends of the bar

ρ: density, E: Young's Modulus, A: sectional area

$$P = \sigma A(x) = EA(x)\varepsilon = EA(x)\frac{\partial u(x,t)}{\partial x}$$ 'constitutive equation'

$$dP = \frac{\partial P}{\partial x}dx = \frac{\partial}{\partial x}\left(EA(x)\frac{\partial u(x,t)}{\partial x}\right)dx$$

$$\frac{\partial}{\partial x}\left(EA(x)\frac{\partial u(x,t)}{\partial x}\right)dx + f(x,t)\,dx = \rho A(x)\,dx\frac{\partial^2 u(x,t)}{\partial t^2}$$

$$if \quad A(x) = A \quad : const.$$

$$EA\frac{\partial^2 u(x,t)}{\partial x^2} + f(x,t) = \rho A\frac{\partial^2 u(x,t)}{\partial t^2}$$

dynamics (vibration)

subjected to

$$\begin{cases} u(x,t=0) = u_0(x), & 0 \le x \le l \\ \dfrac{\partial u}{\partial t}(x,t=0) = \dot{u}_0(x), & 0 \le x \le l \end{cases}$$ → **I.V.P**

$$\begin{cases} u(0,t) = 0, & t > 0 \\ AE\dfrac{\partial u}{\partial x}(l,t) = 0 \text{ or } \dfrac{\partial u}{\partial x}(l,t) = 0, & t > 0 \end{cases}$$ → **B.V.P**

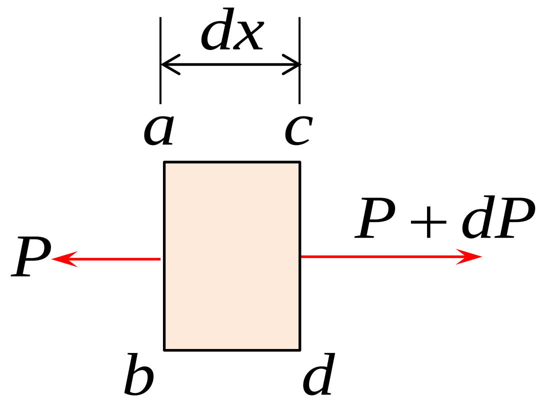
at the free end , axial force

$$\frac{\partial u}{\partial t} = 0$$

$$EA\frac{d^2 u(x)}{dx^2} + f(x) = 0$$ → **statics**

"Longitudinal Vibration of a Bar or Rod : Rao,S.S., Mechanical Vibrations, Fourth Edition, Prentice Hall, 2004, pp597-600

Element : Bar - Differential Eq.



Consider that the tensile force is not constant and there is no other force such as distributed force. Then, this infinitesimal element will move until the tensile forces acting on the opposite side of the element become same. Therefore, if the element is in equilibrium state, the tensile force should be constant.

Memo

1) distributed force인 f(x)가 존재하는 경우
이 경우에 대하여 Rao는 distributed force인 f(x)가 있다고 가정하여
-P + f(x)dx + P + dP = 질량 * 가속도
라는 식을 세웠습니다.

여기에서 statics를 고려해야 하므로, 가속도가 0이라고 가정하면
-P + f(x)dx + P + dP = 0
f(x)dx + dP = 0
라는 식이 되어 P의 변화량인 dP를 f(x)dx가 상쇄시켜줄 수 있습니다.
반대로 해석하면, f(x)dx가 P의 변화인 dP를 야기시킨다고 해석할 수 도 있을 것입니다.

2) distributed force인 f(x)가 존재하지 않는 경우
이 경우에 대하여 Rao가 세운 식에서 f(x)=0이라고 가정하여
-P + P + dP = 질량 * 가속도
라는 식을 세울 수 있으며, statics를 고려하기 위해, 가속도가 0이라고 가정하면
-P + P + dP = 0
dP = 0
이라는 식이 유도 됩니다.
즉, distributed force인 f(x)가 존재하지 않고, statics를 고려하면
dP는 자동으로 0이되어 P=constant라는 결론을 내릴 수 있습니다.

Element : Bar (E elements , E+1 nodes) - Solving D/E using Galerkin's Residual Method

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

Bar

Differential Equation

$$EA\frac{d^2u(x)}{dx^2} = 0 \qquad 0 < x < L$$

Boundary Condition

$$EA\left.\frac{du}{dx}\right|_{x=0} = P \quad , \quad EA\left.\frac{du}{dx}\right|_{x=L} = P$$

Governing equation $A(u) = \mathscr{L}u + p = 0 \quad in \;\; \Omega$

$$EA\frac{d^2u}{dx^2} = 0 \quad \rightarrow \quad \boxed{A(u) = EA\frac{d^2u}{dx^2} = 0 \quad in \;\; \Omega}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

$$A(u) = EA\frac{d^2u}{dx^2} = 0 \quad in \;\; 0 < x < L$$

$$u \approx \hat{u} = \sum_{m=1}^{E+1} u_m N_m$$

,where E is the number of the elements

The residual in domain:

$$\mathbf{R}_\Omega = A(\hat{u}) - \cancel{A(u)} = EA\frac{d^2\hat{u}}{dx^2} \quad in \;\; 0 < x < L$$

The weighted residual form:

$$\int_0^L W_l \mathbf{R}_\Omega dx = 0,\; l = 1, 2, ..., E+1$$

$$\int_0^L W_l \left(EA\frac{d^2\hat{u}}{dx^2} \right) dx = 0,\; l = 1, 2, ..., E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$u \approx \hat{u} = \sum_{m=1}^{E+1} u_m N_m$$

,where E is the number of the elements

$$\int_0^L W_l \left(EA \frac{d^2\hat{u}}{dx^2} \right) dx = 0,\ l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$\int_0^L W_l EA \frac{d^2\hat{u}}{dx^2} dx = 0,\ l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$EA \int_0^L W_l \frac{d^2\hat{u}}{dx^2} dx = 0,\ l = 1, 2, \ldots, E+1$$

