Newton Diagnostics: a New Handy Tool for Failing Initialization Debugging

Francesco Casella

(francesco.casella@polimi.it)

Teus van der Stelt

(teus.vanderstelt@asimptote.nl)

Philip Hannebohm

(philip.hannebohm@hsbi.de)

POLITECNICO MILANO 1863

ASIMPTOTE

ADVANCED SIMULATION CAPABILITIES FOR THE ENERGY MARKET

HSBI Hochschule Bielefeld University of Applied Sciences and Arts

OpenModelica

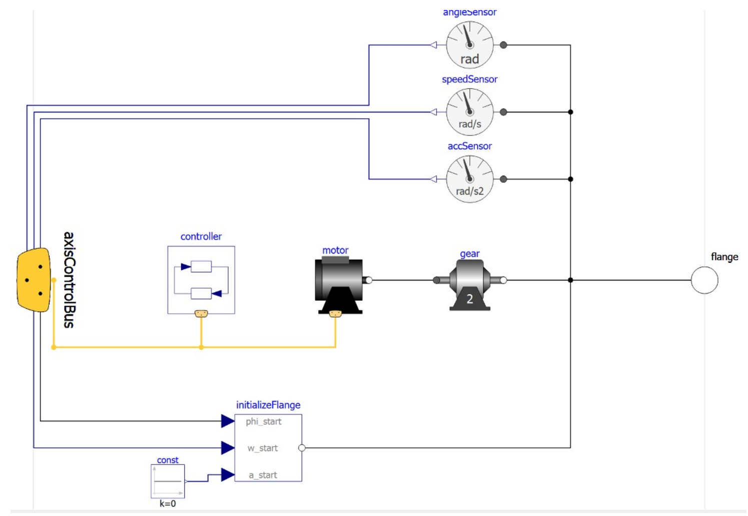
Outline

- Motivation
- Mathematical Background
- Implementation in OpenModelica
- Conclusions & Future Work

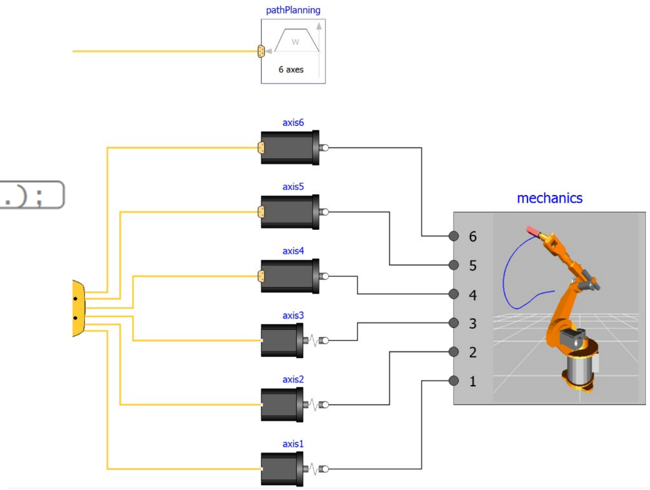
MOTIVATION

EOOMS Is Good!

