Computing multiple solutions of nonlinear problems

Patrick E. Farrell

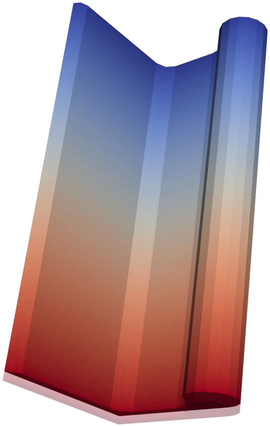
University of Oxford

Can you conduct an experiment twice . . .

. . . and get two different answers?

Can you conduct an experiment twice . . .

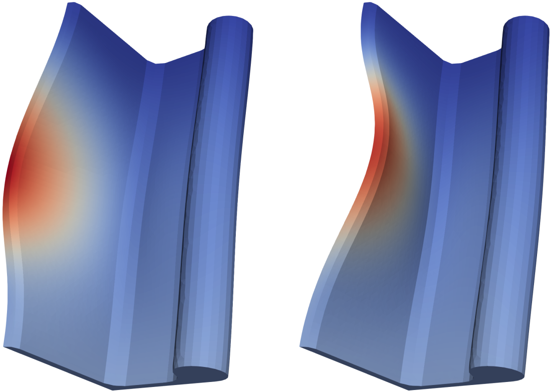
. . . and get two different answers?



Axial displacement test of an Embraer aircraft stiffener.

Can you conduct an experiment twice . . .

. . . and get two different answers?

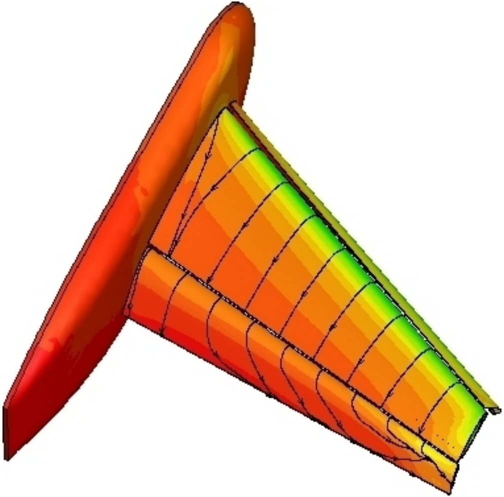


Two different, stable configurations.

When a problem has multiple solutions, it is usually crucial.

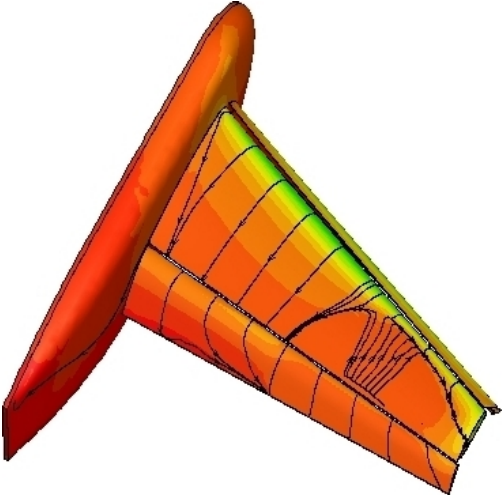
The AIAA/NASA high lift prediction test case (Kamenetskiy et al., 2013)

When a problem has multiple solutions, it is usually crucial.



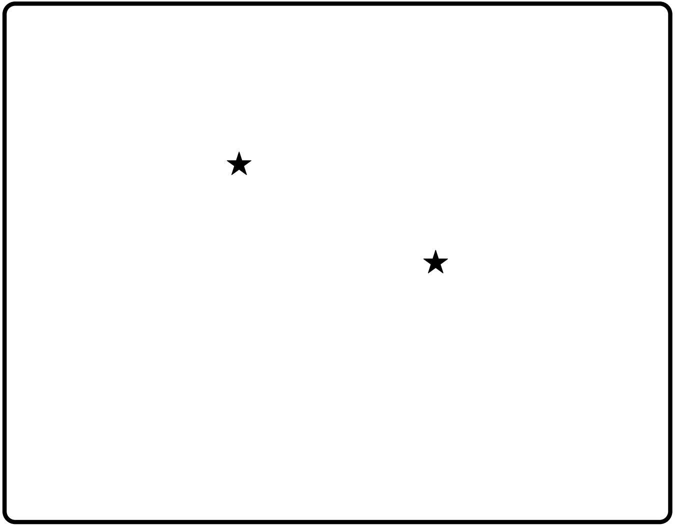
The AIAA/NASA high lift prediction test case (Kamenetskiy et al., 2013)

When a problem has multiple solutions, it is usually crucial.



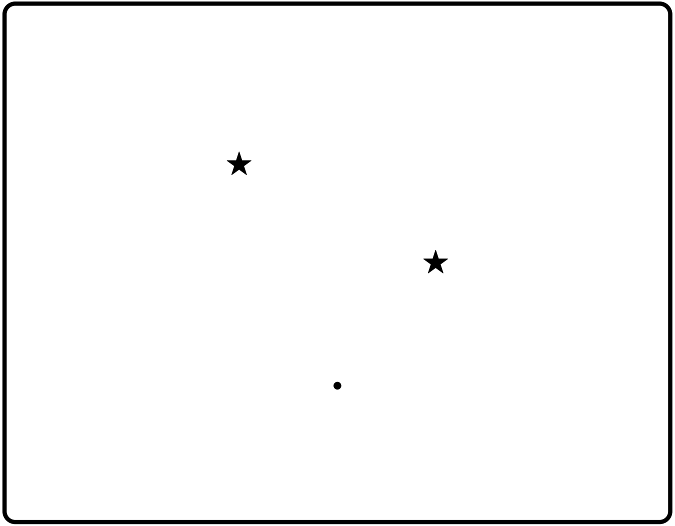
The AIAA/NASA high lift prediction test case (Kamenetskiy et al., 2013)

When a problem has multiple solutions, it is usually crucial.



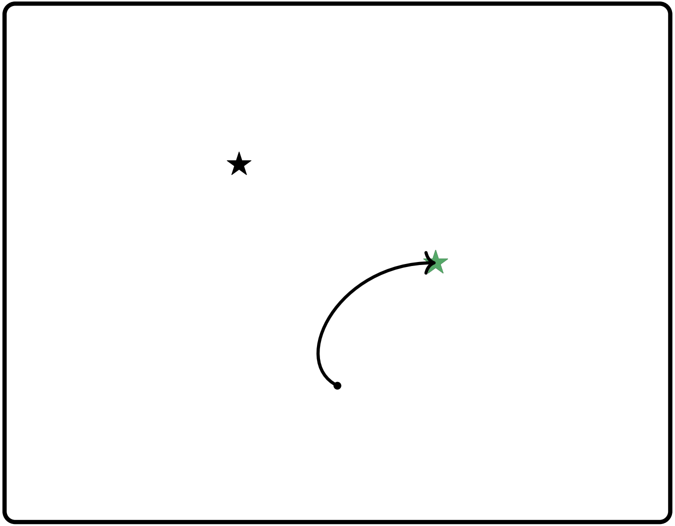
A PDE with two unknown solutions

When a problem has multiple solutions, it is usually crucial.



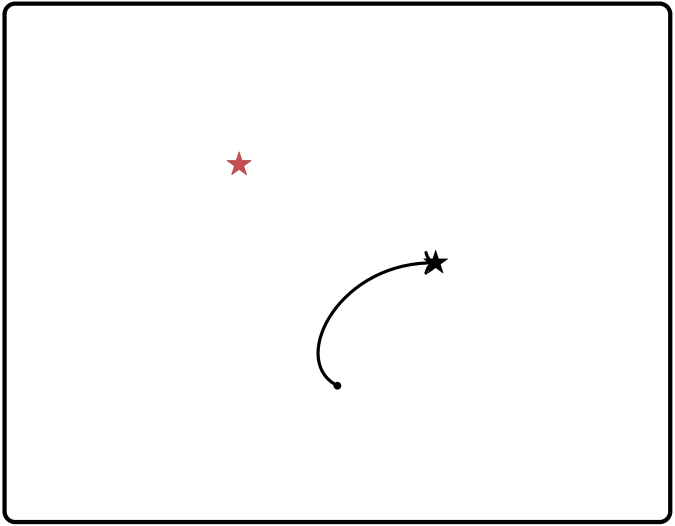
Start from some initial guess

When a problem has multiple solutions, it is usually crucial.



We converge to one solution, our prediction

When a problem has multiple solutions, it is usually crucial.



But nature has chosen another (unknown) solution!

When a problem has multiple solutions, it is usually crucial.

Warning

Your calculations can be correct, but still return the *wrong answer*.

Mathematical formulation

Compute the multiple *solutions* $u^\star$ of a stationary nonlinear equation

$$F(u^\star, \lambda) = 0$$
$$F \in C^1(X \times \mathbb{R}, Y)$$

as a function of a parameter $\lambda \in \mathbb{R}$. (X, Y Banach spaces.)

Mathematical formulation

Compute the multiple *solutions* $u^\star$ of a stationary nonlinear equation

$$F(u^\star, \lambda) = 0$$
$$F \in C^1(X \times \mathbb{R}, Y)$$

as a function of a parameter $\lambda \in \mathbb{R}$. (X, Y Banach spaces.)

Case #1: aircraft stiffener

$u^\star$ displacement (a vector field), λ loading, F hyperelasticity

Case #2: aircraft wing

$u^\star$ velocity and pressure, λ angle of attack, F Navier–Stokes

Mathematical formulation

Compute the multiple *solutions* $u^\star$ of a stationary nonlinear equation

$$F(u^\star, \lambda) = 0$$
$$F \in C^1(X \times \mathbb{R}, Y)$$

as a function of a parameter $\lambda \in \mathbb{R}$. (X, Y Banach spaces.)

Warning

We (usually) can't guarantee to find *all* solutions. But finding many is better than finding one.

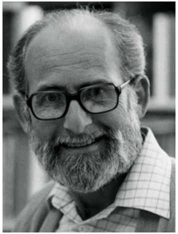
Section 2

Classical techniques

Basic idea of numerical bifurcation analysis:

procedure $\text{ANALYSE}(u_0, \lambda_0)$

end procedure

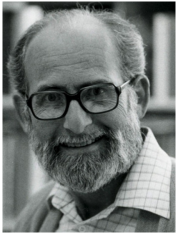


Herbert Keller

Basic idea of numerical bifurcation analysis:

procedure ANALYSE(u_0, λ_0)
    continue branch of solutions;

end procedure



Herbert Keller

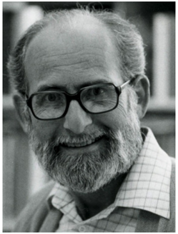
Continuation

Extending our knowledge of the branch to other values of λ.

Basic idea of numerical bifurcation analysis:

procedure ANALYSE(u_0, λ_0)
    continue branch of solutions;
    detect bifurcations on the branch;

end procedure



Herbert Keller

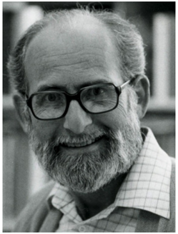
Bifurcation detection

Discovering when a bifurcation has occurred on the branch.

Basic idea of numerical bifurcation analysis:

procedure ANALYSE(u_0, λ_0)
    continue branch of solutions;
    detect bifurcations on the branch;
    localise bifurcations;

end procedure



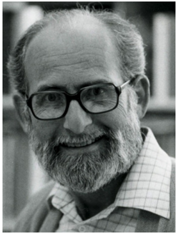
Herbert Keller

Bifurcation localisation

Identifying precisely the bifurcation point.

Basic idea of numerical bifurcation analysis:

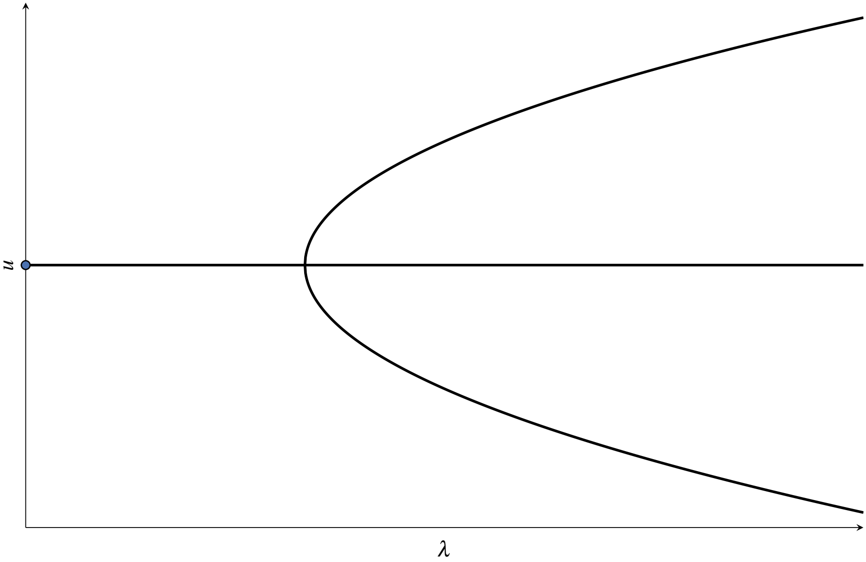
procedure ANALYSE(u_0, λ_0)
continue branch of solutions;
detect bifurcations on the branch;
localise bifurcations;
switch branches at bifurcations, and recurse.
end procedure



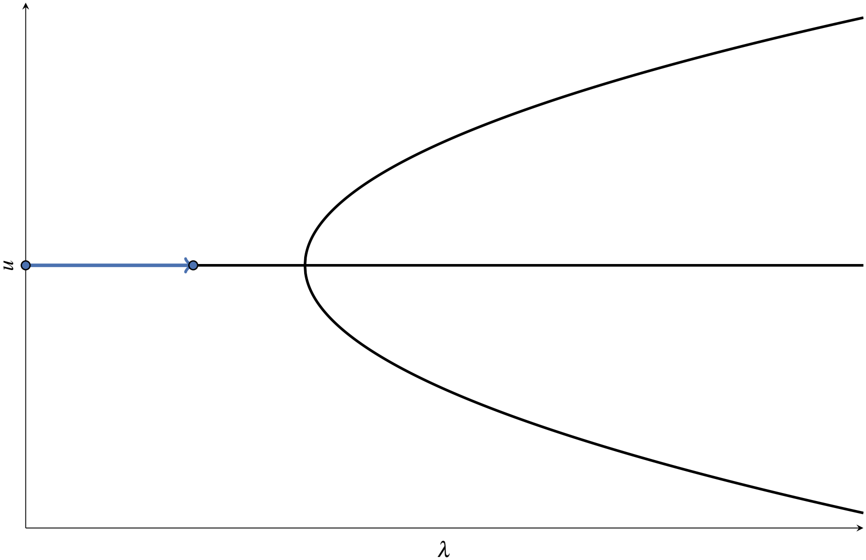
Herbert Keller

Branch switching

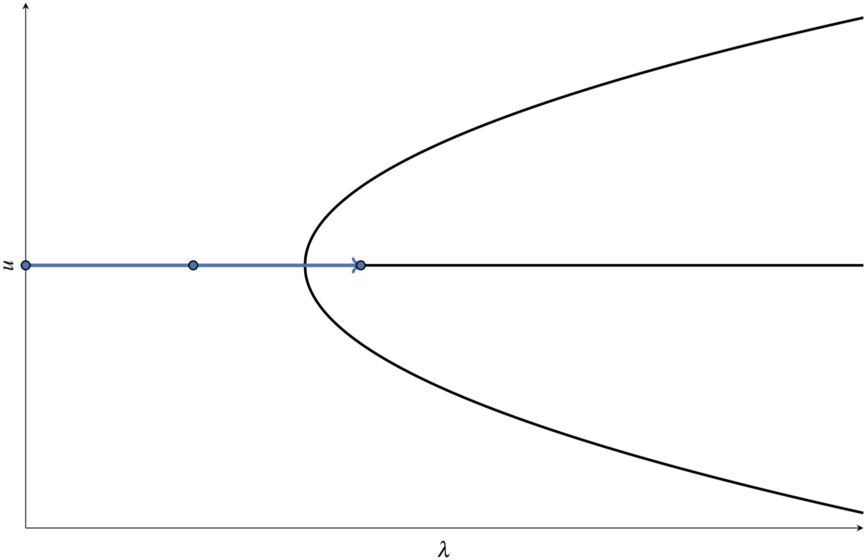
Constructing the emanating branches, and analysing them recursively.



