

Time-Integration and Related Solution Techniques

Tayfun E. Tezduyar
Rice University

Kenji Takizawa
Waseda University

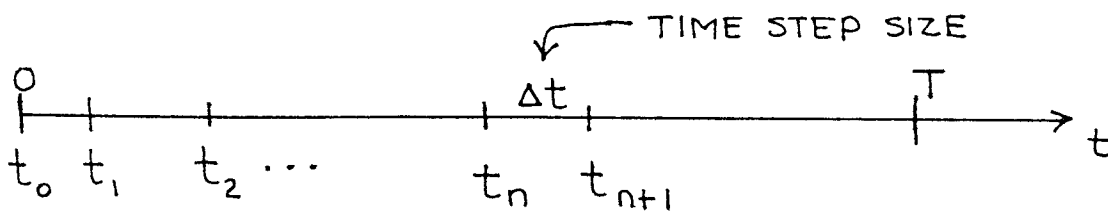
TIME-INTEGRATION

CONSIDER THE FOLLOWING LINEAR, ORDINARY
DIFFERENTIAL EQUATION SYSTEM:

$$\underline{M} \dot{\underline{d}} + \underline{K} \underline{d} = \underline{F}$$

$$\underline{d}(0) = \underline{d}_0$$

TIME-DISCRETIZATION:



n_{ts} : TOTAL NUMBER OF TIME STEPS

$$\underline{d}_n \doteq \underline{d}(t_n)$$

$$\Delta t = t_{n+1} - t_n : \text{ CAN BE VARIABLE}$$

GIVEN: \underline{d}_n , HOW DO WE CALCULATE \underline{d}_{n+1}

THERE ARE MANY WAYS. WE DESCRIBE 3 OF THEM...

$$1) \quad \underline{M} \left(\frac{\underline{d}_{n+1} - \underline{d}_n}{\Delta t} \right) + \underline{K} \left(\underline{d}_n \right) = \underline{F}_n$$

(FORWARD DIFFERENCES)
(FORWARD EULER)

$$2) \quad + \underline{K} \left(\underline{d}_{n+1} \right) = \underline{F}_{n+1}$$

(BACKWARD DIFFERENCES)
(BACKWARD EULER)

$$3) \quad + \underline{K} \left(\frac{\underline{d}_{n+1} + \underline{d}_n}{2} \right) = \frac{1}{2} (\underline{F}_{n+1} + \underline{F}_n)$$

(CENTRAL DIFFERENCES)
(TRAPEZOIDAL RULE)

IN 1), IF WE APPROXIMATE \underline{M} WITH A
DIAGONAL MATRIX, THEN TIME-MARCHING
CAN BE DONE WITHOUT SOLVING A
MATRIX SYSTEM.

A POSSIBLE APPROXIMATION TO \underline{M} IS \underline{M}_L

$$\begin{aligned} \underline{M}_L &= \text{LUMP}(\underline{M}) \\ &= \text{ROW-SUM} \\ &\quad \text{OR} \\ &\quad \text{COLUMN-SUM} \end{aligned} = \begin{bmatrix} \int N_1 d\Omega & & & \\ & \int N_2 d\Omega & & \\ & & \dots & \\ & & & \int N_N d\Omega \end{bmatrix}$$

DEFINITION:

EXPLICIT METHOD

A TIME-MARCHING METHOD THAT DOES NOT
REQUIRE SOLUTION OF A MATRIX SYSTEM.

IMPLICIT METHOD

A TIME-MARCHING METHOD THAT
REQUIRES SOLUTION OF A MATRIX SYSTEM.

NOTE THAT 2) AND 3) ARE IMPLICIT METHODS,
AND, WITH $\underline{\underline{M}} \leftarrow \underline{\underline{M}}_L$, 1) IS AN EXPLICIT METHOD.

