# Solution Algorithms for Nonlinear Problems



1

## Outline

Consistent linearization

Newton and modified Newton methods

Line search

Quasi-Newton methods (“BFGS”)

Arc-length methods

Convergence criteria

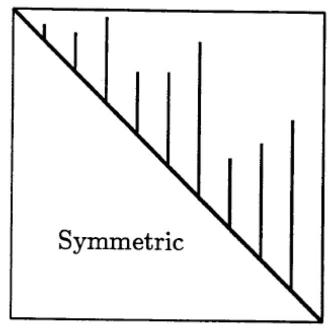
2

## Brief Overview of Linear Finite Element Analysis

**Statics**

$$\underbrace{(S)}_{\text{Strong form of BVP}} \Leftrightarrow \underbrace{(W)}_{\text{Weak}} \approx \underbrace{(G)}_{\text{Gal.}} \Leftrightarrow \underbrace{(M)}_{\text{Matrix (linear algebraic problem)}}$$

$$\underbrace{\underbrace{\boldsymbol{K}}_{\text{Stiff.}}\ \underbrace{\boldsymbol{d}}_{\text{Disp.}}}_{\boldsymbol{F}^{int}} = \underbrace{\underbrace{\boldsymbol{F}}_{\text{Force}}}_{\boldsymbol{F}^{ext}}$$




3

**Dynamics**

$$\underbrace{(S)}_{\text{Strong form of IBVP}} \Leftrightarrow \underbrace{(W)}_{\text{t-cont.}} \approx \underbrace{(G)}_{\text{semi-discretization}} \Leftrightarrow \underbrace{(M)}_{\text{o.d.e.'s (linear)}} \approx \text{time-stepping algorithm}$$

$$\underbrace{\boldsymbol{M}^*}_{\text{Effective mass}}\ \underbrace{\Delta\boldsymbol{a}}_{\text{Acceleration increment}} = \underbrace{\boldsymbol{R}}_{\text{Residual or out-of-balance force}}$$

$+$ update formulas

In each case, a linear algebraic problem

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

4

## Nonlinear Finite Element Analysis

**Nonlinear statics**

$$(S) \Leftrightarrow (W) \approx (G) \Leftrightarrow \text{Nonlinear problem}$$

$$\boldsymbol{F}^{int} = \underbrace{\boldsymbol{F}^{ext}}_{\text{known}}$$

e.g. elastostatics

$$\boldsymbol{F}^{int} = \boldsymbol{N}(\underbrace{\boldsymbol{d}}_{?}) = \text{Nonlin. algebraic ftn. of } \boldsymbol{d}$$

$$\boldsymbol{N} : \mathbb{R}^{n_{eq}} \rightarrow \mathbb{R}^{n_{eq}} \text{ (vect. valued ftn. of vect. } \boldsymbol{d})$$

$$\boldsymbol{N} = \{\underset{\substack{\uparrow \\ \text{eq. no.}}}{N_P}\}, \qquad 1 \le P \le n_{eq}$$

5

**Nonlinear dynamics**

$$(S) \Leftrightarrow (W) \approx (G) \Leftrightarrow \begin{array}{c}\text{Nonlin.}\\ \text{o.d.e.'s}\end{array} \approx \begin{array}{c}\text{time-}\\ \text{stepping}\\ \text{algorithm}\end{array}$$

$$\boldsymbol{M}\ddot{\boldsymbol{d}} + \underbrace{\boldsymbol{F}^{int}}_{\substack{\boldsymbol{N}(\boldsymbol{d}) \\ \text{in elas.}}} = \boldsymbol{F}^{ext}$$

$$\boldsymbol{N}^*(\underbrace{\Delta \boldsymbol{a}}_{?}) = \boldsymbol{R}$$

We need to solve:

**nonlinear algebraic problems**

Reduce to the solution of **linear** algebraic problems by way of incremental/iterative strategies

6

## Residual Methods

Assume the external load is parameterized by “time” (or equivalently, a load parameter).
Try to achieve:

$$0 = \underbrace{\boldsymbol{F}^{ext}}_{\boldsymbol{F}^{ext}(t)} - \underbrace{\boldsymbol{F}^{int}}_{\text{e.g. } \boldsymbol{N}(\boldsymbol{d}(t))} \quad \text{(residual)}$$

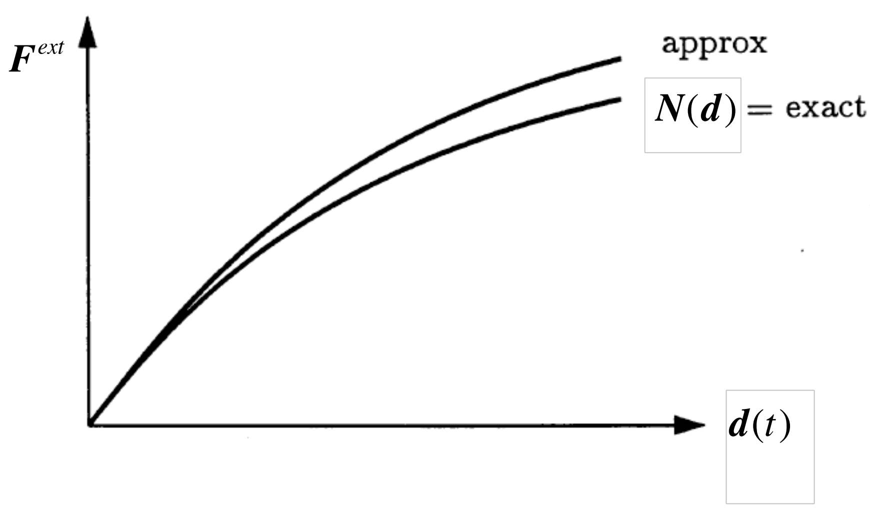
Example: **Incremental load method**

$$\underbrace{\boldsymbol{K}}_{\text{matrix}} \Delta \boldsymbol{d} = \boldsymbol{R}^{n+1} \stackrel{\text{def}}{=} \underbrace{\boldsymbol{F}^{ext}_{n+1}}_{\boldsymbol{F}^{ext}(t_{n+1})} - \underbrace{\boldsymbol{F}^{int}_{n}}_{\boldsymbol{N}(\boldsymbol{d}_n)}$$

$$\boldsymbol{d}_{n+1} = \boldsymbol{d}_n + \Delta \boldsymbol{d}$$

$$n \leftarrow n+1$$

7




Candidate $\boldsymbol{K}$ ?

8

## Consistent Tangents

Example 1. **Nonlinear elastostatics**

$$\begin{aligned} \boldsymbol{F}^{int} &= \boldsymbol{N}(\boldsymbol{d}) \\ \boldsymbol{K}(\boldsymbol{d}) &\overset{\text{def}}{=} \frac{\partial \boldsymbol{N}(\boldsymbol{d})}{\partial \boldsymbol{d}} \quad \text{(tangent stiffness)} \\ \boldsymbol{N} &= \{N_P\} \end{aligned}$$

$$\boldsymbol{K} = \frac{\partial \boldsymbol{N}}{\partial \boldsymbol{d}} = \Bigg[\underbrace{\frac{\partial N_P}{\partial d_Q}}_{K_{PQ}}\Bigg], \qquad 1 \le P,Q \le n_{eq}$$

9

Example 1. **Nonlinear elastostatics** (cont.)

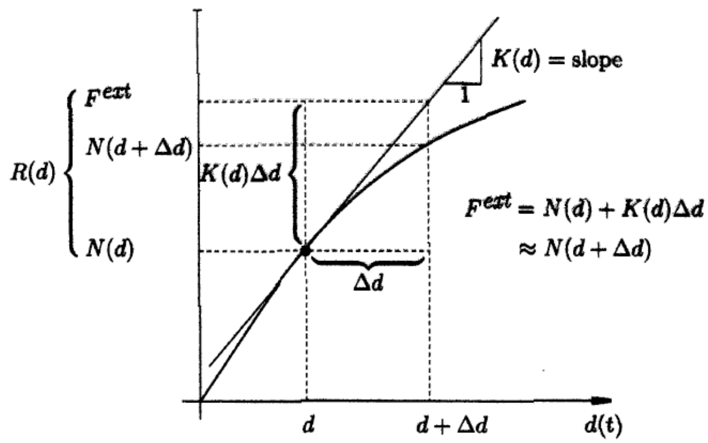
$$\text{(hyper-) elasticity} \Leftrightarrow \exists \overbrace{U(\boldsymbol{d})}^{\text{strain energy}} \text{ such that}$$

$$\boldsymbol{N} = \frac{\partial U}{\partial \boldsymbol{d}}, \quad \text{i.e.} \quad N_P = \frac{\partial U}{\partial d_P}$$

$$\Rightarrow \quad K_{PQ} = \frac{\partial N_P}{\partial d_Q} = \frac{\partial^2 U}{\partial d_Q \partial d_P} = K_{QP}$$

$$\boldsymbol{K} = \boldsymbol{K}^{\mathsf{T}} \quad \text{symmetric}$$

10




11

What's good about consistent tangents?