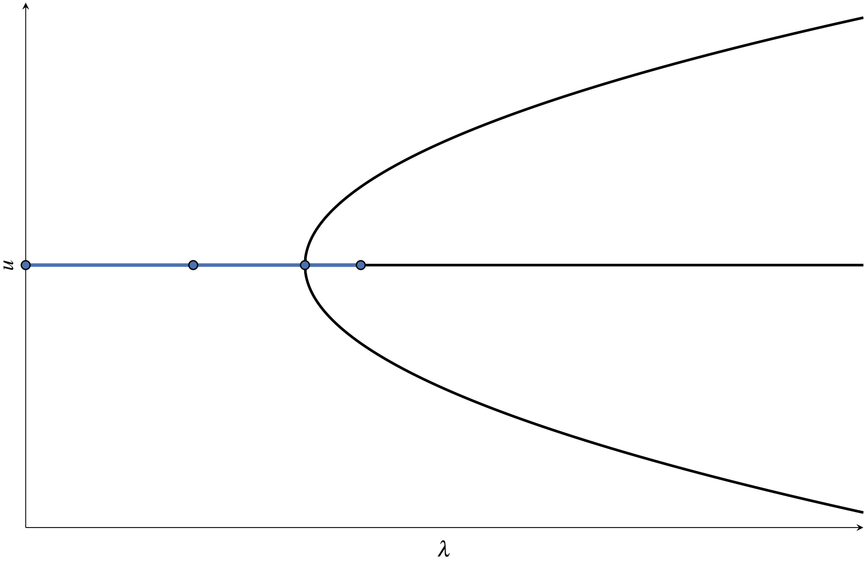
Start with one known solution (u_0, λ_0).



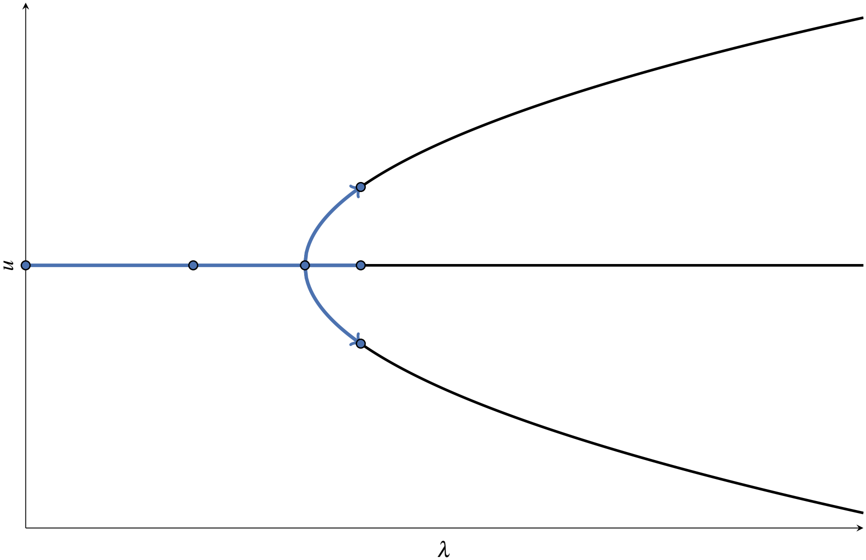
Continue the branch, finding a solution for the next parameter value.



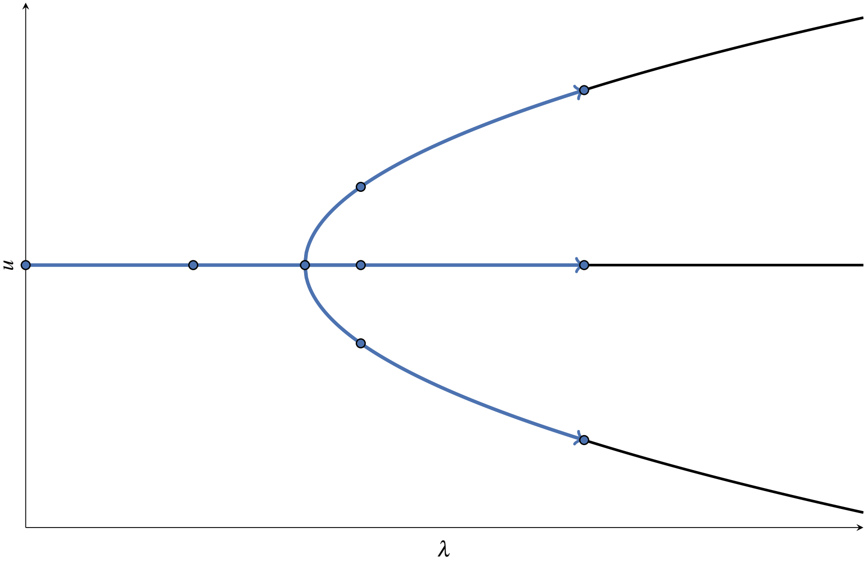
Detect we have passed a bifurcation, where the number of solutions varies.



Localise bifurcation point, by solving a system of equations.



Switch branches.



Apply recursively.

Good news

The combination of continuation and branch switching is very powerful.

Good news

The combination of continuation and branch switching is very powerful.

Bad news

However, it has some disadvantages and weaknesses, too.

Downside A

You have to solve a lot of different problems.

Downside A

You have to solve a lot of different problems.

We work for years to develop a good solver for

$$F(u, \lambda) = 0 \ldots$$

Downside A

You have to solve a lot of different problems.

We work for years to develop a good solver for

$$F(u, \lambda) = 0 \ldots$$

but now we need to solve related but different problems like

$$\begin{bmatrix} F(u, \lambda) \\ p(u, \lambda, s) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad F_u(u, \lambda)v = \lambda v \qquad \begin{bmatrix} F(u, \lambda) \\ F_u(u, \lambda)v \\ \|v\|^2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Downside A

You have to solve a lot of different problems.

We work for years to develop a good solver for

$$F(u, \lambda) = 0 \ldots$$

but now we need to solve related but different problems like

$$\begin{bmatrix} F(u, \lambda) \\ p(u, \lambda, s) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad F_u(u, \lambda)v = \lambda v \qquad \begin{bmatrix} F(u, \lambda) \\ F_u(u, \lambda)v \\ \|v\|^2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Finely-resolved simulations

This is OK when you can afford Gaussian elimination, but not at scale.

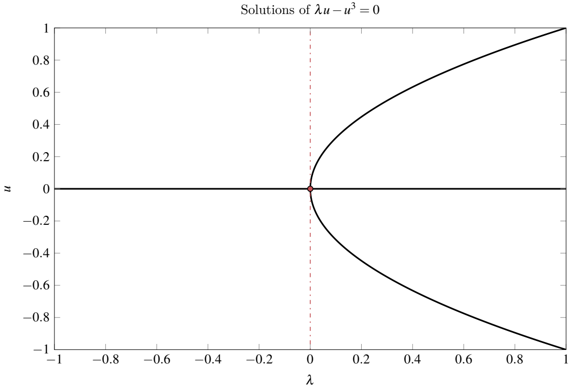
Downside B

We can only find branches *connected* to our initial data.

Downside B

We can only find branches *connected* to our initial data.

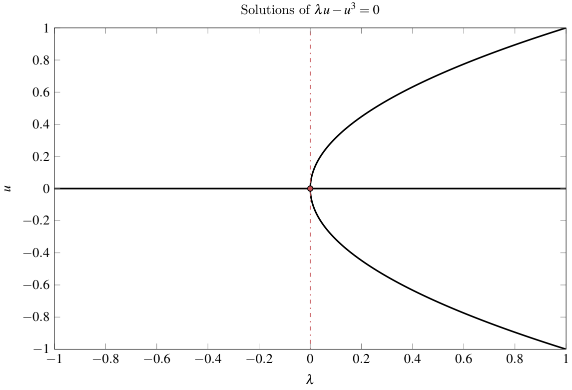
This works fine ...



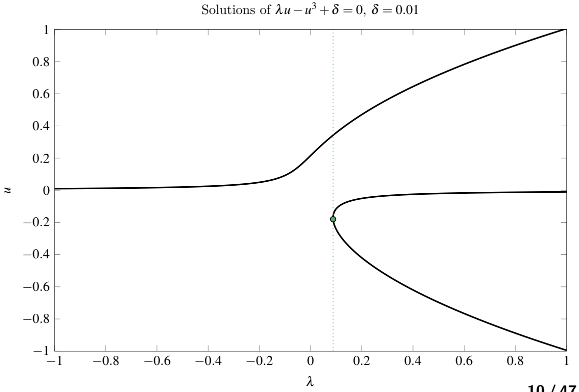
Downside B

We can only find branches *connected* to our initial data.

This works fine . . .



. . . but this does not.



The standard approach to deal with this is to

(a) modify the problem to restore connectedness;

The standard approach to deal with this is to

(a) modify the problem to restore connectedness;

(b) apply continuation + branch switching;

The standard approach to deal with this is to

(a) modify the problem to restore connectedness;

(b) apply continuation + branch switching;

(c) continue the branches you find back to the problem you care about.

The standard approach to deal with this is to

(a) modify the problem to restore connectedness;
(b) apply continuation + branch switching;
(c) continue the branches you find back to the problem you care about.

Problem A

You have to know to look for the missing branches.

The standard approach to deal with this is to

(a) modify the problem to restore connectedness;

(b) apply continuation + branch switching;

(c) continue the branches you find back to the problem you care about.

Problem A

You have to know to look for the missing branches.

Problem B

Executing this is manual and tedious.

The standard approach to deal with this is to

(a) modify the problem to restore connectedness;
(b) apply continuation + branch switching;
(c) continue the branches you find back to the problem you care about.

Problem A

You have to know to look for the missing branches.

Problem B

Executing this is manual and tedious.

Problem C

Restoring connectedness is not always possible!

In this talk I describe a complementary approach.

Disconnected diagrams

An algorithm that can compute **disconnected bifurcation diagrams**.

In this talk I describe a complementary approach.

Disconnected diagrams

An algorithm that can compute **disconnected bifurcation diagrams**.

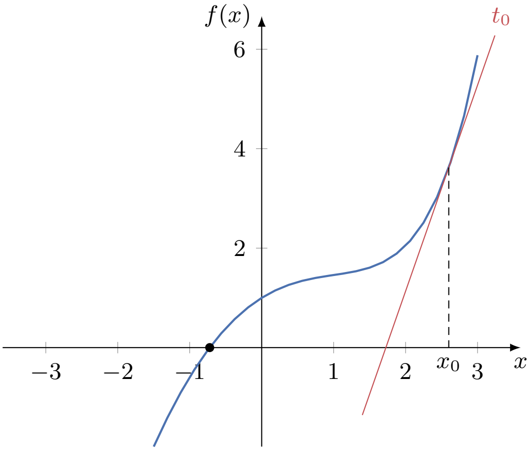
Simplicity & scaling

The computational kernel is exactly the same as Newton's method:
only one kind of problem to solve.

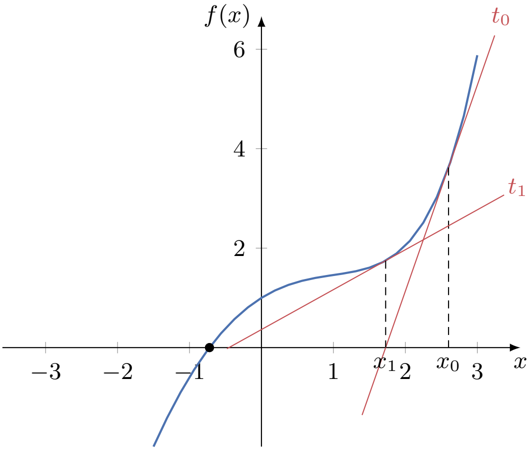
Section 3

Newton's method

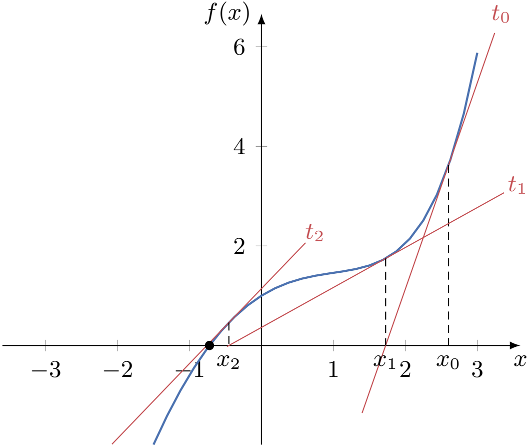
Essential idea: solve a succession of *linearised* rootfinding problems.



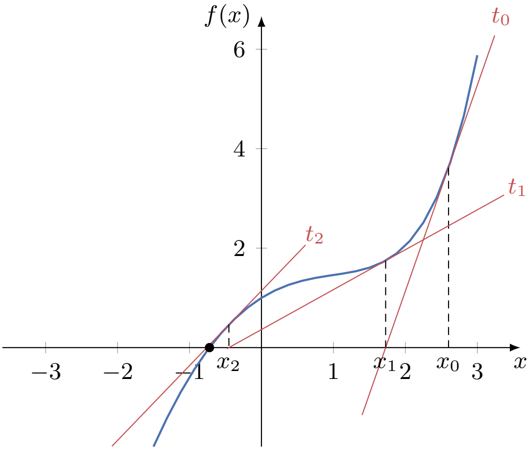
Essential idea: solve a succession of *linearised* rootfinding problems.



Essential idea: solve a succession of *linearised* rootfinding problems.



Essential idea: solve a succession of *linearised* rootfinding problems.



solve $f'(x_k)\delta x_k = -f(x_k)$; update $x_{k+1} = x_k + \delta x_k$.

This extends to $F \in C^1(X;Y)$. Given $u_0 \in X$, Newton's method is to

$$\text{solve } F_u(u_k)\delta u_k = -F(u_k); \text{ update } u_{k+1} = u_k + \delta u_k,$$

where F_u is the Fréchet derivative of F (Jacobian if $X = Y = \mathbb{R}^N$).

This extends to $F \in C^1(X;Y)$. Given $u_0 \in X$, Newton's method is to

$$\text{solve } F_u(u_k)\delta u_k = -F(u_k); \text{ update } u_{k+1} = u_k + \delta u_k,$$

where F_u is the Fréchet derivative of F (Jacobian if $X = Y = \mathbb{R}^N$).

Good news

This converges if u_0 is close to a root, and usually converges *quadratically*.

Section 4

Deflation

Deflation

Fix parameter λ. Given

- a Fréchet differentiable residual $F : X \to Y$
- a solution $u \in X$, $F(u) = 0$, $F_u(u)$ nonsingular

Deflation

Fix parameter λ. Given

- a Fréchet differentiable residual $F : X \to Y$
- a solution $u \in X$, $F(u) = 0$, $F_u(u)$ nonsingular

construct a **new nonlinear problem** $G : X \to Y$ such that:

Deflation

Fix parameter λ. Given

- a Fréchet differentiable residual $F : X \to Y$
- a solution $u \in X$, $F(u) = 0$, $F_u(u)$ nonsingular

construct a **new nonlinear problem** $G : X \to Y$ such that:

- (Preservation of solutions) $F(\tilde{u}) = 0 \iff G(\tilde{u}) = 0 \ \forall \ \tilde{u} \neq u$;

Deflation

Fix parameter λ. Given

- a Fréchet differentiable residual $F : X \to Y$
- a solution $u \in X$, $F(u) = 0$, $F_u(u)$ nonsingular

construct a **new nonlinear problem** $G : X \to Y$ such that:

- (Preservation of solutions) $F(\tilde{u}) = 0 \iff G(\tilde{u}) = 0 \; \forall \; \tilde{u} \neq u$;
- (Deflation property) Newton's method applied to G will never converge to u again, starting from any initial guess.