GENERALIZED TRAPEZOIDAL FAMILY: $(0 \leq \alpha \leq 1)$

$$\underline{\underline{M}} \left(\frac{d_{n+1} - d_n}{\Delta t} \right) + \underline{\underline{K}} \underbrace{(\alpha d_{n+1} + (1-\alpha) d_n)}_{\underline{\underline{def}} \quad d_{n+\alpha}} = \underbrace{\alpha F_{n+1} + (1-\alpha) F_n}_{\underline{\underline{def}} \quad F_{n+\alpha}}$$

$$\Rightarrow \underline{\underline{M}} \left(\frac{d_{n+1} - d_n}{\Delta t} \right) + \underline{\underline{K}} d_{n+\alpha} = F_{n+\alpha}$$

$$\alpha = 0 \Rightarrow 1) \quad \alpha = 1 \Rightarrow 2) \quad \alpha = 1/2 \Rightarrow 3)$$

($\underline{\underline{M}} \leftarrow \underline{\underline{M}}_L$)

RE-ARRANGE THE EQUATION TO COLLECT THE UNKNOWN TERMS ON THE LEFT-HAND-SIDE:

$$\underbrace{(\underline{M} + \alpha \Delta t \underline{K})}_{\text{NEEDS TO BE FORMED AND FACTORIZED IF } \alpha > 0} \underline{d}_{n+1} = \underbrace{(\underline{M} - (1-\alpha) \Delta t \underline{K})}_{\text{THESE MATRICES DO NOT NEED TO BE FORMED. BECAUSE}} \underline{d}_n + \Delta t \underline{F}_{n+\alpha}$$

NEEDS TO BE
FORMED AND
FACTORIZED
IF $\alpha > 0$



IMPLICIT METHOD

THESE MATRICES
DO NOT NEED
TO BE FORMED.
BECAUSE

$$\underline{M} \underline{d}_n = \sum_{e=1}^{nel} (\underline{m}_v)_n^e$$

$$\underline{K} \underline{d}_n = \sum_{e=1}^{nel} (\underline{k}_v)_n^e$$

IF Δt IS SMALL ENOUGH, THEN THERE COULD BE A JUSTIFICATION TO APPROXIMATE $(\underline{M} + \alpha \Delta t \underline{K})$ BY \underline{M}_L .

THEN, THERE IS NO MATRIX FORMATION OR FACTORIZATION.

⇒ EXPLICIT METHOD ⇒ ECONOMICAL

STABILITY : CONDITIONAL

(WILL CONVERGE IF $\Delta t < (\Delta t)_{\text{CRITICAL}}$)

ACCURACY : 1ST-ORDER

(IF YOU HALVE Δt , THE ERROR WILL BE HALVED)

(NOTE:
IN A 2ND-ORDER METHOD: HALVE Δt ⇒ ERROR HALVED TWICE)

PROPERTIES OF THE GENERALIZED TRAPEZOIDAL FAMILY

$\alpha =$	0	1	1/2
NAME OF METHOD	FORWARD DIFF FORWARD EULER	BACKWARD DIFF BACKWARD EULER	CENTRAL DIFF TRAPEZOIDAL RULE
COST OF MATRIX FORMATION & FACTORIZATION	NO	YES	YES
IMPLICIT OR EXPLICIT	EXPLICIT	IMPLICIT	IMPLICIT
STABILITY	CONDITIONAL	UNCONDITIONAL	UNCONDITIONAL
ACCURACY	1ST ORDER	1ST ORDER	2ND ORDER

REMARKS:

