Introduction to Simulation - Lecture 10

Modified Newton Methods

Jacob White

Thanks to Deepak Ramaswamy, Jaime Peraire, Michal Rewienski, and Karen Veroy

Outline

- Damped Newton Schemes
  - Globally Convergent if Jacobian is Nonsingular
  - Difficulty with Singular Jacobians
- Introduce Continuation Schemes
  - Problem with Source/Load stepping
  - More General Continuation Scheme
- Improving Continuation Efficiency
  - Better first guess for each continuation step
- Arc Length Continuation

Multidimensional Newton Method

Newton Algorithm

Newton Algorithm for Solving $F(x) = 0$

$x^0 =$ Initial Guess, $k = 0$

Repeat {

Compute $F(x^k), J_F(x^k)$

Solve $J_F(x^k)(x^{k+1} - x^k) = -F(x^k)$ for x^{k+1}

$k = k + 1$

} Until $\left\|x^{k+1} - x^k\right\|$, $\left\|F(x^{k+1})\right\|$ small enough

Main Theorem

If

$a)\ \left\| J_F^{-1}\left(x^k\right) \right\| \le \beta$ (Inverse is bounded)

$b)\ \left\| J_F\left(x\right) - J_F\left(y\right) \right\| \le \ell \left\| x - y \right\|$ (Derivative is Lipschitz Cont)

Then Newton's method converges given a sufficiently close initial guess

Multidimensional Newton Method

Multidimensional Convergence Theorem

Implications

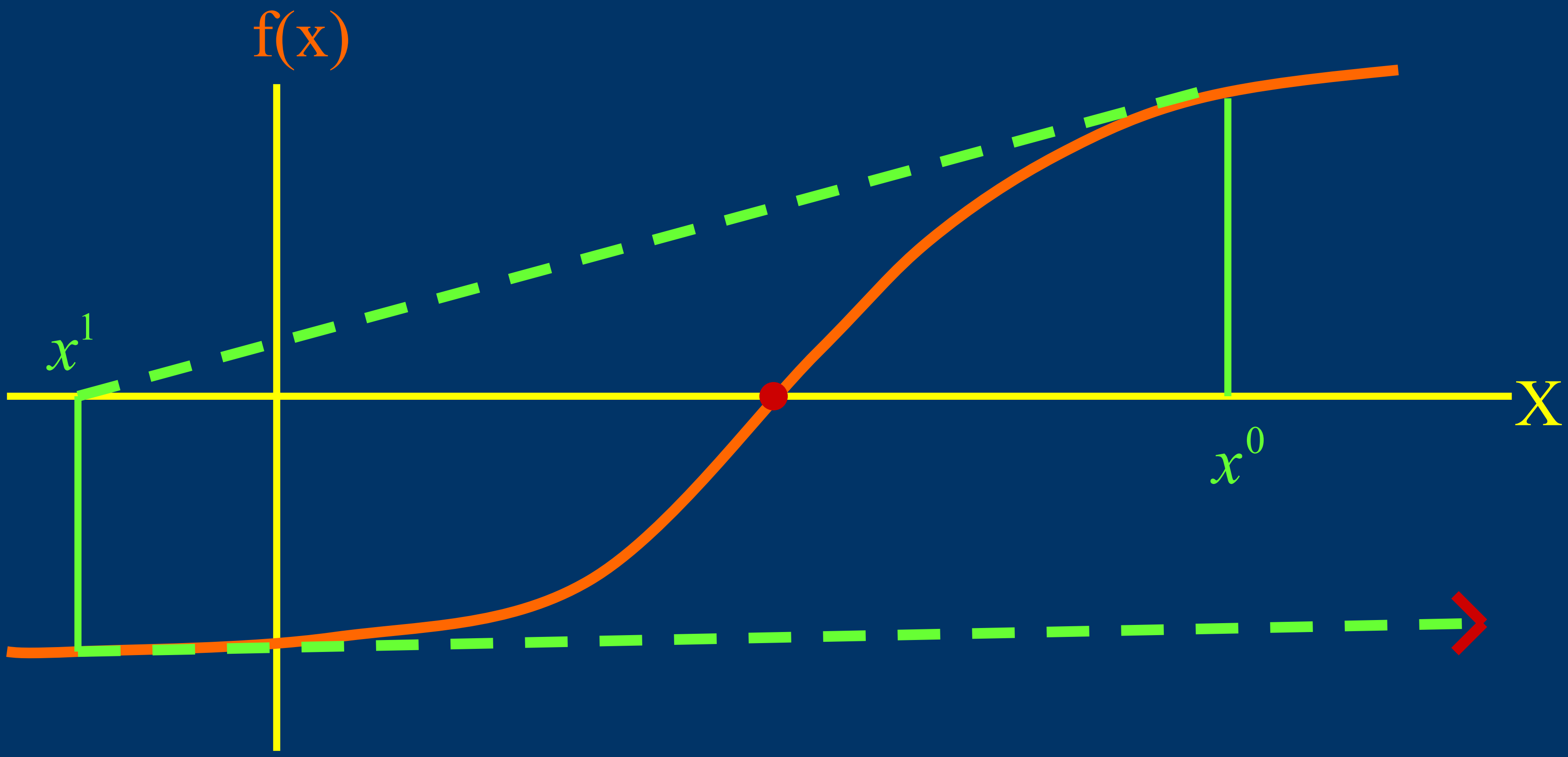
If a function's first derivative never goes to zero, and its second derivative is never too large…

Then Newton's method can be used to find the zero of the function *provided you all ready know the answer.*

Need a way to develop Newton methods which converge regardless of initial guess!

