Forward Mode Automatic Differentiation with an application to solid mechanics

Andrea Vigliotti [1] Ferdinando Auricchio [2]

[1]Italian Aerospace Research Centre

[2]University of Pavia

21/01/2020

Overview

Derivatives and approximations

Automatic differentiation

The dual numbers algebra

AD through dual numbers

Application to solid mechanics

Examples

What is Automatic Differentiation ?

Before what, what for?

```
function foo(x, d0)
   for i = 1:5
      d = sqrt(x[1,i]^2+x[2,i]^2+x[3,i]^2)
      if d > d0
         phi += geez(d)
      else
         phi += baz(d)
      end
   end
   return phi
end
```

Find the derivatives of `foo` with respect to `x` without modifying the source code

What is Automatic Differentiation ?

AD is not Finite Differences

What is Automatic Differentiation ?

AD is not Finite Differences

Finite differences are based on the truncated Taylor series

$$f\left(\boldsymbol{x}+\Delta \boldsymbol{x}\right)=f_0+\frac{\partial f}{\partial x_i}\,\Delta x_i+\mathcal{O}\left(\|\Delta \boldsymbol{x}\|\right)$$

$$\frac{\partial f}{\partial x_i}=\frac{f(\boldsymbol{x}+\Delta x_i\,\boldsymbol{\imath_i})-f(\boldsymbol{x})}{\Delta x_i}+\mathcal{O}\left(|\Delta x_i|\right)$$

What is Automatic Differentiation ?

AD is not Finite Differences

Finite differences are based on the truncated Taylor series

$$f\left(\boldsymbol{x} + \Delta \boldsymbol{x}\right) = f_0 + \frac{\partial f}{\partial x_i} \Delta x_i + \mathcal{O}\left(\|\Delta \boldsymbol{x}\|\right)$$

$$\frac{\partial f}{\partial x_i} = \frac{f(\boldsymbol{x} + \Delta x_i\, \boldsymbol{\imath_i}) - f(\boldsymbol{x})}{\Delta x_i} + \mathcal{O}\left(|\Delta x_i|\right)$$

- The cost is <u>one</u> additional function evaluation per variable
- Subjected to round-off error
- Subjected to truncation error, convergence rate is $\mathcal{O}\left(|\Delta x_i|\right)$

What is Automatic Differentiation ?

AD is not complex step

What is Automatic Differentiation ?

AD is not complex step

Complex step moves along the imaginary axis

$$f\left(\boldsymbol{x} + \imath\,\Delta\boldsymbol{x}\right) = f(\boldsymbol{x}) + \frac{\partial f}{\partial x_i}\frac{\imath\,\Delta x_i}{1!} + \frac{\partial^2 f}{\partial x_i^2}\frac{\left(\imath\,\Delta x_i\right)^2}{2} + \mathcal{O}\left(\|\Delta x\|^2\right)$$

$$\frac{\partial f(x)}{\partial x_i} = \frac{\mathrm{Im}\{f(\boldsymbol{x} + \imath\,\Delta x_i\,\boldsymbol{\imath_i})\}}{\Delta x_i} + \mathcal{O}\left(|\Delta x_i|^2\right).$$

What is Automatic Differentiation ?

AD is not complex step

Complex step moves along the imaginary axis

$$f\left(\boldsymbol{x} + \imath\,\Delta\boldsymbol{x}\right) = f(\boldsymbol{x}) + \frac{\partial f}{\partial x_i}\frac{\imath\,\Delta x_i}{1!} + \frac{\partial^2 f}{\partial x_i^2}\frac{\left(\imath\,\Delta x_i\right)^2}{2} + \mathcal{O}\left(\|\Delta x\|^2\right)$$

$$\frac{\partial f(x)}{\partial x_i} = \frac{\mathrm{Im}\{f(\boldsymbol{x} + \imath\,\Delta x_i\,\boldsymbol{\imath_i})\}}{\Delta x_i} + \mathcal{O}\left(|\Delta x_i|^2\right).$$

- The cost is one additional complex function evaluation per variable
- No round-off error
- Subjected to truncation error, convergence rate is $\mathcal{O}\left(|\Delta x_i|^2\right)$

What is Automatic Differentiation ?

AD is not Symbolic differentiation

What is Automatic Differentiation ?

