Introduction to Simulation - Lecture 11

Newton-Method Case Study – Simulating An Image Smoother

Jacob White

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

Outline

- Image Segmentation Example
  - Large nonlinear system of equations
  - Formulation? Continuation? Linear Solver?
- Newton Iterative Methods
  - Accuracy Theorem
  - Matrix-free idea
- Gershgorin Circle Theorem
  - Lends insight on iterative method convergence
- Arc-Length Continuation

Simple Smoother

Circuit Diagram

Smoothed Output

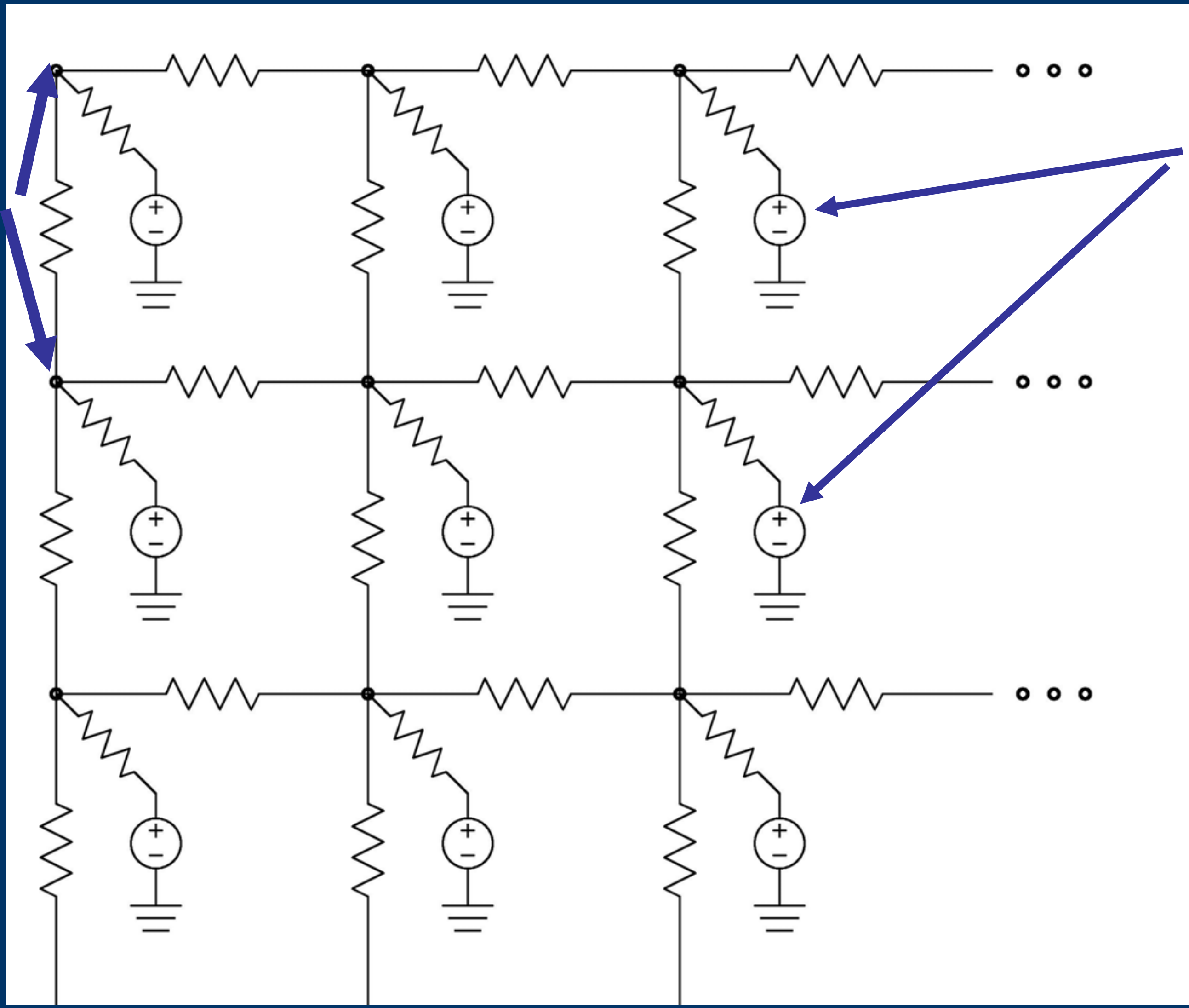
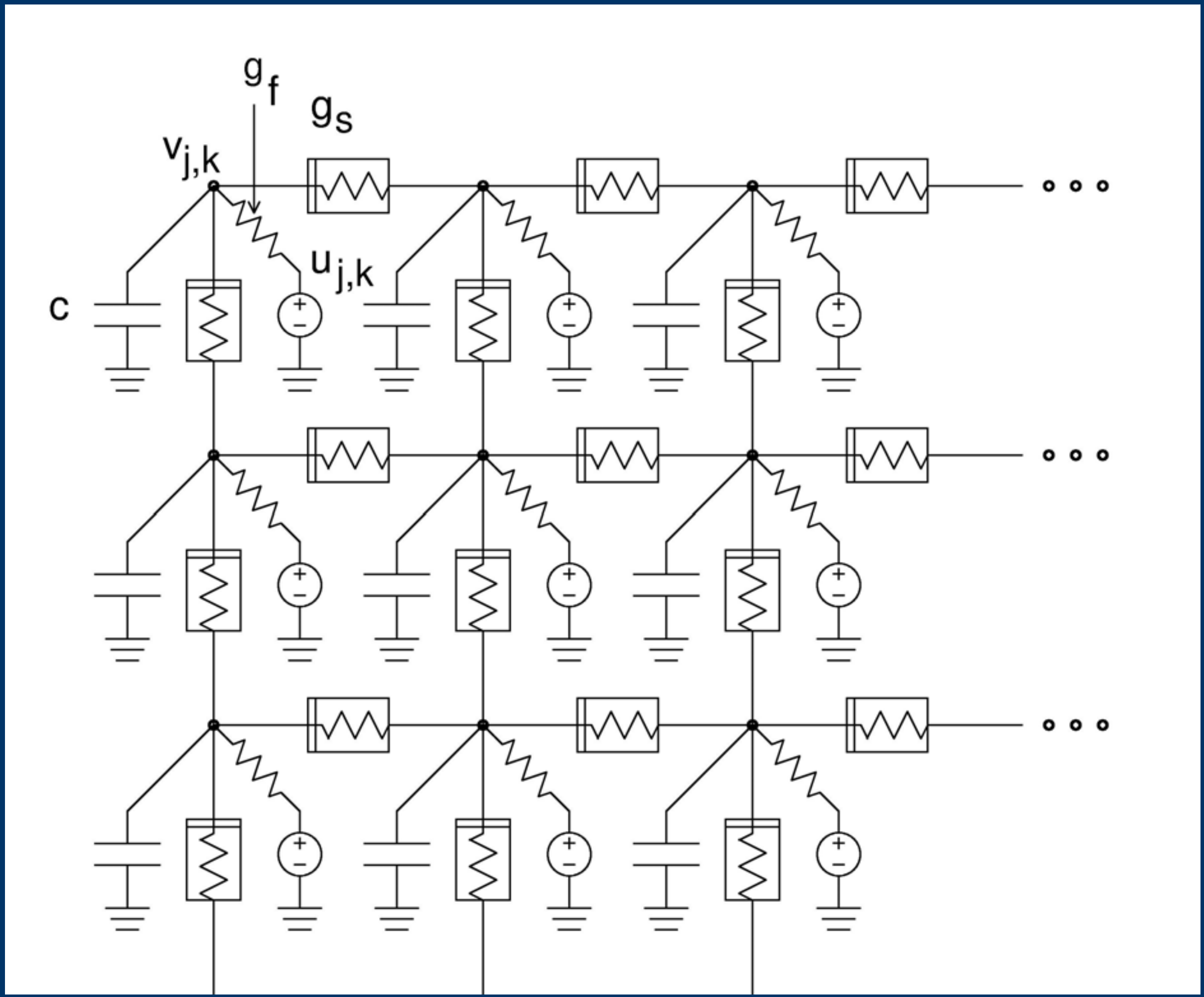