According to Taylor's theorem, the consistent tangent is the **best local approximation** in the sense that

$$\frac{\left\| \boldsymbol{K}(\boldsymbol{d})\Delta \boldsymbol{d} - \left( \boldsymbol{N}(\boldsymbol{d}+\Delta \boldsymbol{d}) - \boldsymbol{N}(\boldsymbol{d}) \right) \right\|}{\left\| \Delta \boldsymbol{d} \right\|} \to 0 \quad \text{as} \quad \left\| \Delta \boldsymbol{d} \right\| \to 0$$

$\boldsymbol{K}(\boldsymbol{d})$ is the unique matrix achieving this property.

**Interpretation:**

$\boldsymbol{N}(\boldsymbol{d}) + \boldsymbol{K}(\boldsymbol{d})\Delta \boldsymbol{d}$ is a **good local** approximation of $\boldsymbol{N}(\boldsymbol{d}+\Delta \boldsymbol{d})$.

12

Example 2. **Nonlinear heat conduction**

$$\boldsymbol{F}^{int} = \underbrace{\hat{\boldsymbol{K}}(\boldsymbol{d})}_{\text{symm. matrix}} \underbrace{\boldsymbol{d}}_{\text{temperature vector}}$$

consistent tangent: $\dfrac{\partial \boldsymbol{F}^{int}}{\partial \boldsymbol{d}} = \left[ \dfrac{\partial F_P^{int}}{\partial d_Q} \right]$

sum

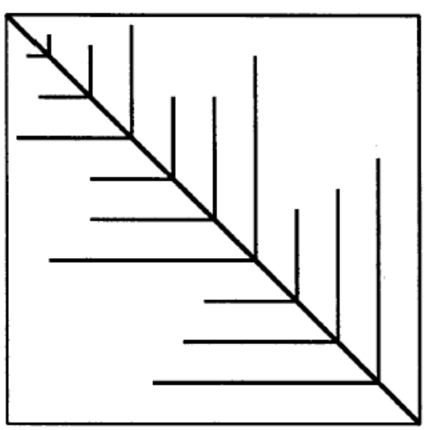
$$\frac{\partial(\hat{K}_{PR} d_R)}{\partial d_Q} = \underbrace{\hat{K}_{PQ}}_{\text{symm.}} + \frac{\partial \hat{K}_{PR}}{\partial d_Q} d_R$$

$$\frac{\partial \hat{K}_{PR}}{\partial d_Q} \overset{?}{=} \frac{\partial \hat{K}_{QR}}{\partial d_P}$$ NO!

**Consistent tangents are not always symmetric**.

13

**Costs increase with the loss of symmetric matrices**

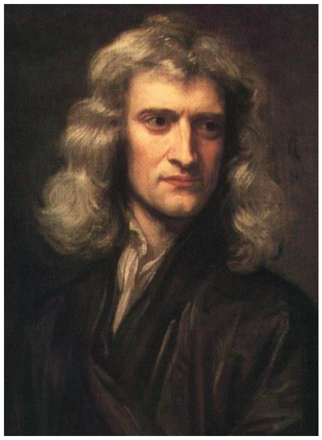


$\Rightarrow$ Storage $\times 2$

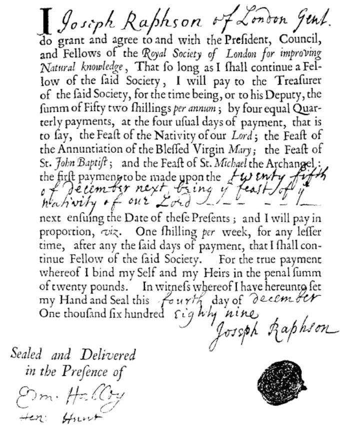
Factor $\times 2$

14

# Newton – Raphson Method



Isaac Newton



I Joseph Raphson of London Gent. do grant and agree to and with the President, Council, and Fellows of the *Royal Society of London for improving Natural knowledge*, That so long as I shall continue a Fellow of the said Society, I will pay to the Treasurer of the said Society, for the time being, or to his Deputy, the summ of Fifty two shillings *per annum*; by four equal Quarterly payments, at the four usual days of payment, that is to say, the Feast of the Nativity of our *Lord*; the Feast of the Annunciation of the Blessed Virgin *Mary*; the Feast of St. *John Baptist*; and the Feast of St. *Michael* the Archangel: the first payment to be made upon the twenty fifth of december next being y^e feast of y^e nativity of our Lord next ensuing the Date of these Presents; and I will pay in proportion, *viz.* One shilling *per* week, for any lesser time, after any the said days of payment, that I shall continue Fellow of the said Society. For the true payment whereof I bind my Self and my Heirs in the penal summ of twenty pounds. In witness whereof I have hereunto set my Hand and Seal this fourth day of december One thousand six hundred eighty nine

Joseph Raphson

*Sealed and Delivered in the Presence of*

Edm. Halley
Hen: Hunt

Joseph Raphson

15

A method based on the consistent tangent

**Newton-Raphson**

inc. loop

$$
\begin{aligned}
i &= 0 \quad \text{(iteration counter)} \\
\boldsymbol{d}^{(i)} &= \boldsymbol{d}_n \\
\boldsymbol{K}(\boldsymbol{d}^{(i)})\Delta\boldsymbol{d}^{(i)} &= \boldsymbol{R}^{(i)} \\
&\overset{\text{def}}{\equiv} \boldsymbol{F}_{n+1}^{ext} - \boldsymbol{F}^{int}(\boldsymbol{d}^{(i)}) \\
\boldsymbol{d}^{(i+1)} &= \boldsymbol{d}^{(i)} + \Delta\boldsymbol{d}^{(i)} \\
\left\|\boldsymbol{R}^{(i+1)}\right\| &\overset{?}{<} \varepsilon\left\|\boldsymbol{R}^{(0)}\right\| \quad \text{(conv. test)}
\end{aligned}
$$

preassigned tol.

it. loop

NO, $i \leftarrow i+1$

YES, $\boldsymbol{d}_{n+1} = \boldsymbol{d}^{(i+1)}, \quad n \leftarrow n+1$

16

where (strictly speaking)

$$\boldsymbol{K} = \frac{\partial \boldsymbol{F}^{int}}{\partial \boldsymbol{d}} = \text{ consistent tangent}$$

In practice, frequently replace consistent tangent, e.g., **modified Newton**, etc.

17

## The order of convergence of an iterative scheme

Consider 1 d.o.f.

Let $$e^{(i)} = d^{(i)} - \underbrace{d}_{\text{exact}}$$

Write $$\left|e^{(i+1)}\right| \leq c\left|e^{(i)}\right|^k$$

$k$ = **order of convergence**

18

Newton-Raphson is such that $k = 2$.

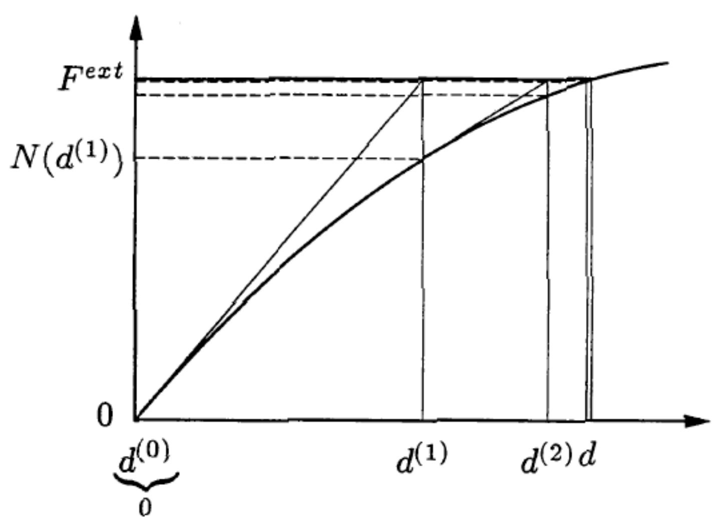
**Interpretation:**

Assume $\left|e^{(0)}\right| = 0.9,\ c = 1,$ then

| $i$ | $\left\vert e^{(i)}\right\vert$ is $\leq$ |
|---|---|
| 1 | .81 |
| 2 | .66 |
| 3 | .44 |
| 4 | .19 |
| 5 | $.\underline{0}36$ |
| 6 | $.\underline{00}13$ |
| 7 | $.\underline{0000}016$ |

quadratic convergence
$\Rightarrow$ the no. of 0's doubles
(at least) in each iteration