AD is not Symbolic differentiation

Symbolic differentiation treats expressions as strings

Cannot handle iterations

In the cases where results are obtained at the end of an iterative process or by accumulation, such as Newton Raphson iterations, symbolic iteration cannot be directly applied

Subject to expression swell

In lengthy calculations simplification algorithms tend to struggle

AD is based on Taylor Series

The sum of two functions

$$f(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = f_0 + f_{,i}\, \Delta x_i + f_{,ij} \frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$g(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = g_0 + g_{,i}\, \Delta x_i + g_{,ij} \frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$h(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = f(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) + g(\boldsymbol{x}_0 + \Delta \boldsymbol{x})$$

$$h(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = (f_0 + g_0) + (f_{,i} + g_{,i})\, \Delta x_i + (f_{,ij} + g_{,ij})\, \frac{\Delta x_i \Delta x_j}{2} + \dots$$

AD is based on Taylor Series

The sum of two functions

$$f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = f_0 + f_{,i}\,\Delta x_i + f_{,ij}\frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$g(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = g_0 + g_{,i}\,\Delta x_i + g_{,ij}\frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$h(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) + g(\boldsymbol{x}_0 + \Delta\boldsymbol{x})$$

$$h(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = \boxed{(f_0 + g_0)} + \boxed{(f_{,i} + g_{,i})}\,\Delta x_i + \boxed{(f_{,ij} + g_{,ij})}\,\frac{\Delta x_i \Delta x_j}{2} + \dots$$

AD is based on Taylor Series

The product of two functions

$$f(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = f_0 + f_{,i}\, \Delta x_i + f_{,ij} \frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$g(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = g_0 + g_{,i}\, \Delta x_i + g_{,ij} \frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$h(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = f(\boldsymbol{x}_0 + \Delta \boldsymbol{x})\, g(\boldsymbol{x}_0 + \Delta \boldsymbol{x})$$

$$\begin{aligned} h(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) = f_0 g_0 + (g_0\, f_{,i} + f_0\, g_{,j})\, \Delta x_i + \\ (g_0 f_{,ij} + f_{,i} g_{,j} + f_{,j} g_{,i} + f_0 g_{,ij})\, \frac{\Delta x_i \Delta x_j}{2} + \dots \end{aligned}$$

AD is based on Taylor Series

The product of two functions

$$f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = f_0 + f_{,i}\,\Delta x_i + f_{,ij}\frac{\Delta x_i \Delta x_j}{2} + \dots$$
$$g(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = g_0 + g_{,i}\,\Delta x_i + g_{,ij}\frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$h(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = f(\boldsymbol{x}_0 + \Delta\boldsymbol{x})\, g(\boldsymbol{x}_0 + \Delta\boldsymbol{x})$$

$$h(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = \boxed{f_0 g_0} + \boxed{(g_0\, f_{,i} + f_0\, g_{,j})}\,\Delta x_i + \boxed{(g_0 f_{,ij} + f_{,i} g_{,j} + f_{,j} g_{,i} + f_0 g_{,ij})}\,\frac{\Delta x_i \Delta x_j}{2} + \dots$$

AD is based on Taylor Series

The general case of an analytic function $g(t)$

$$f(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = f_0 + f_{,i}\,\Delta x_i + f_{,ij}\frac{\Delta x_i \Delta x_j}{2} + \dots$$

$$g(t_0 + \Delta t) = g_0 + \frac{\mathrm{d}g}{\mathrm{d}t}\Delta t + \frac{\mathrm{d}^2 g}{\mathrm{d}t^2}\frac{\Delta t^2}{2} + \dots$$

$$h = g\left(f\left(\boldsymbol{x}\right)\right) \quad \Leftrightarrow \quad \begin{cases} t = f(\boldsymbol{x}) \\ h = g(t) \end{cases}$$

$$h(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = g\left(f_0\right) + \frac{\mathrm{d}g}{\mathrm{d}t}\, f_{,i}\,\Delta x_i + \left(\frac{\mathrm{d}^2 g}{\mathrm{d}t^2}\, f_{,i} f_{,j} + \frac{\mathrm{d}g}{\mathrm{d}t}\, f_{,ij}\right)\frac{\Delta x_i \Delta x_j}{2} + \dots$$

How computers evaluate expressions

$$y = x_1^3\, x_2^2 + x_3^2$$

$$\begin{array}{l}
\hline
h_1 = x_1 \\
h_2 = x_2 \\
h_3 = x_3 \\
\hline
h_4 = h_1^3 \\
h_5 = h_2^2 \\
h_6 = h_3^2 \\
\hline
h_7 = h_5\, h_6 \\
\hline
h_8 = h_7 + h_6 \\
\hline
y = h_8
\end{array}$$