Image Input

Nonlinear Smoother

Circuit Diagram



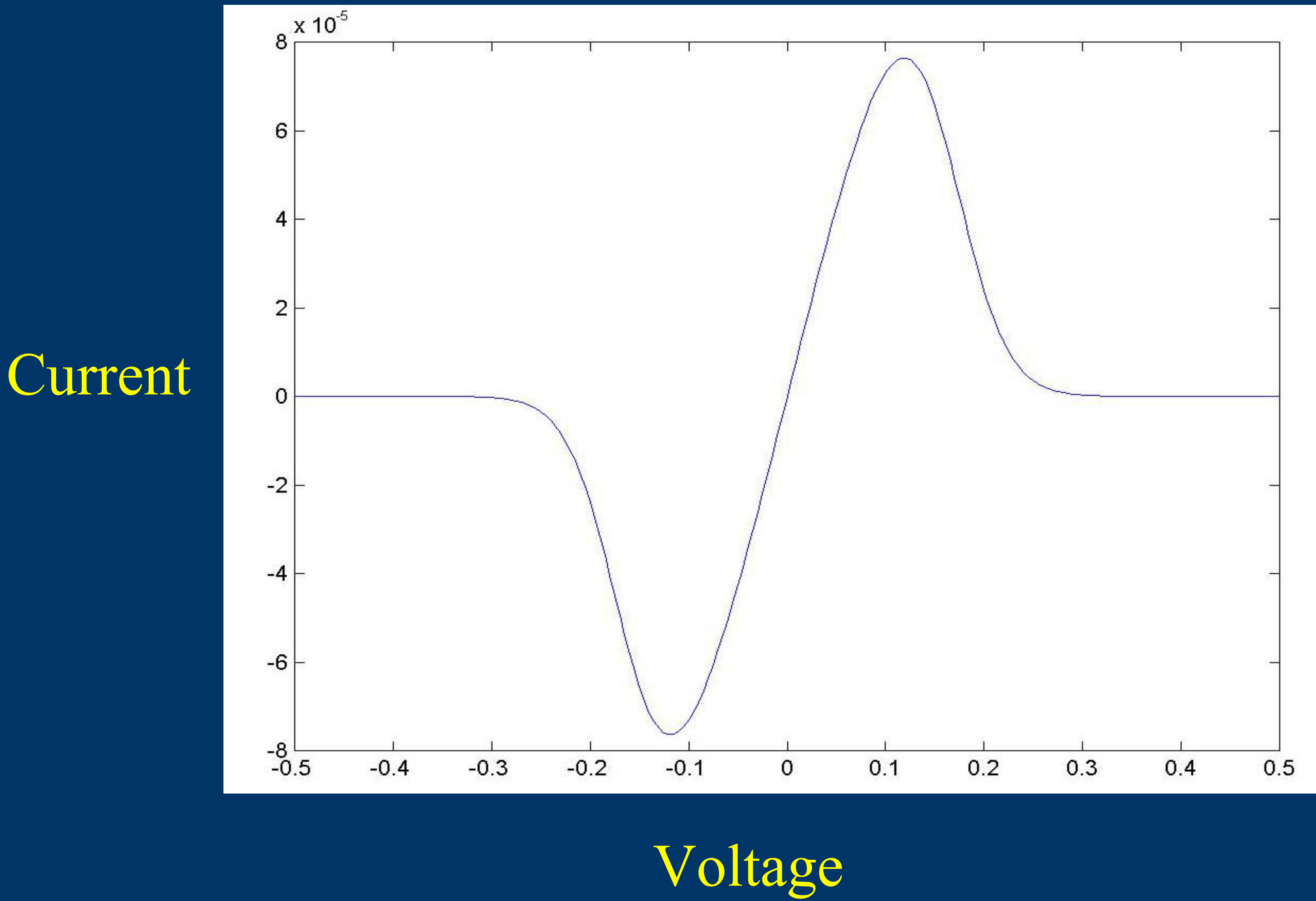
Nonlinear Smoother

Nonlinear Resistor Constitutive Equation

$$i(v) = \frac{\alpha v}{1 + e^{-\beta(\gamma - \alpha v^2)}}$$

Nonlinear Smoother

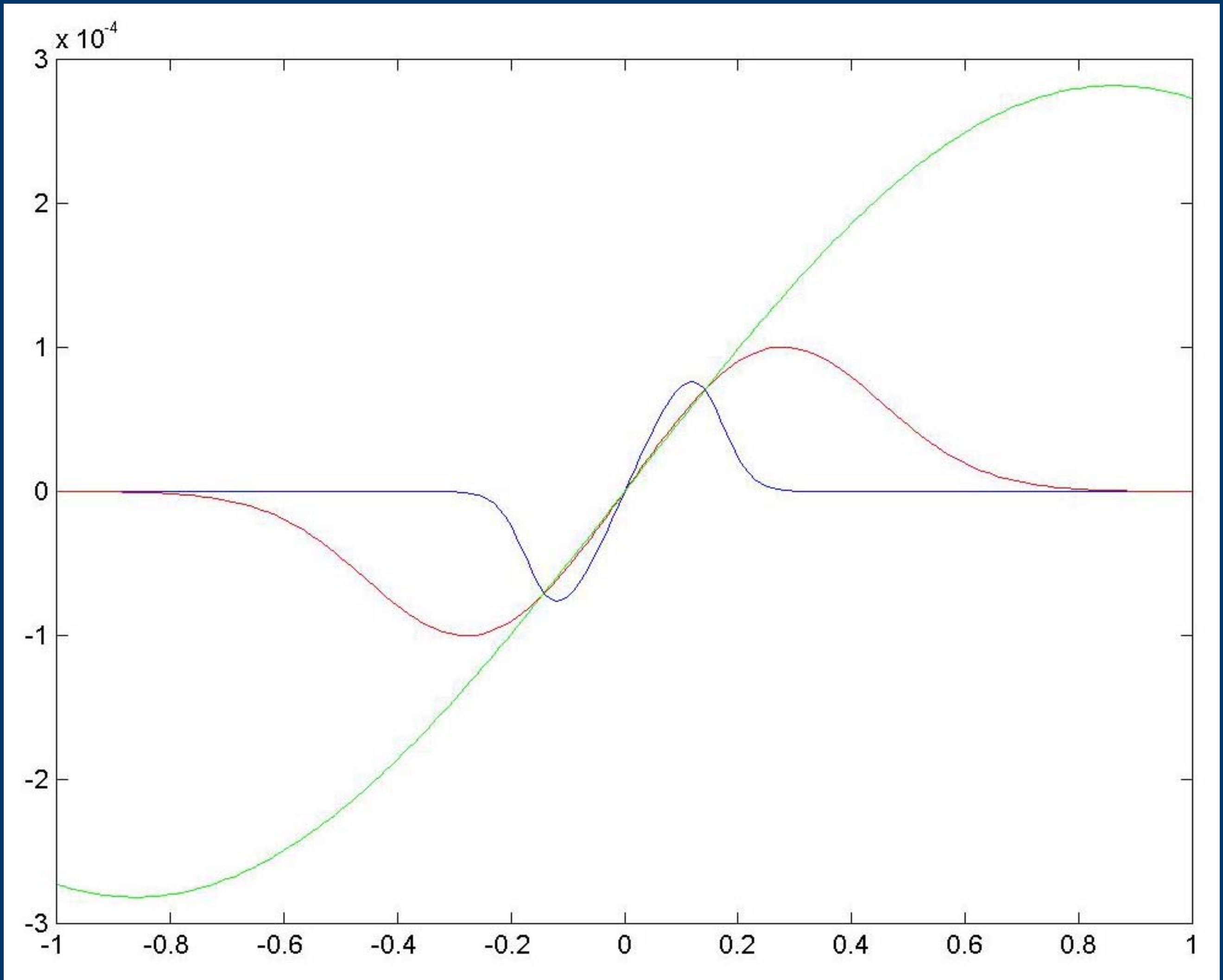
Nonlinear Resistor Constitutive Equation