Non-converging Case

1-D Picture



Limiting the changes in X might improve convergence

Newton Algorithm for Solving $F(x) = 0$

$x^0 =$ Initial Guess, $k = 0$

Repeat {

Compute $F(x^k), J_F(x^k)$

Solve $J_F(x^k)\Delta x^{k+1} = -F(x^k)$ for Δx^{k+1}

$x^{k+1} = x^k + \text{limited}(\Delta x^{k+1})$

$k = k+1$

} Until $\left\|\Delta x^{k+1}\right\|, \left\|F(x^{k+1})\right\|$ small enough

Newton Method with Limiting

Damped Newton Scheme

General Damping Scheme

Solve $J_F\left(x^k\right)\Delta x^{k+1} = -F\left(x^k\right)$ for Δx^{k+1}

$$x^{k+1} = x^k + \alpha^k \Delta x^{k+1}$$

Key Idea: Line Search

Pick α^k to minimize $\left\|F\left(x^k + \alpha^k \Delta x^{k+1}\right)\right\|_2^2$

$$\left\|F\left(x^k + \alpha^k \Delta x^{k+1}\right)\right\|_2^2 \equiv F\left(x^k + \alpha^k \Delta x^{k+1}\right)^T F\left(x^k + \alpha^k \Delta x^{k+1}\right)$$

Method Performs a one-dimensional search in Newton Direction

Newton Method with Limiting

Damped Newton

Convergence Theorem

If

$a) \quad \left\| J_F^{-1}\left(x^k\right) \right\| \leq \beta \qquad \left(\text{Inverse is bounded}\right)$

$b) \quad \left\| J_F\left(x\right) - J_F\left(y\right) \right\| \leq \ell \left\| x - y \right\| \qquad \left(\text{Derivative is Lipschitz Cont}\right)$

Then

There exists a set of α^k's $\in \left(0,1\right]$ such that

$$\left\| F\left(x^{k+1}\right) \right\| = \left\| F\left(x^k + \alpha^k \Delta x^{k+1}\right) \right\| < \gamma \left\| F\left(x^k\right) \right\| \text{ with } \gamma < 1$$

Every Step reduces F-- Global Convergence!

Newton Method with Limiting

Damped Newton

Nested Iteration

$x^0 =$ Initial Guess, $k = 0$

Repeat {

Compute $F\left(x^k\right), J_F\left(x^k\right)$

Solve $J_F\left(x^k\right)\Delta x^{k+1} = -F\left(x^k\right)$ for Δx^{k+1}

Find $\alpha^k \in (0,1]$ such that $\left\|F\left(x^k + \alpha^k \Delta x^{k+1}\right)\right\|$ is minimized

$x^{k+1} = x^k + \alpha^k \Delta x^{k+1}$

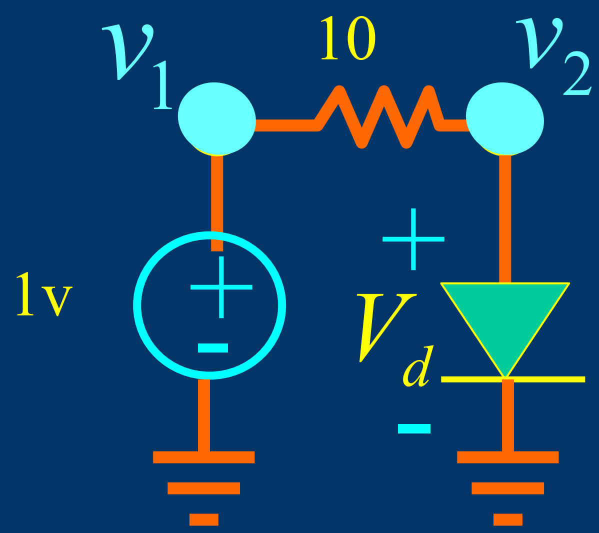
$k = k+1$

} Until $\left\|\Delta x^{k+1}\right\|, \left\|F\left(x^{k+1}\right)\right\|$ small enough

Newton Method with Limiting

Damped Newton

Example



$$I_r - \frac{1}{10} V_r = 0$$

$$I_d - I_s\left(e^{V_d / V_t} - 1\right) = 0$$

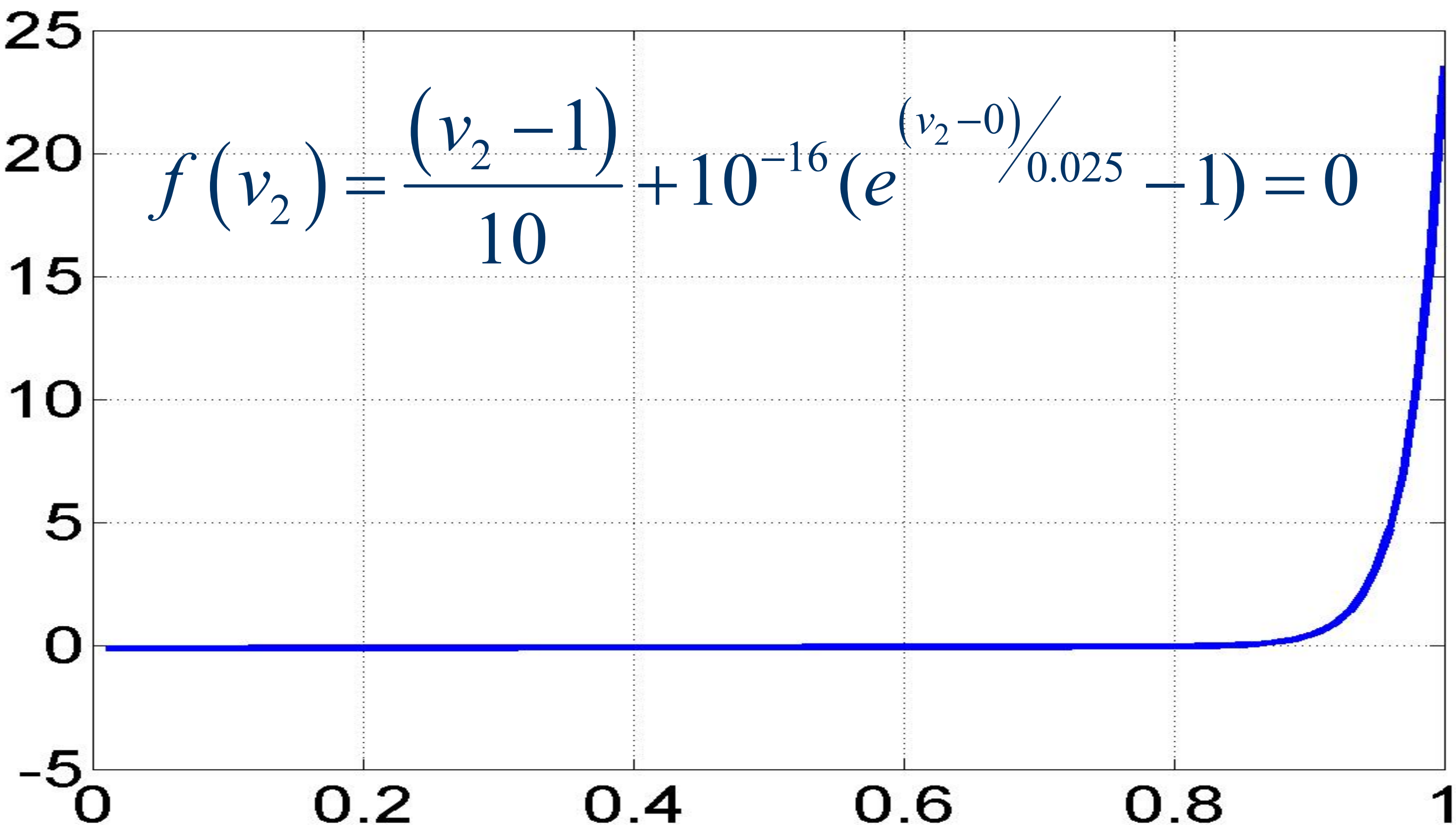
<u>Nodal Equations with Numerical Values</u>