How computers evaluate expressions

$$y = x_1^3\, x_2^2 + x_3^2$$

$h_1 = x_1$	$\delta h_1 = \delta x_1 \leftarrow \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$
$h_2 = x_2$	$\delta h_2 = \delta x_2 \leftarrow \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$
$h_3 = x_3$	$\delta h_3 = \delta x_3 \leftarrow \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$
$h_4 = h_1^3$	$\delta h_4 = 3h_1^2\, \delta h_1$
$h_5 = h_2^2$	$\delta h_5 = 2h_2^2\, \delta h_2$
$h_6 = h_3^2$	$\delta h_6 = 2h_3^2\, \delta h_3$
$h_7 = h_5\, h_6$	$\delta h_7 = h_5\, \delta h_6 + h_6\, \delta h_5$
$h_8 = h_7 + h_6$	$\delta h_8 = \delta h_7 + \delta h_6$
$y = h_8$	$\delta y = \delta h_8$

Dual numbers - definition

[1] Shenitzer, A. and Kantor, I.L. and Solodovnikov, A.S, *Hypercomplex Numbers: An Elementary Introduction to Algebras*, 2011, Springer New York

Dual numbers - definition

Complex numbers have a scalar real and imaginary part

$$\boldsymbol{x} \equiv a + \imath b$$

with the product rule:

$$\imath^2 = -1$$

[1] Shenitzer, A. and Kantor, I.L. and Solodovnikov, A.S, *Hypercomplex Numbers: An Elementary Introduction to Algebras*, 2011, Springer New York

Dual numbers - definition

Complex numbers have a scalar real and imaginary part

$$\boldsymbol{x} \equiv a + \imath b$$

with the product rule:

$$\imath^2 = -1$$

Dual numbers [1] have multiple higher dimensional parts.

$$\boldsymbol{x} \equiv x_0 + x_i \boldsymbol{\imath}_i + x_{ij} \boldsymbol{\imath}_{ij}$$

where:

$$\boldsymbol{\imath}_{ij} \equiv \boldsymbol{\imath}_i \otimes \boldsymbol{\imath}_j + \boldsymbol{\imath}_j \otimes \boldsymbol{\imath}_i$$

with the product rule:

$$\boldsymbol{\imath}_i \boldsymbol{\imath}_j \equiv \boldsymbol{\imath}_{ij} \qquad \text{and} \qquad \boldsymbol{\imath}_{ij} \boldsymbol{\imath}_k \equiv 0$$

[1] Shenitzer, A. and Kantor, I.L. and Solodovnikov, A.S, *Hypercomplex Numbers: An Elementary Introduction to Algebras*, 2011, Springer New York

Dual numbers - equality

given:

$$\boldsymbol{x} \equiv x_0 + x_i \boldsymbol{\imath}_i + x_{ij} \boldsymbol{\imath}_{ij}$$

$$\boldsymbol{y} \equiv y_0 + y_i \boldsymbol{\imath}_i + y_{ij} \boldsymbol{\imath}_{ij}$$

$$\boldsymbol{x} = \boldsymbol{y} \iff \begin{cases} y_0 = x_0 \\ y_i = x_i \\ y_{ij} = x_{ij} \end{cases} \qquad \forall\, i, j$$

Dual numbers - sum

given:

$$\boldsymbol{x} \equiv x_0 + x_i \boldsymbol{\imath}_i + x_{ij} \boldsymbol{\imath}_{ij}$$
$$\boldsymbol{y} \equiv y_0 + y_i \boldsymbol{\imath}_i + y_{ij} \boldsymbol{\imath}_{ij}$$

$$\boldsymbol{x} + \boldsymbol{y} = x_0 + (x_i + y_i) \boldsymbol{\imath}_i + (x_{ij} + y_{ij}) \boldsymbol{\imath}_{ij}$$

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} \iff \begin{cases} z_0 = x_0 + y_0 \\ z_i = x_i + y_i \\ z_{ij} = x_{ij} + y_{ij} \end{cases}$$

The neutral element is

$$0 + 0\, \boldsymbol{\imath}_i + 0\, \boldsymbol{\imath}_{ij} \qquad \forall\, i, j$$