$\downarrow$ **Integration by parts**

$$-EA \int_0^L \frac{dW_l}{dx} \frac{d\hat{u}}{dx} dx + EA \left[W_l \frac{d\hat{u}}{dx} \right]_0^L = 0,\ l = 1, 2, \ldots, E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$-EA\int_0^L \frac{dW_l}{dx}\frac{d\hat{u}}{dx}dx + EA\left[W_l \frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, \ldots, E+1$$

$$\downarrow$$

$$EA\int_0^L \frac{dW_l}{dx}\frac{d\sum_{m=1}^{E+1} u_m N_m}{dx}dx - EA\left[W_l \frac{d\hat{u}}{dx}\right]_0^L = 0,\ l = 1, 2, \ldots, E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\int_0^L \frac{dW_l}{dx}\frac{d\sum_{m=1}^{E+1}u_m N_m}{dx}dx - EA\left[W_l\frac{d\hat{u}}{dx}\right]_0^L = 0,\ l=1,2,\ldots,E+1$$

$\downarrow$ **Galerkin methods** $W_l = N_l$

$$EA\int_0^L \frac{dN_l}{dx}\frac{d\sum_{m=1}^{E+1}u_m N_m}{dx}dx - EA\left[N_l\frac{d\hat{u}}{dx}\right]_0^L = 0,\ l=1,2,\ldots,E+1$$

$\downarrow$

$$EA\sum_{m=1}^{E+1}\int_0^L u_m \frac{dN_l}{dx}\frac{dN_m}{dx}dx - \left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L = 0,\ l=1,2,\ldots,E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\sum_{m=1}^{E+1}\int_0^L u_m \frac{dN_l}{dx}\frac{dN_m}{dx}dx-\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L=0,\ l=1,2,\ldots,E+1$$

$$\downarrow \quad m=1,2,\ldots,E+1$$

$$EA\left(\int_0^L u_1\frac{dN_l}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2\frac{dN_l}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1}\frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right)$$

$$-\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L=0,\quad l=1,2,\ldots,E+1$$

$$\downarrow$$

$$EA\left(\int_0^L u_1\frac{dN_l}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2\frac{dN_l}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1}\frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right)=$$

$$+\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L,\quad l=1,2,\ldots,E+1$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^L u_1 \frac{dN_l}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_l}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1} \frac{dN_l}{dx}\frac{dN_{E+1}}{dx}dx\right)=$$

$$+\left[N_l EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\downarrow\ l=1,2,\ldots,E+1$$

$$EA\left(\int_0^L u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx\right)=\left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$EA\left(\int_0^L u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx\right)=\left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\vdots$$

$$EA\left(\int_0^L u_1 \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx+\int_0^L u_2 \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx+\ldots+\int_0^L u_{E+1} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx\right)=\left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

$$EA\left(\int_0^L u_1 \frac{dN_1}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_1}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx\right) = \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$EA\left(\int_0^L u_1 \frac{dN_2}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_2}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx\right) = \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\vdots$$

$$EA\left(\int_0^L u_1 \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx + \int_0^L u_2 \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx + \ldots + \int_0^L u_{E+1} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx\right) = \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L$$

$$\downarrow$$

$$\underbrace{EA\begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{E+1} \end{bmatrix}}_{\mathbf{d}} = \underbrace{\begin{bmatrix} \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L \\ \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L \\ \vdots \\ \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L \end{bmatrix}}_{\mathbf{F}}$$

Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form:

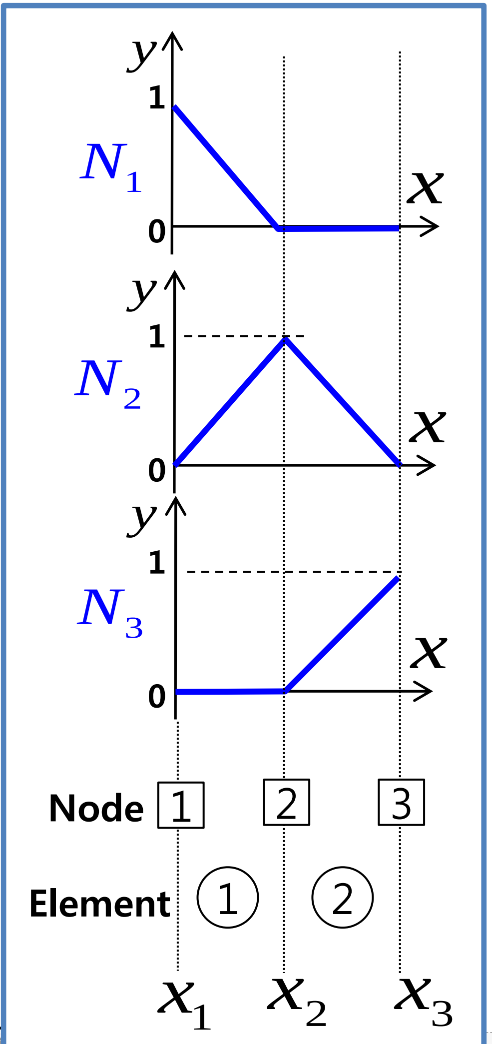
$$\mathbf{Kd} = \mathbf{F}$$

$$\mathbf{K} = EA\begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{E+1} \end{bmatrix} \qquad \mathbf{F} = \begin{bmatrix} \left[N_1 EA\frac{d\hat{u}}{dx}\right]_0^L \\ \left[N_2 EA\frac{d\hat{u}}{dx}\right]_0^L \\ \vdots \\ \left[N_{E+1} EA\frac{d\hat{u}}{dx}\right]_0^L \end{bmatrix}$$

N_i is corresponding to the 1st order B-spline basis functions

The number of the elements E is 2



Element : Bar (2 elements , 3 nodes)
- Solving D/E using Galerkin's Residual Method