Nonlinear Smoother

Nonlinear Resistor Constitutive Equation

Varying Beta



Questions

- What Equation Formulation?
  - Node-Branch or Nodal?
- What Newton Method?
  - Standard, Damped, or Continuation?
  - What kind of Continuation?
- What Linear Solver?
  - Sparse Gaussian Elimination or Krylov?
  - Will Krylov converge rapidly?
  - How will formulation, Newton choices interact?

Newton-Iterative Method

Basic Algorithm

Nested Iteration

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

Repeat {

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

Solve (Using GCR)

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

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

$$k = k + 1$$

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

How Accurately Should We Solve with GCR?

Newton-Iterative Method

Basic Algorithm

Solve Accuracy Required

After l steps of GCR

$$J_F\left(x^k\right)\underbrace{\Delta x^{k+1,l}}_{\substack{\text{Newton}\\\text{delta from}\\ l\text{ GCR Steps}}}=-F\left(x^k\right)+\underbrace{r^{k,l}}_{\substack{\text{GCR}\\\text{Residual}}}$$

If

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

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

$c)$ $\left\|r^{k,l}\right\|\leq C\left\|F\left(x^k\right)\right\|^2$ (More accurate near convergence)

Then

The Newton-Iterative Method Converges Quadratically

Newton-Iterative Method

Basic Algorithm

Convergence Proof

By definition of the Newton-Iterative Method

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

Multidimensional Mean Value Lemma

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

Combining

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

Newton-Iterative Method

Basic Algorithm

Convergence Proof Cont.

Canceling the Jacobian and its inverse on the previous slide

$$\left\| F\left(x^{k+1}\right) - F\left(x^{k}\right) + F\left(x^{k}\right) + r^{k,l} \right\| \leq \frac{\ell}{2} \left\| J_F\left(x^{k}\right)^{-1} \left(F\left(x^{k}\right) + r^{k,l}\right) \right\|^2$$

Combining terms and using the triangle inequality

$$\left\| F\left(x^{k+1}\right) \right\| \leq \frac{\ell}{2} \left\| J_F\left(x^{k}\right)^{-1} \left(F\left(x^{k}\right) + r^{k,l}\right) \right\|^2 + \left\| r^{k,l} \right\|$$

Using the Jacobian Bound and the triangle inequality

$$\left\| F\left(x^{k+1}\right) \right\| \leq \frac{\beta^2 \ell}{2} \left\| F\left(x^{k}\right) \right\|^2 + \left(1 + \frac{\beta^2 \ell \left\| r^{k,l} \right\|}{2}\right) \left\| r^{k,l} \right\|$$

Newton-Iterative Method

Basic Algorithm

Convergence Proof Cont. II

Using the bound on the iterative solver error

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

And combining terms yields

$$\left\|F\left(x^{k+1}\right)\right\| \le \underbrace{\left(\frac{\beta^2 \ell}{2} + \left(1 + \frac{\beta^2 \ell \left\|F\left(x^k\right)\right\|^2}{2}\right) C\right)}_{\text{Easily Bounded}} \left\|F\left(x^k\right)\right\|^2$$

Newton-Iterative Method

Matrix-Free Idea

Consider Applying GCR to The Newton Iterate Equation

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

At each iteration GCR forms a matrix-vector product

$$J_F\left(x^k\right)p^l \approx \frac{1}{\varepsilon}\left(F\left(x^k+\varepsilon p^l\right)-F\left(x^k\right)\right)$$

It is possible to use Newton-GCR without Jacobians!