Deflation

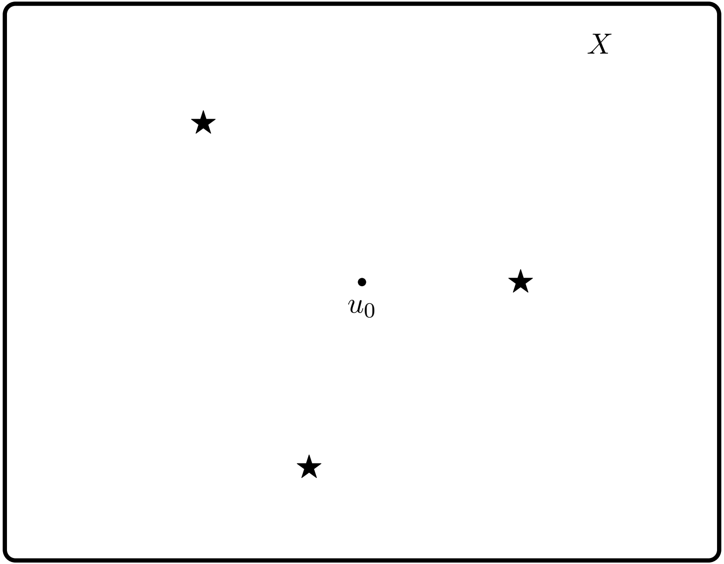
Fix parameter λ. Given

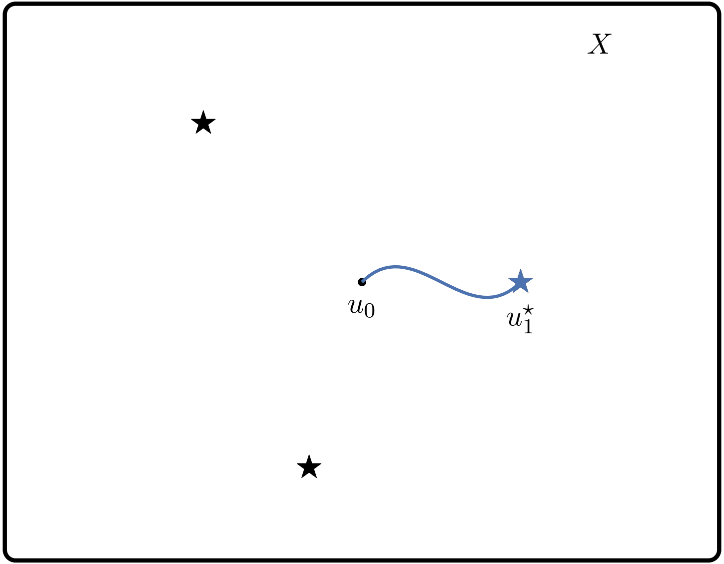
- a Fréchet differentiable residual $F : X \to Y$
- a solution $u \in X$, $F(u) = 0$, $F_u(u)$ nonsingular

construct a **new nonlinear problem** $G : X \to Y$ such that:

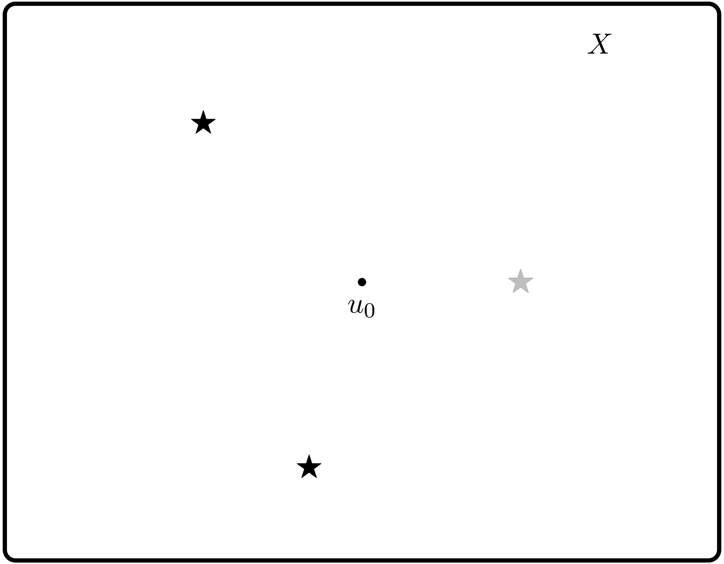
- (Preservation of solutions) $F(\tilde{u}) = 0 \iff G(\tilde{u}) = 0 \;\forall\; \tilde{u} \neq u$;
- (Deflation property) Newton's method applied to G will never converge to u again, starting from any initial guess.

Find more solutions, starting from the same initial guess.

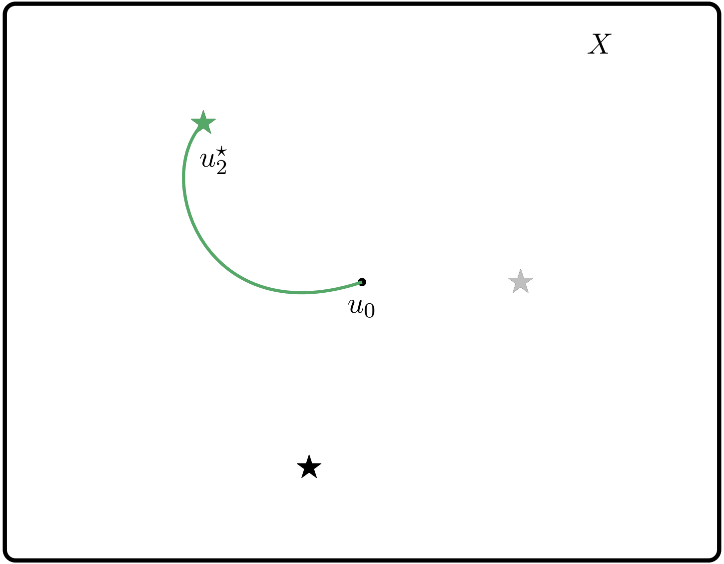

X
u_0



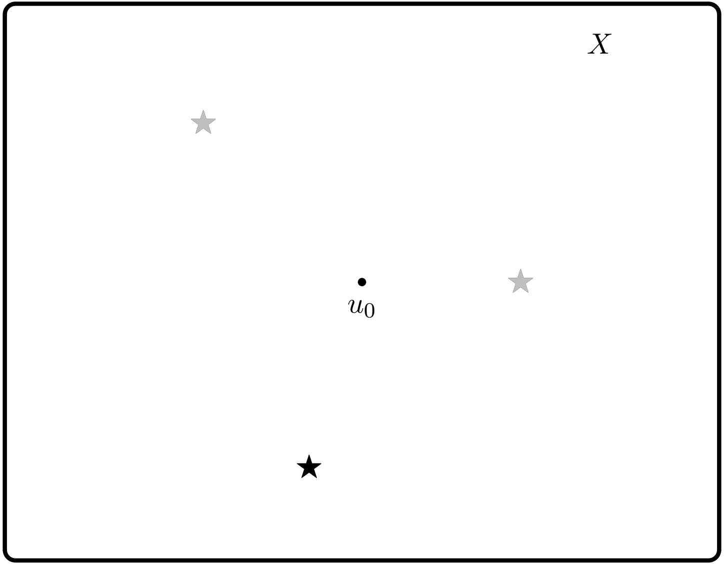
Newton from initial guess.



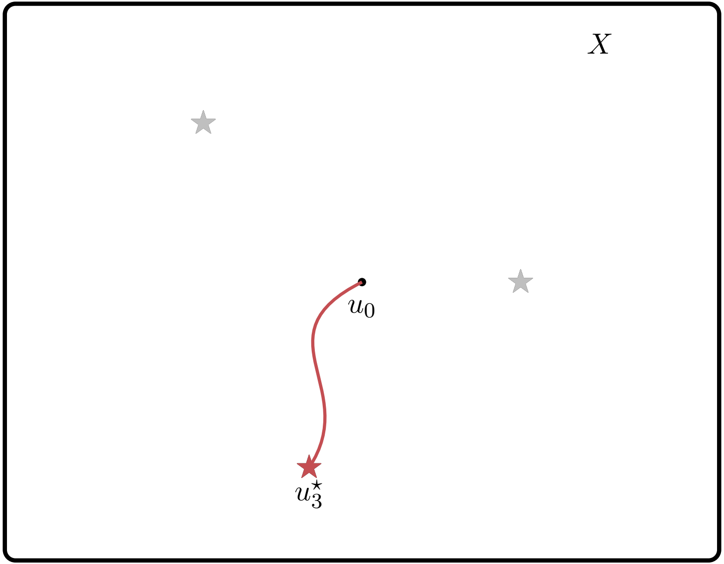
Deflate solution found.



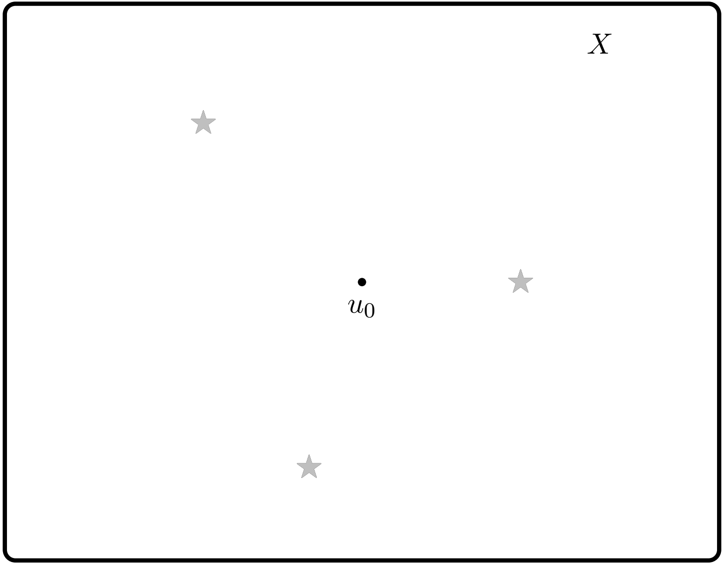
Newton from initial guess.



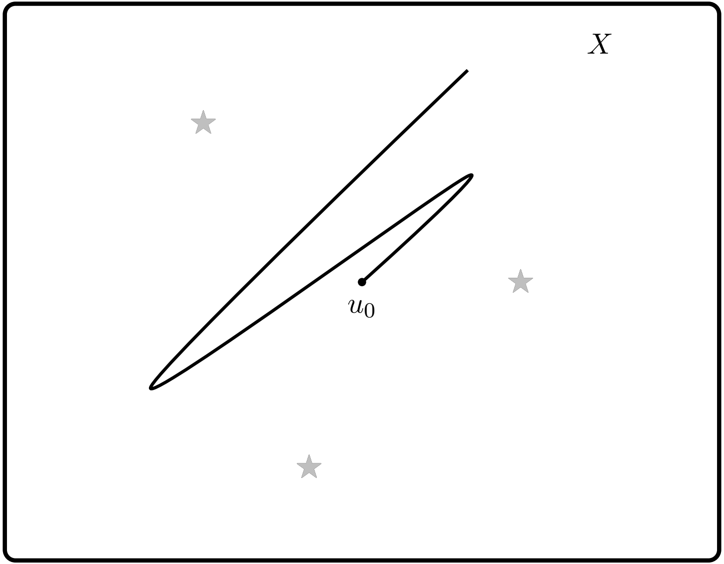
Deflate solution found.



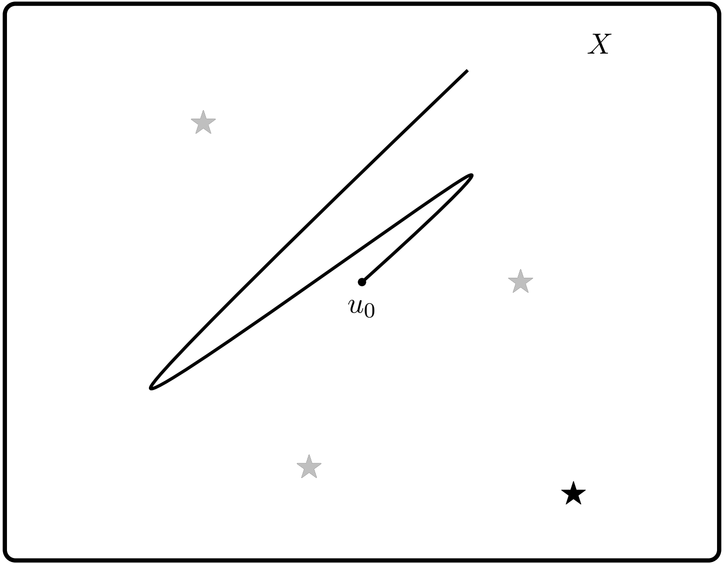
Newton from initial guess.



Deflate solution found.



Terminate on nonconvergence.



Terminate on nonconvergence.

How can you do it?

Definition

We say that $M(u; u^\star) \in C^1(X, Y)$ is a *deflation operator* if

$$\liminf_{u \to u^\star} \|G(u)\| := \liminf_{u \to u^\star} \|M(u; u^\star) F(u)\| > 0.$$

How can you do it?

Definition

We say that $M(u;u^\star) \in C^1(X,Y)$ is a *deflation operator* if

$$\liminf_{u \to u^\star} \|G(u)\| := \liminf_{u \to u^\star} \|M(u;u^\star)F(u)\| > 0.$$

Newton's method is looking for sequences (u_k) that send $G(u_k) \to 0$. But a deflation operator hides the existence of the root as $u \to u^\star$.

The deflation operator we propose

$$M(u; u^\star) := \left(\frac{1}{\|u - u^\star\|^p} + 1 \right), \quad p \geq 1.$$

Ásgeir Birkisson, 1985–

Simon Funke, 1983–

The deflation operator we propose

$$M(u; u^\star) := \left(\frac{1}{\|u - u^\star\|^p} + 1\right), \quad p \geq 1.$$

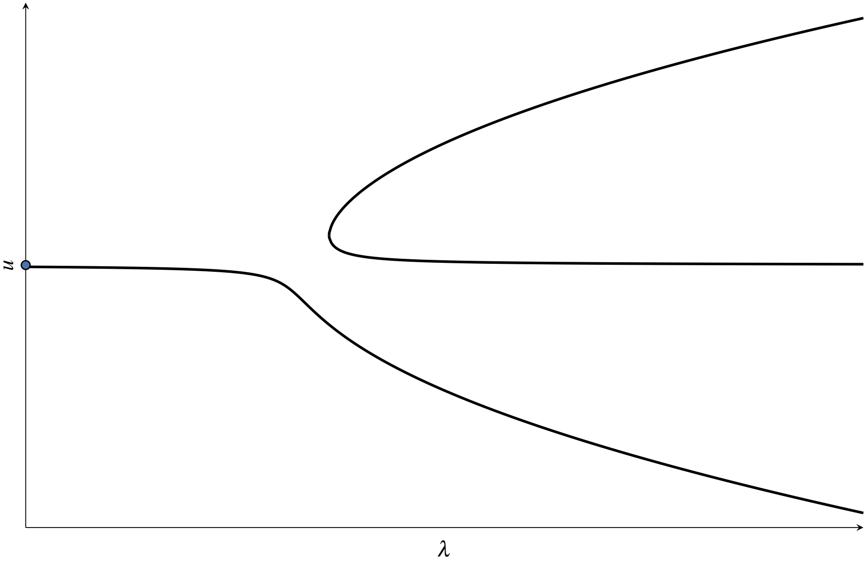
This has the right behaviour *both* as

$$\|u - u^\star\| \to 0,$$
$$\|u - u^\star\| \to \infty,$$

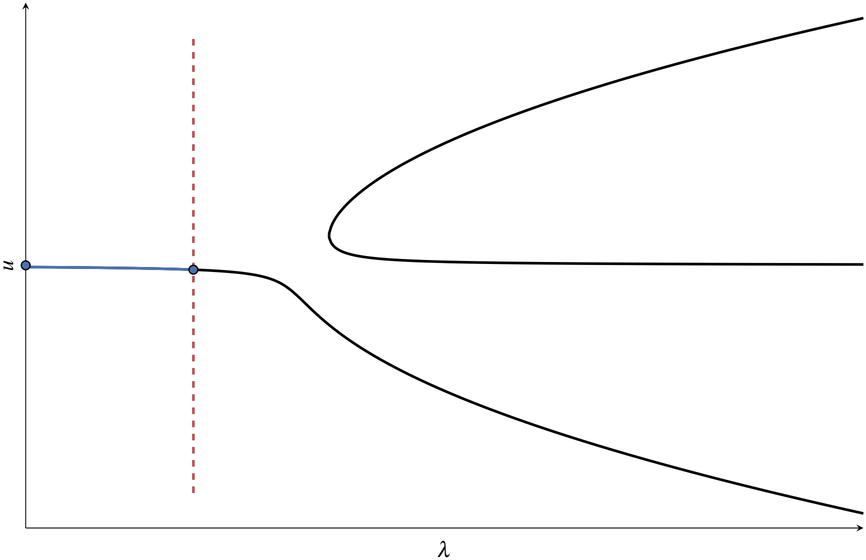
which makes deflation much more reliable.