The weighted residual form: $\mathbf{Kd} = \mathbf{F}$

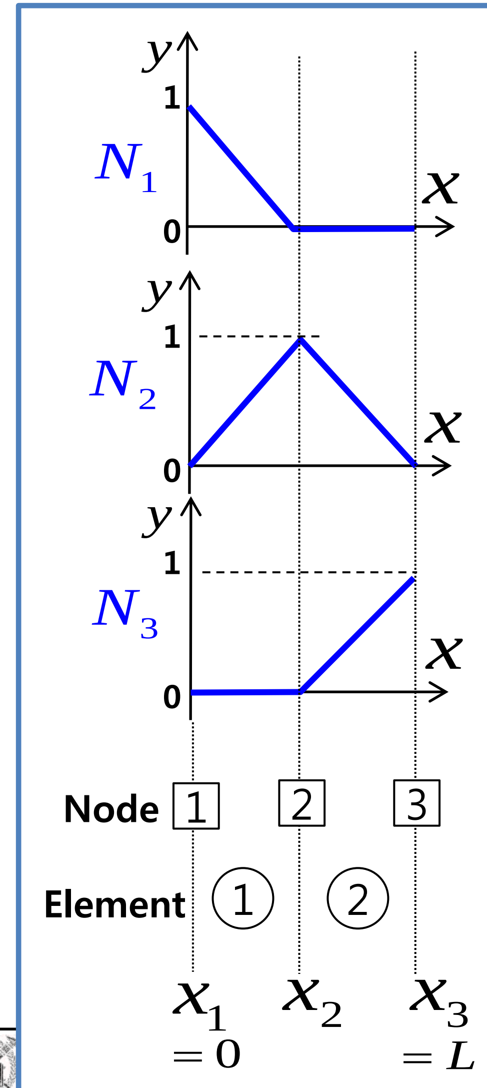
The number of the elements E is 2

$$\mathbf{K} = EA\begin{bmatrix} \int_0^L \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_0^L \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \int_0^L \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix}$$

$$\mathbf{K} = \mathbf{K}_1 + \mathbf{K}_2$$

$$\mathbf{K}_1 = EA\begin{bmatrix} \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_1}^{x_2} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_1}^{x_2} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_{x_1}^{x_2} \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_{x_1}^{x_2} \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \boxed{\int_{x_1}^{x_2}} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_1}^{x_2}}$$

$$\mathbf{K}_2 = EA\begin{bmatrix} \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_2}^{x_3} \frac{dN_1}{dx}\frac{dN_{E+1}}{dx}dx \\ \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_2}{dx}dx & \cdots & \int_{x_2}^{x_3} \frac{dN_2}{dx}\frac{dN_{E+1}}{dx}dx \\ \vdots & \vdots & \ddots & \vdots \\ \int_{x_2}^{x_3} \frac{dN_{E+1}}{dx}\frac{dN_1}{dx}dx & \int_{x_2}^{x_3} \frac{dN_{E+1}}{dx}\frac{dN_2}{dx}dx & \cdots & \boxed{\int_{x_2}^{x_3}} \frac{dN_{E+1}}{dx}\frac{dN_{E+1}}{dx}dx \end{bmatrix} \quad \boxed{\int_{x_2}^{x_3}}$$



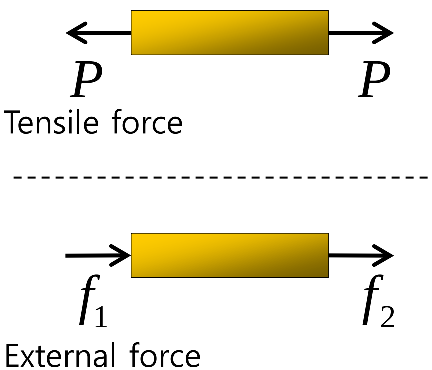
SUMMARY

Element : Bar (1 element , 2 nodes)

- Comparison between the Solutions of D/E using Galerkin's Residual Method and direct equilibrium approach

Solutions of D/E using Galerkin's Residual Method

Bar



Differential Equation

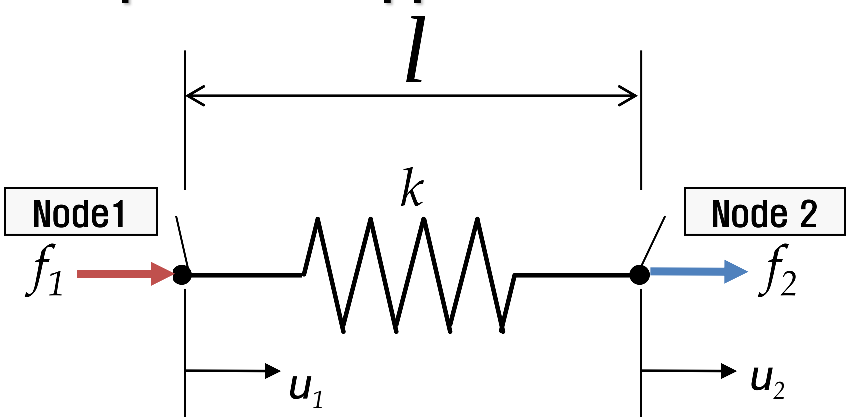
$$EA\frac{d^2u(x)}{dx^2}=0$$

$$u(x)=a_0+a_1x$$

Galerkin's residual method

$$\frac{EA}{l}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}=\begin{bmatrix}-P\\ P\end{bmatrix}=\begin{bmatrix}f_1\\ f_2\end{bmatrix}$$

Direct equilibrium approach



$$k=\frac{EA}{l}$$

Forces exerted on the nodes

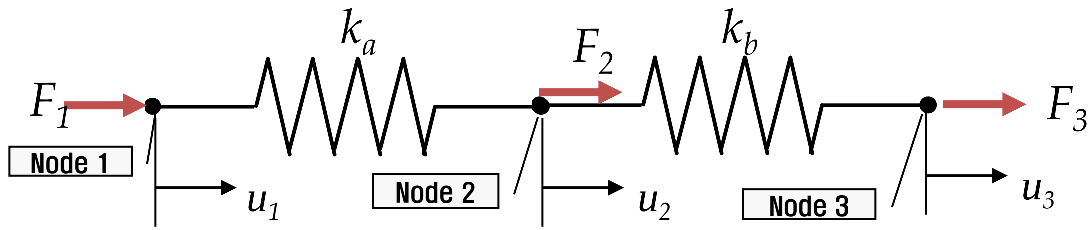
$$\begin{bmatrix}f_1\\ f_2\end{bmatrix}=\begin{bmatrix}k & -k\\ -k & k\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