$$f(v_2) = \frac{(v_2 - 1)}{10} + 10^{-16}\left(e^{(v_2 - 0)/0.025} - 1\right) = 0$$

Newton Method with Limiting

Damped Newton

Example cont.



$$f\left(v_2\right)=\frac{\left(v_2-1\right)}{10}+10^{-16}\left(e^{\left(v_2-0\right)/0.025}-1\right)=0$$

Newton Method with Limiting

Damped Newton

Nested Iteration

$x^0 =$ Initial Guess, $k = 0$

Repeat {

Compute $F\left(x^k\right), J_F\left(x^k\right)$

Solve $J_F\left(x^k\right)\Delta x^{k+1} = -F\left(x^k\right)$ for Δx^{k+1}

Find $\alpha^k \in (0,1]$ such that $\left\|F\left(x^k + \alpha^k \Delta x^{k+1}\right)\right\|$ is minimized

$x^{k+1} = x^k + \alpha^k \Delta x^{k+1}$

$k = k+1$

} Until $\left\|\Delta x^{k+1}\right\|$, $\left\|F\left(x^{k+1}\right)\right\|$ small enough

How can one find the damping coefficients?

Newton Method with Limiting

Damped Newton

Theorem Proof

By definition of the Newton Iteration

$$x^{k+1} = x^k - \alpha^k \underbrace{J_F\left(x^k\right)^{-1} F\left(x^k\right)}_{\text{Newton Direction}}$$

Multidimensional Mean Value Lemma

$$\left\| F(x) - F(y) - J_F(y)(x-y) \right\| \leq \frac{\ell}{2} \left\| x - y \right\|^2$$

Combining

$$\left\| F\left(x^{k+1}\right) - F\left(x^k\right) + J_F\left(x^k\right) \left[\alpha^k J_F\left(x^k\right)^{-1} F\left(x^k\right) \right] \right\| \leq \frac{\ell}{2} \left\| \alpha^k J_F\left(x^k\right)^{-1} F\left(x^k\right) \right\|^2$$

Newton Method with Limiting

Damped Newton

Theorem Proof-Cont

From the previous slide

$$\left\| F\left(x^{k+1}\right) - F\left(x^k\right) + J_F\left(x^k\right)\left[\alpha^k J_F\left(x^k\right)^{-1} F\left(x^k\right)\right]\right\| \le \frac{\ell}{2}\left\|\alpha^k J_F\left(x^k\right)^{-1} F\left(x^k\right)\right\|^2$$

Combining terms and moving scalars out of norms

$$\left\| F\left(x^{k+1}\right) - \left(1-\alpha^k\right) F\left(x^k\right)\right\| \le \left(\alpha^k\right)^2 \frac{\ell}{2}\left\| J_F\left(x^k\right)^{-1} F\left(x^k\right)\right\|^2$$

Using the Jacobian Bound and splitting the norm

$$\left\| F\left(x^{k+1}\right)\right\| \le \left[\left(1-\alpha^k\right)\left\| F\left(x^k\right)\right\| + \left(\alpha^k\right)^2 \frac{\beta^2 \ell}{2}\left\| F\left(x^k\right)\right\|^2\right]$$

Yields a quadratic in the damping coefficient

Newton Method with Limiting

Damped Newton

Theorem Proof-Cont-II

Simplifying quadratic from previous slide

$$\left\|F\left(x^{k+1}\right)\right\| \le \left[1-\alpha^k+\left(\alpha^k\right)^2\frac{\beta^2\ell}{2}\left\|F\left(x^k\right)\right\|\right]\left\|F\left(x^k\right)\right\|$$

Two Cases:

1) $\dfrac{\beta^2\ell}{2}\left\|F\left(x^k\right)\right\| < \dfrac{1}{2}$ Pick $\alpha^k = 1$ (Standard Newton)

$$\Rightarrow\left(1-\alpha^k+\left(\alpha^k\right)^2\frac{\beta^2\ell}{2}\left\|F\left(x^k\right)\right\|\right) < \frac{1}{2}$$

2) $\dfrac{\beta^2\ell}{2}\left\|F\left(x^k\right)\right\| > \dfrac{1}{2}$ Pick $\alpha^k = \dfrac{1}{\beta^2\ell\left\|F\left(x^k\right)\right\|}$

$$\Rightarrow\left(1-\alpha^k+\left(\alpha^k\right)^2\frac{\beta^2\ell}{2}\left\|F\left(x^k\right)\right\|\right) < 1-\frac{1}{2\beta^2\ell\left\|F\left(x^k\right)\right\|}$$

Newton Method with Limiting

Damped Newton

Theorem Proof-Cont-III

Combining the results from the previous slide

$$\left\| F\left(x^{k+1}\right) \right\| \le \gamma^k \left\| F\left(x^k\right) \right\|$$ not good enough, need γ independent from k

The above result does imply

$$\left\| F\left(x^{k+1}\right) \right\| \le \left\| F\left(x^0\right) \right\|$$ not yet a convergence theorem

For the case where $\frac{\beta^2 \ell}{2} \left\| F\left(x^k\right) \right\| > \frac{1}{2}$

$$1 - \frac{1}{2\beta^2 \ell \left\| F\left(x^k\right) \right\|} \le 1 - \frac{1}{2\beta^2 \ell \left\| F\left(x^0\right) \right\|} \le \gamma^0$$