Dual numbers - product

given:

$$\boldsymbol{x} \equiv x_0 + x_i \boldsymbol{\imath}_i + x_{ij} \boldsymbol{\imath}_{ij}$$
$$\boldsymbol{y} \equiv y_0 + y_i \boldsymbol{\imath}_i + y_{ij} \boldsymbol{\imath}_{ij}$$

$$\begin{aligned} \boldsymbol{xy} = x_0 y_0 + & (y_0 x_i + x_0 y_i) \boldsymbol{\imath}_i + \\ & (x_{ij} y_0 + x_i y_j + x_j y_i + x_0 y_{ij}) \boldsymbol{\imath}_{ij} \end{aligned}$$

$$\boldsymbol{z} = \boldsymbol{xy} \iff \begin{cases} z_0 = x_0 y_0 \\ z_i = x_i y_0 + x_0 y_i \\ z_{ij} = x_{ij} y_0 + x_i y_j + x_j y_i + x_0 y_{ij} \end{cases}$$

The neutral element is

$$1 + 0\, \boldsymbol{\imath}_i + 0\, \boldsymbol{\imath}_{ij} \qquad \forall i, j$$

Dual numbers - integer power

By recursively applying the product rule we can define the integer power of a Dual number as

given:

$$\boldsymbol{x} \equiv x_0 + x_i \boldsymbol{\imath}_i + x_{ij} \boldsymbol{\imath}_{ij}$$

$$\boldsymbol{y} = \boldsymbol{x}^n \iff \begin{cases} y_0 = x_0^n \\ y_i = n\, x_0^{n-1}\, x_i \\ y_{ij} = n\,(n-1)\; x_0^{n-2} x_i x_j + n\, x_0^{n-1}\, x_{ij} \end{cases}$$

Dual numbers are an Algebra

We can define analytic functions over Dual numbers

Exponential:

$$\mathrm{e}^{\boldsymbol{x}} = 1 + \boldsymbol{x} + \frac{\boldsymbol{x}^2}{2} + \cdots + \frac{\boldsymbol{x}^n}{n!} + \cdots$$

sin:

$$\sin(\boldsymbol{x}) = \boldsymbol{x} - \frac{\boldsymbol{x}^3}{3!} + \cdots + \frac{(-1)^n}{(2n+1)!}\boldsymbol{x}^{2n+1} + \cdots$$

etc.

$$\vdots$$

One numerical example

$$y(x_1, x_2, x_3) = x_1^3\, x_2^2 + x_3^2$$

One numerical example

$$y(x_1, x_2, x_3) = x_1^3\, x_2^2 + x_3^2$$

$$\boldsymbol{x}_1 = x_1 + \boldsymbol{\imath}_1\,, \qquad \boldsymbol{x}_2 = x_2 + \boldsymbol{\imath}_2\,, \qquad \boldsymbol{x}_3 = x_3 + \boldsymbol{\imath}_3$$

One numerical example

$$y(x_1, x_2, x_3) = x_1^3\, x_2^2 + x_3^2$$

$$\boldsymbol{x}_1 = x_1 + \boldsymbol{\imath}_1\,, \qquad \boldsymbol{x}_2 = x_2 + \boldsymbol{\imath}_2\,, \qquad \boldsymbol{x}_3 = x_3 + \boldsymbol{\imath}_3$$

$$\boldsymbol{y} = (x_1 + \boldsymbol{\imath}_1)^3 (x_2 + \boldsymbol{\imath}_2)^2 + (x_3 + \boldsymbol{\imath}_3)^2 \;=\; y_0 + y_i \boldsymbol{\imath}_i + y_{ij} \boldsymbol{\imath}_{ij}$$

One numerical example

$$y(x_1, x_2, x_3) = x_1^3\, x_2^2 + x_3^2$$

$$\boldsymbol{x}_1 = x_1 + \boldsymbol{\imath}_1\,, \qquad \boldsymbol{x}_2 = x_2 + \boldsymbol{\imath}_2\,, \qquad \boldsymbol{x}_3 = x_3 + \boldsymbol{\imath}_3$$

$$\boldsymbol{y} = (x_1 + \boldsymbol{\imath}_1)^3 (x_2 + \boldsymbol{\imath}_2)^2 + (x_3 + \boldsymbol{\imath}_3)^2 \;=\; y_0 + y_i \boldsymbol{\imath}_i + y_{ij} \boldsymbol{\imath}_{ij}$$