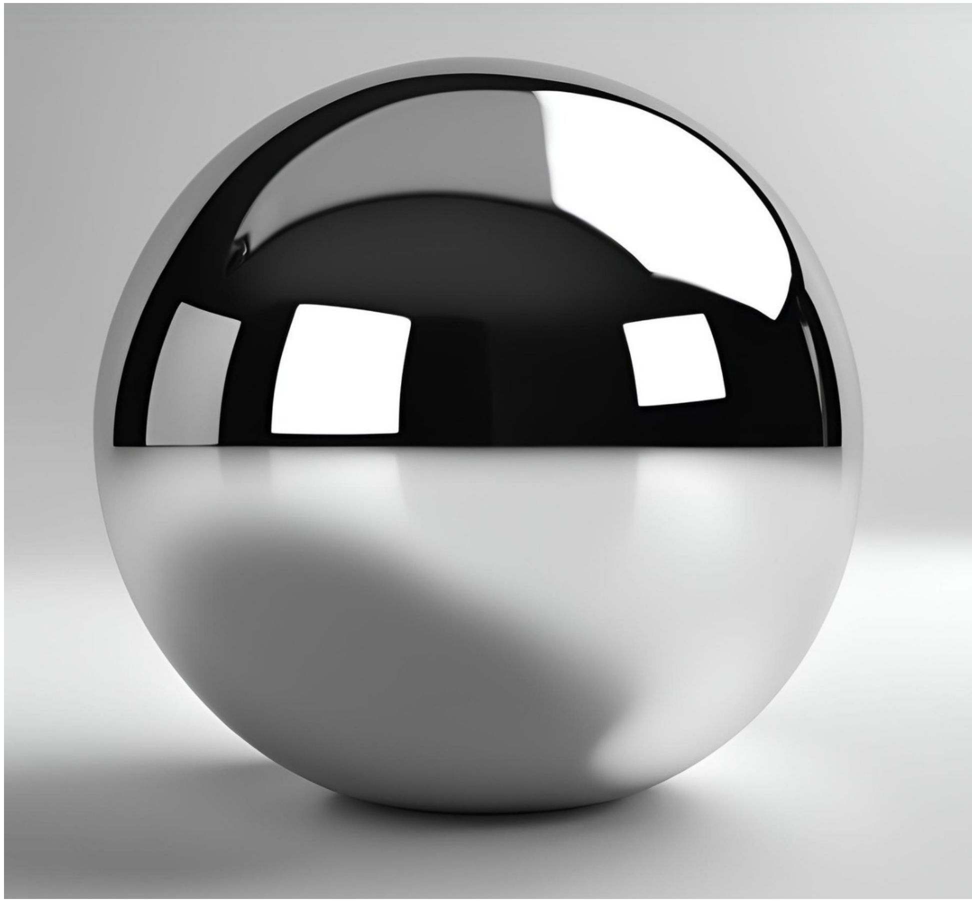
- High-level modelling
- Declarative equations
- Self-documenting
- A-causal model decomposition
- Inheritance
- Model Reuse
- Model-solver separation




```
within Modelica.Thermal.HeatTransfer.Components;
model HeatCapacitor "Lumped thermal element storing heat"
  parameter SI.HeatCapacity C
    "Heat capacity of element (= cp*m)";
  SI.Temperature T(start=293.15, displayUnit="degC")
    "Temperature of element";
  SI.TemperatureSlope der_T(start=0)
    "Time derivative of temperature (= der(T))";
  Interfaces.HeatPort_a port annotation (Placement(transformation( ...);
equation
  T = port.T;
  der_T = der(T);
  C*der(T) = port.Q_flow;
  annotation ( ...);
end HeatCapacitor;
```



EOOMS Is Good!



BUT...

The Dark Side of EOOMS

```
While solving non-linear system an assertion failed during initialization.
  The non-linear solver tries to solve the problem that could take some time.
  It could help to provide better start-values for the iteration variables.
  For more information simulate with -lv LOG_NLS_V
nonlinear system 3431 fails: at t=0
```

Debug more

```
proper start-values for some of the following iteration variables might help
  [1] Real HEX.pipe_2.mediums[1].T(start=300, nominal=500)
  [2] Real HEX.pipe_2.mediums[2].T(start=300, nominal=500)
  [3] Real HEX.pipe_2.mediums[3].T(start=300, nominal=500)
  [4] Real HEX.pipe_2.mediums[4].T(start=300, nominal=500)
  [5] Real HEX.pipe_2.mediums[5].T(start=300, nominal=500)
  [6] Real HEX.pipe_2.mediums[6].T(start=300, nominal=500)
  [7] Real HEX.pipe_2.mediums[7].T(start=300, nominal=500)
  [8] Real HEX.pipe_2.mediums[8].T(start=300, nominal=500)
  [9] Real HEX.pipe_2.mediums[9].T(start=300, nominal=500)
  [10] Real HEX.pipe_2.mediums[10].T(start=300, nominal=500)
  [11] Real HEX.pipe_2.mediums[11].T(start=300, nominal=500)
  [12] Real HEX.pipe_2.mediums[12].T(start=300, nominal=500)
  [13] Real HEX.pipe_2.mediums[13].T(start=300, nominal=500)
  [14] Real HEX.pipe_2.mediums[14].T(start=300, nominal=500)
  ...
  [59] Real HEX.pipe_1.mediums[18].T(start=304, nominal=500)
  [60] Real HEX.pipe_1.mediums[20].T(start=304, nominal=500)
  [61] Real HEX.pipe_1.mediums[17].T(start=304, nominal=500)
  [62] Real HEX.pipe_1.mediums[4].T(start=304, nominal=500)
Solving non-linear system 3431 failed at time=0. For more information please use -lv LOG_NLS.
```