Ásgeir Birkisson, 1985–

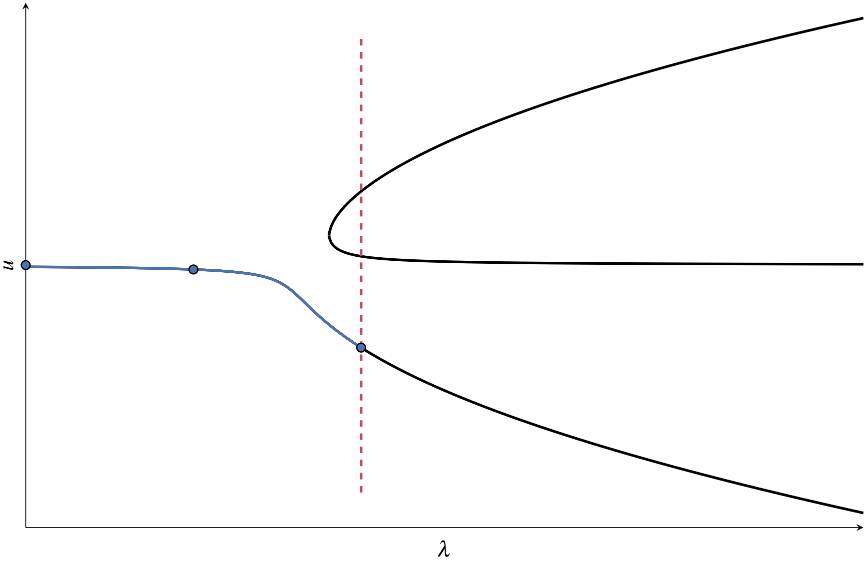
Simon Funke, 1983–



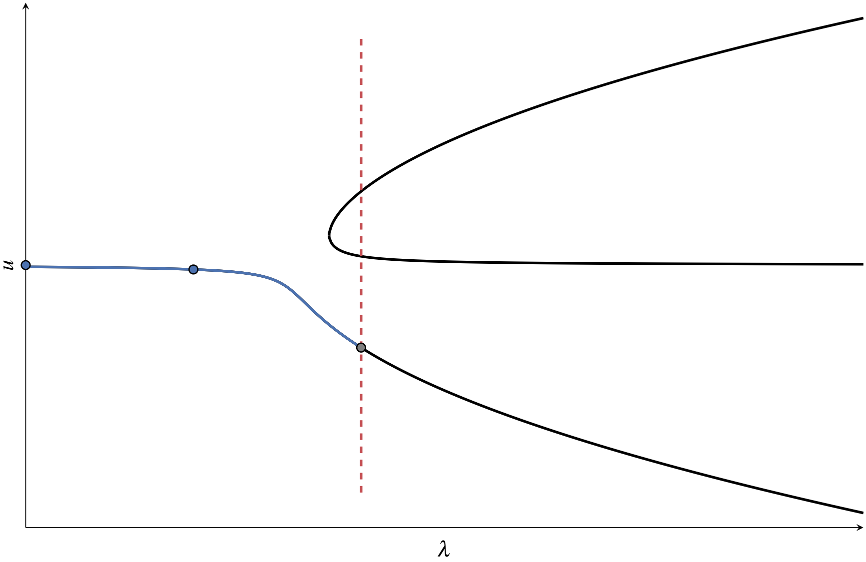
Start with (u_0, λ_0).



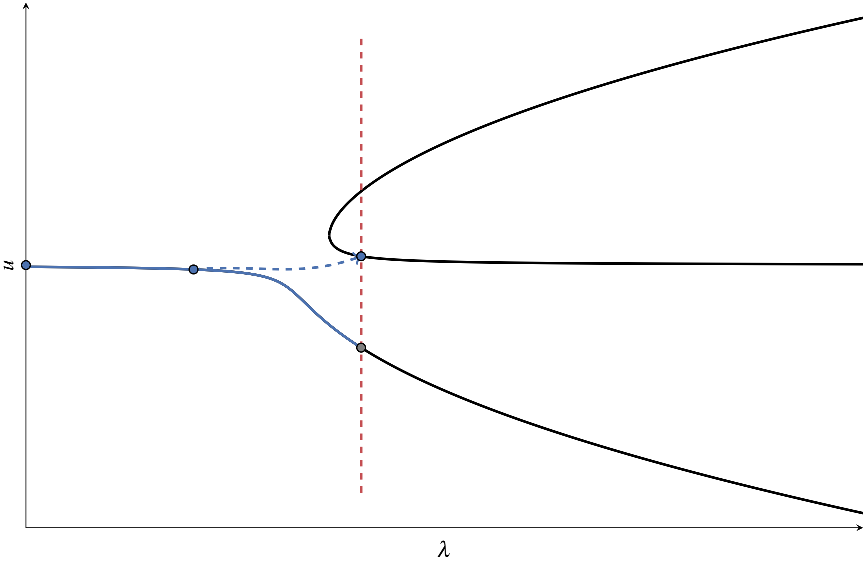
Perform a continuation step.



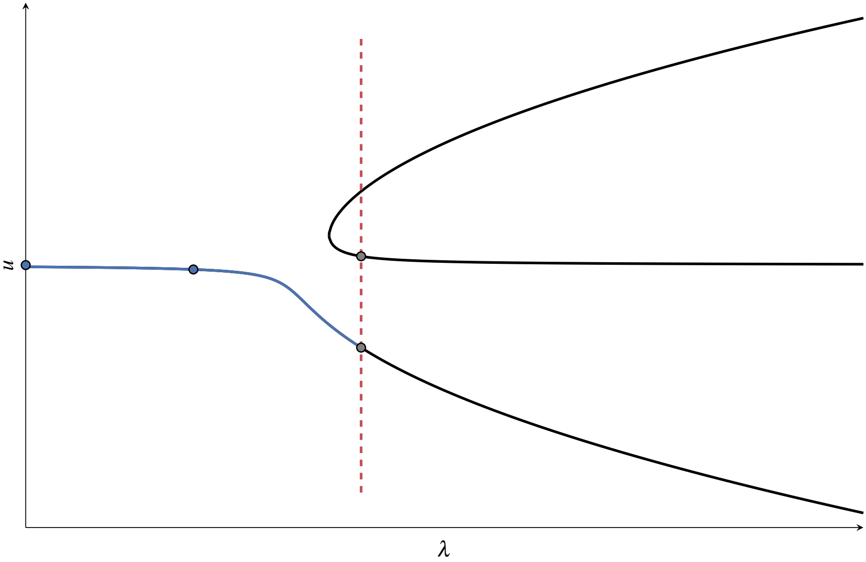
Perform another continuation step.



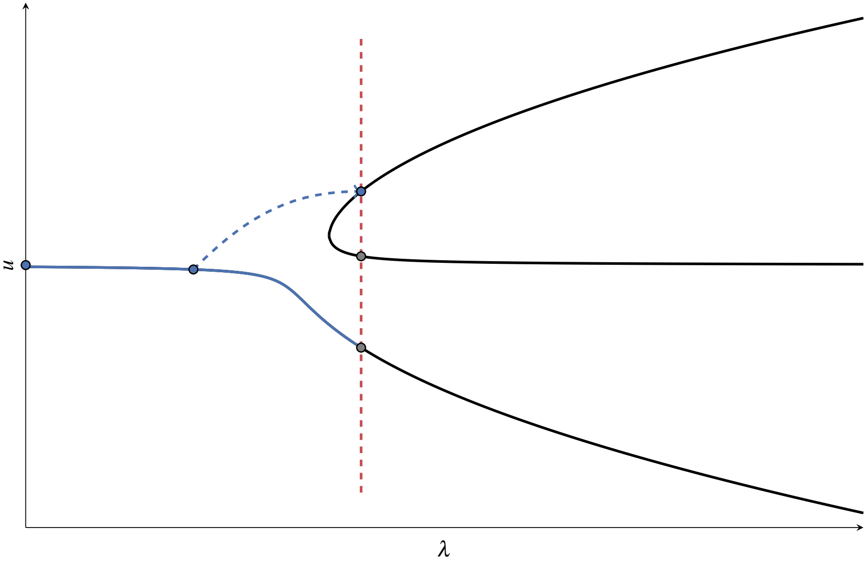
Deflate the solution found.



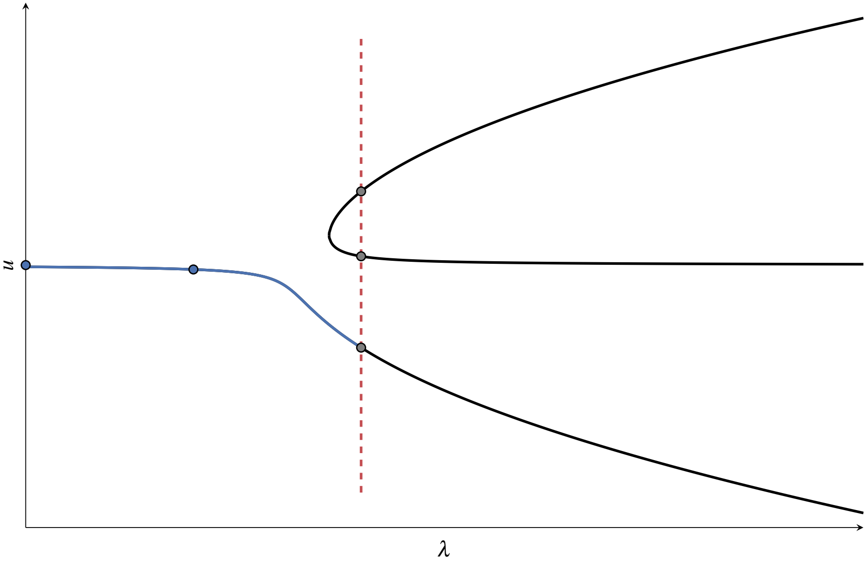
Solve again.



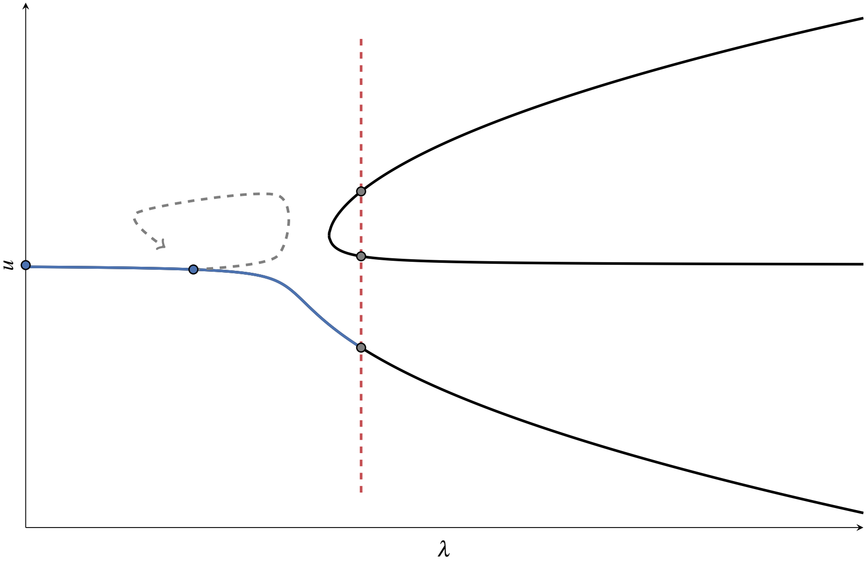
Deflate the solution found.



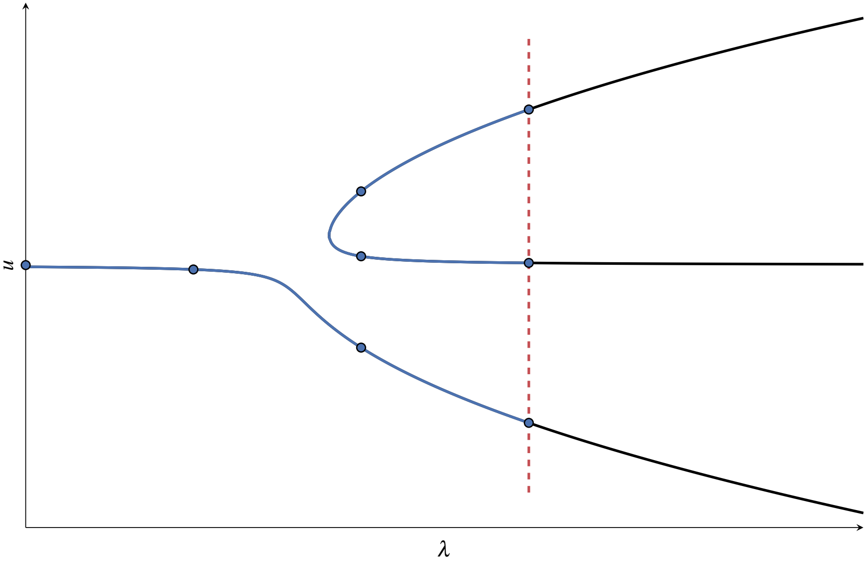
Solve again.



Deflate the solution found.



Search again, unsuccessfully.



Repeat.

Good news

Deflation lets us discover disconnected branches!

Section 5

Solving the deflated problem

We assume we have a good solver for our discretised Newton step

$$F_u(u, \lambda)\delta u_F = -F(u, \lambda), \quad F \in C^1(\mathbb{R}^N \times \mathbb{R}, \mathbb{R}^N).$$

We assume we have a good solver for our discretised Newton step

$$F_u(u,\lambda)\delta u_F = -F(u,\lambda), \quad F \in C^1(\mathbb{R}^N \times \mathbb{R}, \mathbb{R}^N).$$

We now want to solve

$$G_u(u,\lambda)\delta u_G = -G(u,\lambda)$$

where

$$G(u,\lambda) = M(u;u_1)M(u;u_2)\cdots M(u;u_n)F(u,\lambda) =: M(u)F(u,\lambda).$$

We assume we have a good solver for our discretised Newton step

$$F_u(u,\lambda)\delta u_F = -F(u,\lambda), \quad F \in C^1(\mathbb{R}^N \times \mathbb{R}, \mathbb{R}^N).$$

We now want to solve

$$G_u(u,\lambda)\delta u_G = -G(u,\lambda)$$

where

$$G(u,\lambda) = M(u;u_1)M(u;u_2)\cdots M(u;u_n)F(u,\lambda) =: M(u)F(u,\lambda).$$

Good news

You can compute δu_G easily from δu_F! (Sherman–Morrison)

By the product rule,

$$G_u(u, \lambda) = M(u)F_u(u, \lambda) + F(u, \lambda)M_u^\top.$$

By the product rule,

$$G_u(u,\lambda) = M(u)F_u(u,\lambda) + F(u,\lambda)M_u^\top.$$

By the product rule,

$$G_u(u,\lambda) = M(u)F_u(u,\lambda) + F(u,\lambda)M_u^\top.$$

At first this looks bad. The deflated Jacobian is dense, as it is a rank-one update of a sparse matrix.

By the product rule,

$$G_u(u, \lambda) = M(u) F_u(u, \lambda) + F(u, \lambda) M_u^\top.$$

At first this looks bad. The deflated Jacobian is dense, as it is a rank-one update of a sparse matrix.