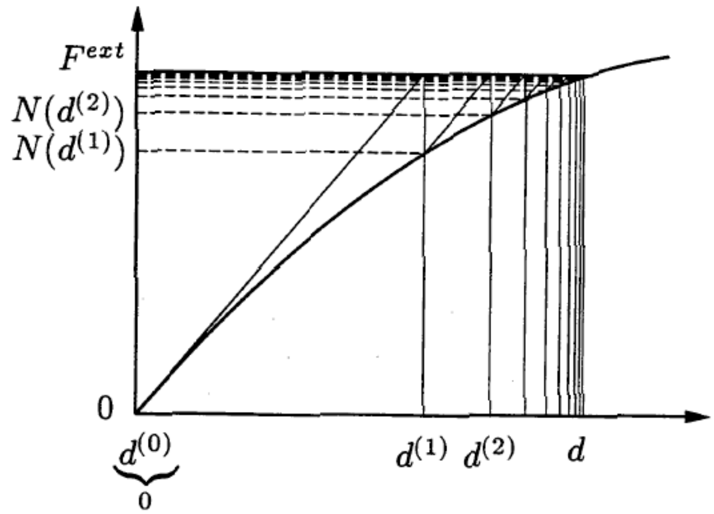
19




Newton-Raphson

$K = N'(d^{(i)})$

20

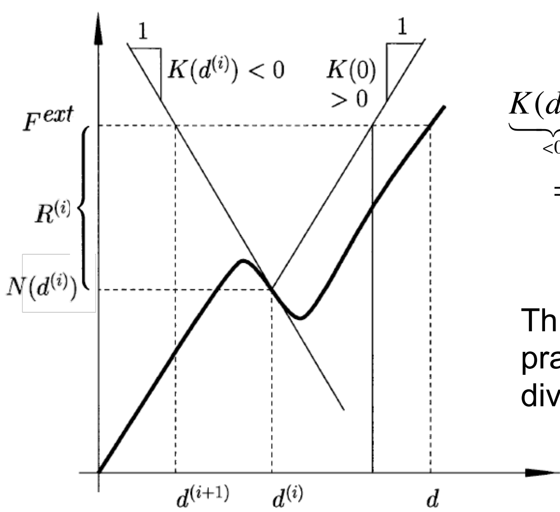



Modified Newton-Raphson

$$K = N'(\underbrace{d^{(0)}}_{0})$$

21

Globally, or far away from the solution, consistent tangents are frequently very poor.




$$\underbrace{K(d^{(i)})}_{<0}\Delta d^{(i)} = \underbrace{F^{ext} - N(d^{(i)})}_{>0}$$

$$\Rightarrow d^{(i+1)} - d^{(i)} = \Delta d^{(i)} < 0$$

This will recover, but in practical situations divergence may occur.

23

Developing an understanding of some aspects of nonlinear equation solving:

1. **Soft/stiff tangent behavior**

   1 d.o.f. linear model problem: $Kd = F, \quad K > 0$

Newton-Raphson-type algorithm

$$\underbrace{\tilde{K}}_{\alpha K} \Delta d^{(i)} = F - Kd^{(i)} \quad (= R^{(i)})$$

$$\alpha < 1, \quad \text{soft}$$

$$\alpha > 1, \quad \text{stiff}$$

$$\begin{aligned} \alpha K(d^{(i+1)} - d^{(i)}) &= F - Kd^{(i)} \\ -\alpha K(\underbrace{d \quad - d}_{0}) &= -(\underbrace{F - Kd}_{0}) \\ \hline \alpha K(e^{(i+1)} - e^{(i)}) &= -Ke^{(i)} \end{aligned}$$

24

where $e^{(i)} = d^{(i)} - d$

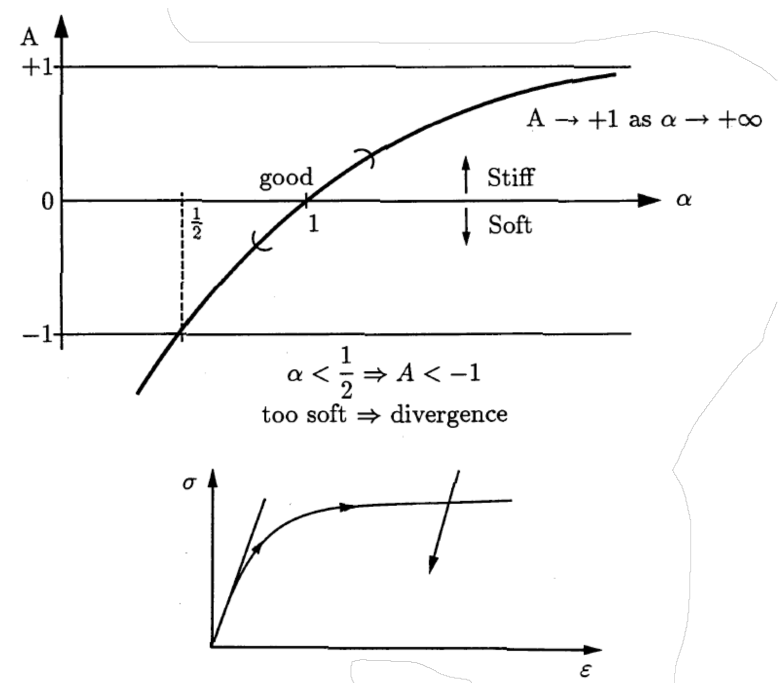
$$e^{(i+1)} = Ae^{(i)}$$

$$A = \frac{\alpha - 1}{\alpha} = 1 - \frac{1}{\alpha}$$

convergence $\Leftrightarrow \ |A| < 1$

divergence $\Leftrightarrow \ |A| > 1$

25




26

2. **Linear systems reduce to 1 d.o.f. model problem**

Consider

$$\tilde{\boldsymbol{K}}\Delta\boldsymbol{d}^{(i)} = \boldsymbol{F} - \boldsymbol{K}\boldsymbol{d}^{(i)} \qquad (*)$$

$$\boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + \Delta\boldsymbol{d}^{(i)}$$

Assume $\boldsymbol{K} = \boldsymbol{K}^{\mathsf{T}} > 0$

$$\tilde{\boldsymbol{K}} = \tilde{\boldsymbol{K}}^{\mathsf{T}} > 0$$

The associated eigenproblem:

$$(\boldsymbol{K} - \lambda\tilde{\boldsymbol{K}})\psi = 0$$

Solve for $\{\lambda,\psi\}_1^{n_{eq}}$; leads to uncoupled scalar versions of $(*)$:

$$\Delta d_l^{(i)} = F_l - \lambda_l d_l^{(i)}, \qquad l = 1,\ldots,n_{eq}$$

no sum

27

Drop "$l$" subscript, proceed as before:

$$
\begin{aligned}
d^{(i+1)} - d^{(i)} \quad &= \quad F - \lambda d^{(i)} \\
-(d \quad - d) \quad &= -(F - \lambda d) \underset{\text{defines } d}{\equiv} 0 \\
\hline
e^{(i+1)} \quad &= \quad (\underbrace{1-\lambda}_{A})e^{(i)}
\end{aligned}
$$

**Remarks**

1. $\lambda = \dfrac{1}{\alpha}$ (cf. previous analysis) $\Rightarrow$ each "mode" reduces to previous analysis.
2. A good $\tilde{\boldsymbol{K}}$ is such that $\lambda$'s are as close to unity as possible.
3. Soft modal components of $\tilde{\boldsymbol{K}}$ will cause the iteration to **diverge**.

28

# Line Search

**I. Linear case**

$$
\tilde{\boldsymbol{K}} \underbrace{\Delta \boldsymbol{d}^{(i)}}_{\substack{\text{search} \\ \text{dir. (vect.)}}} = \boldsymbol{F} - \boldsymbol{K}\boldsymbol{d}^{(i)} = \boldsymbol{R}^{(i)}
$$

$$
\boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + \underbrace{s^{(i)}}_{\substack{\text{search} \\ \text{param. (scalar)}}} \Delta \boldsymbol{d}^{(i)}
$$

What is the proper $s^{(i)}$?