Debug more

```
Failed to solve the initialization problem with global homotopy with equidistant step size.
Unable to solve initialization problem.
simulation terminated by an assertion at initialization
Simulation process failed. Exited with code 0xffffffffffffffff.
```

The Dark Side of EOOMS

The Dark Side of EOOMS

- *Proper start values for some of these variables could help (list of 62 iteration variables follows)*

Could you please, please help me?

Yes, we can!

MATHEMATICAL BACKGROUND

It All Started Here

Forte dei Marmi Beach, Tuscany, Summer 2017

The Mathematical Foundations

ELSEVIER

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

On the choice of initial guesses for the Newton-Raphson algorithm

Check for updates

Francesco Casella*, Bernhard Bachmann

[a] *Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy*
[b] *Faculty of Engineering and Mathematics, University of Applied Sciences Bielefeld, Germany*

ARTICLE INFO

Article history:
Received 19 December 2019
Revised 6 October 2020
Accepted 5 January 2021

Keywords:
Newton-Raphson's algorithm
Convergence
Nonlinear equations
Equation-based modelling

ABSTRACT

The initialization of equation-based differential-algebraic system models, and more in general the solution of many engineering and scientific problems, require the solution of systems of nonlinear equations. Newton-Raphson's method is widely used for this purpose; it is very efficient in the computation of the solution if the initial guess is close enough to it, but it can fail otherwise. In this paper, several criteria are introduced to analyze the influence of the initial guess on the evolution of Newton-Raphson's algorithm and to identify which initial guesses need to be improved in case of convergence failure. In particular, indicators based on first and second derivatives of the residual function are introduced, whose values allow to assess how much the initial guess of each variable can be responsible for the convergence failure. The use of such criteria, which are based on rigorously proven results, is successfully demonstrated in three exemplary test cases.

Problem Set-Up

$$f(x) = 0 \qquad \begin{aligned} x &\in \mathbb{R}^m \\ f &: \mathbb{R}^m \to \mathbb{R}^m \end{aligned}$$

$$x = \begin{bmatrix} w \\ z \end{bmatrix} \qquad f_x(x) = J(w) \qquad \begin{aligned} w &\in \mathbb{R}^q \\ z &\in \mathbb{R}^{m-q} \end{aligned}$$

$$f(x) = \begin{bmatrix} n(x) \\ l(x) \end{bmatrix} \qquad \begin{aligned} n &: \mathbb{R}^m \to \mathbb{R}^p \\ l &: \mathbb{R}^m \to \mathbb{R}^{m-p} \end{aligned}$$

$$f\left(\begin{bmatrix} w \\ z \end{bmatrix}\right) = g(w) + f_z z$$

$$f_x(x_{j-1})(x_j - x_{j-1}) = -f(x_{j-1}), \qquad j = 1, 2, \cdots$$

initial guess x_0

Theorem 1: Superlinear Convergence Close to Solution

$$f_x(x_{j-1})(x_j - x_{j-1}) = -f(x_{j-1}), \qquad j = 1, 2, \cdots$$

initial guess x_0

If the Jacobian $f_x(\bar{x})$ is non-singular in the solution $\bar{x}$ and Lipschitz-continuous in a neighbourhood of $\bar{x}$,

for all x_0 sufficiently close to $\bar{x}$,

the sequence $\{x_j\}$ converges not less than quadratically to $\bar{x}$.

Theorem 2: Convergence - Linear Case

$$f_x(x_{j-1})(x_j - x_{j-1}) = -f(x_{j-1}), \qquad j = 1, 2, \cdots \quad (1)$$

initial guess x_0

If Eq. (1) is linear and f_x is non-singular,

NR's algorithm converges in one step,

irrespective of the chosen initial guess x_0.