stiffness matrix

$$[f]=[K][u]$$

External force and internal force for two spring assemblage

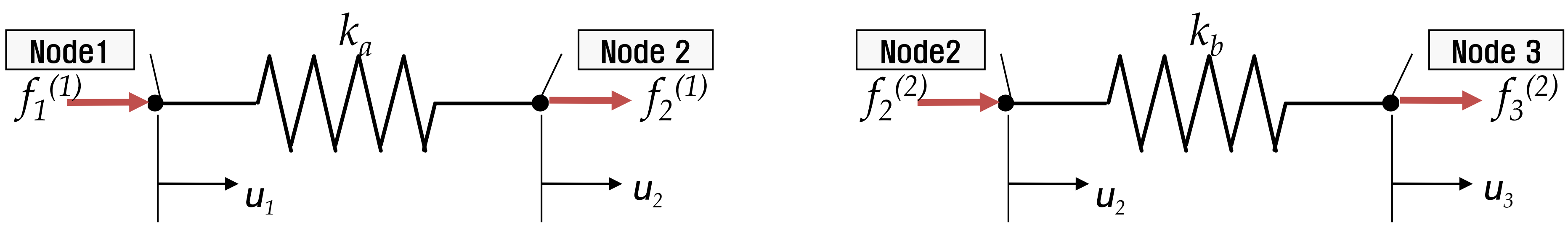
We will consider two spring assemblage.



F_1, F_2, and F_3 are **<u>external forces</u>** which are applied at node 1, 2, and 3 respectively.

<u>Free-body diagrams</u> of each element and nodes are shown as follows

element #1 **element #2**

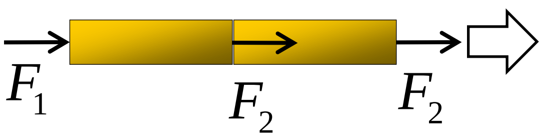


$f_1^{(1)}$, $f_2^{(1)}$, $f_2^{(2)}$, and $f_3^{(2)}$ are **<u>internal forces</u>**.

Based on the free-body diagrams, however, ***"f"*** can be regarded as external forces **<u>for each element</u>**.

Solutions of D/E using Galerkin's Residual Method

Bar



Differential Equation

$$EA\frac{d^2u(x)}{dx^2}=0$$

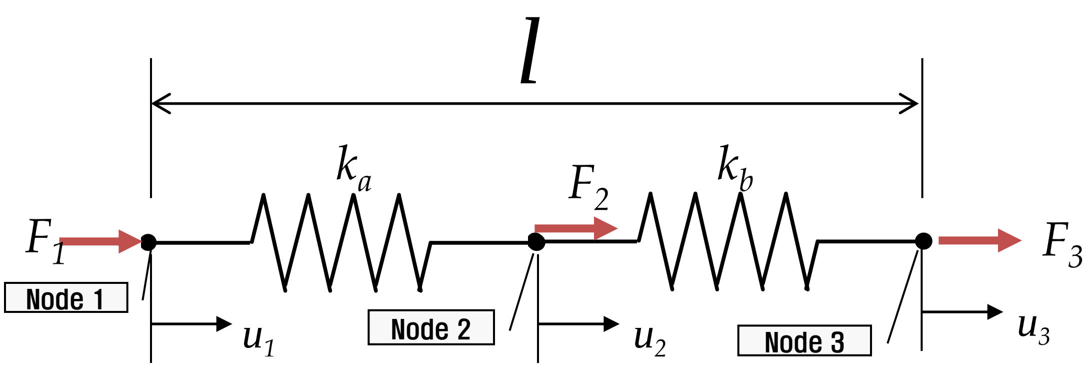
Galerkin's residual method

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \frac{EA}{l/2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1+1 & -1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

External force

Direct equilibrium approach

$$k=\frac{EA}{l/2}$$

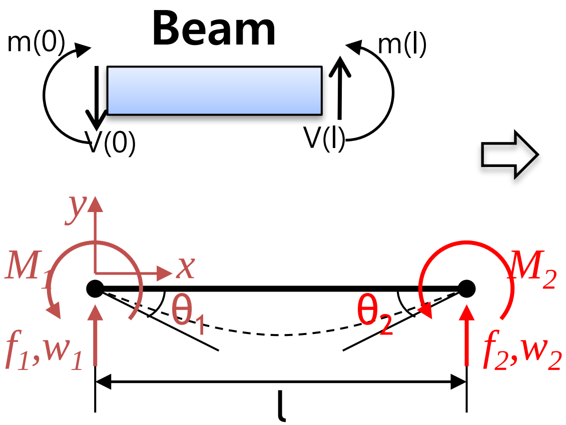


$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} k_a & -k_a & 0 \\ -k_a & k_a + k_b & -k_b \\ 0 & -k_b & k_b \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

F_1, F_2, F_3: Applied external force at each node.

※ superposition of stiffness matrix

Solutions of D/E using Galerkin's Residual Method



Differential Equation

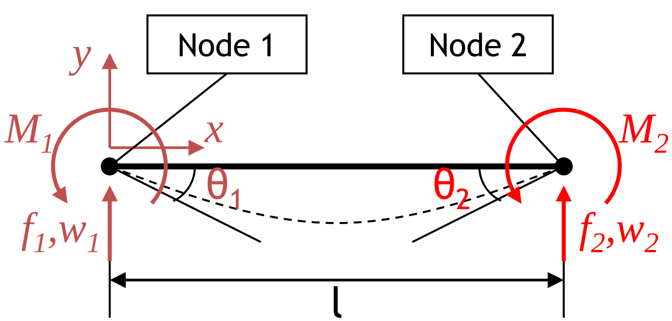
$$EI\frac{d^4w(x)}{dx^4}=0$$

$$w(x)=b_0+b_1x+b_2x^2+b_3x^3$$

Galerkin's residual method

$$\frac{2EI}{l^3}\begin{bmatrix}6 & 3l & -6 & 3l\\ 3l & 2l^2 & -3l & l^2\\ -6 & -3l & 6 & -3l\\ 3l & l^2 & -3l & 2l^2\end{bmatrix}\begin{bmatrix}w_1\\ \theta_1\\ w_2\\ \theta_2\end{bmatrix}=\begin{bmatrix}V(0)\\ -m(0)\\ -V(l)\\ m(l)\end{bmatrix}=\begin{bmatrix}f_1\\ M_1\\ f_2\\ M_2\end{bmatrix}$$