29

**Two derivations**

1. Select $s^{(i)}$ such that the potential energy is minimized

$$p(s^{(i)}) = P(\boldsymbol{d}^{(i)} + s^{(i)}\Delta\boldsymbol{d}^{(i)})$$

where $$P(\boldsymbol{d}) = \frac{1}{2}\boldsymbol{d}^{\mathsf{T}}\boldsymbol{K}\boldsymbol{d} - \boldsymbol{d}^{\mathsf{T}}\boldsymbol{F}$$

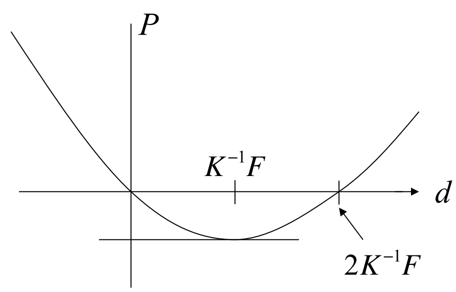
$$\frac{dp}{ds} = 0 \quad \Rightarrow$$

$$\boxed{s^{(i)} = \frac{\Delta\boldsymbol{d}^{(i)} \cdot \overset{\tilde{\boldsymbol{K}}\Delta\boldsymbol{d}^{(i)}}{\boldsymbol{R}^{(i)}}}{\Delta\boldsymbol{d}^{(i)} \cdot \boldsymbol{K}\Delta\boldsymbol{d}^{(i)}}}$$

$$> 0, \text{ if } \Delta\boldsymbol{d} \neq 0$$

30

$$Kd = F$$
$$P(d) = \tfrac{1}{2}d^T Kd - d^T F$$
$$P'(d) = 0 \Rightarrow Kd = F$$




1 d.o.f. :

$$s^{(i)} = \frac{\cancel{\Delta d}^{(i)^T}\tilde{K}\cancel{\Delta d}^{(i)}}{\cancel{\Delta d}^{(i)^T}K\cancel{\Delta d}^{(i)}}$$
$$= \alpha\frac{K}{K} = \alpha$$

31

**Two derivations** (cont.)

2. Select $s^{(i)}$ such that $\boldsymbol{R}^{(i+1)} = \boldsymbol{F} - \boldsymbol{K}\boldsymbol{d}^{(i+1)}$ has zero component in the direction of $\Delta\boldsymbol{d}^{(i)}$, viz.

$$0 = \Delta\boldsymbol{d}^{(i)} \cdot \boldsymbol{R}^{(i+1)} \Rightarrow s^{(i)} \text{ above}$$

This criterion is very general and may be applied when no potential exists.

32

**II. Nonlinear cases**

$$\tilde{\boldsymbol{K}}\Delta\boldsymbol{d}^{(i)} = \boldsymbol{F}^{ext} - \boldsymbol{F}^{int}(\boldsymbol{d}^{(i)})$$

$$\boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + s^{(i)}\Delta\boldsymbol{d}^{(i)}$$

Derivation 2 may be used:

$$(*) \qquad \underbrace{0 = G(s^{(i)})}_{\text{Nonlin. scalar eq. for } s^{(i)}} = \Delta\boldsymbol{d}^{(i)} \cdot \left(\boldsymbol{F}^{ext} - \boldsymbol{F}^{int}(\boldsymbol{d}^{(i)} + s^{(i)}\Delta\boldsymbol{d}^{(i)})\right)$$

Strang-Matthies suggest accepting $s^{(i)}$ if

$$\left|G(s^{(i)})\right| \leq \frac{1}{2}\left|G(0)\right|$$

For details see H. Matthies and G. Strang, "The solution of Nonlinear F.E. Equations," *IJNME*, Vol. 14, 1613-1626 (1979).

33

**Remarks**

1. If $\exists$ a potential $U(\boldsymbol{d})$ such that

$$\boldsymbol{F}^{int} = \boldsymbol{N} = \frac{\partial U}{\partial \boldsymbol{d}},$$

then derivation 1 also results in ($*$). This is the case for nonlinear elastostatics where

$$P(\boldsymbol{d}) = U(\boldsymbol{d}) - \boldsymbol{d}^{\mathsf{T}}\boldsymbol{F}$$

2. For situations in which the consistent tangent is not symmetric, positive definite, a minimization of the residual is frequently used to define $s$ :

$$\min_{s} \; \boldsymbol{R}(\boldsymbol{d} + s\Delta\boldsymbol{d}) \cdot \boldsymbol{R}(\boldsymbol{d} + s\Delta\boldsymbol{d})$$

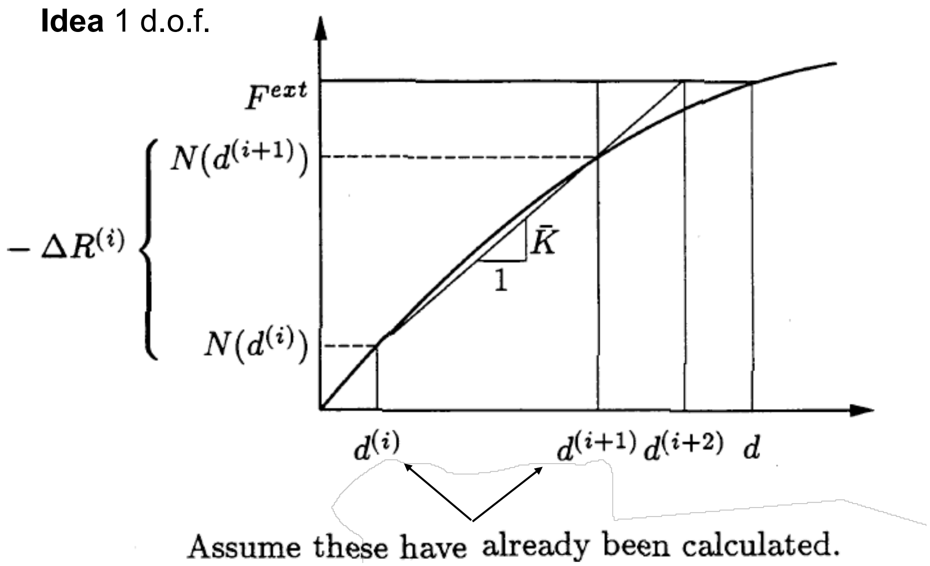
**Caution**: Structural bending elements, multiphysics, etc.

34

## Quasi-Newton Updates

(Also referred to as **secant updates**)

**Idea** 1 d.o.f.




35

$$\Delta R^{(i)} = R^{(i+1)} - R^{(i)}$$
$$= \left(\cancel{F}^{ext} - N(d^{(i+1)}) - \left(\cancel{F}^{ext} - N(d^{(i)})\right)\right)$$
$$= -\left(N(d^{(i+1)}) - N(d^{(i)})\right)$$

The **secant** $\bar{K}$ would suffice as a good approximate l.h.s. matrix

$$\bar{K} \equiv \frac{-\Delta R^{(i)}}{(d^{(i+1)} - d^{(i)})} \qquad (*)$$

$$\bar{K}(d^{(i+2)} - d^{(i+1)}) = R^{(i+1)} \equiv F^{ext} - N(d^{(i+1)})$$

defines $d^{(i+2)}$

36

For $n_{eq} = 1,$ $(*)$ uniquely defines $\bar{K}$. The generalization of $(*)$ for $n_{eq} > 1$ is

$$\bar{\boldsymbol{K}}(\underbrace{\boldsymbol{d}^{(i+1)} - \boldsymbol{d}^{(i)}}_{\text{known}}) = -\underbrace{\Delta \boldsymbol{R}^{(i)}}_{\text{known}}$$

This is called the **quasi - Newton equation**. It does not uniquely define $\bar{\boldsymbol{K}}$ for $n_{eq} > 1,$ but places a restriction on its definition.

Quasi-Newton updates provide definitions for $\bar{\boldsymbol{K}}$ such that the quasi-Newton equation is satisfied.

37

Quasi-Newton updates are often used with **line searches**.

The quasi-Newton equation becomes

$$\boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + s^{(i)} \Delta \boldsymbol{d}^{(i)}$$

$$\bar{\boldsymbol{K}} s^{(i)} \Delta \boldsymbol{d}^{(i)} = -\Delta \boldsymbol{R}^{(i)}$$

38

Perhaps the most widely used in F. E. analysis is BFGS, which we shall describe.