Need to Select a good ε

Gerschgorin Circle Theorem

Theorem Statement

Given a matrix

$$M = \begin{bmatrix} m_{1,1} & \cdots & m_{1,N} \\ \vdots & \ddots & \vdots \\ m_{N,1} & \cdots & m_{N,N} \end{bmatrix}$$

For each eigenvalue of M there exists an i, 1< i < N such that

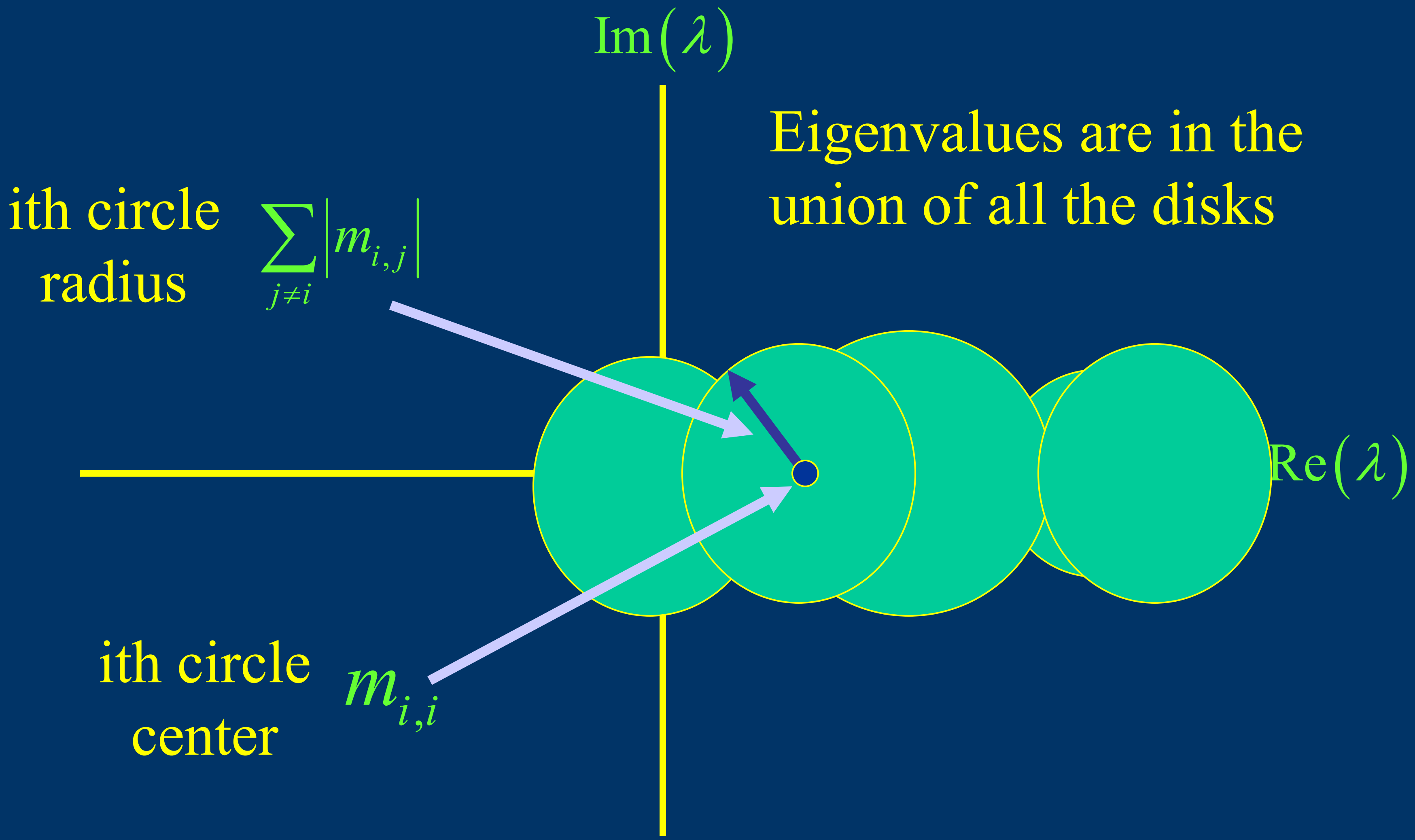
$$\left|\lambda - m_{i,i}\right| \leq \sum_{j \neq i} \left|m_{i,j}\right|$$

We say that the eigenvalues are contained in the union of the Gerschgorin circles