1. IF $(\underline{M} + \alpha \Delta t \underline{K})$ IS LEFT AS IT IS, THEN IT TAKES ONLY ONE ITERATION TO CONVERGE.
2. WE CAN APPROXIMATE $(\underline{M} + \alpha \Delta t \underline{K})$ ANYWAY WE WISH, PROVIDED THAT THE ITERATIONS CONVERGE.
3. THE OBJECTIVE IS TO APPROXIMATE $(\underline{M} + \alpha \Delta t \underline{K})$ WITH SOMETHING SUFFICIENTLY SIMPLE, BUT NOT TOO SIMPLE TO CONVERGE WITHIN REASONABLE NUMBER OF ITERATIONS.
4. EXAMPLE: $(\underline{M} + \alpha \Delta t \underline{K}) \leftarrow \underline{M}_L$
 - LIMIT ON Δt
 - CONVERGES SLOWLY
5. EXAMPLES: $(\underline{M} + \alpha \Delta t \underline{K}) \leftarrow (\underline{M}_L + \alpha \Delta t \text{DIAG}(\underline{K}))$
 $(\underline{M} + \alpha \Delta t \underline{K}) \leftarrow \text{DIAG}(\underline{M} + \alpha \Delta t \underline{K})$

SOLUTION TECHNIQUES FOR NONLINEAR EQUATIONS

CONSIDER THE FOLLOWING SYSTEM OF NONLINEAR ORDINARY DIFFERENTIAL EQUATIONS IN TIME :

$$\underline{M} \dot{\underline{U}} + \underline{N}(\underline{U}) + \underline{K} \underline{U} = \underline{F}$$

$$\underline{U}(0) = \underline{U}_0$$

THIS CAN BE DISCRETIZED AS FOLLOWS :

$$\underbrace{\underline{M} \left(\frac{\underline{U}_{n+1} - \underline{U}_n}{\Delta t} \right)}_{\text{LINEAR TERMS}} + \underbrace{\underline{N}(\alpha \underline{U}_{n+1} + (1-\alpha) \underline{U}_n)}_{\text{NONLINEAR VECTOR FUNCTION}} + \underline{K} (\alpha \underline{U}_{n+1} + (1-\alpha) \underline{U}_n) = \underline{F}_{n+\alpha}$$

THIS MEANS THAT IN NONLINEAR PROBLEMS,

INSTEAD OF A LINEAR EQUATION SYSTEM

$$(\underline{M} + \alpha \Delta t \underline{K}) \underline{d}_{n+1} = \dots ,$$

WE NEED TO SOLVE AT EVERY TIME STEP

A NONLINEAR EQUATION SYSTEM $\underline{N}(\underline{d}_{n+1}) = \dots$

FOR NOTATIONAL CONVENIENCE, DROP THE SUBSCRIPT $n+1$ AND CALL THE R.H.S. \underline{F} . THEN,

$$\boxed{\underline{N}(\underline{d}) = \underline{F}}$$

IS WHAT NEEDS TO BE SOLVED AT EVERY TIME STEP.

NEWTON-RAPHSON METHOD

IN SOLVING

$$\underline{N}(\underline{d}) = \underline{F} \quad ,$$

\uparrow
 UNKNOWN

START WITH AN INITIAL GUESS \underline{d}^0 ← ZERO TH ITERATION
 AND AT EACH ITERATION, GIVEN \underline{d}^i , CALCULATE
 A CORRECTION $\Delta \underline{d}^i$, SUCH THAT

$$\underline{N}(\underline{d}^i + \Delta \underline{d}^i) = \underline{F} \quad .$$

TO CALCULATE $\Delta \underline{d}^i$, CONSIDER THE FOLLOWING TRUNCATED
 TAYLOR SERIES EXPANSION AROUND \underline{d}^i :

$$\underline{N}(\underline{d}^i + \Delta \underline{d}^i) = \underline{N}(\underline{d}^i) + \left. \frac{\partial \underline{N}}{\partial \underline{d}} \right|_{\underline{d}^i} \Delta \underline{d}^i$$

THEN SOLVE

$$\left. \frac{\partial \underline{N}}{\partial \underline{d}} \right|_{\underline{d}^i} \Delta \underline{d}^i = \underline{F} - \underline{N}(\underline{d}^i) \quad ,$$

$\underline{K}_T(\underline{d}^i)$: "TANGENT STIFFNESS" MATRIX

AND UPDATE :

$$\underline{d}^{i+1} = \underline{d}^i + \Delta \underline{d}^i$$

NEWTON-RAPHSON METHOD CONVERGES QUADRATICALLY,
PROVIDED THAT

- 1) THE INITIAL GUESS \underline{d}^0 IS SUFFICIENTLY CLOSE TO THE SOLUTION, AND
- 2) INVERSE OF $\frac{\partial \underline{N}}{\partial \underline{d}}$ EXISTS.