Direct equilibrium approach



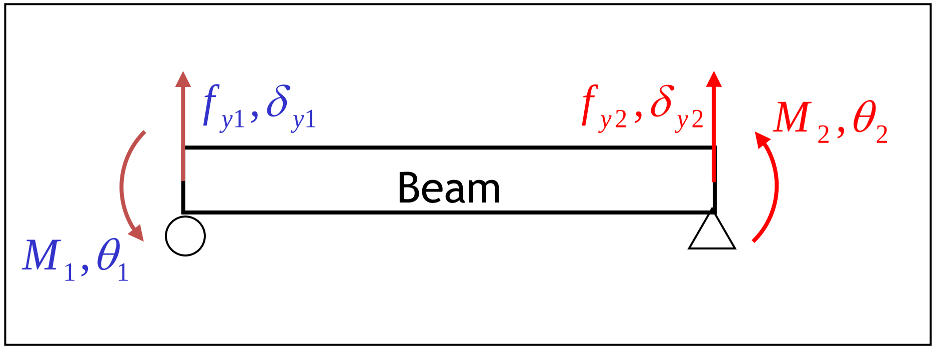
$$\frac{2EI}{l^3}\begin{bmatrix}6 & 3l & -6 & 3l\\ 3l & 2l^2 & -3l & l^2\\ -6 & -3l & 6 & -3l\\ 3l & l^2 & -3l & 2l^2\end{bmatrix}\begin{bmatrix}w_1\\ \theta_1\\ w_2\\ \theta_2\end{bmatrix}=\begin{bmatrix}f_1\\ M_1\\ f_2\\ M_2\end{bmatrix}$$

※ superposition of stiffness matrix

SUMMARY

Element: Beam
- Definition

- Beam[1)2)]
  - Structural members subjected to lateral loads, that is, forces or moments having their vectors perpendicular to the axis of the member.
  - Lateral loads occur shear forces and bending moments in beams.
  - Beam elements have two degrees of freedom at each end: a rotation about axis perpendicular to the plane of the mean and a translation perpendicular the axis of the beam.
  - Axial deformation is neglected.



1) Gere, J. M., Goodno, B. J, Mechanics of Meterials, 7th edition, Cengage Learing, 2009, p.306
2) Sennett, R. E., Matrix analysis of structures, Prentice hall, 1994, p.59

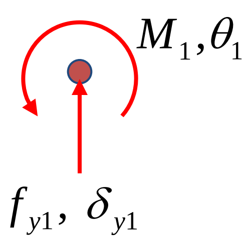
Element: Beam
- Boundary Condition

- Boundary condition

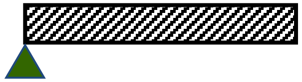
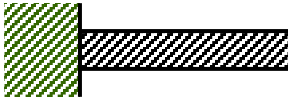
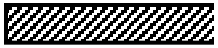
  ① If the loads are given, the deformations are unknown variables.

  ② If the deformations are given, the loads are unknown variables.

Force, moment, and deformations



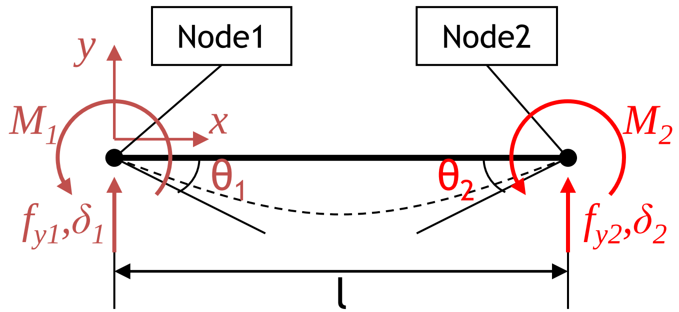
Examples of left end of a beam

	Simple Support 	Fixed End 	Free-End 
Known	$\delta_{y1},\ M_1$ (=0)	$\delta_{y1},\ \theta_1$ (=0)	$f_{y1},\ M_1$ (Given)
Unknown	$f_{y1},\ \theta_1$	$f_{y1},\ M_1$	$\delta_{y1},\ \theta_1$

Element: Beam

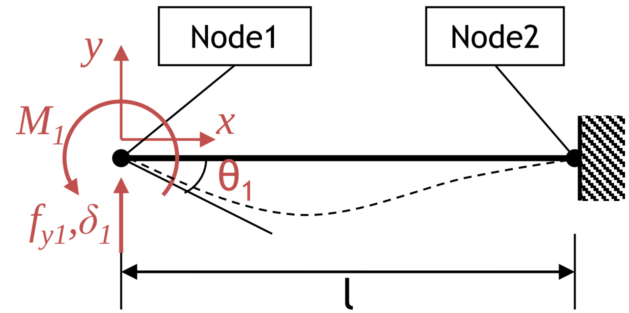
1) 박재형, “구조해석 꼭 어려워야 하나” , 기문당, 2006, pp.99~103

- Derivation of the beam elemental stiffness matrix[1)]



Deflection curve of beam: $EI\dfrac{d^2y}{dx^2} = -M$

(1) Case #1: The node2 is fixed supported $(\delta_{y2}=0,\ \theta_2=0)$



$$EI\frac{d^2y}{dx^2} = -M_1 + f_{y1}x \quad \textbf{-(1)}$$

① Integration of Eq.(1)

$$EI\frac{dy}{dx} = -M_1x + \frac{f_{y1}}{2}x^2 + C_1 \quad \text{-(2)}$$

$$\left.\frac{dy}{dx}\right]_{x=l} = \theta_2 = 0 \quad \text{at} \quad x=l$$

$$-M_1l + \frac{f_{y1}}{2}l^2 + C_1 = 0 \qquad \therefore C_1 = M_1l - \frac{f_{y1}}{2}l^2$$

② Integration of Eq.(2)

$$EIy = -\frac{M_1}{2}x^2 + \frac{f_{y1}}{6}x^3 + C_1x + C_2 \quad \text{-(3)}$$

$$y\big]_{x=l} = \delta_{y2} = 0 \quad \text{at} \quad x=l$$

$$-\frac{M_1}{2}l^2 + \frac{f_{y1}}{6}l^3 + C_1l + C_2 = 0$$

$$\therefore C_2 = \frac{M_1}{2}l^2 - \frac{f_{y1}}{6}l^3 - C_1l = -\frac{M_1}{2}l^2 + \frac{f_{y1}}{3}l^3$$