Gerschgorin Circle Theorem

Theorem Statement

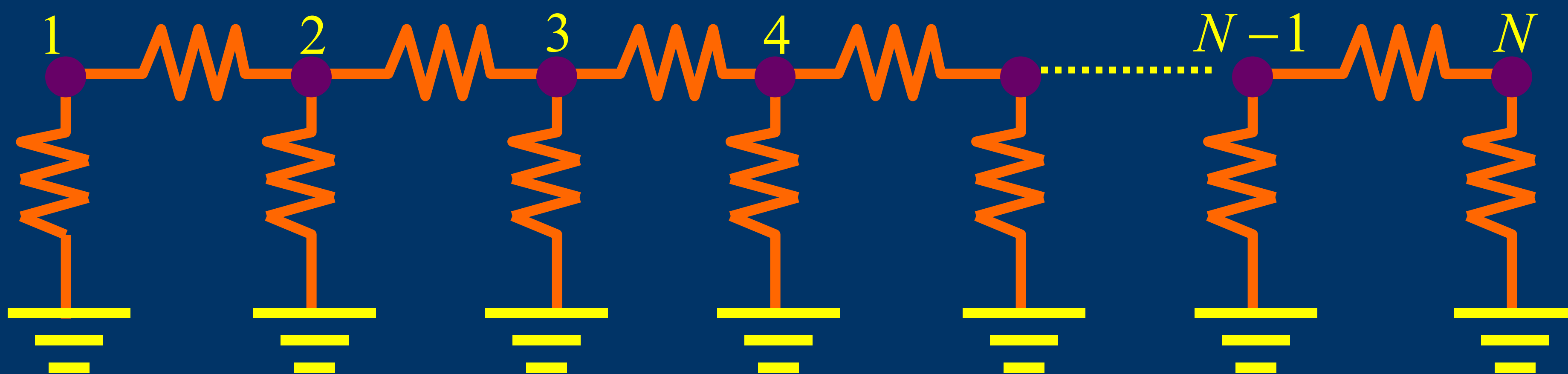
Picture of Gerschgorin



Gerschgorin Circle Theorem

Grounded Resistor Line

Nodal Matix



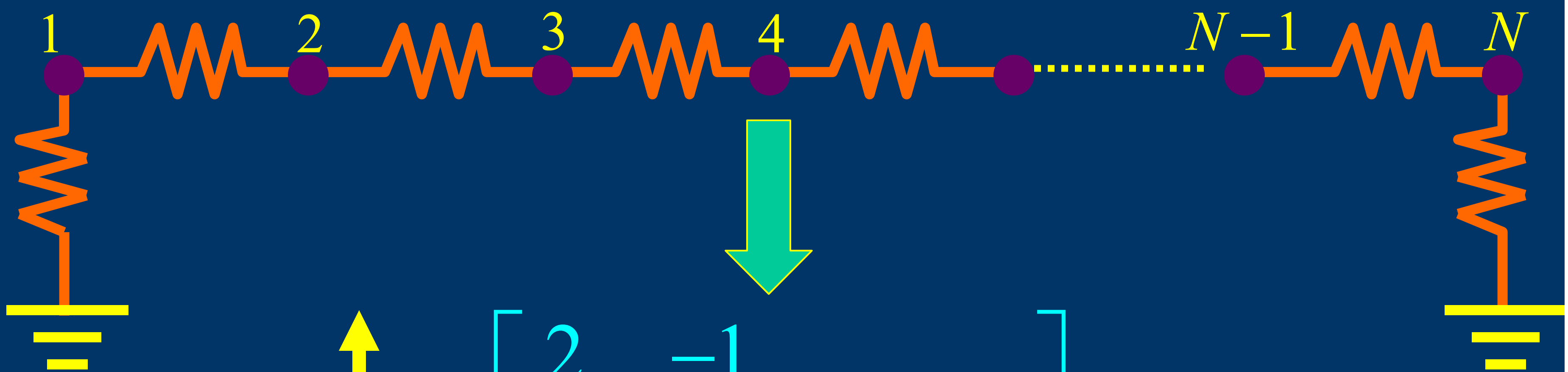
$$
N \updownarrow \underbrace{\begin{bmatrix} 2.1 & -1 & & \\ -1 & 2.1 & & \\ & & \ddots & -1 \\ & & -1 & 2.1 \end{bmatrix}}_{M}
$$

Nodal Equation Form

Gerschgorin Circle Theorem

Resistor Line

Nodal Matix



$$\underbrace{\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & & \\ & & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}}_{M} \quad \updownarrow N$$