HOW DO WE DECIDE WHEN TO STOP ITERATIONS ?

ANSWER : WHEN

- a) $\underline{F} - \underline{N}(\underline{d}^i)$ IS SUFFICIENTLY CLOSE TO $\underline{0}$, OR
- b) A PREDETERMINED NUMBER OF ITERATIONS HAVE BEEN PERFORMED.

IN BOTH CASES, THE RESIDUAL $\underline{F} - \underline{N}(\underline{d}^i)$ NEED TO BE MONITORED. THE RESIDUAL AT ITERATION i :

$$\underline{R}^i = \underline{F} - \underline{N}(\underline{d}^i).$$

CHECK TO SEE IF

$$\|\underline{R}^i\| \leq \epsilon \leftarrow \text{A PREDETERMINED, SMALL VALUE}$$

EXAMPLE FOR THE NORM $\|\cdot\|$:

$$\|\underline{x}\| = (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2}.$$

TO CHECK CONVERGENCE IN A WAY THAT MAKES SENSE,
 WE NEED TO SCALE $\|\underline{R}^i\|$. ONE WAY IS TO
 SCALE IT WITH $\|\underline{R}^0\|$.
 \uparrow
 INITIAL RESIDUAL

THEREFORE, THE CONVERGENCE CHECK TAKES

THE FORM :

$$\frac{\|\underline{R}^i\|}{\|\underline{R}^0\|} \stackrel{?}{\leq} \epsilon$$

QUITE OFTEN, IT IS MORE PRACTICAL, ESPECIALLY
 IN SOLVING TIME-DEPENDENT PROBLEMS, TO PERFORM
 A PREDETERMINED NUMBER OF ITERATIONS AT EACH
 TIME STEP, BUT AT THE SAME TIME KEEP AN
 EYE ON

$$\frac{\|\underline{R}^i\|}{\|\underline{R}^0\|},$$

TO MAKE SURE THAT IT DOES NOT EXCEED A
 CERTAIN LEVEL.

MODIFIED NEWTON-RAPHSON METHOD

$$\underline{K}_T(\underline{d}^i) \Delta \underline{d}^i = \underline{F} - \underline{N}(\underline{d}^i)$$

IN THE "FULL" NEWTON-RAPHSON METHOD

THIS GETS UPDATED AT EVERY ITERATION

$$\text{BY } \underline{K}_T(\underline{d}^i) = \frac{\partial \underline{N}}{\partial \underline{d}} \bigg|_{\underline{d}^i}$$

IN THE MODIFIED NEWTON-RAPHSON METHOD

\underline{K}_T IS NOT UPDATED EVERY ITERATION.

AT SOME OF THE ITERATIONS, OLD VALUES

OF \underline{K}_T ARE USED.

THIS APPROACH

- CUTS COST OF COMPUTATION
- BUT MAY REDUCE THE CONVERGENCE RATE.

INCREMENTAL NEWTON-RAPHSON METHOD

THE NEWTON-RAPHSON METHOD MAY HAVE CONVERGENCE DIFFICULTIES IF THE INITIAL GUESS \underline{d}^0 IS TOO "FAR" FROM THE SOLUTION.

ONE APPROACH: INCREMENTAL "LOADING" (ie. "LOAD" RAMPING)

$$N(\underline{d}) = \underline{F}$$

- 1) \uparrow START WITH A FRACTION OF THIS. THIS MEANS A "LOAD" SMALLER THAN THE ACTUAL ONE.
- 2) GET A CONVERGED SOLUTION FOR THIS LOAD LEVEL
- 3) USE THAT SOLUTION AS THE INITIAL GUESS FOR THE NEXT LEVEL.
- 4) REPEAT 2) AND 3) UNTIL THE FULL LOAD LEVEL IS REACHED.