Sherman–Morrison–Woodbury formula

$$\left(A + uv^\top\right)^{-1} = A^{-1} - \left(\frac{A^{-1} u v^\top A^{-1}}{1 + v^\top A^{-1} u}\right).$$

Maurice Bartlett

By the product rule,

$$G_u(u, \lambda) = M(u)F_u(u, \lambda) + F(u, \lambda)M_u^\top.$$

At first this looks bad. The deflated Jacobian is dense, as it is a rank-one update of a sparse matrix.

Sherman–Morrison–Woodbury formula

$$\left(A + uv^\top\right)^{-1} = A^{-1} - \left(\frac{A^{-1}uv^\top A^{-1}}{1 + v^\top A^{-1}u}\right).$$

Maurice Bartlett

At first it looks like applying this to a vector w requires two solves with A: $A^{-1}u$ and $A^{-1}w$. But something magical happens . . .

Applying the Sherman–Morrison–Woodbury formula, we have

$$\delta u_G = -[G_u]^{-1} G = -\left(M F_u + F M_u^\top \right)^{-1} (MF)$$

Applying the Sherman–Morrison–Woodbury formula, we have

$$\begin{aligned}
\delta u_G = -[G_u]^{-1} G &= -\left(MF_u + FM_u^\top\right)^{-1}(MF) \\
&= -\left[M^{-1}F_u^{-1} - \frac{M^{-1}F_u^{-1}FM_u^\top M^{-1}F_u^{-1}}{1 + M_u^\top M^{-1}F_u^{-1}F}\right](MF)
\end{aligned}$$

Applying the Sherman–Morrison–Woodbury formula, we have

$$
\begin{aligned}
\delta u_G = -[G_u]^{-1} G &= -\left(M F_u + F M_u^\top \right)^{-1} (MF) \\
&= -\left[M^{-1} F_u^{-1} - \frac{M^{-1} F_u^{-1} F M_u^\top M^{-1} F_u^{-1}}{1 + M_u^\top M^{-1} F_u^{-1} F} \right] (MF) \\
&= -F_u^{-1} F + \frac{F_u^{-1} F M_u^\top M^{-1} F_u^{-1} F}{1 + M_u^\top M^{-1} F_u^{-1} F}
\end{aligned}
$$

Applying the Sherman–Morrison–Woodbury formula, we have

$$
\begin{aligned}
\delta u_G = -[G_u]^{-1}G &= -\left(MF_u + FM_u^\top\right)^{-1}(MF) \\
&= -\left[M^{-1}F_u^{-1} - \frac{M^{-1}F_u^{-1}FM_u^\top M^{-1}F_u^{-1}}{1 + M_u^\top M^{-1}F_u^{-1}F}\right](MF) \\
&= -F_u^{-1}F + \frac{F_u^{-1}FM_u^\top M^{-1}F_u^{-1}F}{1 + M_u^\top M^{-1}F_u^{-1}F} \\
&= \left(1 - \frac{M^{-1}M_u^\top F_u^{-1}F}{1 + M_u^\top M^{-1}F_u^{-1}F}\right)\left(-F_u^{-1}F\right)
\end{aligned}
$$

Applying the Sherman–Morrison–Woodbury formula, we have

$$
\begin{aligned}
\delta u_G = -[G_u]^{-1} G &= -\left(M F_u + F M_u^\top\right)^{-1} (MF) \\
&= -\left[M^{-1} F_u^{-1} - \frac{M^{-1} F_u^{-1} F M_u^\top M^{-1} F_u^{-1}}{1 + M_u^\top M^{-1} F_u^{-1} F}\right] (MF) \\
&= -F_u^{-1} F + \frac{F_u^{-1} F M_u^\top M^{-1} F_u^{-1} F}{1 + M_u^\top M^{-1} F_u^{-1} F} \\
&= \left(1 - \frac{M^{-1} M_u^\top F_u^{-1} F}{1 + M_u^\top M^{-1} F_u^{-1} F}\right) \left(-F_u^{-1} F\right) \\
&= \left(1 + \frac{M^{-1} M_u^\top \delta u_F}{1 - M^{-1} M_u^\top \delta u_F}\right) \delta u_F.
\end{aligned}
$$

Applying the Sherman–Morrison–Woodbury formula, we have

$$
\begin{aligned}
\delta u_G = -[G_u]^{-1} G &= -\left(MF_u + FM_u^\top\right)^{-1}(MF) \\
&= -\left[M^{-1}F_u^{-1} - \frac{M^{-1}F_u^{-1}FM_u^\top M^{-1}F_u^{-1}}{1 + M_u^\top M^{-1}F_u^{-1}F}\right](MF) \\
&= -F_u^{-1}F + \frac{F_u^{-1}FM_u^\top M^{-1}F_u^{-1}F}{1 + M_u^\top M^{-1}F_u^{-1}F} \\
&= \left(1 - \frac{M^{-1}M_u^\top F_u^{-1}F}{1 + M_u^\top M^{-1}F_u^{-1}F}\right)\left(-F_u^{-1}F\right) \\
&= \left(1 + \frac{M^{-1}M_u^\top \delta u_F}{1 - M^{-1}M_u^\top \delta u_F}\right)\delta u_F.
\end{aligned}
$$

So we only need to solve one system with F_u!

Solving the deflated problem

To solve

$$G_u \delta u_G = -G,$$

do the following:

Solving the deflated problem

To solve

$$G_u \delta u_G = -G,$$

do the following:

1. Solve

$$F_u \delta u_F = -F.$$

Solving the deflated problem

To solve

$$G_u \delta u_G = -G,$$

do the following:

1. Solve
$$F_u \delta u_F = -F.$$
2. Evaluate
$$p = M_u^\top \delta u_F.$$

Solving the deflated problem

To solve

$$G_u \delta u_G = -G,$$

do the following:

1. Solve
$$F_u \delta u_F = -F.$$
2. Evaluate
$$p = M_u^\top \delta u_F.$$
3. Evaluate
$$\tau = 1 + \frac{M^{-1} p}{1 - M^{-1} p}.$$

Solving the deflated problem

To solve

$$G_u \delta u_G = -G,$$

do the following:

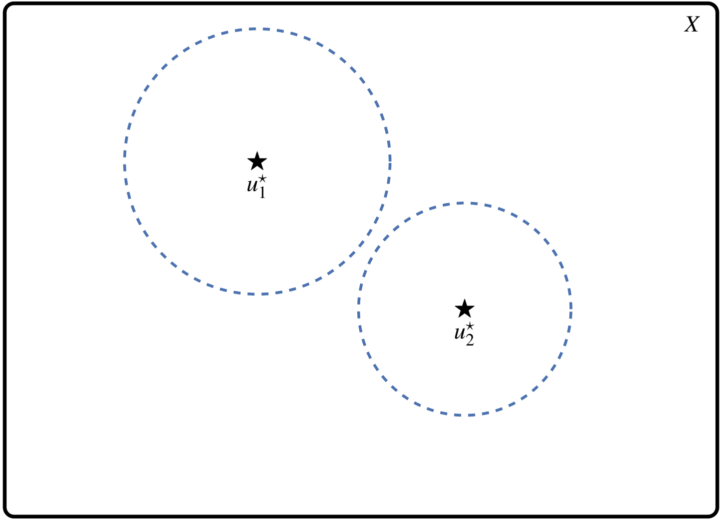
1. Solve
$$F_u \delta u_F = -F.$$
2. Evaluate
$$p = M_u^\top \delta u_F.$$
3. Evaluate
$$\tau = 1 + \frac{M^{-1} p}{1 - M^{-1} p}.$$
4. Return
$$\delta u_G = \tau \delta u_F.$$

Section 6

Convergence of deflation

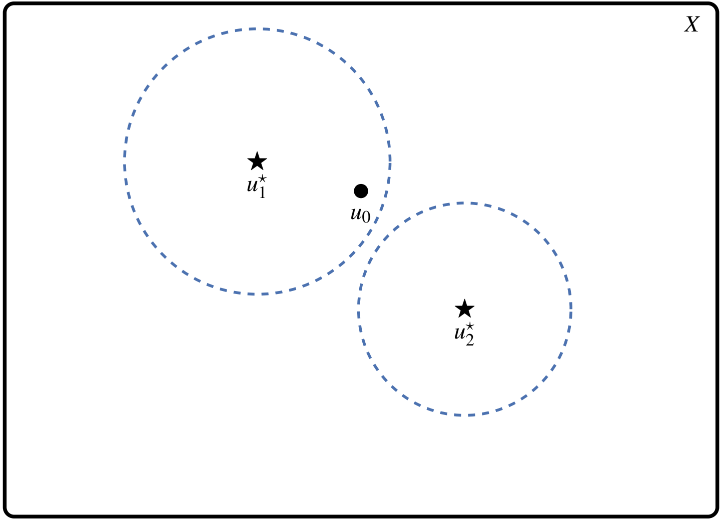
It is possible to give sufficient conditions for deflation to find two roots.

It is possible to give sufficient conditions for deflation to find two roots.



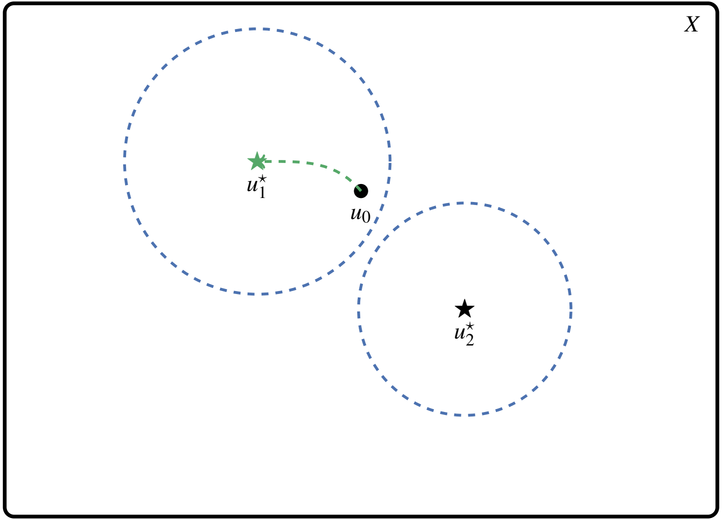
Two solutions, with Rall–Rheinboldt balls.

It is possible to give sufficient conditions for deflation to find two roots.



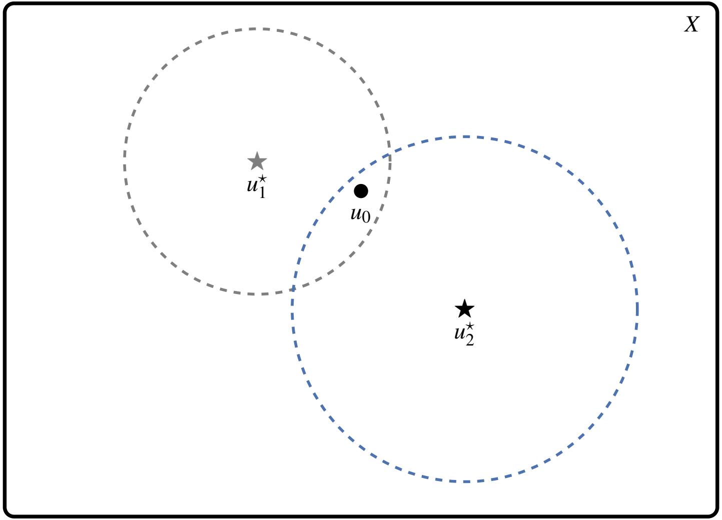
Start with an initial guess within a ball.

It is possible to give sufficient conditions for deflation to find two roots.



Converge to that solution.

It is possible to give sufficient conditions for deflation to find two roots.



Deflate that solution; the other Rall–Rheinboldt ball expands.

Section 7

Examples

Allen–Cahn equation

$$F(u, \lambda) = -\lambda^2 \nabla^2 u + u^3 - u = 0, \quad u = g \text{ on } \partial\Omega.$$

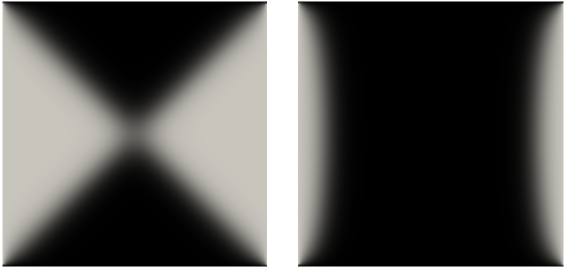
Allen–Cahn equation

$$F(u, \lambda) = -\lambda^2 \nabla^2 u + u^3 - u = 0, \quad u = g \text{ on } \partial\Omega.$$

Solutions found starting from $u = 0$ for $\lambda = 0.04$.

Allen–Cahn equation

$$F(u, \lambda) = -\lambda^2 \nabla^2 u + u^3 - u = 0, \quad u = g \text{ on } \partial\Omega.$$



Solutions found starting from $u = 0$ for $\lambda = 0.04$.

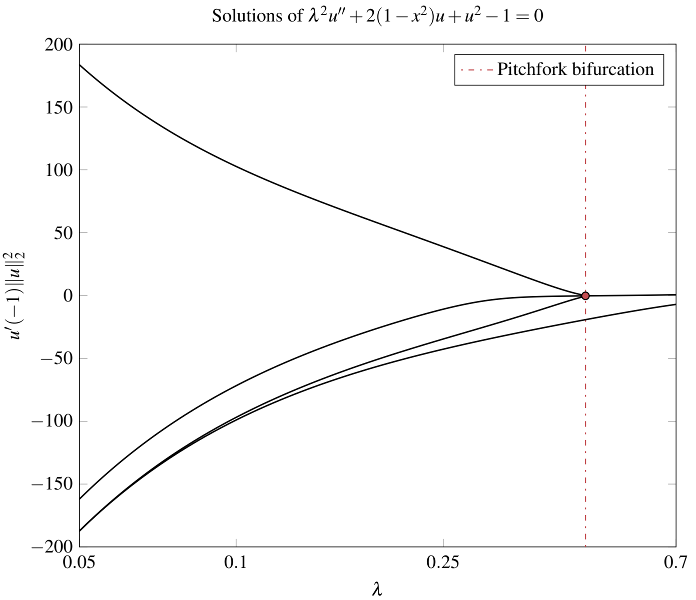
Allen–Cahn equation

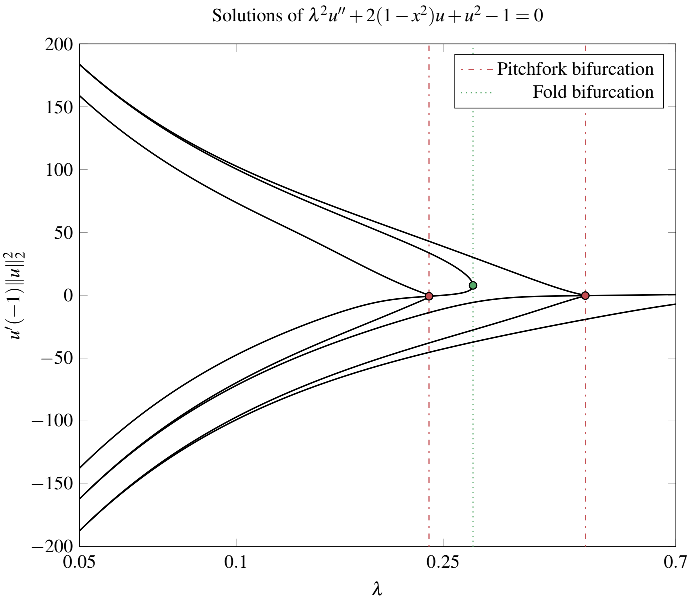
$$F(u, \lambda) = -\lambda^2 \nabla^2 u + u^3 - u = 0, \quad u = g \text{ on } \partial\Omega.$$

Solutions found starting from $u = 0$ for $\lambda = 0.04$.