Note the proof technique

First – Show that the iterates do not increase

Second – Use the non-increasing fact to prove convergence

Newton Method with Limiting

Damped Newton

Nested Iteration

$x^0 =$ Initial Guess, $k = 0$

Repeat {

Compute $F\left(x^k\right), J_F\left(x^k\right)$

Solve $J_F\left(x^k\right)\Delta x^{k+1} = -F\left(x^k\right)$ for Δx^{k+1}

Find $\alpha^k \in (0,1]$ such that $\left\|F\left(x^k + \alpha^k \Delta x^{k+1}\right)\right\|$ is minimized

$x^{k+1} = x^k + \alpha^k \Delta x^{k+1}$

$k = k + 1$

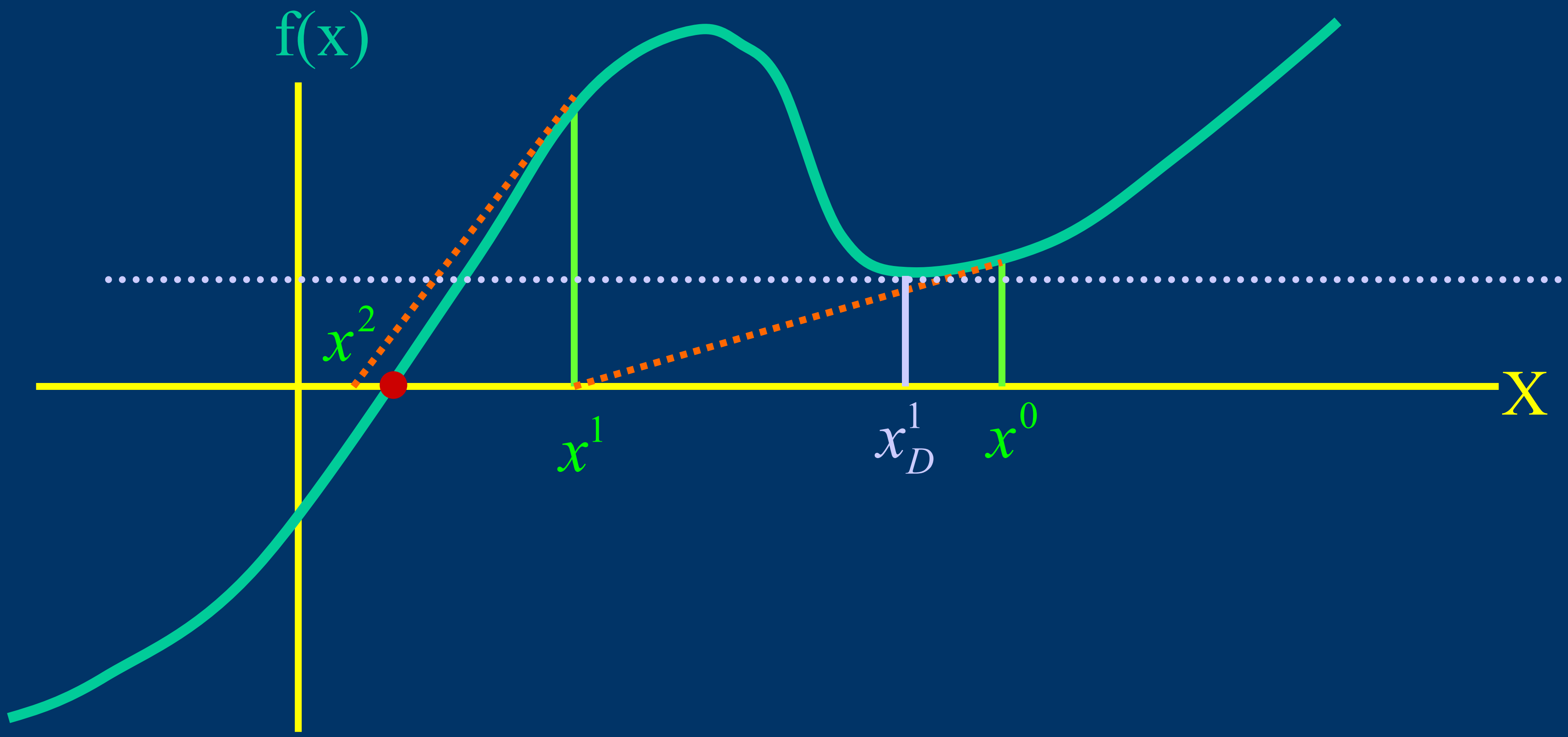
} Until $\left\|\Delta x^{k+1}\right\|, \left\|F\left(x^{k+1}\right)\right\|$ small enough

Many approaches to finding α^k

Newton Method with Limiting

Damped Newton

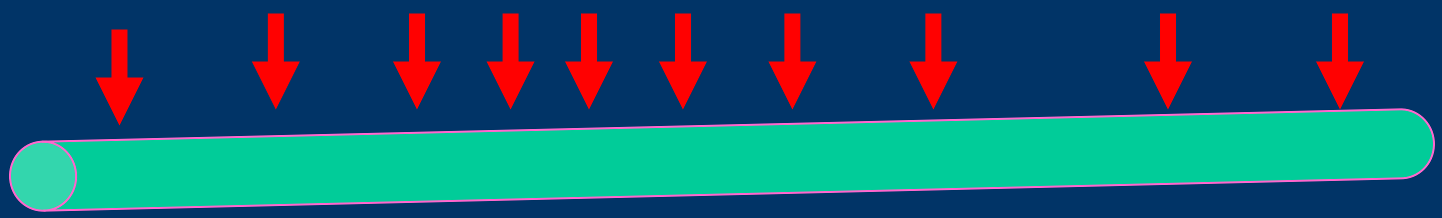
Singular Jacobian Problem



Damped Newton Methods "push" iterates to local minimums
Finds the points where Jacobian is Singular

Source or Load-Stepping

- Newton converges given a close initial guess
  - Generate a sequence of problems
  - Make sure previous problem generates guess for next problem
- Heat-conducting bar example