EXAMPLE: START WITH $\frac{1}{10} \underline{F} \xrightarrow{\text{GET}} \underline{d}_{1/10}$

USE $\underline{d}_{1/10}$ AS INITIAL GUESS FOR $\frac{2}{10} \underline{F} \xrightarrow{\text{GET}} \underline{d}_{2/10}$

USE $\underline{d}_{2/10}$ " " " " $\frac{3}{10} \underline{F} \xrightarrow{\text{GET}} \underline{d}_{3/10}$

⋮

USE $\underline{d}_{9/10}$ " " " " $\underline{F} \xrightarrow{\text{GET}} \underline{d}$

RAMPING UP THE REYNOLDS NUMBER

IN FLUID DYNAMICS PROBLEMS, TYPICALLY, THE NONLINEAR TERMS ARE PROPORTIONAL TO THE REYNOLDS NUMBER. ONE CAN START WITH A FRACTION OF THE ACTUAL REYNOLDS NUMBER,

- 2) GET A CONVERGED SOLUTION FOR THAT REYNOLDS NUMBER
- 3) USE THAT SOLUTION AS THE INITIAL GUESS FOR THE NEXT LEVEL OF THE REYNOLDS NUMBER
- 4) REPEAT 2) AND 3) UNTIL THE FULL REYNOLDS NUMBER IS REACHED.

EXAMPLE : START WITH $Re \times 10^{-5} \xrightarrow{\text{GET}} \underline{d}_{\times 10^{-5}}$

USE $\underline{d}_{\times 10^{-5}}$ AS INITIAL GUESS FOR $Re \times 10^{-4} \xrightarrow{\text{GET}} \underline{d}_{\times 10^{-4}}$

USE $\underline{d}_{\times 10^{-4}}$ " " " " $Re \times 10^{-3} \xrightarrow{\text{GET}} \underline{d}_{\times 10^{-3}}$

:

USE $\underline{d}_{\times 10^{-1}}$ " " " " $Re \xrightarrow{\text{GET}} \underline{d}$.

NOTE : IN A TIME-DEPENDENT PROBLEM, RAMPING THE REYNOLDS NUMBER NEEDS TO BE EXERCISED ONLY IN THE INITIAL STAGES OF THE TIME-MARCHING.

ITERATIVE SOLUTION TECHNIQUES FOR LINEAR EQUATION SYSTEMS

$$\underline{A} \underline{x} = \underline{b}$$

WHERE DO THESE SYSTEMS COME FROM?

EXAMPLE :

$$\left. \begin{array}{l} \text{NONLINEAR} \\ \text{PROBLEMS} \end{array} \right\} \underline{K}_T(\underline{d}^i) \Delta \underline{d}^i = \underline{F} - \underline{N}(\underline{d}^i)$$

$$\underline{A} \underline{x} = \underline{b}$$

$$\left. \begin{array}{l} \text{LINEAR} \\ \text{PROBLEMS} \end{array} \right\} (\underline{M} + \alpha \Delta t \underline{K}) \Delta \underline{d}_{n+1}^{(i)} = \dots$$

IN MOST CASES, ESPECIALLY THOSE IN 3D,
THE MATRIX \underline{A} IS TOO LARGE, AND THEREFORE
WE CANNOT USE A "DIRECT" SOLUTION METHOD
SUCH AS THE GAUSSIAN ELIMINATION METHOD OR
OTHER FACTORIZATION TECHNIQUES, BECAUSE WE
CANNOT AFFORD TO

- a) FACTORIZE \underline{A}
- b) STORE THE FACTORIZED \underline{A} .