## BFGS (inverse) update

Broyden-Fletcher-Goldfarb-Shanno

$$\bar{\boldsymbol{K}}^{-1} = (\boldsymbol{I} + \boldsymbol{v}\boldsymbol{w}^{\mathsf{T}})\tilde{\boldsymbol{K}}^{-1}(\boldsymbol{I} + \boldsymbol{w}\boldsymbol{v}^{\mathsf{T}})$$

$\boldsymbol{v}, \boldsymbol{w}$ are vectors

**Properties**

1. $\tilde{\boldsymbol{K}}^{-1}$ symmetric $\Rightarrow \bar{\boldsymbol{K}}^{-1}$ symmetric
2. $\tilde{\boldsymbol{K}}^{-1} > 0$ (and $\boldsymbol{v} \cdot \boldsymbol{w} \neq -1$), then $\bar{\boldsymbol{K}}^{-1} > 0$
3. "Rank-2 update"

39

**Calculations** (from right to left)

$$
\begin{aligned}
\Delta\boldsymbol{d}^{(i+1)} &= \bar{\boldsymbol{K}}^{-1}\boldsymbol{R}^{(i+1)} \\
&= \underbrace{(\boldsymbol{I}+\boldsymbol{v}\boldsymbol{w}^{\mathsf{T}})\underbrace{\underbrace{\tilde{\boldsymbol{K}}^{-1}\underbrace{\underbrace{(\boldsymbol{I}+\boldsymbol{w}\boldsymbol{v}^{\mathsf{T}})\boldsymbol{R}^{(i+1)}}_{\boldsymbol{R}^{(i+1)}+(\boldsymbol{v}\cdot\boldsymbol{R}^{(i+1)})\boldsymbol{w}}}_{\bar{\boldsymbol{R}}}}_{\text{fwd. red./back sub.}}}_{=\Delta\bar{\boldsymbol{d}}}} \\
&= \Delta\bar{\boldsymbol{d}} + (\boldsymbol{w}\cdot\Delta\bar{\boldsymbol{d}})\boldsymbol{v}
\end{aligned}
$$

The factored $\tilde{\boldsymbol{K}}$ is never altered during the calculations

40

**Series of updates** ($i$ updates)

$\tilde{\boldsymbol{K}}^{-1}$ represents the last formed and factored stiffness.

$$
\begin{aligned}
\bar{\boldsymbol{K}}^{-1} = {} & (\boldsymbol{I}+\boldsymbol{v}^{(i)}\boldsymbol{w}^{(i)\mathsf{T}})(\boldsymbol{I}+\boldsymbol{v}^{(i-1)}\boldsymbol{w}^{(i-1)\mathsf{T}})\ldots \\
& (\boldsymbol{I}+\boldsymbol{v}^{(1)}\boldsymbol{w}^{(1)\mathsf{T}})\tilde{\boldsymbol{K}}^{-1}(\boldsymbol{I}+\boldsymbol{w}^{(1)}\boldsymbol{v}^{(1)\mathsf{T}})\times \\
& (\boldsymbol{I}+\boldsymbol{w}^{(2)}\boldsymbol{v}^{(2)\mathsf{T}})\ldots(\boldsymbol{I}+\boldsymbol{w}^{(i)}\boldsymbol{v}^{(i)\mathsf{T}})
\end{aligned}
$$

41

$$
\begin{aligned}
\Delta \boldsymbol{d}^{(i+1)} &= \tilde{\boldsymbol{K}}^{-1}\boldsymbol{R}^{(i+1)} \\
&= \underbrace{\overset{i}{(\ \ )}\,\overset{i-1}{(\ \ )}\,\cdots\,\overset{1}{(\ \ )}}\ \Big|\,\tilde{\boldsymbol{K}}^{-1}\,\underbrace{\overset{1}{(\ \ )}\,\overset{2}{(\ \ )}\,\cdots\,\overset{i}{(\ \ )}}\ \boldsymbol{R}^{(i+1)}
\end{aligned}
$$

$$
\begin{cases}
1. & \bar{\boldsymbol{R}}^{(1)} = \boldsymbol{R}^{(i+1)} + \left(\boldsymbol{v}^{(i)}\cdot \boldsymbol{R}^{(i+1)}\right)\boldsymbol{w}^{(i)} \\
2. & \bar{\boldsymbol{R}}^{(2)} = \bar{\boldsymbol{R}}^{(1)} + \left(\boldsymbol{v}^{(i-1)}\cdot \bar{\boldsymbol{R}}^{(1)}\right)\boldsymbol{w}^{(i-1)} \\
\vdots & \quad\ \ \vdots \\
i. & \bar{\boldsymbol{R}}^{(i)} = \bar{\boldsymbol{R}}^{(i-1)} + \left(\boldsymbol{v}^{(1)}\cdot \bar{\boldsymbol{R}}^{(i-1)}\right)\boldsymbol{w}^{(1)}
\end{cases}
$$

$$
\begin{cases}
\textbf{summary:} \quad \text{do} \quad k = 1,\ldots,i \\
\quad \bar{\boldsymbol{R}}^{(k)} = \bar{\boldsymbol{R}}^{(k-1)} + \left(\boldsymbol{v}^{(i-k+1)}\cdot \bar{\boldsymbol{R}}^{(k-1)}\right)\boldsymbol{w}^{(i-k+1)} \\
\text{where} \qquad \bar{\boldsymbol{R}}^{(0)} \equiv \boldsymbol{R}^{(i+1)}
\end{cases}
$$

$$
\text{Solve} \qquad \Delta \bar{\boldsymbol{d}}^{(0)} = \tilde{\boldsymbol{K}}^{-1}\bar{\boldsymbol{R}}^{(i)}
$$

$$
\Delta \boldsymbol{d}^{(i+1)} = \underbrace{\overset{i}{(\ \ )}\,\overset{i-1}{(\ \ )}\cdots\overset{1}{(\ \ )}\,\Delta \bar{\boldsymbol{d}}^{(0)}}
$$

42

$$
\begin{cases}
1. & \Delta \bar{\boldsymbol{d}}^{(1)} = \Delta \bar{\boldsymbol{d}}^{(0)} + \left(\boldsymbol{w}^{(1)}\cdot \Delta \bar{\boldsymbol{d}}^{(0)}\right)\boldsymbol{v}^{(1)} \\
2. & \Delta \bar{\boldsymbol{d}}^{(2)} = \Delta \bar{\boldsymbol{d}}^{(1)} + \left(\boldsymbol{w}^{(2)}\cdot \Delta \bar{\boldsymbol{d}}^{(1)}\right)\boldsymbol{v}^{(2)} \\
\vdots & \quad\ \ \vdots \\
i. & \Delta \bar{\boldsymbol{d}}^{(i)} = \Delta \bar{\boldsymbol{d}}^{(i-1)} + \left(\boldsymbol{w}^{(i)}\cdot \Delta \bar{\boldsymbol{d}}^{(i-1)}\right)\boldsymbol{v}^{(i)}
\end{cases}
$$

$$
\begin{cases}
\textbf{summary:} \quad \text{do} \quad k = 1,\ldots,i \\
\qquad \Delta \bar{\boldsymbol{d}}^{(k)} = \Delta \bar{\boldsymbol{d}}^{(k-1)} + \left(\boldsymbol{w}^{(k)}\cdot \Delta \bar{\boldsymbol{d}}^{(k-1)}\right)\boldsymbol{v}^{(k)} \\
\qquad \text{define} \quad \Delta \boldsymbol{d}^{(i+1)} = \Delta \bar{\boldsymbol{d}}^{(i)}
\end{cases}
$$

43

**Storage** $= \text{no. of BFGS updates} \times 2 \times n_{eq}$

$= \{\boldsymbol{v}^{(i)}, \boldsymbol{w}^{(i)}\}_{i=1}^{\text{no. updates}}$

**Definitions of BFGS vectors:**

$$\boldsymbol{v}^{(k)} = \frac{\Delta \boldsymbol{d}^{(k)}}{(\Delta \boldsymbol{d}^{(k)} \cdot \Delta \boldsymbol{R}^{(k)})}$$

$$\boldsymbol{w}^{(k)} = -\Delta \boldsymbol{R}^{(k)} + \alpha^{(k)} \boldsymbol{R}^{(k)}$$

where $$\alpha^{(k)} = \sqrt{\frac{-s^{(k)} \Delta \boldsymbol{R}^{(k)} \cdot \Delta \boldsymbol{d}^{(k)}}{\boldsymbol{R}^{(k)} \cdot \Delta \boldsymbol{d}^{(k)}}}$$

44

**Alternative definitions of BFGS vectors commonly used in practice.**

**Recall:** $$G^{(k)}(s^{(k)}) = \Delta \boldsymbol{d}^{(k)} \cdot \left( \boldsymbol{R}(\boldsymbol{d}^{(k)} + s^{(k)} \Delta \boldsymbol{d}^{(k)} \right)$$

$$G^{(k)}(0) = \Delta \boldsymbol{d}^{(k)} \cdot \left( \boldsymbol{R}(\boldsymbol{d}^{(k)}) \right)$$

$$G^{(k)}(s^{(k)}) - G^{(k)}(0) = \Delta \boldsymbol{d}^{(k)} \cdot \Delta \boldsymbol{R}^{(k)}$$

---

$$\boldsymbol{v}^{(k)} = \frac{\Delta \boldsymbol{d}^{(k)}}{(G^{(k)}(s^{(k)}) - G^{(k)}(0))}$$

$$\boldsymbol{w}^{(k)} = -\Delta \boldsymbol{R}^{(k)} + \alpha^{(k)} \boldsymbol{R}^{(k)}$$

where $$\alpha^{(k)} = \sqrt{\frac{-s^{(k)} \left( G^{(k)}(s^{(k)}) - G^{(k)}(0) \right)}{G^{(k)}(0)}}$$

(See Strang and Matthies for further details.)