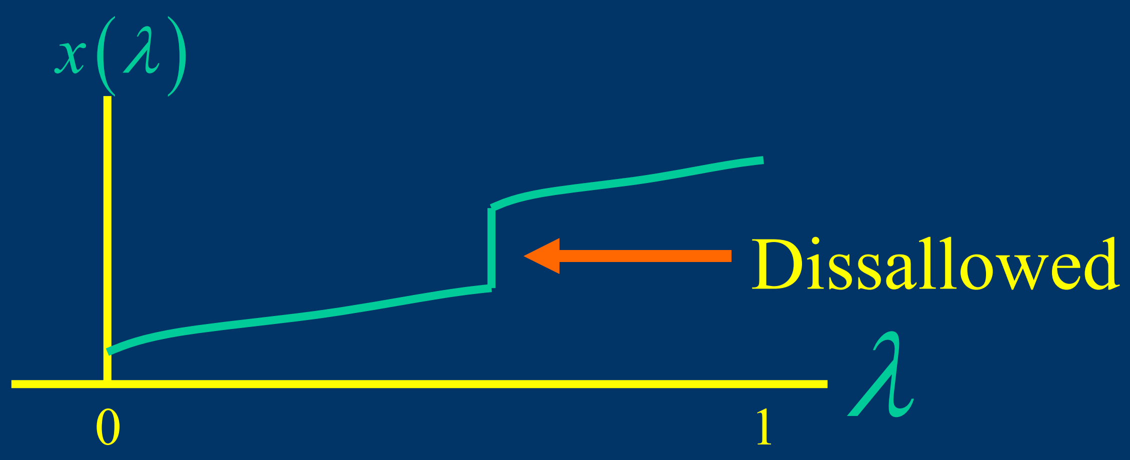
1. Start with heat off, T= 0 is a very close initial guess
2. Increase the heat slightly, T=0 is a good initial guess
3. Increase heat again

Solve $\tilde{F}\left(x(\lambda),\lambda\right)=0$ where:

a) $\tilde{F}\left(x(0),0\right)=0$ is easy to solve   Starts the continuation

b) $\tilde{F}\left(x(1),1\right)=F(x)$ Ends the continuation

c) $x(\lambda)$ is sufficiently smooth   Hard to insure!



Template Algorithm

Solve $\tilde{F}\left(x(0),0\right), \quad x\left(\lambda_{prev}\right)=x(0)$

$\delta\lambda$=0.01, $\lambda=\delta\lambda$

While $\lambda<1$ {

$x^0(\lambda)=x\left(\lambda_{prev}\right)$

Try to Solve $\tilde{F}\left(x(\lambda),\lambda\right)=0$ with Newton

If Newton Converged

$x\left(\lambda_{prev}\right)=x(\lambda), \;\; \lambda=\lambda+\delta\lambda, \;\; \delta\lambda=2\delta\lambda$

Else

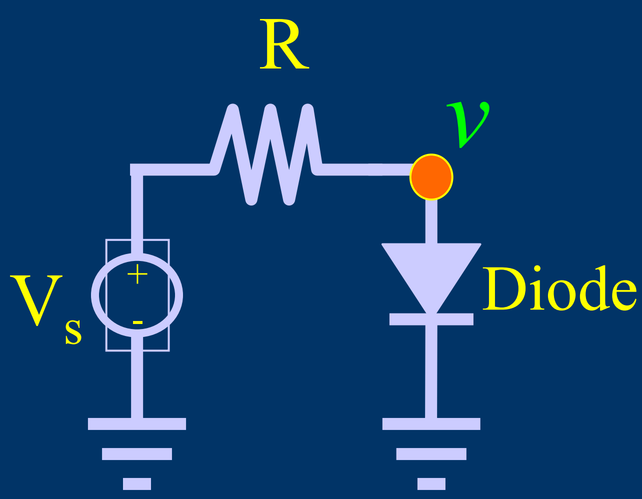
$\delta\lambda=\frac{1}{2}\delta\lambda, \quad \lambda=\lambda_{prev}+\delta\lambda$

}

Continuation Schemes

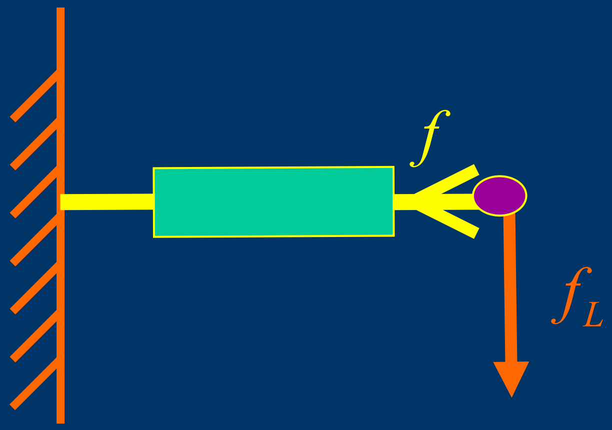
Basic Concepts

Source/Load Stepping Examples



$$\tilde{f}\left(v(\lambda),\lambda\right) = i_{diode}\left(v\right) + \frac{1}{R}\left(v - \lambda V_s\right) = 0$$

$$\frac{\partial \tilde{f}\left(v,\lambda\right)}{\partial v} = \frac{\partial i_{diode}\left(v\right)}{\partial v} + \frac{1}{R} \leftarrow \text{Not } \lambda \text{ dependent!}$$



$$\tilde{F}\left(\vec{x},\lambda\right) = \begin{array}{l} f_x\left(x,y\right) = 0 \\ f_y\left(x,y\right) + \lambda f_l = 0 \end{array}$$

Source/Load Stepping Does Not Alter Jacobian

Continuation Schemes

Jacobian Altering Scheme

Description

$$\tilde{F}\left(x(\lambda),\lambda\right) = \lambda F\left(x(\lambda)\right) + (1-\lambda)\,x(\lambda)$$

Observations

$\lambda=0$ $\quad \tilde{F}\left(x(0),0\right) = x(0) = 0$

$$\frac{\partial \tilde{F}\left(x(0),0\right)}{\partial x} = I$$

Problem is easy to solve and Jacobian definitely nonsingular.