IN ITERATIVE SOLUTION OF

$$\underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{b}},$$

↑
UNKNOWN

START WITH AN INITIAL GUESS $\underline{\underline{x}}_0 \leftarrow$ ZEROth ITERATION

AND AT EACH ITERATION, GIVEN $\underline{\underline{x}}_m$, CALCULATE

A CORRECTION $\Delta \underline{\underline{x}}_m$, SUCH THAT

$$\underline{\underline{A}} (\underline{\underline{x}}_m + \Delta \underline{\underline{x}}_m) = \underline{\underline{b}}$$

OR

$$\underline{\underline{A}} \Delta \underline{\underline{x}}_m = \underline{\underline{b}} - \underline{\underline{A}} \underline{\underline{x}}_m$$

AND UPDATE

$$\underline{\underline{x}}_{m+1} = \underline{\underline{x}}_m + \Delta \underline{\underline{x}}_m.$$

IF WE LEAVE $\underline{\underline{A}}$ AS IT IS, THIS PROCEDURE IS

ESSENTIALLY NO DIFFERENT THAN SOLVING $\underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{b}}$,

AND IT CONVERGES IN ONE ITERATION TO THE

SOLUTION OF $\underline{\underline{A}} \underline{\underline{x}} = \underline{\underline{b}}$. FOR EXAMPLE, IF WE

START WITH $\underline{\underline{x}}_0 = \underline{\underline{0}}$, THEN $\Delta \underline{\underline{x}}_1 = \underline{\underline{x}}$. BUT....

IN MOST CASES, WE CANNOT AFFORD LEAVING \underline{A} AS IT IS. THEREFORE WE APPROXIMATE \underline{A} WITH A "PRECONDITIONING" MATRIX \underline{P} , AND SOLVE AT EVERY ITERATION

$$\underline{P} \Delta \underline{y}_m = \underline{b} - \underline{A} \underline{x}_m$$

AND UPDATE \underline{x}_m BY USING AN UPDATE METHOD MORE SOPHISTICATED THAN SIMPLY ADDING $\Delta \underline{y}_m$ TO \underline{x}_m .

THERE ARE 3 MAIN ISSUES HERE.

- 1) HOW TO COMPUTE $\underline{b} - \underline{A} \underline{x}_m$ IN THE MOST EFFICIENT WAY.
- 2) HOW TO DESIGN \underline{P} IN A WAY THAT KEEPS IT SIMPLE WITHOUT SLOWING THE CONVERGENCE TOO MUCH.
- 3) HOW TO UPDATE \underline{x}_m IN A WAY THAT IS ECONOMICAL BUT ALSO HELPS CONVERGENCE.

1) HOW TO COMPUTE $\underline{\underline{b}} - \underline{\underline{A}} \underline{\underline{x}}_m$

a) SPARSE-MATRIX-BASED :

STORE $\underline{\underline{A}}$ BY USING A SPARSE-MATRIX STORAGE TECHNIQUE, AND PERFORM THE COMPUTATION WITH GLOBAL MATRIX-VECTOR PRODUCTS.

b) ELEMENT-MATRIX-BASED :

NOTE THAT:
$$\underline{\underline{A}} \underline{\underline{x}} = \left(\sum_{e=1}^{nel} \underline{\underline{A}}^e \right) \underline{\underline{x}}$$

BUT WE DO NOT NEED THIS

BECAUSE

$$= \sum_{e=1}^{nel} \left(\underline{\underline{A}}^e \underline{\underline{x}} \right)$$

\swarrow ELEMENT MATRIX
 \nwarrow ELEMENT VECTOR

c) ELEMENT-VECTOR-BASED :

$$\underline{\underline{A}} \underline{\underline{x}}_m = \frac{\partial \underline{\underline{N}}}{\partial \underline{\underline{d}}} \underline{\underline{x}}_m = \lim_{\epsilon \rightarrow 0} \frac{\underline{\underline{N}}(\underline{\underline{d}} + \epsilon \underline{\underline{x}}_m) - \underline{\underline{N}}(\underline{\underline{d}})}{\epsilon}$$

$$\Rightarrow \underline{\underline{A}} \underline{\underline{x}}_m = \frac{\sum_{e=1}^{nel} \underline{\underline{N}}^e(\underline{\underline{d}} + \epsilon \underline{\underline{x}}_m) - \sum_{e=1}^{nel} \underline{\underline{N}}^e(\underline{\underline{d}})}{\epsilon}$$

WHERE ϵ IS A SMALL NUMBER.