45

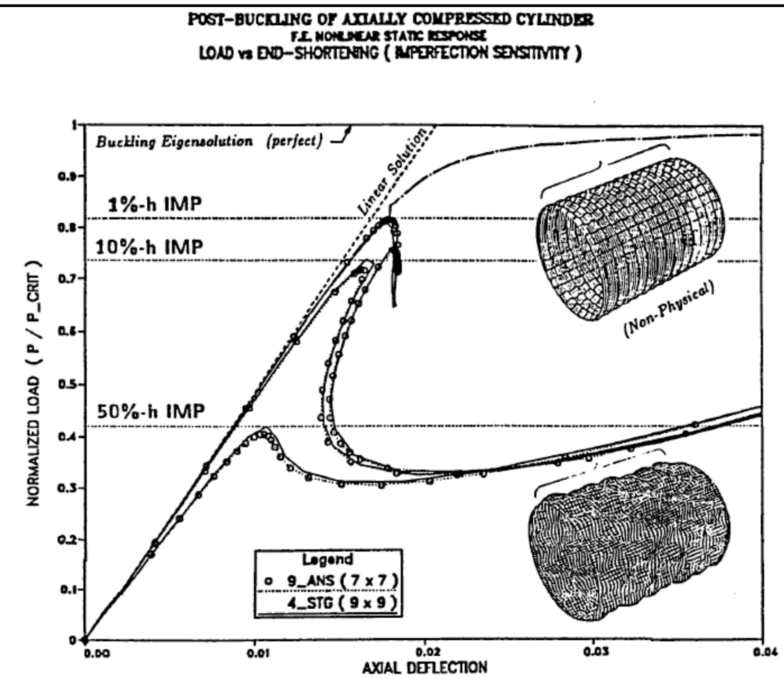



Figure 6.16 Post-Buckling of Axially Compressed Cylinder: Imperfection Sensitivity with Fixed Mesh

46

**Arc-length strategies** (written in collaboration with R.M. Ferencz)

So far in this lecture external loading has been assumed to be a **given** function of time. Now we wish to trace load-displacement processes in which there is no **a priori** knowledge of the load-time function.

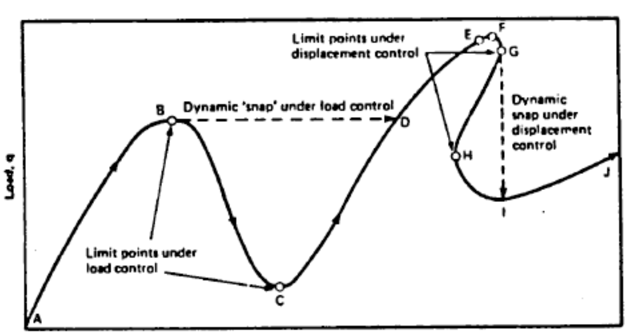



Fig. 1. "Snap buckling". (Crisfield 1981)

47

Better methods are clearly needed. In structures, the basic ideas are due to Riks, Wempner, Ramm, Bergan and Crisfield. Some useful references follow. In the mathematical literature, these techniques are referred to as **continuation methods**, in structures they are called **arc-length methods**.

**References**


- M.A. Crisfield, "A Fast Incremental/Iterative Solution Procedure that Handles 'Snap-through'," *Computers and Structures*, Vol. 13, 55-62, (1981).
- M.A. Crisfield, "An Arc-length Method Including Line Searches and Accelerations," *IJNME*, Vol. 19, 1269-1289, (1983).


48

**References** (cont.)


- K.H. Schweizerhof and P. Wriggers, "Consistent Linearization for Path Following Methods in Nonlinear F.E. Analysis," *CMAME*, Vol. 59, 261-279, (1986)

- C.A. Felippa, "Solution of Nonlinear Static Equations," Handbook Series on Computer Methods in Mechanics, Volume on Large Deflection and Stability of Structures.


49

The essential idea in arc-length strategies is to control the arc-length of the solution curve in force-displacement space. This will accommodate automatic unloading as is needed to trace the solution curves in the above fig. A scalar arc-length function (i.e. constraint equation) is introduced along with an additional unknown, the loading parameter, to automatically control the amplitude of applied loading. The continuous version of the problem may be stated as follows:

Find $\boldsymbol{d}(t)$ and $\lambda(t)$ such that

$$(*) \quad \begin{cases} \boldsymbol{F}^{int}(\boldsymbol{d}(t)) = \lambda(t)\boldsymbol{F}^{ext} \\ f(\dot{\boldsymbol{d}}, \dot{\lambda}) = \dot{a} = \text{ given} \end{cases}$$

where

$a =$ arc-length

$f =$ arc-length function

$$(\dot{\ }) = \frac{\partial(\ )}{\partial t}$$

50

$f$ should satisfy the definition of a norm on the $n_{eq} + 1$ dimensional space consisting of pairs $\dot{\boldsymbol{d}},\ \dot{\lambda}$.

The discrete form of $(*)$ is seen more frequently in the literature:

$$\left.\begin{aligned} \boldsymbol{F}^{int}(\underbrace{\boldsymbol{d}_{n+1}}_{\text{unknown}}) &= \underbrace{\lambda_{n+1}}_{\text{unknown}} \boldsymbol{F}^{ext} \\ f(\delta\boldsymbol{d}, \delta\lambda) &\equiv f(\boldsymbol{d}_{n+1} - \boldsymbol{d}_n, \lambda_{n+1} - \lambda_n) \\ &= \delta a = \text{ given} \end{aligned}\right\} \qquad (**)$$

51

As a candidate $f$ consider:

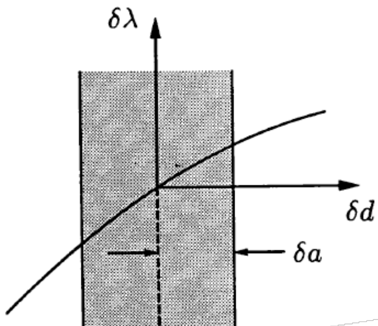
$$f(\delta \boldsymbol{d}, \delta\lambda) = \sqrt{(c\delta\boldsymbol{d}^{\mathsf{T}}\mathrm{diag}\tilde{\boldsymbol{K}}_0\delta\boldsymbol{d} + b\delta\lambda^2}$$

where

$$\begin{aligned} c &= \frac{(1-b)}{(\boldsymbol{q}^{\mathsf{T}}\mathrm{diag}\tilde{\boldsymbol{K}}_0\boldsymbol{q})} \\ \boldsymbol{q} &= \tilde{\boldsymbol{K}}_0^{-1}\boldsymbol{F}^{ext} \\ b &\in [0,1] \end{aligned}$$

and $\tilde{\boldsymbol{K}}_0$ is the first formed/factorized l.h.s. matrix.

52

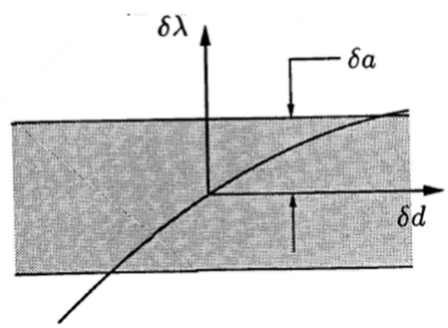



$$b = 0 \quad \text{(disp. control)}$$

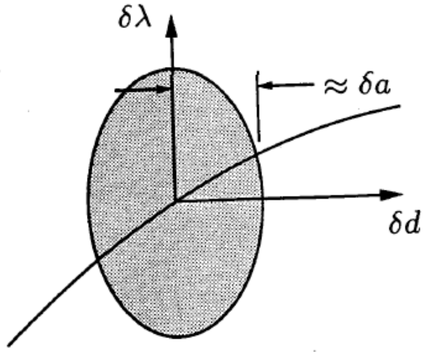
$$\delta d = \sqrt{\frac{\delta\boldsymbol{d}^T\ \mathrm{diag}\tilde{\boldsymbol{K}}_0\,\delta\boldsymbol{d}}{\boldsymbol{q}^T\mathrm{diag}\tilde{\boldsymbol{K}}_0\,\boldsymbol{q}}}$$

$b = 1$

(load control)