$\lambda=1$ $\quad \tilde{F}\left(x(1),1\right) = F\left(x(1)\right)$

$$\frac{\partial \tilde{F}\left(x(0),0\right)}{\partial x} = \frac{\partial F\left(x(1)\right)}{\partial x}$$

Back to the original problem and original Jacobian

Continuation Schemes

Jacobian Altering Scheme

Basic Algorithm

Solve $\tilde{F}\left(x(0),0\right), \quad x\left(\lambda_{prev}\right)=x(0)$

$\delta\lambda$=0.01, $\lambda=\delta\lambda$

While $\lambda<1$ {

$x^0(\lambda)=x\left(\lambda_{prev}\right)+?$

Try to Solve $\tilde{F}\left(x(\lambda),\lambda\right)=0$ with Newton

If Newton Converged

$x\left(\lambda_{prev}\right)=x(\lambda), \quad \lambda=\lambda+\delta\lambda, \quad \delta\lambda=2\delta\lambda$

Else

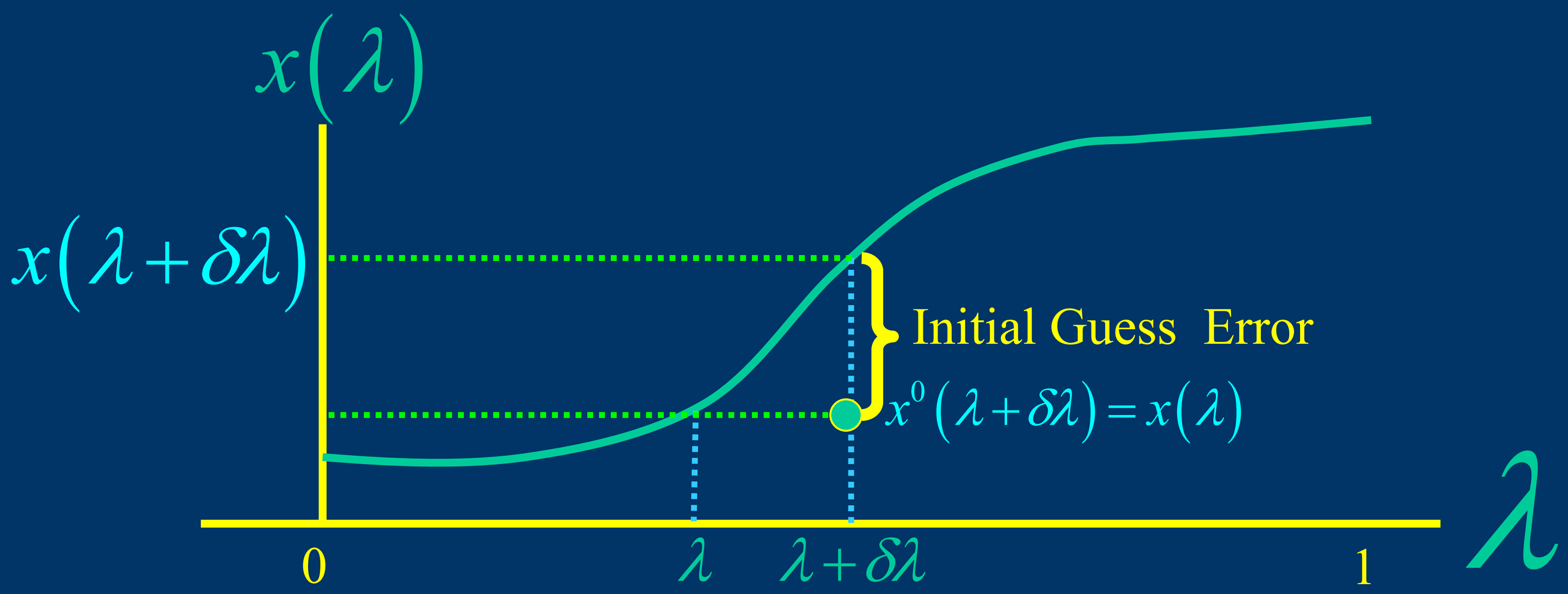
$\delta\lambda=\frac{1}{2}\delta\lambda, \qquad \lambda=\lambda_{prev}+\delta\lambda$

}

Continuation Schemes

Jacobian Altering Scheme

Initial Guess for each step.



Continuation Schemes

Jacobian Altering Scheme

Update Improvement

$$\tilde{F}\left(x(\lambda+\delta\lambda),\lambda+\delta\lambda\right) \approx \cancelto{0}{\tilde{F}\left(x(\lambda),\lambda\right)} + \frac{\partial \tilde{F}\left(x(\lambda),\lambda\right)}{\partial x}\left(x(\lambda+\delta\lambda)-x(\lambda)\right) + \frac{\partial \tilde{F}\left(x(\lambda),\lambda\right)}{\partial \lambda}\delta\lambda$$

$$\Rightarrow \underbrace{\frac{\partial \tilde{F}\left(x(\lambda),\lambda\right)}{\partial x}}_{\text{Have From last step's Newton}}\left(\underbrace{x^{0}(\lambda+\delta\lambda)}_{\text{Better Guess for next step's Newton}}-x(\lambda)\right) = -\frac{\partial \tilde{F}\left(x(\lambda),\lambda\right)}{\partial \lambda}\delta\lambda$$

Update Improvement Cont.

If

$$\tilde{F}\left(x(\lambda),\lambda\right) = \lambda F\left(x(\lambda)\right) + (1-\lambda)\,x(\lambda)$$

Then

$$\frac{\partial \tilde{F}(x,\lambda)}{\partial \lambda} = \underbrace{F(x) - x(\lambda)}_{\text{Easily Computed}}$$