Element: Beam
- Derivation of the beam elemental stiffness matrix

③ Using Eq. (2) and (3), translation in y direction(δ_{y1}) and rotation(θ_1) can be calculated at $x=0$

$$EIy = -\frac{M_1}{2}x^2 + \frac{f_{y1}}{6}x^3 + C_1 x + C_2 \quad \text{-(3)}$$

$$EI\frac{dy}{dx} = -M_1 x + \frac{f_{y1}}{2}x^2 + C_1 \quad \text{-(2)}$$

$$C_1 = M_1 l - \frac{f_{y1}}{2}l^2 \quad , \quad C_2 = -\frac{M_1}{2}l^2 + \frac{f_{y1}}{3}l^3$$

Translation in y direction(δ_{y1}) at $x=0$

$$\delta_{y1} = y\big|_{x=0} = \frac{C_2}{EI} = -\frac{M_1}{2EI}l^2 + \frac{f_{y1}}{3EI}l^3$$

Rotation(θ_1) can be calculated at $x=0$

$$\theta_1 = \frac{dy}{dx}\bigg|_{x=0} = \frac{C_1}{EI} = \frac{M_1}{EI}l - \frac{f_{y1}}{2EI}l^2$$

▪ Matrix form

$$\begin{bmatrix}\delta_{y1}\\ \theta_1\end{bmatrix} = \begin{bmatrix}\frac{l^3}{3EI} & -\frac{l^2}{2EI}\\ -\frac{l^2}{2EI} & \frac{l}{EI}\end{bmatrix}\begin{bmatrix}f_{y1}\\ M_1\end{bmatrix}$$

$$\therefore \begin{bmatrix}f_{y1}\\ M_1\end{bmatrix} = \begin{bmatrix}\frac{l^3}{3EI} & -\frac{l^2}{2EI}\\ -\frac{l^2}{2EI} & \frac{l}{EI}\end{bmatrix}^{-1}\begin{bmatrix}\delta_{y1}\\ \theta_1\end{bmatrix} = \begin{bmatrix}\frac{12EI}{l^3} & \frac{6EI}{l^2}\\ \frac{6EI}{l^2} & \frac{4EI}{l}\end{bmatrix}\begin{bmatrix}\delta_{y1}\\ \theta_1\end{bmatrix}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix

$$\begin{bmatrix} f_{y1} \\ M_1 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \end{bmatrix}$$

④ Using force and moment equilibrium at $x=L$, f_{y2} and M_2 can be calculated

Force equilibrium: $\sum f_y = f_{y1} + f_{y2} = 0$

Moment equilibrium: $\sum M_2 = -f_{y1}l + M_1 + M_2 = 0$

- Force at node 2

$$f_{y2} = -f_{y1}$$

- Moment at node 2

$$M_2 = f_{y1}l - M_1$$

- Matrix form

$$\begin{bmatrix} f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ l & -1 \end{bmatrix} \begin{bmatrix} f_{y1} \\ M_1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ l & -1 \end{bmatrix} \begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \end{bmatrix}$$

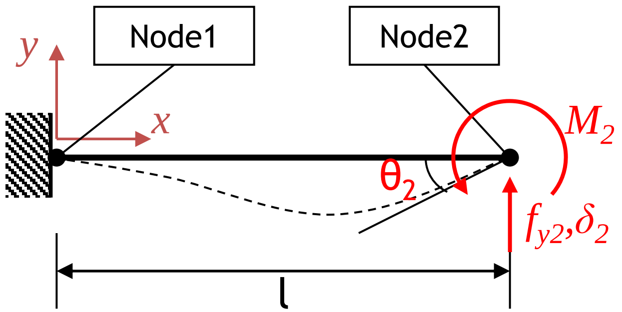
$$= \begin{bmatrix} -\frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \end{bmatrix}$$

$$\begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{L^2} & \frac{4EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \end{bmatrix}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix

(2) Case #2: The node1 is fixed supported

$(\delta_{y1} = 0,\ \theta_1 = 0)$



$$EI\frac{d^2y}{dx^2} = M_2 + f_{y2}x \quad \text{-(4)}$$

① Integration of Eq.(4)

$$EI\frac{dy}{dx} = M_2 x + \frac{f_{y2}}{2}x^2 + C_1$$

$$\left.\frac{dy}{dx}\right]_{x=0} = \theta_1 = 0 \quad \text{at} \quad x = 0$$

$$\therefore C_1 = 0 \quad \Rightarrow \quad EI\frac{dy}{dx} = M_2 x + \frac{f_{y2}}{2}x^2 \quad \text{-(5)}$$

② Integration of Eq.(5)

$$EIy = \frac{M_2}{2}x^2 + \frac{f_{y2}}{6}x^3 + C_2$$

$$y\big]_{x=0} = \delta_{y1} = 0 \quad \text{at} \quad x = 0$$

$$\therefore C_2 = 0 \quad \Rightarrow \quad EIy = \frac{M_2}{2}x^2 + \frac{f_{y2}}{6}x^3 \quad \text{-(6)}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix

③ Using Eq. (5) and (6), translation in y direction(δ_{y2}) and rotation(θ_2) can be calculated at $x=l$

$$EIy = \frac{M_2}{2}x^2 + \frac{f_{y2}}{6}x^3 \quad \text{-(6)}$$

$$EI\frac{dy}{dx} = M_2 x + \frac{f_{y2}}{2}x^2 \quad \text{-(5)}$$

Translation in y direction(δ_{y1}) at $x=l$

$$\delta_{y2} = y\big|_{x=L} = \frac{M_2}{2EI}l^2 + \frac{f_{y2}}{3EI}l^3$$

Rotation(θ_1) can be calculated at $x=l$

$$\theta_2 = \frac{dy}{dx}\bigg|_{x=l} = \frac{M_2}{EI}l + \frac{f_{y2}}{2EI}l^2$$