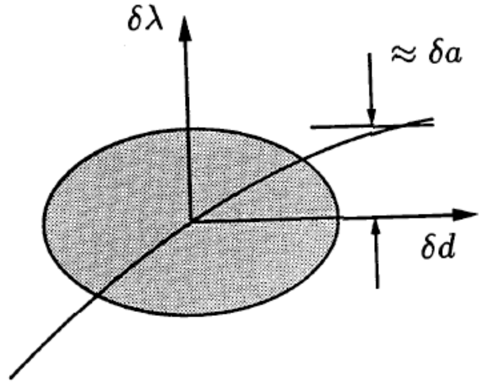



53




$$b = \varepsilon$$

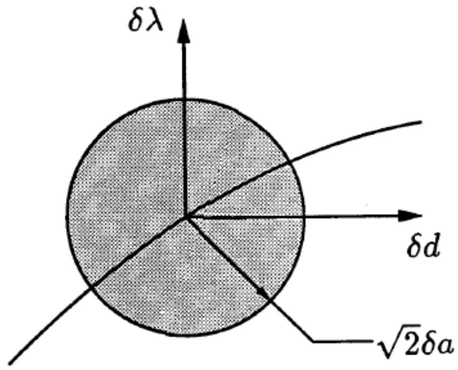
$$0 < \varepsilon \ll 1$$




$$b = 1 - \varepsilon$$

$$0 < \varepsilon \ll 1$$

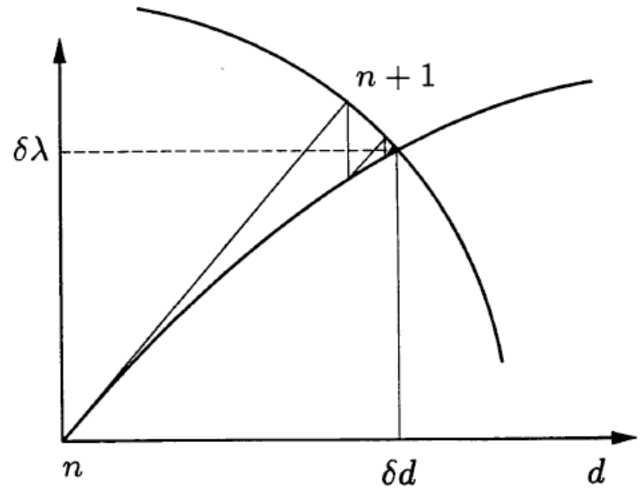
54




$$b = \frac{1}{2}$$

Experience seems to favor displacement control ($b = 0$).

55




56

**Solution of** $(**)$

Consistently linearize $(**)$ w.r.t. $\boldsymbol{d}_{n+1}$ and $\lambda_{n+1}$. That is

$$\boldsymbol{d}_{n+1} \leftarrow \boldsymbol{d}_{n+1}^{(i)} + \Delta\boldsymbol{d}^{(i)}$$
$$\lambda_{n+1} \leftarrow \lambda_{n+1}^{(i)} + \Delta\lambda^{(i)}$$

in $(**)$ and linearize w.r.t. $\Delta\boldsymbol{d}^{(i)}$ and $\Delta\lambda^{(i)}$. Using Taylor's formula:

$$\begin{aligned}\boldsymbol{F}^{int}(\boldsymbol{d}_{n+1}^{(i)}) + \boldsymbol{K}(\boldsymbol{d}_{n+1}^{(i)})\Delta\boldsymbol{d}^{(i)} &\cong \boldsymbol{F}^{int}(\boldsymbol{d}_{n+1}^{(i)} + \Delta\boldsymbol{d}^{(i)}) \\ &= (\lambda_{n+1}^{(i)} + \Delta\lambda^{(i)})\boldsymbol{F}^{ext}\end{aligned}$$

57

$$f_{n+1}^{(i)} + \left(\frac{\partial f}{\partial \delta \boldsymbol{d}}\right)_{n+1}^{(i)} \cdot \Delta \boldsymbol{d}^{(i)} + \left(\frac{\partial f}{\partial \delta \lambda}\right)_{n+1}^{(i)} \cdot \Delta \lambda^{(i)}$$
$$\cong f(\boldsymbol{d}_{n+1}^{(i)} + \Delta \boldsymbol{d}^{(i)} - \boldsymbol{d}_n, \lambda_{n+1}^{(i)} + \Delta \lambda^{(i)} - \lambda_n)$$
$$= \delta a$$

where

$$\begin{aligned} f_{n+1}^{(i)} &= f(\delta \boldsymbol{d}^{(i)}, \delta \lambda^{(i)}) \\ \left(\frac{\partial f}{\partial \delta \boldsymbol{d}}\right)_{n+1}^{(i)} &= \frac{\partial f}{\partial \delta \boldsymbol{d}}(\delta \boldsymbol{d}^{(i)}, \delta \lambda^{(i)}) \\ \left(\frac{\partial f}{\partial \delta \lambda}\right)_{n+1}^{(i)} &= \frac{\partial f}{\partial \delta \lambda}(\delta \boldsymbol{d}^{(i)}, \delta \lambda^{(i)}) \\ \delta \boldsymbol{d}^{(i)} &= \boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n \\ \delta \lambda^{(i)} &= \lambda_{n+1}^{(i)} - \lambda_n \end{aligned}$$

58

$n_{eq} + 1$ **equation matrix problem :**

$$\begin{bmatrix} \boldsymbol{K} & -\boldsymbol{F}^{ext} \\ \dfrac{\partial f}{\partial \delta \boldsymbol{d}} & \dfrac{\partial f}{\partial \delta \lambda} \end{bmatrix} \begin{Bmatrix} \Delta \boldsymbol{d}^{(i)} \\ \Delta \lambda^{(i)} \end{Bmatrix} = \begin{Bmatrix} \overbrace{\lambda_{n+1}^{(i)} \boldsymbol{F}^{ext} - \boldsymbol{F}^{int}(\boldsymbol{d}_{n+1}^{(i)})}^{\boldsymbol{R}_{n+1}^{(i)}} \\ \delta a - f_{n+1}^{(i)} \end{Bmatrix}$$

**Remarks**

1. $\boldsymbol{K}$ can be replaced by any convenient l.h.s. matrix, say $\bar{\boldsymbol{K}}$, accounting, e.g., for quasi-Newton updates.
2. Line searches may also be accounted for by updating

$$\boldsymbol{d}_{n+1}^{(i+1)} = \boldsymbol{d}_{n+1}^{(i)} + s^{(i)} \Delta \boldsymbol{d}^{(i)}$$

59

where $s^{(i)}$ is defined by

$$
\begin{aligned}
0 &= G(s^{(i)}) \\
&\equiv \Delta \boldsymbol{d}^{(i)} \cdot \left( \lambda_{n+1}^{(i)} \boldsymbol{F}^{ext} - \boldsymbol{F}^{int} ( \boldsymbol{d}_{n+1}^{(i)} + s^{(i)} \Delta \boldsymbol{d}^{(i)} ) \right)
\end{aligned}
$$

3. It is not expensive to calculate $\dfrac{\partial f}{\partial \delta \boldsymbol{d}}$ and $\dfrac{\partial f}{\partial \delta \lambda}$, so normally the latest updated values are used.
4. Explicit formulae:

$$
\left( \frac{\partial f}{\partial \delta \boldsymbol{d}} \right)_{n+1}^{(i)} = \frac{c}{f_{n+1}^{(i)}} \operatorname{diag} \tilde{\boldsymbol{K}}_0 \delta \boldsymbol{d}^{(i)}
$$

$$
\left( \frac{\partial f}{\partial \delta \lambda} \right)_{n+1}^{(i)} = \frac{b}{f_{n+1}^{(i)}} \delta \lambda^{(i)}
$$

60

**Solution** (Note: $\dfrac{\partial f}{\partial \delta \lambda}$ may be zero!)

1. Solve first eq. for $\Delta \boldsymbol{d}^{(i)}$ as a function of $\Delta \lambda^{(i)}$ (unknown).
2. Substitute this result into second eq. and solve for $\Delta \lambda^{(i)}$.

**Results**

$$
\begin{aligned}
\Delta \lambda^{(i)} &= \frac{\left( \delta a - f_{n+1}^{(i)} \right) - \left( \dfrac{\partial f}{\partial \delta \boldsymbol{d}} \right)_{n+1}^{(i)} \cdot \Delta \bar{\boldsymbol{d}}}{\left( \left( \dfrac{\partial f}{\partial \delta \boldsymbol{d}} \right)_{n+1}^{(i)} \cdot \bar{\boldsymbol{q}} + \left( \dfrac{\partial f}{\partial \delta \lambda} \right)_{n+1}^{(i)} \right)} \\
\Delta \boldsymbol{d}_{n+1}^{(i)} &= \Delta \lambda^{(i)} \bar{\boldsymbol{q}} + \Delta \bar{\boldsymbol{d}} \\
\Delta \bar{\boldsymbol{d}} &= \bar{\boldsymbol{K}}^{-1} \boldsymbol{R}_{n+1}^{(i)} \\
\bar{\boldsymbol{q}} &= \bar{\boldsymbol{K}}^{-1} \boldsymbol{F}^{ext}
\end{aligned}
$$

61

**Initilization**

To get the arc-length method started in each step requires **specification** of $\delta\lambda^{(0)}$ and $\delta\boldsymbol{d}^{(0)}$.