2) HOW TO DESIGN \underline{P}

$\underline{P} = \text{DIAG}(\underline{A})$ IS THE SIMPLEST CHOICE.

↑
DIAGONAL OR NODAL-BLOCK-DIAGONAL

CONVERGES SLOWLY, ESPECIALLY FOR
INCOMPRESSIBLE FLOWS.

- OTHER POSSIBILITIES :
- CLUSTERED ELEMENT-BY-ELEMENT PRECONDITIONING
 - MIXED PRECONDITIONING

3) HOW TO UPDATE \underline{x}_m

AS A WAY JUST ONE STEP MORE SOPHISTICATED
THAN SIMPLY ADDING $\Delta \underline{y}_m$ TO \underline{x}_m ,

WE UPDATE \underline{x}_m BY USING THE EXPRESSION

$$\underline{x}_{m+1} = \underline{x}_m + s \Delta \underline{y}_m$$

↑
A SCALAR "SEARCH"
PARAMETER

WE DETERMINE THE "BEST" "S" BY MINIMIZING

THE UPDATED RESIDUAL $\underline{r}_{m+1} = \underline{b} - \underline{A} \underline{x}_{m+1}$

WITH RESPECT TO "S".

$$\underline{r}_{m+1} = \underline{b} - \underline{A} (\underline{x}_m + s \Delta \underline{y}_m) = \underbrace{\underline{b} - \underline{A} \underline{x}_m}_{\underline{r}_m} - s \underline{A} \Delta \underline{y}_m$$

$$\underbrace{\underline{\zeta}_{m+1}}_{\substack{\text{MINIMIZE} \\ \text{THIS}}} = \underbrace{\underline{\zeta}_m}_{\text{KNOWN}} - s \underbrace{\underline{A} \Delta \underline{y}_m}_{\text{KNOWN}}$$

↑
WITH RESPECT TO THIS

MINIMIZE $\|\underline{\zeta}_{m+1}\|$ WITH RESPECT TO "s". \Leftrightarrow

MINIMIZE $\|\underline{\zeta}_{m+1}\|^2$ WITH RESPECT TO "s". \Leftrightarrow

$$\frac{d}{ds} \|\underline{\zeta}_{m+1}\|^2 = 0 \quad \Leftrightarrow \quad \frac{d}{ds} (\underline{\zeta}_{m+1} \cdot \underline{\zeta}_{m+1}) = 0$$

$$\Rightarrow \underline{\zeta}_{m+1} \cdot \frac{d}{ds} (\underline{\zeta}_{m+1}) = 0$$

$$\Rightarrow \underline{\zeta}_{m+1} \cdot (-\underline{A} \Delta \underline{y}_m) = 0$$

$$\Rightarrow (\underline{\zeta}_m - s \underline{A} \Delta \underline{y}_m) \cdot (\underline{A} \Delta \underline{y}_m) = 0$$

$$\Rightarrow s = \frac{\underline{\zeta}_m \cdot (\underline{A} \Delta \underline{y}_m)}{(\underline{A} \Delta \underline{y}_m) \cdot (\underline{A} \Delta \underline{y}_m)}$$

OR

$$= \frac{\underline{\zeta}_m \cdot (\underline{A} \Delta \underline{y}_m)}{\|\underline{A} \Delta \underline{y}_m\|^2}$$

THERE ARE MANY OTHERWAYS, INCLUDING METHODS WITH MULTIPLE "SEARCH" PARAMETERS AND DIRECTIONS.