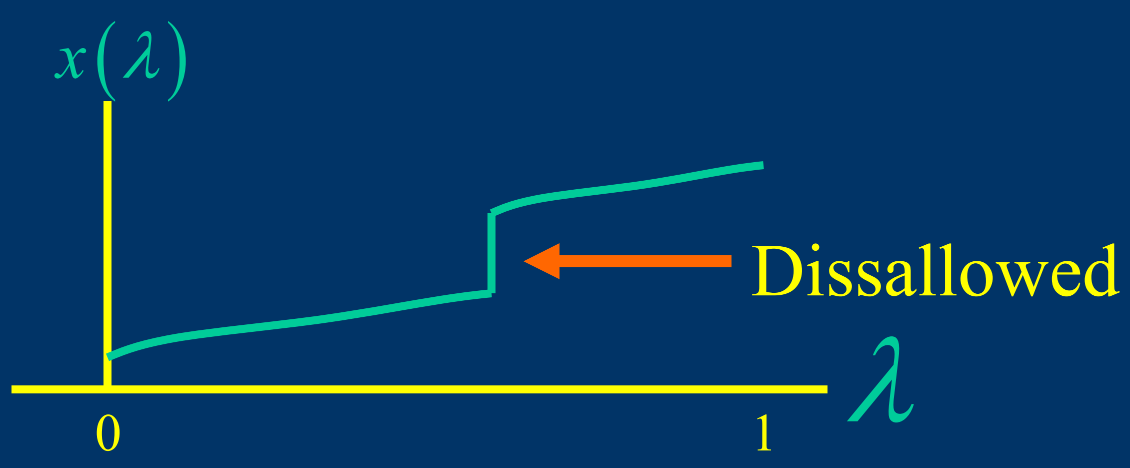
Nodal Equation Form

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

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

Else

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

}

Continuation Schemes

Jacobian Altering Scheme

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

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 $\lambda < 1 \quad \{$

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

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

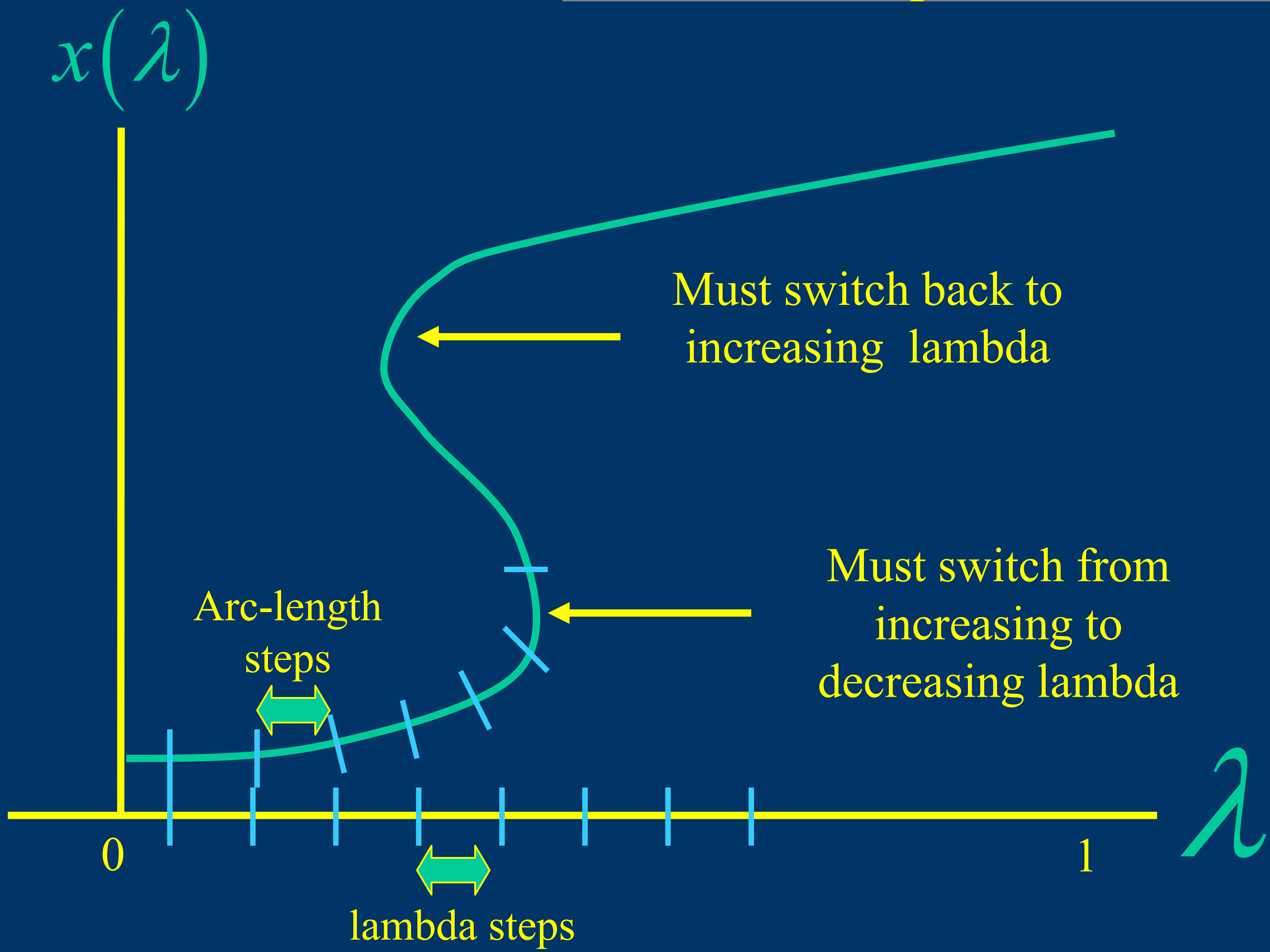
$$\delta\lambda = \frac{1}{2}\delta\lambda, \quad \lambda = \lambda_{prev} + \delta\lambda$$