Let

$$\begin{aligned}\tilde{\boldsymbol{q}} &= \tilde{\boldsymbol{K}}^{-1}(\boldsymbol{d}_n)\boldsymbol{F}^{ext} \\ \delta\boldsymbol{d}^{(0)} &= \delta\lambda^{(0)}\tilde{\boldsymbol{q}}\end{aligned}$$

Thus all that is required is to define $\delta\lambda^{(0)}$. The arc-length constraint determines the **magnitude** of $\delta\lambda^{(0)}$.

62

$$\begin{aligned}\delta a &= f(\delta\boldsymbol{d}^{(0)},\delta\lambda^{(0)}) \\ &= f(\delta\lambda^{(0)}\tilde{\boldsymbol{q}},\delta\lambda^{(0)}) \\ &= \left|\delta\lambda^{(0)}\right| f(\tilde{\boldsymbol{q}},1)\end{aligned}$$

$$\Rightarrow \quad \left|\delta\lambda^{(0)}\right| = \frac{\delta a}{f(\tilde{\boldsymbol{q}},1)}$$

To determine the **sign** of $\delta\lambda^{(0)} = \delta\lambda_{n+1}^{(0)}$, we proceed as follows:

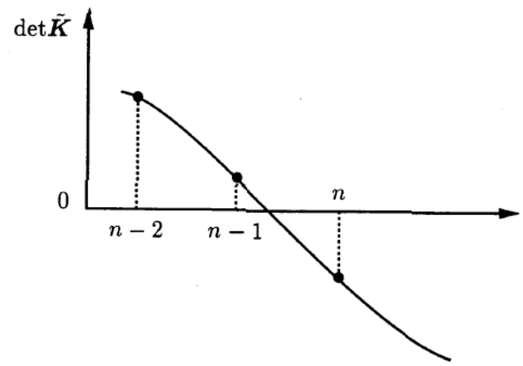
63

If the following two conditions are true we set

$$\operatorname{sign} \delta\lambda^{(0)} = -\operatorname{sign} \delta\lambda_n^{(0)} :$$

(i) $\det \tilde{\boldsymbol{K}}(\boldsymbol{d}_n) = -\det \tilde{\boldsymbol{K}}(\boldsymbol{d}_{n-1})$

(ii) The determinant has passed through **zero**; for example:




64

If **either** of the above is **not** true we set

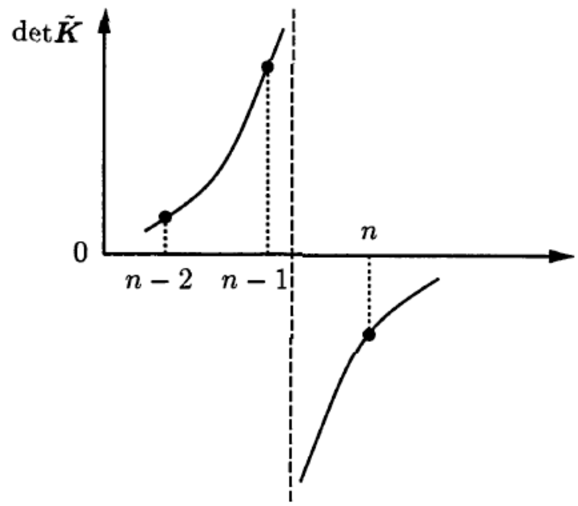
$$\operatorname{sign} \delta\lambda^{(0)} = \operatorname{sign} \delta\lambda_n^{(0)}$$

**Remarks**

To understand the rationale behind the strategy, consider the scalar case. The condition that the determinant passed through zero corresponds to a **maximum** or a **minimum** in the load-displacement diagram:

65

Note that the determinant can change signs by passing through **infinity**; for example:




66

This corresponds to **turning points** in the load-deflection curve:

**Remarks**

1. An adaptive strategy for increasing/decreasing the size of $\delta a$ is essential for good performance of the algorithm.
2. Arc-length strategies have not yet progressed to the point where they are 100% reliable.

67

**Convergence**

def. $\|\boldsymbol{R}\|^2 = \boldsymbol{R}^{\mathsf{T}}(\operatorname{diag}\tilde{\boldsymbol{K}}_0)^{-1}\boldsymbol{R}$

(i) Insist $\left\|\boldsymbol{R}_{n+1}^{(i)}\right\| \leq \varepsilon_R\left\|\boldsymbol{R}^{(0)}\right\|$

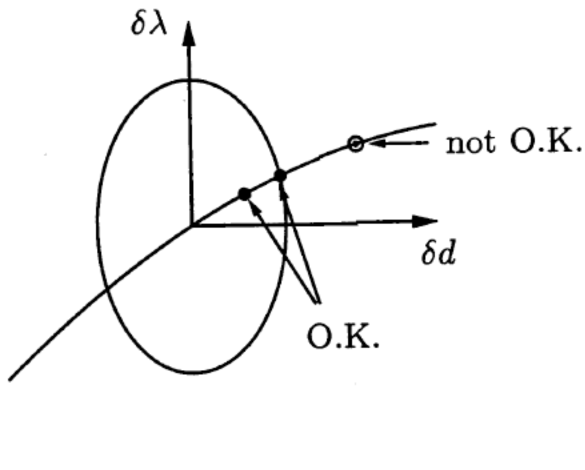
where $\varepsilon_R$ is a user-specified tolerance.

(ii) In addition, we must ensure arc-length constraint is maintained:

$$f_{n+1}^{(i)} - \delta a \leq \varepsilon_f \delta a$$

68

**Interpretation:** $\boldsymbol{d}_{n+1}^{(i)}, \lambda_{n+1}^{(i)}$ is inside the ellipse.




69

Some codes also employ a displacement criterion:

Let $\|\boldsymbol{d}\|^2 = \boldsymbol{d}^{\mathsf{T}} \operatorname{diag} \tilde{\boldsymbol{K}}_0 \boldsymbol{d}$

Insist,

$$\left\|\Delta \boldsymbol{d}^{(i)}\right\| \le \varepsilon_d \left\|\delta \boldsymbol{d}^{(i)}\right\| = \varepsilon_d \left\|\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n\right\|$$

As long as $b \neq 1$ (i.e., not load control), the displacement criterion is redundant with the arc-length constraint. In other situation (e.g., Newton-like methods), the displacement criterion should generally be used.

The next slide shows an application of the arc-length method to a shell buckling problem. From G. Stanley, Ph.D. Thesis, Stanford, 1985.

70

POST–BUCKLING OF AXIALLY COMPRESSED CYLINDER
F.E. NONLINEAR STATIC RESPONSE
LOAD vs END–SHORTENING ( IMPERFECTION SENSITIVITY )

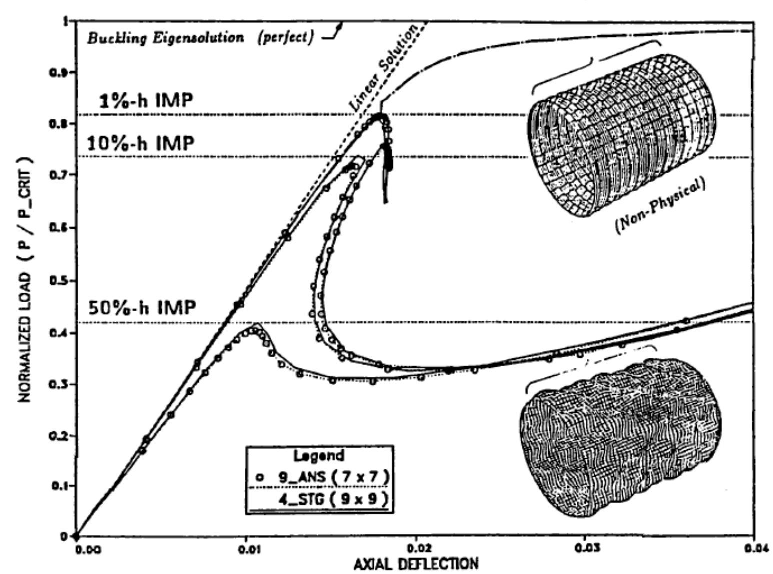



Figure 6.16 Post-Buckling of Axially Compressed Cylinder: Imperfection Sensitivity with Fixed Mesh

71