▪ Matrix form

$$\begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \frac{l^3}{3EI} & \frac{l^2}{2EI} \\ \frac{l^2}{2EI} & \frac{l}{EI} \end{bmatrix} \begin{bmatrix} f_{y2} \\ M_2 \end{bmatrix}$$

$$\therefore \begin{bmatrix} f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} \frac{l^3}{3EI} & \frac{l^2}{2EI} \\ \frac{l^2}{2EI} & \frac{l}{EI} \end{bmatrix}^{-1} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix

$$\begin{bmatrix} f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

④ Using force and moment equilibrium at $x=0$, f_{y1} and M_1 can be calculated

Force equilibrium: $\sum f_y = f_{y1} + f_{y2} = 0$

Moment equilibrium: $\sum M_1 = f_{y2} l + M_1 + M_2 = 0$

- Force at node 1

$$f_{y1} = -f_{y2}$$

- Moment at node 1

$$M_1 = -f_{y2} l - M_2$$

- Matrix form

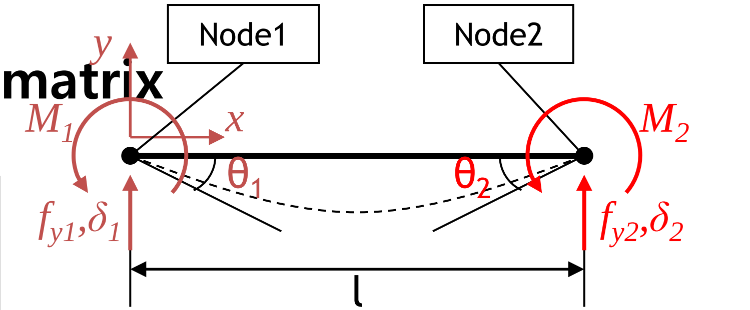
$$\begin{bmatrix} f_{y1} \\ M_1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -l & -1 \end{bmatrix} \begin{bmatrix} f_{y2} \\ M_2 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ -l & -1 \end{bmatrix} \begin{bmatrix} \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{2EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

$$\begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{2EI}{l} \\ \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

Element: Beam
- Derivation of the beam elemental stiffness matrix



(1) Case #1: The node2 is fixed supported

$$\begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \end{bmatrix}$$

(1) Case #2: The node1 is fixed supported

$$\begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{2EI}{l} \\ \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y2} \\ \theta_2 \end{bmatrix}$$

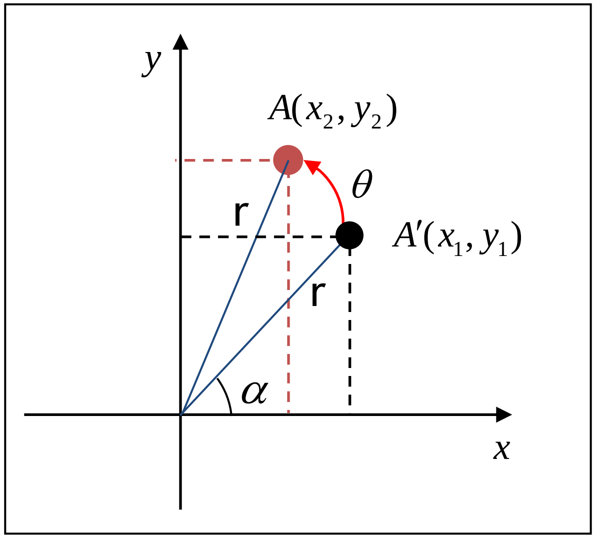
$$\begin{bmatrix} f_{y1} \\ M_1 \\ f_{y2} \\ M_2 \end{bmatrix} = \begin{bmatrix} \frac{12EI}{l^3} & \frac{6EI}{l^2} & -\frac{12EI}{l^3} & \frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{4EI}{l} & -\frac{6EI}{l^2} & \frac{2EI}{l} \\ -\frac{12EI}{l^3} & -\frac{6EI}{l^2} & \frac{12EI}{l^3} & -\frac{6EI}{l^2} \\ \frac{6EI}{l^2} & \frac{2EI}{l} & -\frac{6EI}{l^2} & \frac{4EI}{l} \end{bmatrix} \begin{bmatrix} \delta_{y1} \\ \theta_1 \\ \delta_{y2} \\ \theta_2 \end{bmatrix}$$

- **the beam elemental stiffness matrix**

TEMPORARY REFERENCE SLIDE(LATER DELETE !)

Rotational Transformation : Point

- Rotational Transformation : Point



① trigonometric identities : angle sum

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$$

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\cos\alpha\cos\beta$$

② components of point A

$$x_2 = r\cos(\alpha+\theta)$$

$$y_2 = r\sin(\alpha+\theta)$$

③ by using the angle sum identities

$$\begin{aligned} x_2 &= r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ &= (r\cos\alpha)\cos\theta-(r\sin\alpha)\sin\theta \\ &= x_1\cos\theta - y_1\sin\theta \end{aligned}$$

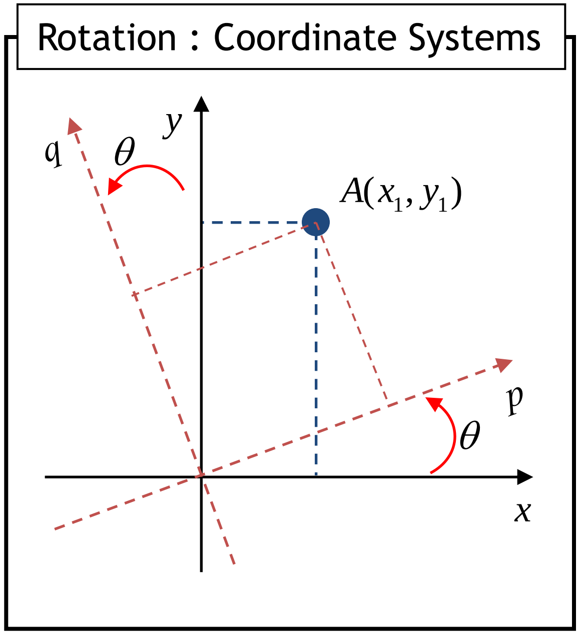
$$\begin{aligned} y_2 &= r\sin\alpha\cos\theta + r\cos\alpha\sin\theta \\ &= (r\sin\alpha)\cos\theta+(r\cos\alpha)\sin\theta \\ &= y_1\cos\theta + x_1\sin\theta \end{aligned}$$