$\}$

Continuation Schemes

Jacobian Altering Scheme

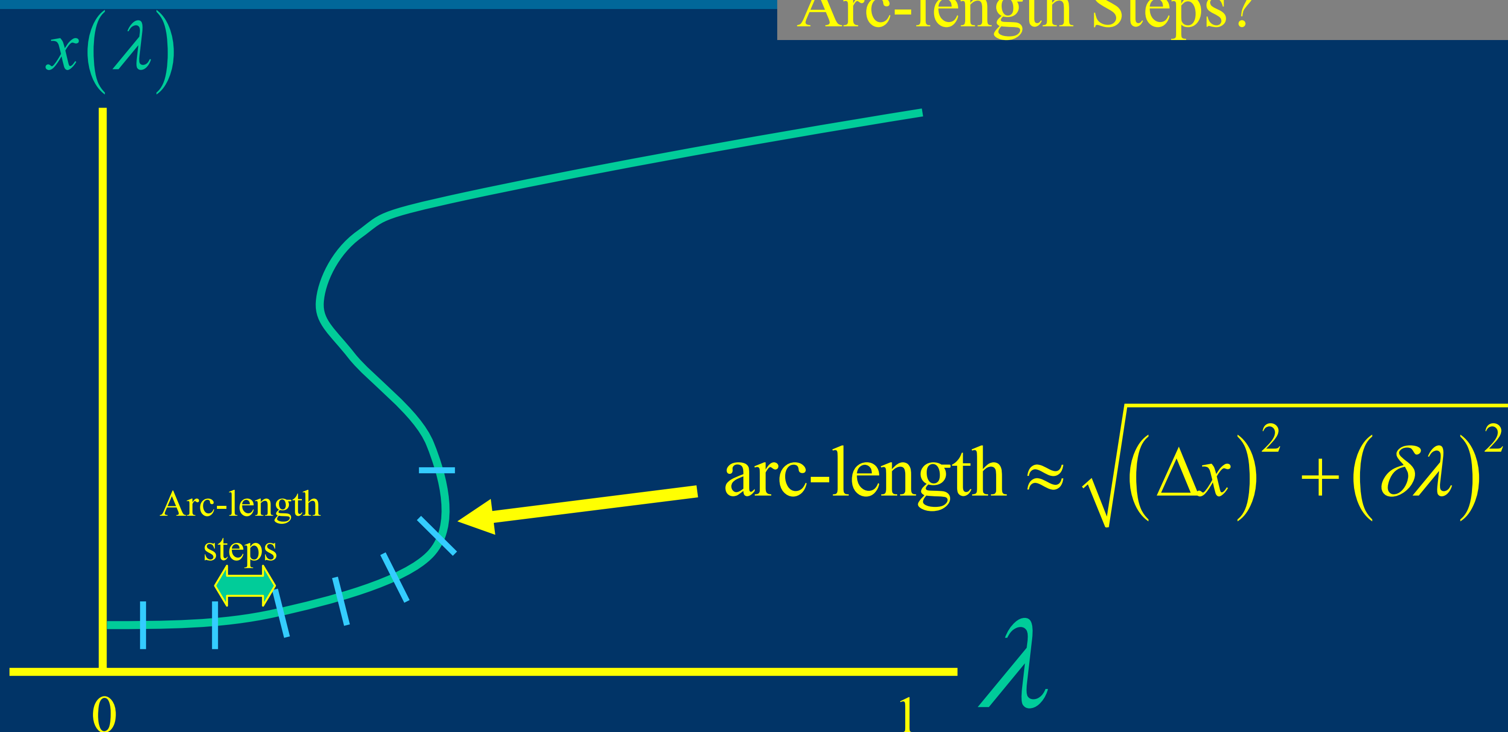
Still can have problems



Continuation Schemes

Jacobian Altering Scheme

Arc-length Steps?



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

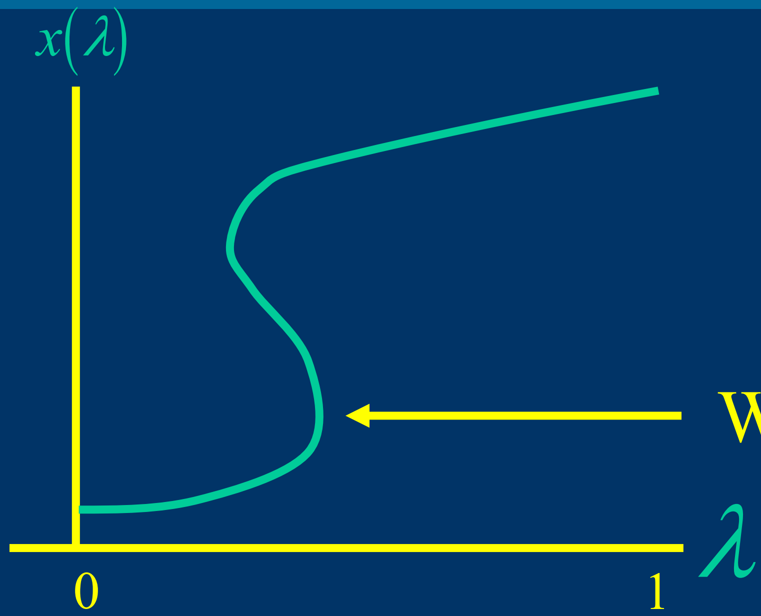
$$\begin{bmatrix} \dfrac{\partial \tilde{F}\left(x^k,\lambda^k\right)}{\partial x} & \dfrac{\partial \tilde{F}\left(x^k,\lambda^k\right)}{\partial \lambda} \\ 2\left(x^k - x\left(\lambda_{prev}\right)\right)^T & 2\left(\lambda^k - \lambda_{prev}\right) \end{bmatrix} \begin{bmatrix} x^{k+1}-x^k \\ \lambda^{k+1}-\lambda^k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

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

Continuation Schemes

Jacobian Altering Scheme

Arc-length Turning point



What happens here?

Upper left-hand Block is singular

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

Summary

- Image Segmentation Example
  - Large nonlinear system of equations
  - Examined issues in selecting numerical methods
- Newton Iterative Methods
  - Do not need to solve iteration equations exactly
- Gershgorin Circle Theorem
  - Sometimes gives useful bounds on eigenvalues
- Arc-Length Continuation