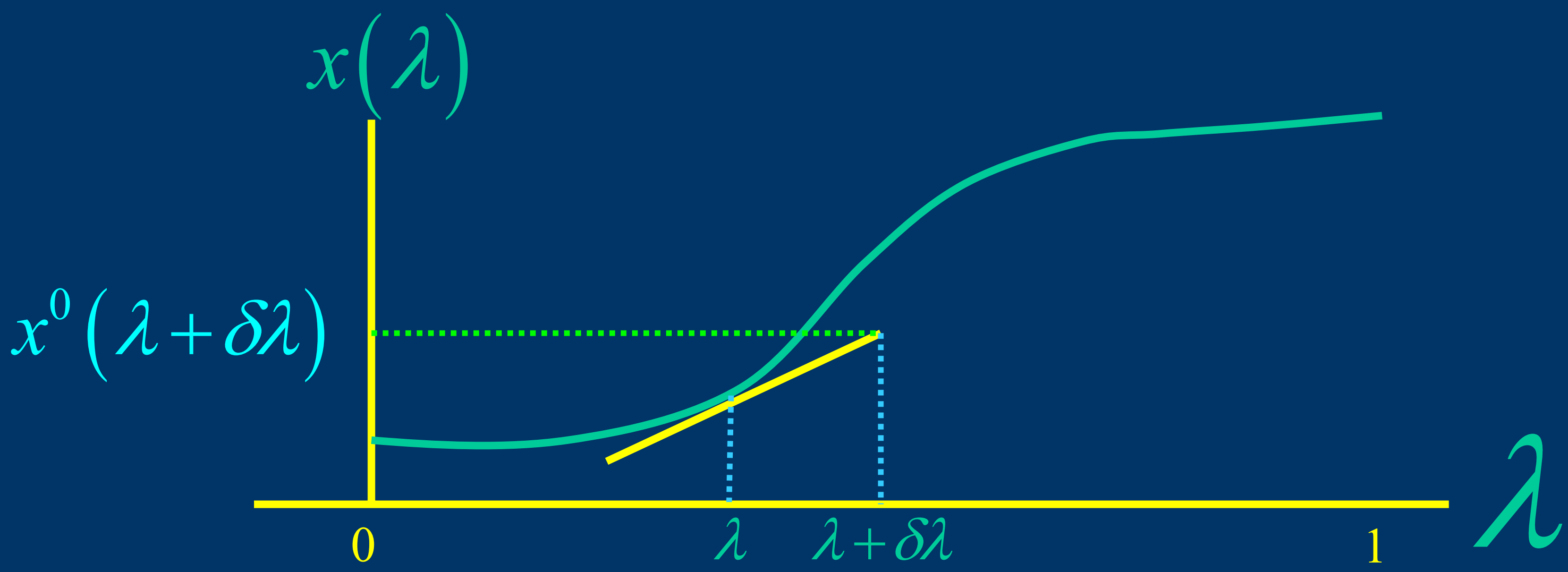
Continuation Schemes

Jacobian Altering Scheme

Update Improvement Cont. II.

$$x^0(\lambda+\delta\lambda) = x(\lambda) - \left(\frac{\partial \tilde{F}(x(\lambda),\lambda)}{\partial x}\right)^{-1} \frac{\partial \tilde{F}(x(\lambda),\lambda)}{\partial \lambda}\delta\lambda$$