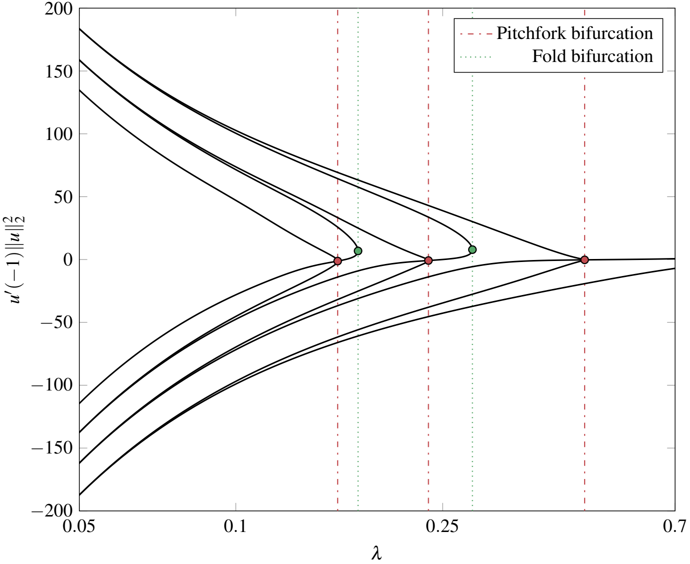
Carrier's equation

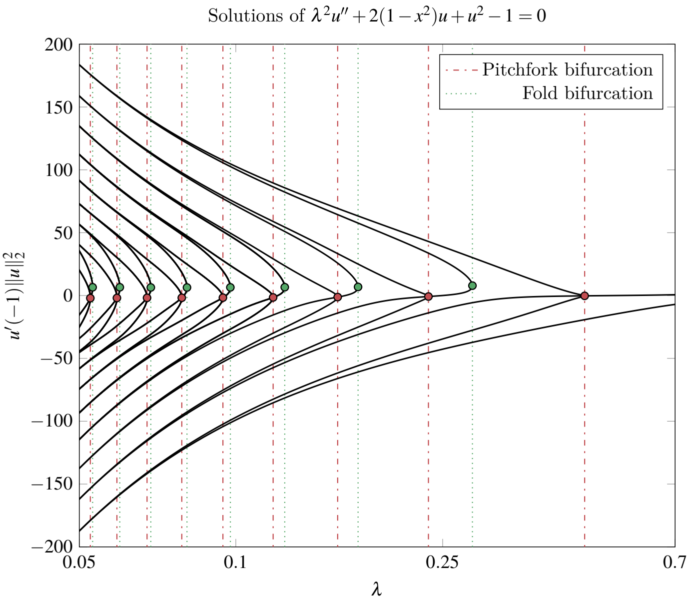
$$F(u,\lambda) = \lambda^2 u'' + 2(1-x^2)u + u^2 - 1 = 0, \quad u(-1) = 0 = u(1).$$


Solutions of $\lambda^2 u'' + 2(1-x^2)u + u^2 - 1 = 0$
Pitchfork bifurcation
$u'(-1)\|u\|_2^2$
200
150
100
50
0
-50
-100
-150
-200
0.05
0.1
0.25
0.7
λ


Solutions of $\lambda^2 u'' + 2(1-x^2)u + u^2 - 1 = 0$
200
150
100
50
0
−50
−100
−150
−200
0.05
0.1
0.25
0.7
λ
$u'(-1)\|u\|_2^2$
Pitchfork bifurcation
Fold bifurcation

Solutions of $\lambda^2 u'' + 2(1-x^2)u + u^2 - 1 = 0$




Solutions of $\lambda^2 u'' + 2(1-x^2)u + u^2 - 1 = 0$
Pitchfork bifurcation
Fold bifurcation
$u'(-1)\,\|u\|_2^2$
λ
200
150
100
50
0
−50
−100
−150
−200
0.05
0.1
0.25
0.7

Section 8

Symmetries

Symmetries

What if the equation has a continuous symmetry group?

Symmetries

What if the equation has a continuous symmetry group?

Philosophy

The fundamental structures are the distinct **orbits** of solutions.

Symmetries

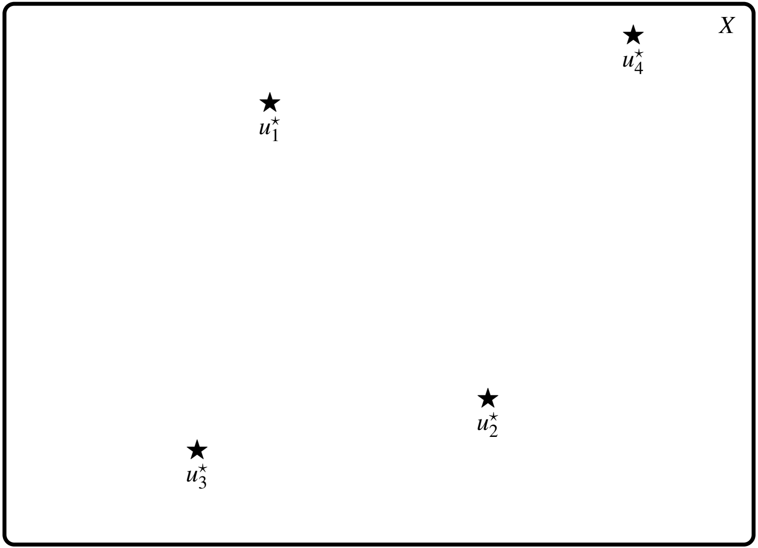
What if the equation has a continuous symmetry group?

Philosophy

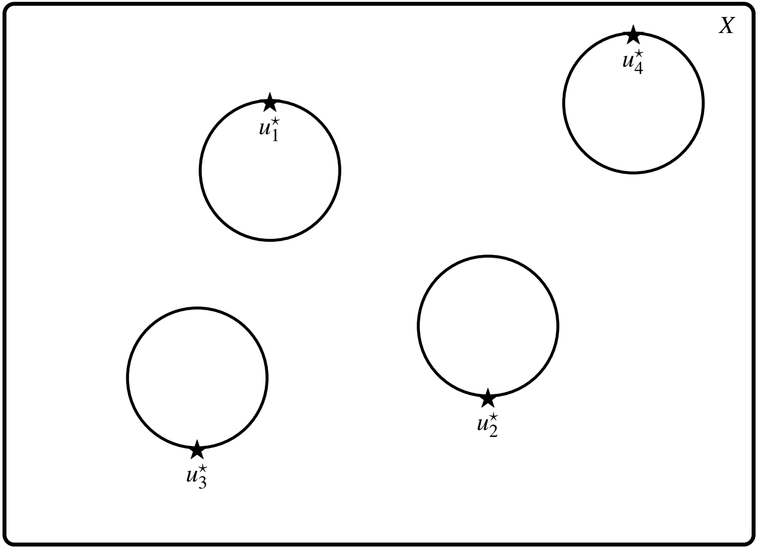
The fundamental structures are the distinct **orbits** of solutions.

Key idea

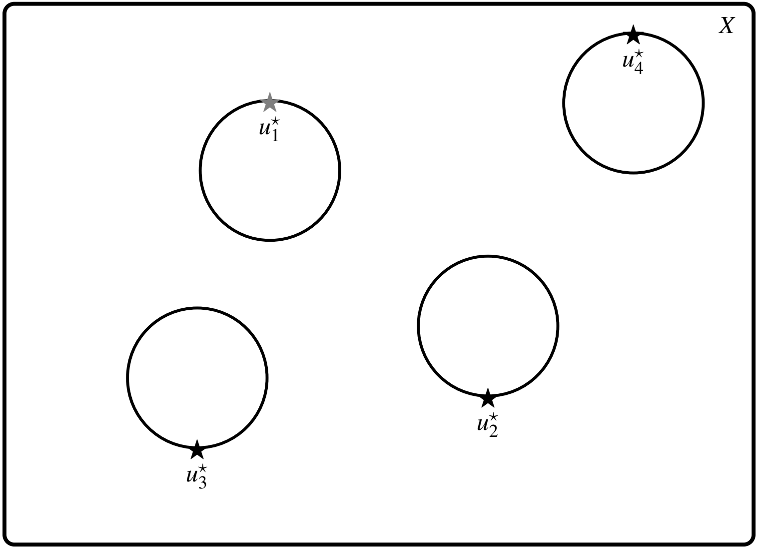
Construct a deflation operator invariant under the action of the Lie group.



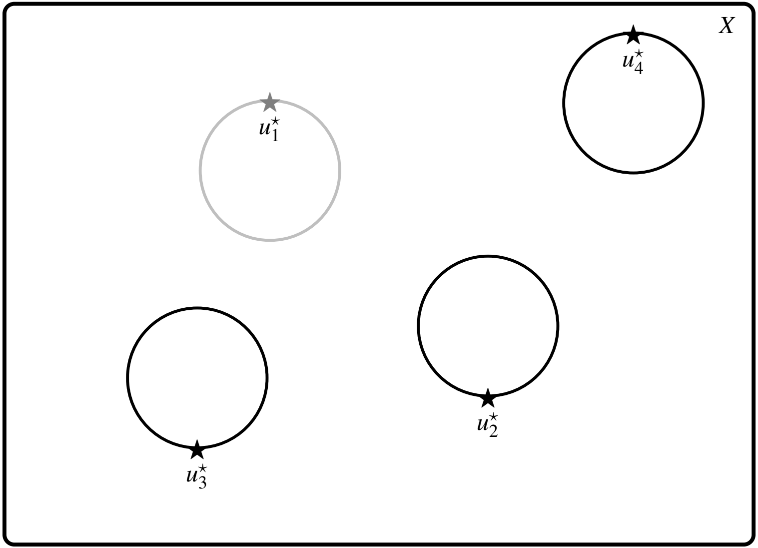
Four solutions, not related by the symmetry group.



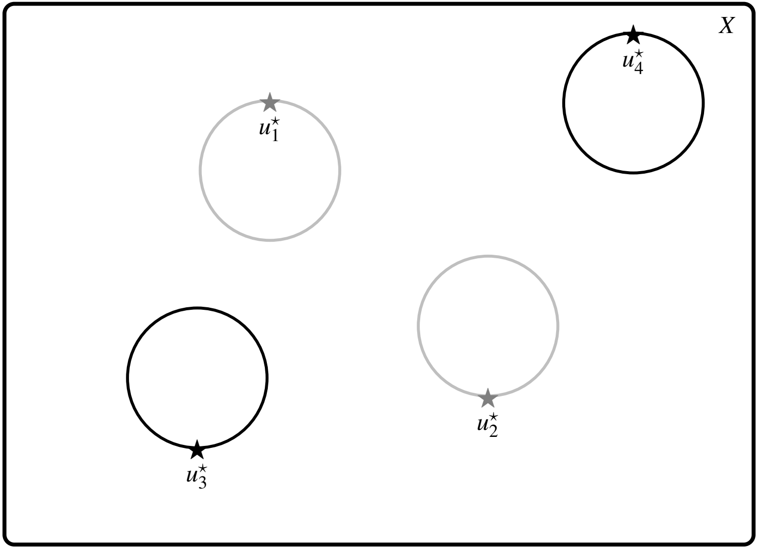
Each solution induces a *group orbit* of solutions, related by symmetry.



Not enough to deflate the solution—must deflate the entire orbit.



Design a deflation operator that deflates the entire orbit.



Design a deflation operator that deflates the entire orbit.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

First symmetry group $\mathrm{SO}(2)$: phase shifts

$$u(\vec{x}) \mapsto e^{i\theta}u(\vec{x}), \quad \theta \in \mathbb{R}.$$

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

First symmetry group $\mathrm{SO}(2)$: phase shifts

$$u(\vec{x}) \mapsto e^{i\theta} u(\vec{x}), \quad \theta \in \mathbb{R}.$$

Invariant deflation operator

$$M(u;r) = \left\| |u|^2 - |r|^2 \right\|^{-2} + 1.$$

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Second symmetry group $\mathrm{SO}(3)$: spatial rotations

$$u(\vec{x}) \mapsto u(R\vec{x}), \quad R^{-1} = R^T, \ \det(R) = 1.$$

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2 + y^2 + z^2}{2} u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Second symmetry group $\mathrm{SO}(3)$: spatial rotations

$$u(\vec{x}) \mapsto u(R\vec{x}), \quad R^{-1} = R^T, \;\; \det(R) = 1.$$

Invariant deflation operator

$$M(u; r) = \|\bar{u} - \bar{r}\|^{-2} + 1,$$

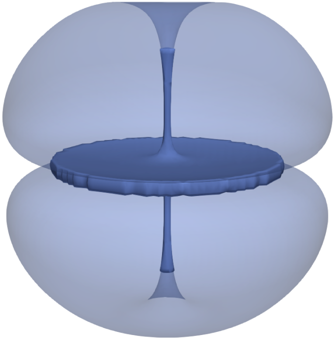
where

$\bar{u}(r, \theta, \psi)$ averages u over the sphere of radius r.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

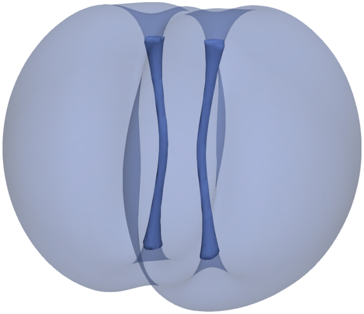


A vortex line and a planar dark soliton.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

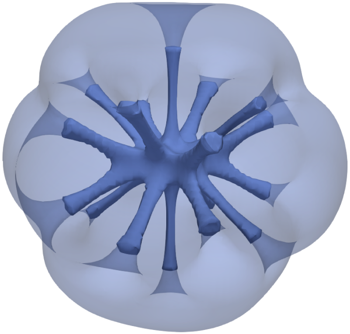


A pair of vortex lines.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

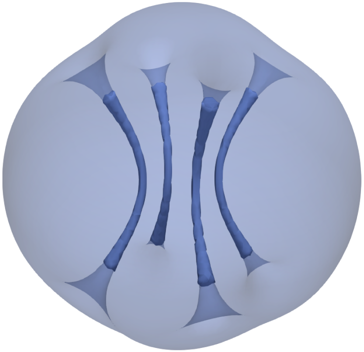


A vortex star.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

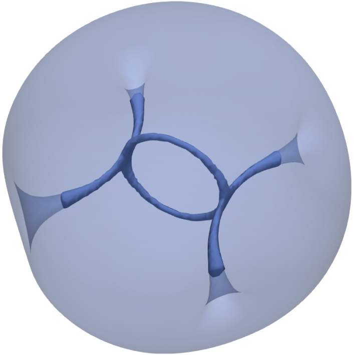


Four vortex lines of alternating charge.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

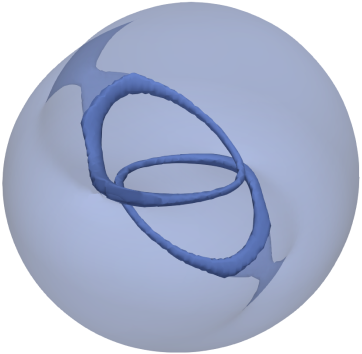


A vortex ring with two “handles”.

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

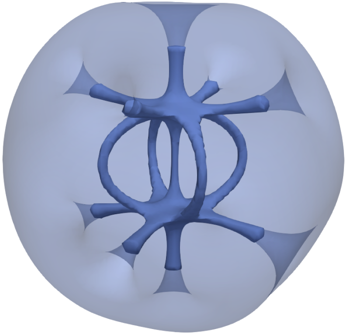


Two bent vortex rings?

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.

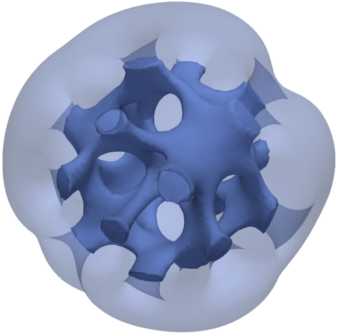


Two vortex rings and five lines?

Gross–Pitaevskii equation

$$-\frac{1}{2}\Delta u + \frac{x^2+y^2+z^2}{2}u - \mu u + |u|^2 u = 0, \qquad u|_{\partial\Omega} = 0.$$

Solutions for $\mu = 6$.



A vortex ring cage?

Section 9

Semismooth problems

Many problems feature inequality constraints.

Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$

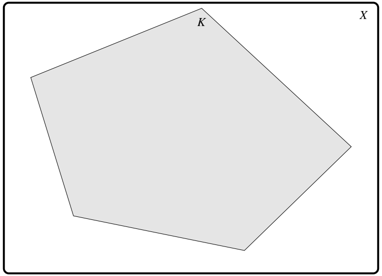
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



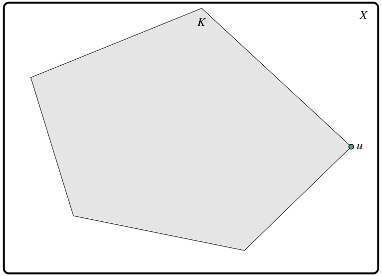
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$.
The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



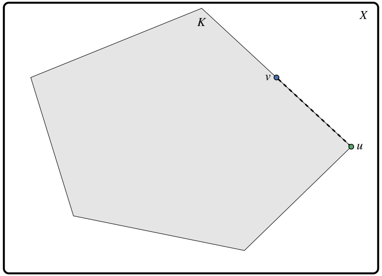
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



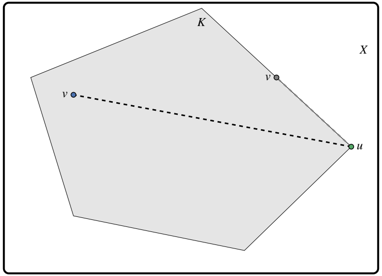
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



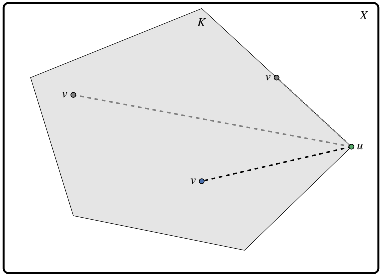
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



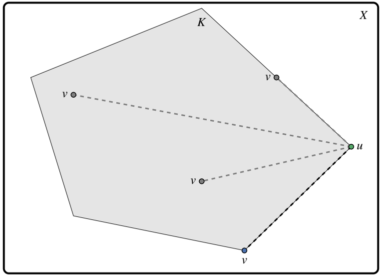
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$.
The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$



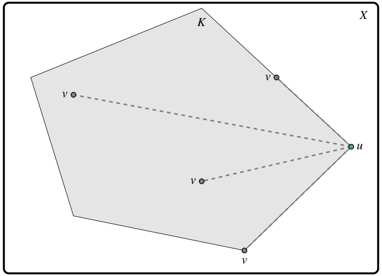
Many problems feature inequality constraints.

The natural language for formulating these is as a *variational inequality*.

$\mathrm{VI}(Q, K)$

Let X be a real reflexive Banach space, $K \subset X$ a closed convex subset, and $Q : K \to X^*$. The task is to

$$\text{find } u^\star \in K \text{ such that } \langle Q(u^\star), v - u^\star \rangle \geq 0 \text{ for all } v \in K.$$

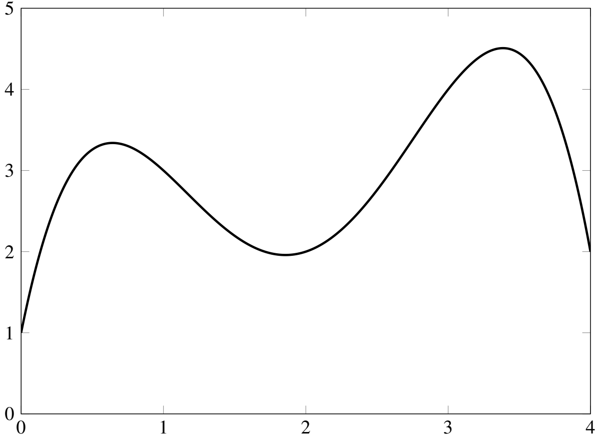


For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$

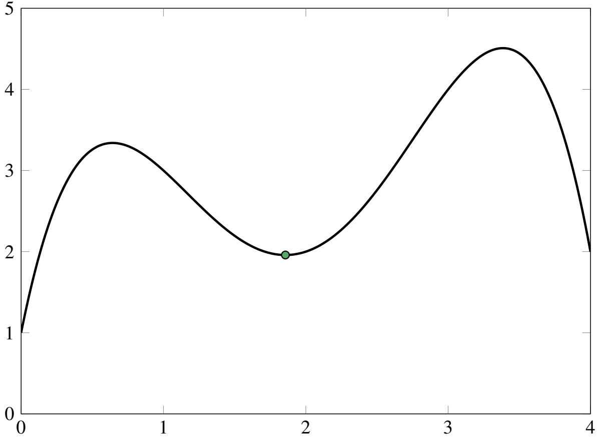
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



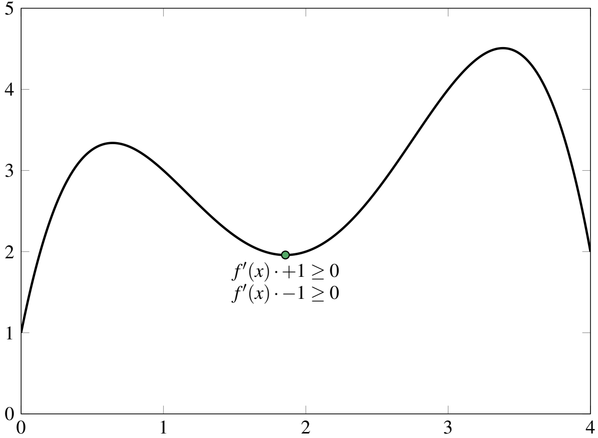
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



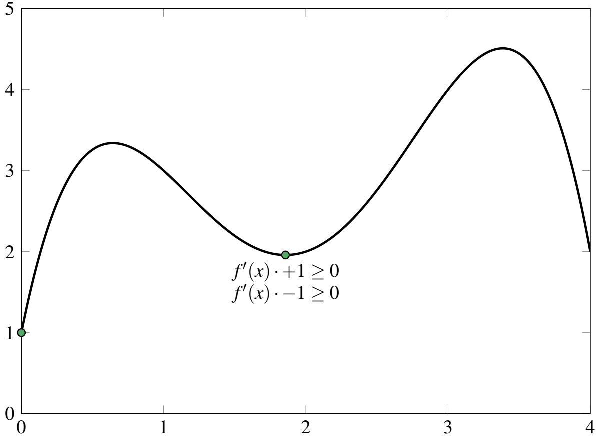
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



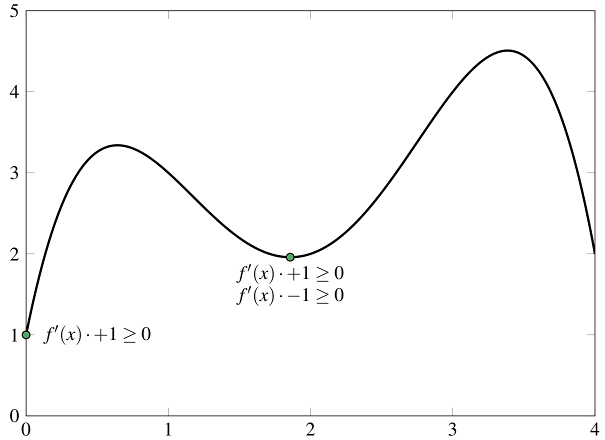
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



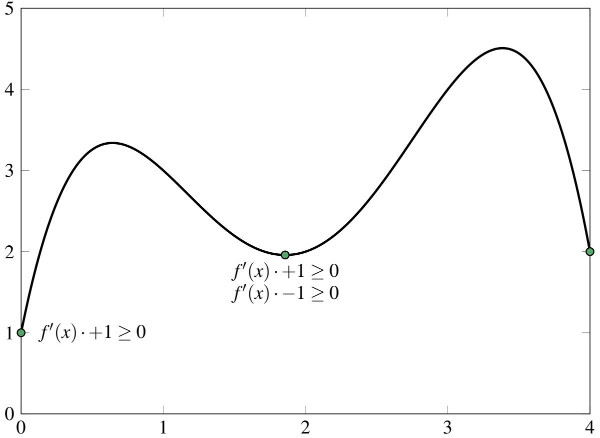
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



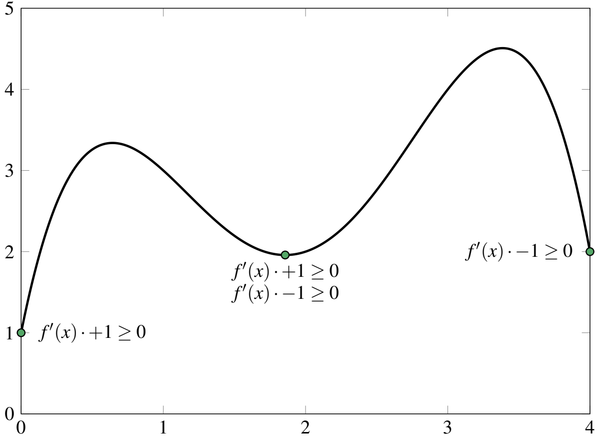
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



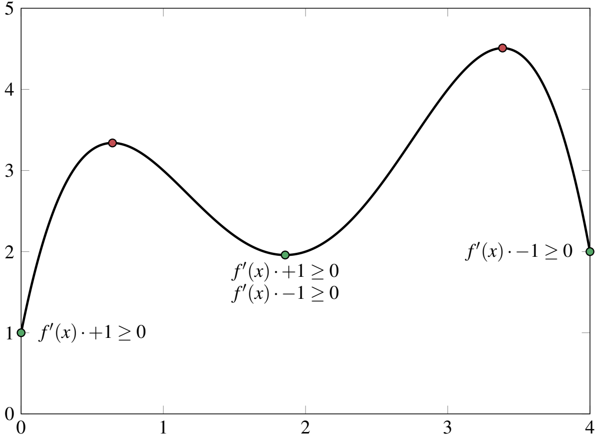
For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



For example, if you want to minimise $f \in C^1(\mathbb{R}, \mathbb{R})$ over a closed interval $I \subset \mathbb{R}$, the necessary optimality condition is

$$\mathrm{VI}(f', I).$$



The main way of solving variational inequalities is to reformulate them as a system of equations.

The main way of solving variational inequalities is to reformulate them as a system of equations.

For example, $\mathrm{VI}(Q, K)$ with

$$K = \{x \in \mathbb{R} : x \geq 0\}$$

is equivalent to

$$S(x) \coloneqq \sqrt{x^2 + [Q(x)]^2} - x - Q(x) = 0.$$

The main way of solving variational inequalities is to reformulate them as a system of equations.

For example, $\mathrm{VI}(Q, K)$ with

$$K = \{x \in \mathbb{R} : x \geq 0\}$$

is equivalent to

$$S(x) := \sqrt{x^2 + [Q(x)]^2} - x - Q(x) = 0.$$

The price we pay . . .

. . . is that S is not smooth.

Good news

S is *just smooth enough* to define a Newton-type method with superlinear convergence.

Michael Hintermüller

Michael Ulbrich

Good news

S is *just smooth enough* to define a Newton-type method with superlinear convergence.

Semismooth Newton works just like normal:

$$u_{i+1} = u_i - [H(u_i)]^{-1} S(u_i),$$

where H is the so-called Newton derivative.

This algorithm usually converges superlinearly.

Michael Hintermüller

Michael Ulbrich

Good news

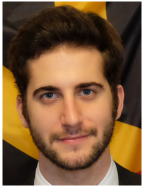
Deflation works for semismooth problems.

Good news

Deflation works for semismooth problems.

Theorem (F., Croci, Surowiec, 2020)

Under the same assumptions that are required for superlinear convergence of semismooth Newton, deflation works the same.



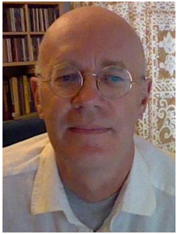
Matteo Croci

Thomas Surowiec

Gould gives an example where the central path is ill-behaved:

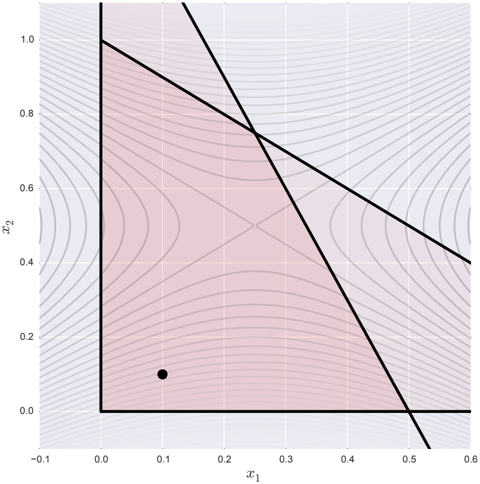
Nonconvex quadratic programming problem

$$
\begin{aligned}
\underset{x\in\mathbb{R}^2}{\text{minimise}} \quad & -2(x_1-0.25)^2+2(x_2-0.5)^2 \\
\text{subject to} \quad & x_1+x_2 \le 1 \\
& 3x_1+x_2 \le 1.5 \\
& x_1 \ge 0 \\
& x_2 \ge 0
\end{aligned}
$$



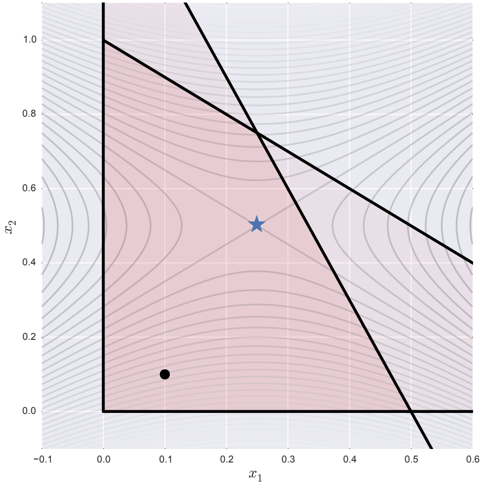
Nick Gould

Gould gives an example where the central path is ill-behaved:



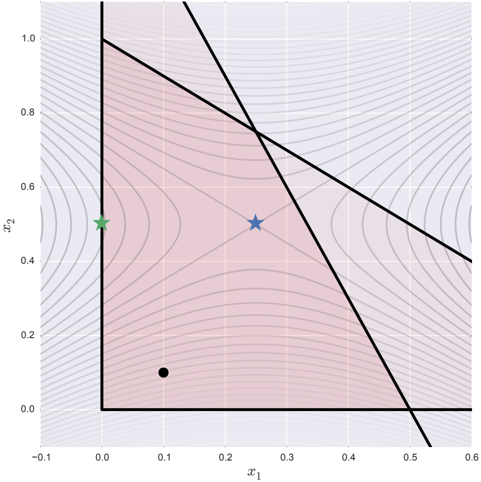
Deflation finds both minima and the saddle point.

Gould gives an example where the central path is ill-behaved:



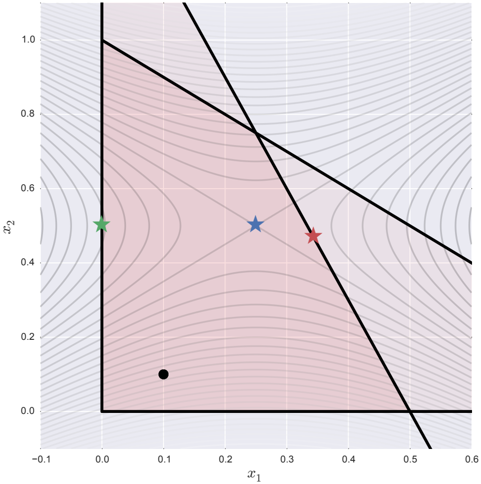
Deflation finds both minima and the saddle point.

Gould gives an example where the central path is ill-behaved:



Deflation finds both minima and the saddle point.

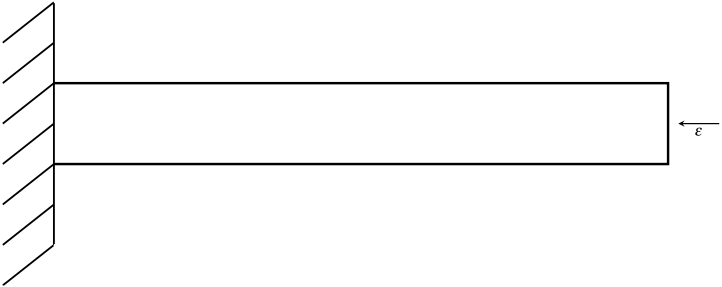
Gould gives an example where the central path is ill-behaved:



Deflation finds both minima and the saddle point.