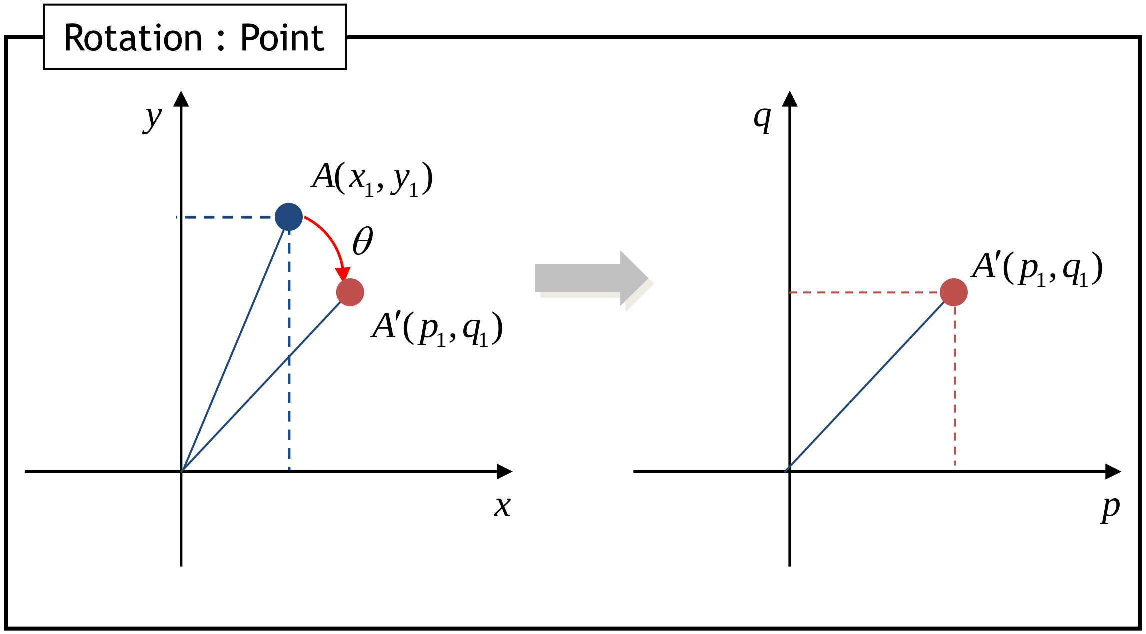
④ in matrix form

$$\boxed{\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}}$$

Rotational Transformation : Coordinate System

▪ Rotational Transformation : Coordinate System





※ Rotational Transformation : Point

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Rotation of point by $-\theta$

$$\begin{bmatrix} p_1 \\ q_1 \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Rotational Transformation : Coordinate System

- 2 Dimension $(xy \rightarrow pq)$

$$\begin{bmatrix} p_1 \\ q_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

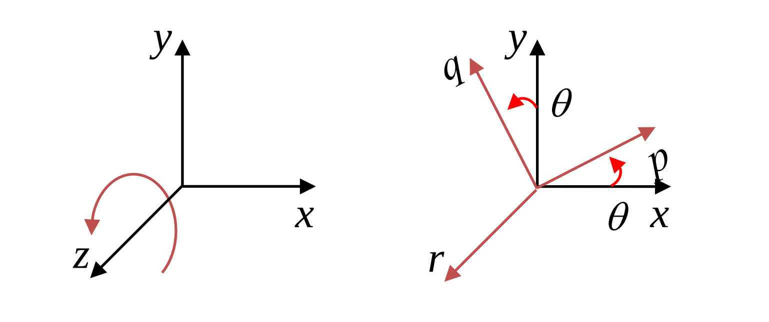
- 3 Dimension $(xyz \rightarrow pqr)$

$$\begin{bmatrix} p_1 \\ q_1 \\ r_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$$: rotation about z axis

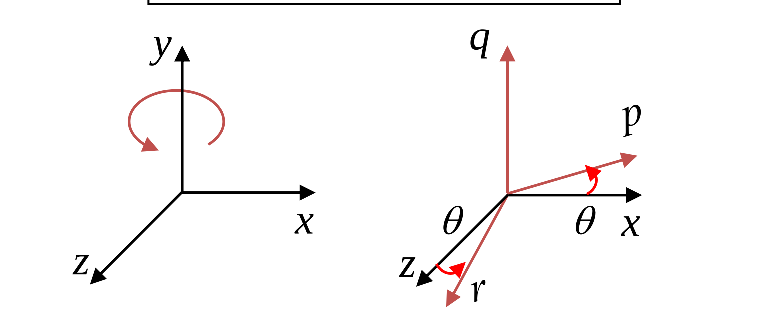
$$\begin{bmatrix} p_1 \\ q_1 \\ r_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$$: rotation about y axis

$$\begin{bmatrix} p_1 \\ q_1 \\ r_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$$: rotation about x axis

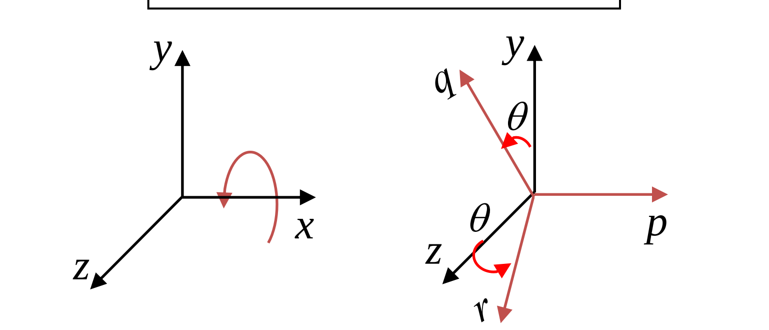
rotation about z axis by θ



rotation about y axis by θ



rotation about x axis by θ



Seoul National University
Naval Architecture & Ocean Engineering
Advanced Ship Design Automation Laboratory

TEL. (02) 880-8378
FAX. (02) 886-4920
http://asdal.snu.ac.kr