Theorem 3: Convergence - Mixed Linear-Nonlinear Case

$$f_x(x_{j-1})(x_j - x_{j-1}) = -f(x_{j-1}), \qquad j = 1, 2, \cdots \qquad (1)$$

initial guess x_0

If NR's algorithm is initialized with a first guess

$$x_0 = \begin{bmatrix} w_0 \\ z_0 \end{bmatrix},$$

the values of the approximated solution x_j at each step $j > 0$
only depend on the guess values w_0
regardless of the choice of z_0.

Theorem 4: Linear Residuals After the First Step

$$f_x(x_{j-1})(x_j - x_{j-1}) = -f(x_{j-1}), \qquad j = 1, 2, \cdots \qquad (1) \qquad f(x) = \begin{bmatrix} n(x) \\ l(x) \end{bmatrix}$$

initial guess x_0

The residuals of the linear equations after the first iteration of NR's algorithm are zero,

$$l(x_1) = 0,$$

regardless of the initial guess values x_0.

Take-Home Message

Theorem 3 + Theorem 4

*Only **nonlinear variable start values** and **nonlinear equations** matter!*

Don't bother about start attributes for the linear variables z

How Do We Understand NR is Close to Convergence?

~~$\| f(x_1) \| << \| f(x_0) \|$?~~

$$f(x) = \begin{bmatrix} n(x) \\ l(x) \end{bmatrix}$$

Trivial $z_0 = 0$ *means large* $l(x_0)$
that becomes $l(x_1) = 0$ *!*

We define the nonlinear residual $r(x_0)$

$$r(x_0) = f(x_0) + f_z(z_1 - z_0)$$

$\| f(x_1) \| << \| r(x_0) \|$ means close to convergence

Theorem 5: Is NR Close to Convergence?

$$f^i(x_1) = f^i(x_0) + f^i_x(x_0)(x_1 - x_0) + \frac{1}{2}(x_1 - x_0)' f^i_{xx}(x_0)(x_1 - x_0) + h^i(x_1, x_0)$$

$$|h^i(x_1, x_0)| = \alpha_i ||r(x_0)||_\infty$$

$$\alpha_i = \frac{\left| f^i(x_1) - \frac{1}{2}(x_1 - x_0)' f^i_{xx}(x_0)(x_1 - x_0) \right|}{||r(x_0)||_\infty}$$

$$\alpha = max(\alpha_i)$$

$$\Gamma_{ijk} = \left| \frac{1}{2} \frac{\partial^2 g^i(w_0)}{\partial w_j \partial w_k} \frac{(w_{1,k} - w_{0,k})(w_{1,j} - w_{0,j})}{||r(x_0)||_\infty} \right|$$

$$\boxed{\sum_{jk} \Gamma_{ijk} \leq \beta \qquad \forall i = 1, \cdots, p \quad \Rightarrow \quad \|f(x_1)\|_\infty \leq (\alpha + \beta)\|r(x_0)\|_\infty}$$

Main (Heuristic) Idea

$$\alpha_i = \frac{\left| f^i(x_1) - \frac{1}{2}(x_1 - x_0)' f^i_{xx}(x_0)(x_1 - x_0) \right|}{||r(x_0)||_\infty} \qquad \alpha = max(\alpha_i)$$

$$\Gamma_{ijk} = \left| \frac{1}{2} \frac{\partial^2 g^i(w_0)}{\partial w_j \partial w_k} \frac{(w_{1,k} - w_{0,k})(w_{1,j} - w_{0,j})}{||r(x_0)||_\infty} \right|$$

$$\sum_{jk} \Gamma_{ijk} \leq \beta \qquad \forall i = 1, \cdots, p \quad \Rightarrow \quad \|f(x_1)\|_\infty \leq (\alpha + \beta)\|r(x_0)\|_\infty$$

If all the α_i and Γ_{ijk} are small, α and β are small

$\rightarrow || f(x_1)||_\infty$ << $|| r(x_0)|_\infty|$

$\rightarrow$ *NR is close to convergence!*

$\Downarrow$

If NR does not converge, the culprits are likely related to the large α_i and Γ_{ijk} !