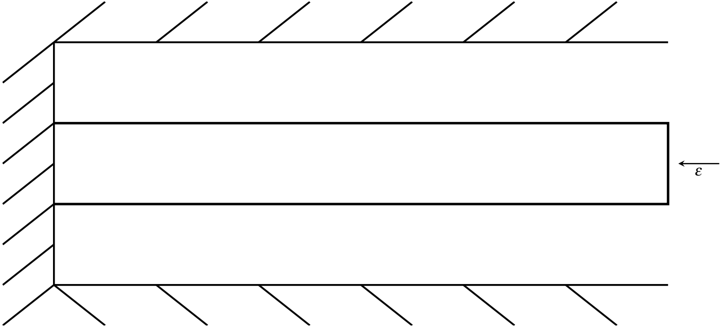
Buckling of a hyperelastic beam with contact constraints

$$\begin{aligned}\underset{u \in H^1(\Omega;\mathbb{R}^2)}{\text{minimise}} \quad & \Pi(u) = \int_\Omega \psi(u)\ \mathrm{d}x - \int_\Omega B \cdot u\ \mathrm{d}x \\ \text{subject to} \quad & u|_{\text{left}} = (0,0),\ \ u|_{\text{right}} = (-\varepsilon, 0), \\ & \operatorname{tr}(u_y) \in [a,b] \text{ a.e. in } \Gamma_{\text{top}}, \Gamma_{\text{bottom}}.\end{aligned}$$



Buckling of a hyperelastic beam with contact constraints

$$\begin{aligned} \underset{u \in H^1(\Omega;\mathbb{R}^2)}{\text{minimise}} \quad & \Pi(u) = \int_\Omega \psi(u) \, \mathrm{d}x - \int_\Omega B \cdot u \, \mathrm{d}x \\ \text{subject to} \quad & u|_{\text{left}} = (0,0), \;\; u|_{\text{right}} = (-\varepsilon, 0), \\ & \mathrm{tr}(u_y) \in [a,b] \text{ a.e. in } \Gamma_{\text{top}}, \Gamma_{\text{bottom}}. \end{aligned}$$



Buckling of a hyperelastic beam with contact constraints

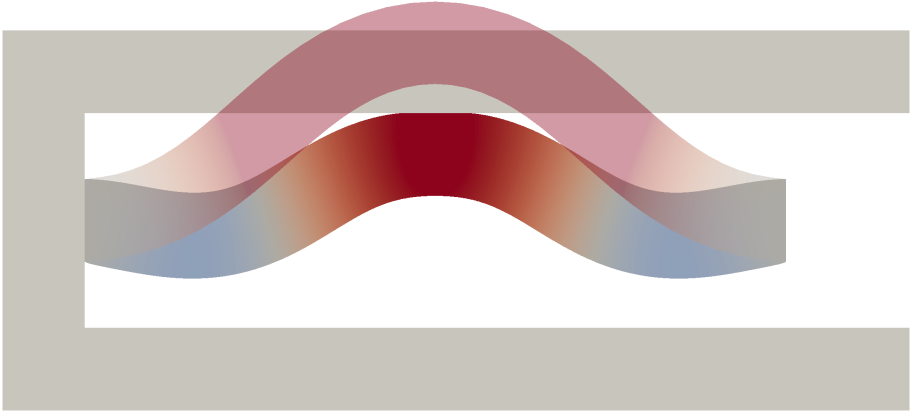
$$
\begin{aligned}
\underset{u \in H^1(\Omega;\mathbb{R}^2)}{\text{minimise}} \quad & \Pi(u) = \int_\Omega \psi(u)\ \mathrm{d}x - \int_\Omega B \cdot u\ \mathrm{d}x \\
\text{subject to} \quad & u|_{\text{left}} = (0,0),\ \ u|_{\text{right}} = (-\varepsilon, 0), \\
& \operatorname{tr}(u_y) \in [a,b] \text{ a.e. in } \Gamma_{\text{top}}, \Gamma_{\text{bottom}}.
\end{aligned}
$$

Neo–Hookean strain energy density

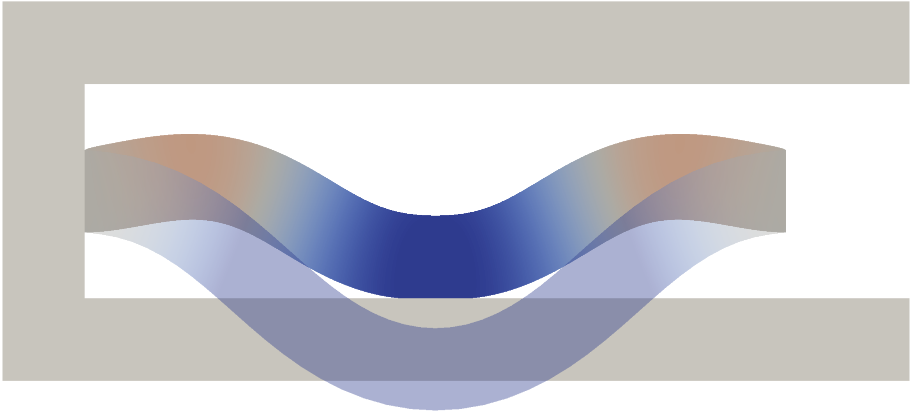
$$
\psi(u) = \frac{\mu}{2}(\operatorname{tr}(C) - 2) - \mu \log(\det(C)) + \frac{\lambda}{2}\log(\det(C))^2,
$$

where

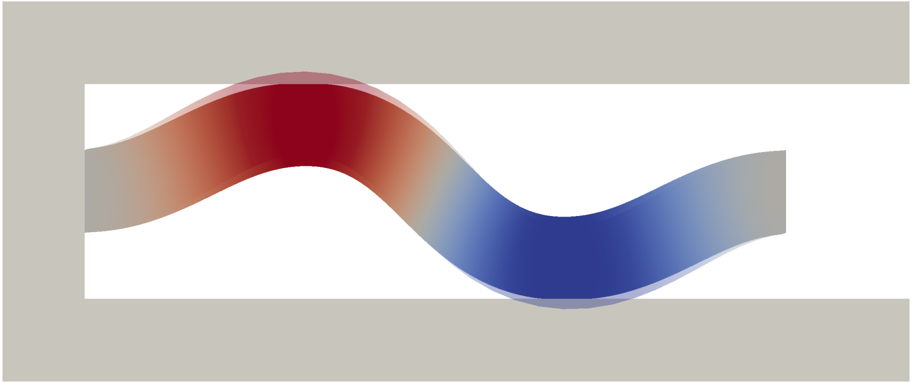
$$
C = (I + \nabla u)^\top (I + \nabla u).
$$



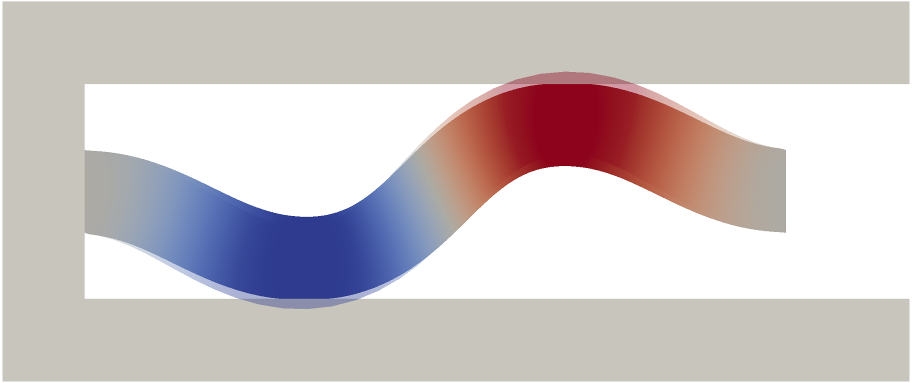
Multiple solutions of the beam with contact constraints.



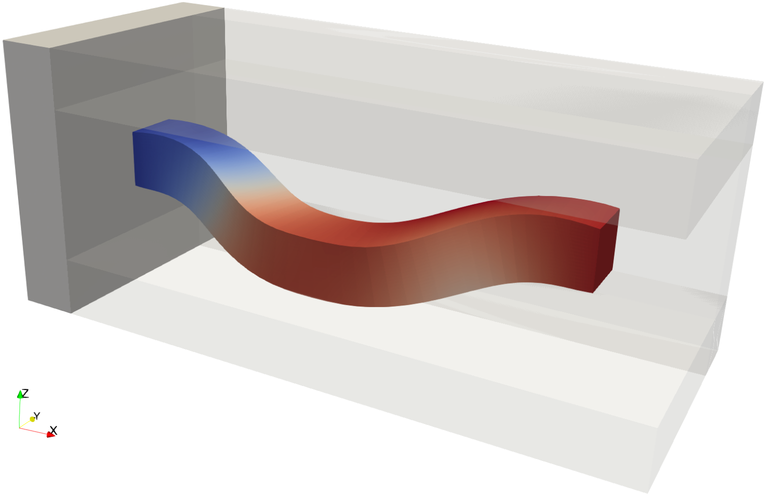
Multiple solutions of the beam with contact constraints.



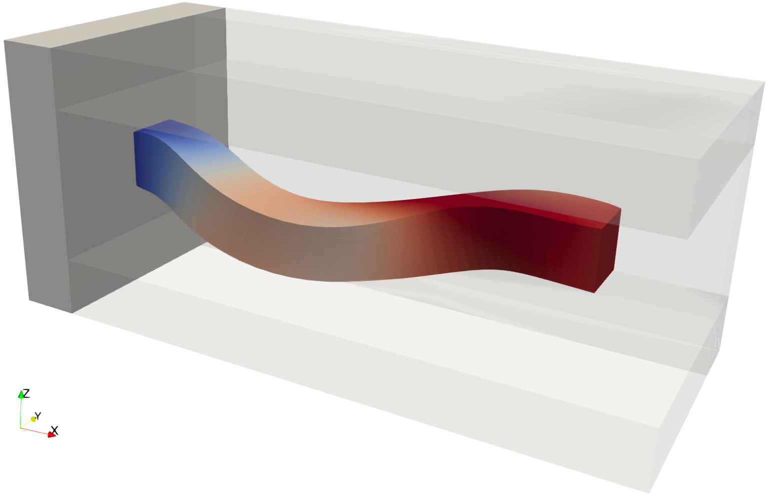
Multiple solutions of the beam with contact constraints.



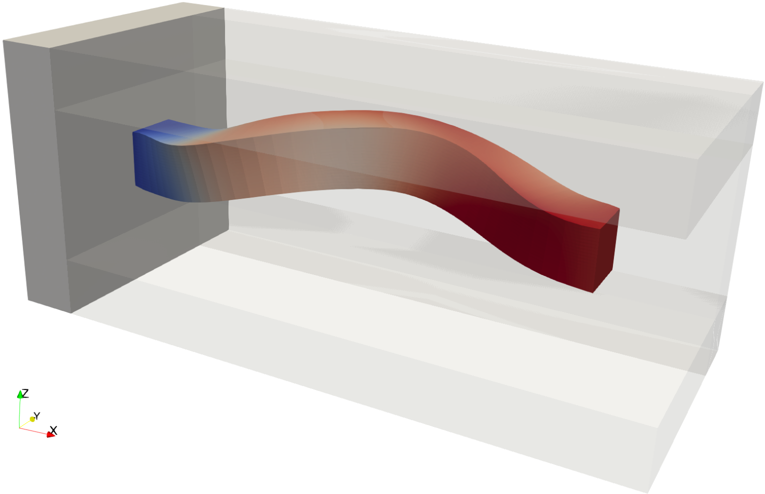
Multiple solutions of the beam with contact constraints.



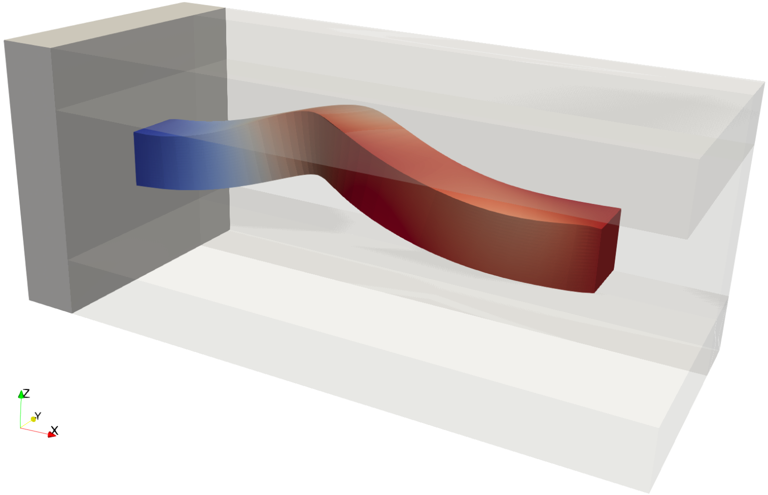
Multiple solutions of the beam with contact constraints.



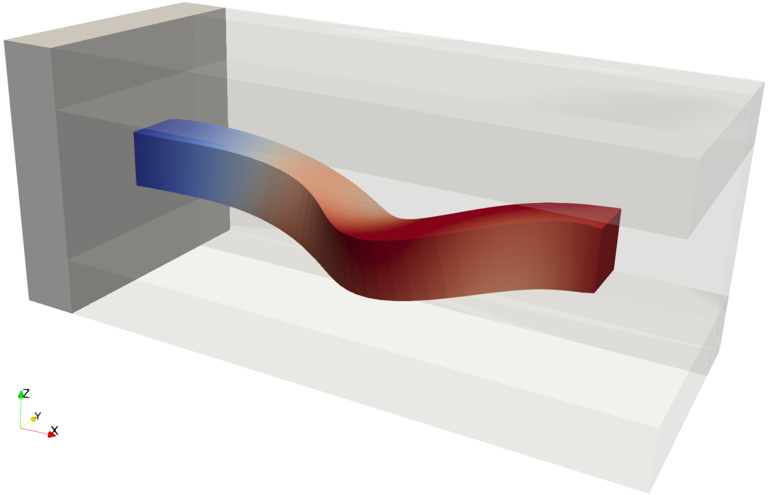
Multiple solutions of the beam with contact constraints.



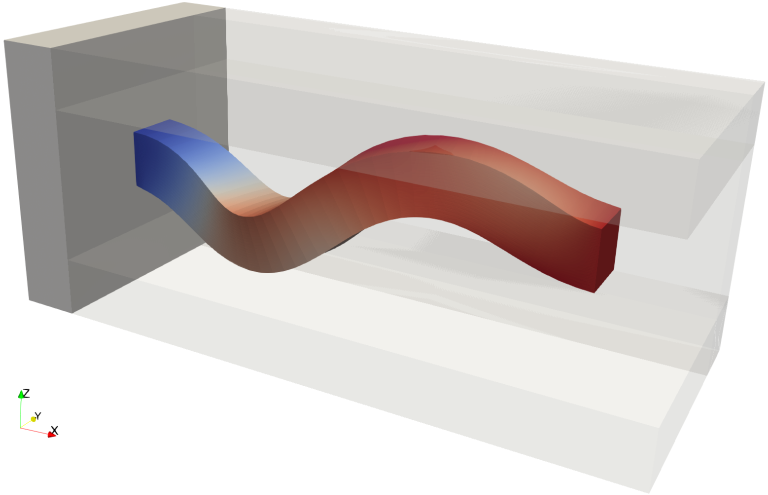
Multiple solutions of the beam with contact constraints.



Multiple solutions of the beam with contact constraints.



Multiple solutions of the beam with contact constraints.



Multiple solutions of the beam with contact constraints.

Conclusions!

Conclusions!

Main message

When solving nonlinear problems, think about multiple solutions!

Conclusions!

Main message

When solving nonlinear problems, think about multiple solutions!

Algorithms

Deflation is a very powerful technique for computing multiple solutions.