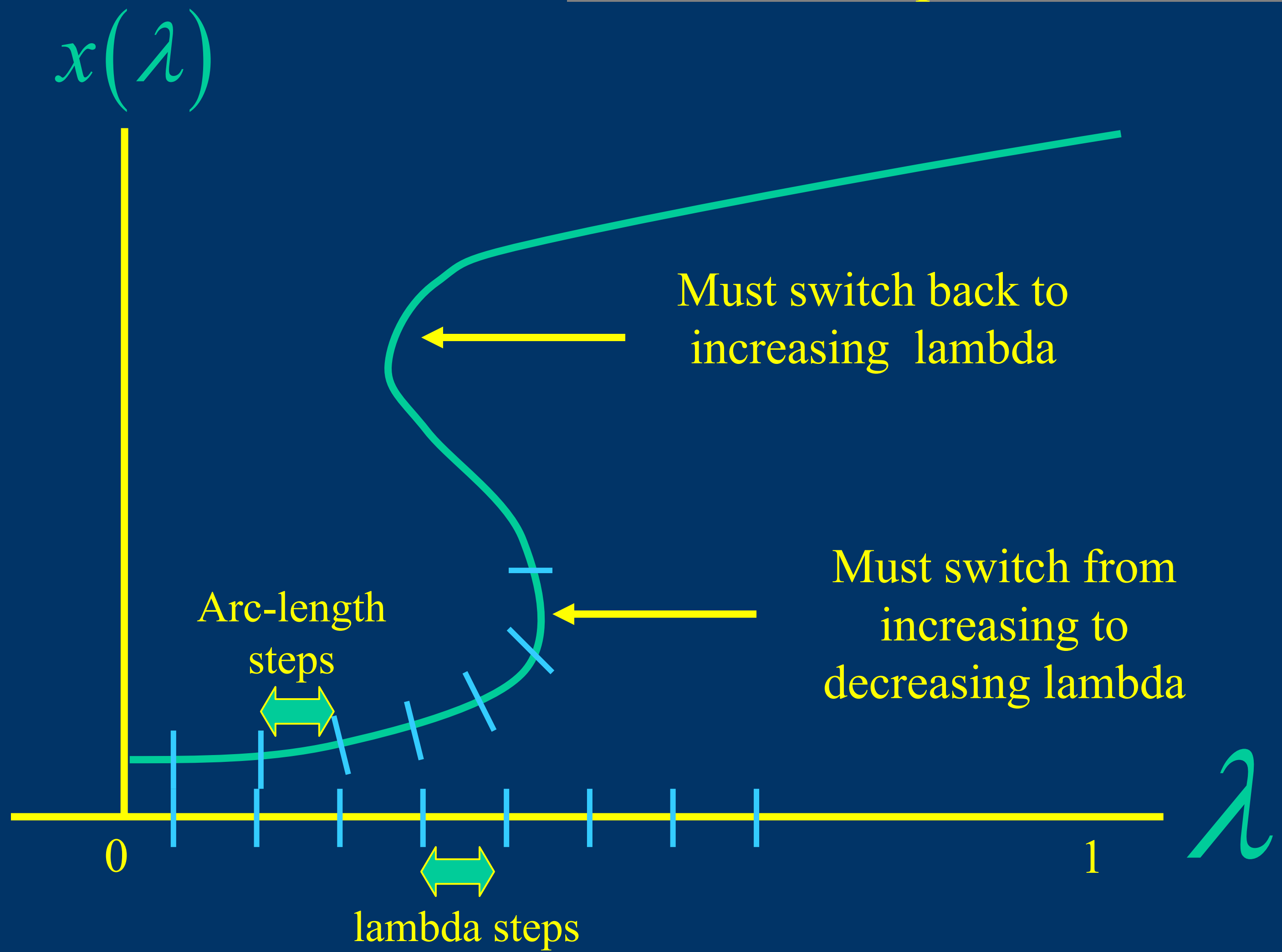
Graphically



Continuation Schemes

Jacobian Altering Scheme

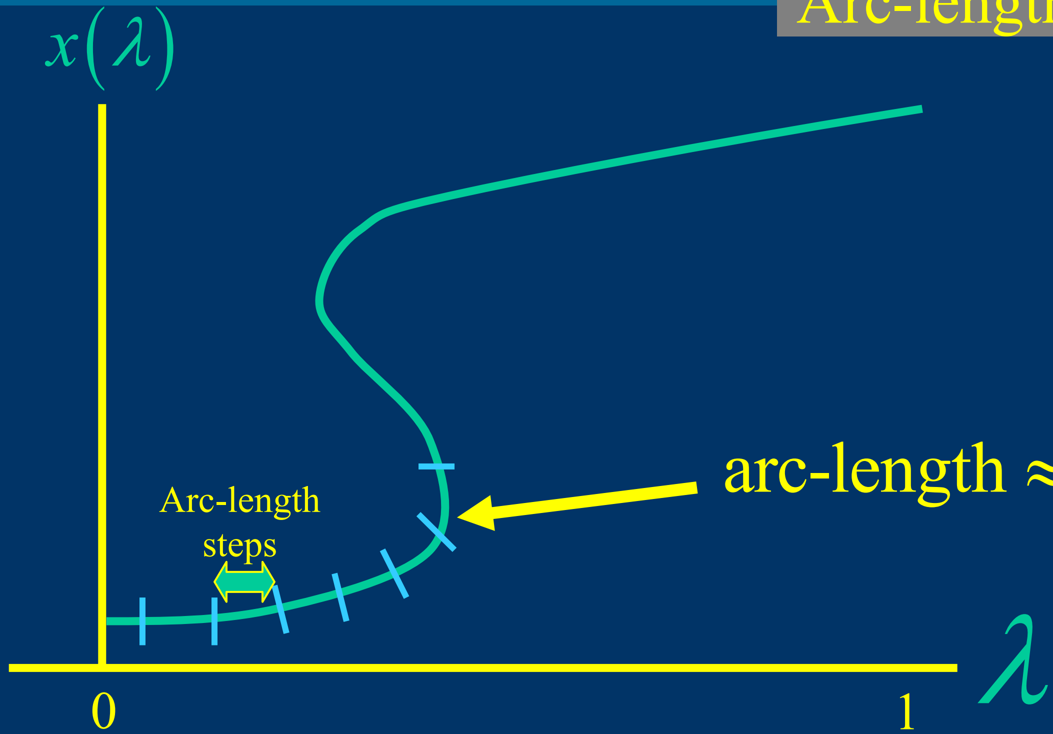
Still can have problems



Continuation Schemes

Jacobian Altering Scheme

Arc-length Steps?



$$\text{arc-length} \approx \sqrt{(\Delta x)^2 + (\delta\lambda)^2}$$

Must Solve For Lambda

$$\tilde{F}(x,\lambda) = 0$$

$$(\lambda - \lambda_{prev})^2 + \left\| x - x(\lambda_{prev}) \right\|_2^2 - arc^2 = 0$$

Continuation Schemes

Jacobian Altering Scheme

Arc-length steps by Newton

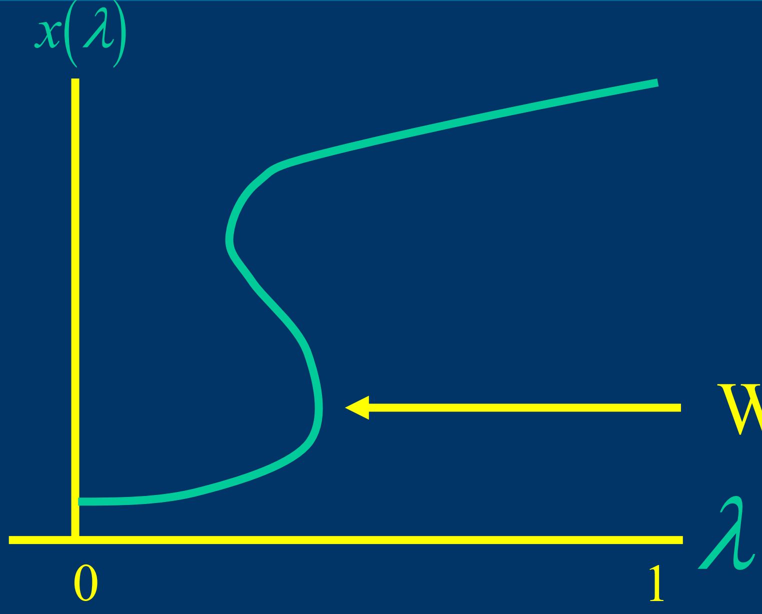
$$\begin{bmatrix} \dfrac{\partial \tilde{F}\left(x^k, \lambda^k\right)}{\partial x} & \dfrac{\partial \tilde{F}\left(x^k, \lambda^k\right)}{\partial \lambda} \\ 2\left(x^k - x\left(\lambda_{prev}\right)\right)^T & 2\left(\lambda^k - \lambda_{prev}\right) \end{bmatrix} \begin{bmatrix} x^{k+1} - x^k \\ \lambda^{k+1} - \lambda^k \end{bmatrix} =$$

$$-\begin{bmatrix} \tilde{F}\left(x^k, \lambda^k\right) \\ \left(\lambda^k - \lambda_{prev}\right)^2 + \left\| x^k - x\left(\lambda_{prev}\right) \right\|_2^2 - arc^2 \end{bmatrix}$$

Continuation Schemes

Jacobian Altering Scheme

Arc-length Turning point



What happens here?

Upper left-hand Block is singular

$$\begin{bmatrix} \dfrac{\partial \tilde{F}\left(x^k, \lambda^k\right)}{\partial x} & \dfrac{\partial \tilde{F}\left(x^k, \lambda^k\right)}{\partial \lambda} \\ 2\left(x^k - x\left(\lambda_{prev}\right)\right)^T & 2\left(\lambda^k - \lambda_{prev}\right) \end{bmatrix}$$

Summary

- Damped Newton Schemes
  - Globally Convergent if Jacobian is Nonsingular
  - Difficulty with Singular Jacobians
- Introduce Continuation Schemes
  - Problem with Source/Load stepping
  - More General Continuation Scheme
- Improving Efficiency
  - Better first guess for each continuation step
- Arc-length Continuation