Main (Heuristic) Idea - Cont'd

$$\Gamma_{ijk} = \left| \frac{1}{2} \frac{\partial^2 g^i(w_0)}{\partial w_j \partial w_k} \frac{(w_{1,k} - w_{0,k})(w_{1,j} - w_{0,j})}{||r(x_0)||_\infty} \right|$$

$$\alpha_i = \frac{\left| f^i(x_1) - \frac{1}{2}(x_1 - x_0)' f^i_{xx}(x_0)(x_1 - x_0) \right|}{||r(x_0)||_\infty}$$

- *Large Γ_{ijk} points to initial guesses $w_{0,j}$ and $w_{0,k}$ creating trouble in equation i*

- *Large α_i points to nonlinear equation creating trouble*

- *Γ_{ijk} considers 2nd-order nonlinear effects*
  α_i considers higher-order nonlinear effects

Additional (Heuristic) Idea

Theorem 1 $\rightarrow$ *if* x_0 *is close to the solution,* x_1 *will be almost there, no matter what the value of* x_0

Compute $\dfrac{\partial x_1}{\partial x_0} = \Sigma$

If NR is close to convergence, $\Sigma \approx 0$

If NR does not converge, the culprits are likely related to the large diagonal values σ_{ii} *of* Σ*!*

How To Compute Σ

$$H_i = (w_1 - w_0)' f^i_{ww}(w_0)$$

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \cdots \\ H_p \end{bmatrix}$$

$$\frac{\partial x_1}{\partial x_0} = \Sigma$$

$$\Sigma = -[f_x(w_0)]^{-1} \begin{bmatrix} H_{p\times q} & 0_{p\times(m-q)} \\ 0_{(m-p)\times q} & 0_{(m-p)\times(m-q)} \end{bmatrix}$$

➡ *Combines 1st-order and 2nd-order information*

Summary

In case of NR converge failure:

- Compute one Newton step
- Compute the α_i, Γ_{ijk}, σ_{ii} indicators
- Rank them in descending order
- Variables with potentially problematic start values are
  - Found in equations with large α_i
  - Pointed by j and k indeces of large Γ_{ijk}
  - Pointed by i indeces of large σ_{ii}
- Equations with large α_i and Γ_{ijk} may be made less nonlinear by homotopy

IMPLEMENTATION IN OMC

Implementation in OpenModelica

- Implementation in the C runtime by Teus van der Stelt started in 2021 with help from Karim Abdelhak and Andreas Heuermann
- The development stalled numerous times for various reasons during 2021 and 2022
- It was eventually resumed during late 2024 and finalized January 2025.

- Available in version 1.25.0
- Activated with runtime flag `-lv=NLS_LOG_NEWTON_DIAGNOSTICS`
- Tested successfully on the examples shown in the paper

Example Output

- Mixed linear/nonlinear electrical circuit model
- Known analytic solution
- Start values selected 10% away from the solution
  $\rightarrow$ solver fails
- Diagnostic output clearly points out
  - the variable v_d whose start attribute must be fixed to achieve convergence
  - the most problematic equation (the diode)

$$i - \left(i_s e^{v_d/v_t} - 1\right) = 0$$

$$vi - P = 0$$

$$v - \sum_{j=1}^{N} v_j - v_d = 0$$

$$v_j - Ri = 0$$

```
By variable
  Var no.  Var name                                  Initial guess  max(Gamma,sigma)
  -------  ----------------------------------------  -------------  ----------------
        4                                       v_d           0.63    14.99
        3                                         i            0.9     0.07
        5                                         v           9.63     0.05
By equation
  Eq no.  Eq idx    max(alpha,Gamma)
  ------  ------    ----------------
       4      10  1.31e+05
       5       9   0.03
```

What Still Needs To Be Done

- Handling of systems with only numerical Jacobians
- Scaling of residuals
  (essential to handle $||r(x_0)||_\infty$ with residual in SI units)
- Testing in real failure cases
- Integration in the Equation-Based Debugger

Conclusions

- Nonlinear Newton-Raphson solver failures are a big problem in EOOMS
- Identifying which start attributes need to be improved to achieve convergence *can help a lot*
- Identifying the most strongly nonlinear equations can also help, e.g. to point out where using homotopy could be beneficial

- The LOG_NLS_NEWTON_DIAGNOSTICS method now provides this information
- Available in OMC 1.25.0

Thank you for your kind attention!