$$y_0 = x_1^3 x_2^2 + x_3^2$$

$$y_i\, \boldsymbol{\imath}_i = 3x_1^2 x_2^2\, \boldsymbol{\imath}_1 + 2x_1^3 x_2\, \boldsymbol{\imath}_2 + 2x_3 \boldsymbol{\imath}_3 \equiv \begin{bmatrix} 3\, x_1^2 x_2^2 \\ 2\, x_1^3 x_2 \\ 2\, x_3 \end{bmatrix}$$

$$y_{ij}\, \boldsymbol{\imath}_{ij} = 3x_1 x_2^2\, \boldsymbol{\imath}_{11} + 6x_1^2 x_2\, \boldsymbol{\imath}_{12} + x_1^3\, \boldsymbol{\imath}_{22} + \boldsymbol{\imath}_{33} \equiv \begin{bmatrix} 6\, x_1 x_2^2 & 6\, x_1^2 x_2 & 0 \\ 6x_1^2 x_2 & 2\, x_1^3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

a few remarks

- AD derivatives are exact (within the number representation accuracy)
- AD proceeds by accumulation
- AD is insensitive to computer programming constructs such as `if-then-else`, `for-loops`, etc.
- Forward mode AD can be implemented by operators overloading
- Forward mode AD suits FE

Solid mechanics with FE

Solid mechanics with FE

The Virtual Work Principle is

$$\int_{V_0} \left(P_{ij} \frac{\partial F_{ij}}{\partial u_k} - b_{0_k} \right) \delta u_k \, \mathrm{d}V_0 - \int_{S_0} t_{0_k} \, \delta u_k \, \mathrm{d}S_0 = 0 \qquad \begin{array}{c} \forall \, \delta u_k \\ k = 1, 2, 3 \end{array}$$

Solid mechanics with FE

The Virtual Work Principle is

$$\int_{V_0} \left(P_{ij} \, \frac{\partial F_{ij}}{\partial u_k} - b_{0_k} \right) \, \delta u_k \, \mathrm{d}V_0 - \int_{S_0} t_{0_k} \, \delta u_k \, \mathrm{d}S_0 = 0 \qquad \begin{array}{c} \forall \, \delta u_k \\ k = 1, 2, 3 \end{array}$$

the residual force vector is

$$\boldsymbol{r} = \sum_{k=1}^{N_{BE}} \sum_{l=1}^{N_{BW}^k} w_l^k \, \left[P_{ij} \frac{\partial F_{ij}}{\partial \boldsymbol{u}} - \boldsymbol{b}_0 \right]_{r_l^k} - \sum_{k=1}^{N_{SE}} \sum_{l=1}^{N_{SW}^k} v_l^k \, [\boldsymbol{t}_0]_{r_l^k} = \boldsymbol{0}$$

Solid mechanics with FE

The Virtual Work Principle is

$$\int_{V_0} \left(P_{ij} \frac{\partial F_{ij}}{\partial u_k} - b_{0_k} \right) \delta u_k \, \mathrm{d}V_0 - \int_{S_0} t_{0_k} \, \delta u_k \, \mathrm{d}S_0 = 0 \qquad \begin{array}{c} \forall \, \delta u_k \\ k = 1, 2, 3 \end{array}$$

the residual force vector is

$$\boldsymbol{r} = \sum_{k=1}^{N_{BE}} \sum_{l=1}^{N_{BW}^k} w_l^k \left[P_{ij} \frac{\partial F_{ij}}{\partial \boldsymbol{u}} - \boldsymbol{b}_0 \right]_{r_l^k} - \sum_{k=1}^{N_{SE}} \sum_{l=1}^{N_{SW}^k} v_l^k \, [\boldsymbol{t}_0]_{r_l^k} = \boldsymbol{0}$$

the tangent stiffness matrix is

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}} = \sum_{k=1}^{N_{BE}} \sum_{i=1}^{N_{BW}^k} w_l^k \left[\frac{\partial P_{ij}}{\partial F_{hk}} \frac{\partial F_{hk}}{\partial \boldsymbol{u}} \frac{\partial F_{ij}}{\partial \boldsymbol{u}} - \frac{\partial \boldsymbol{b}_0}{\partial \boldsymbol{u}} \right]_{r_l^k} - \sum_{k=1}^{N_{SE}} \sum_{i=1}^{N_{SW}^k} v_l^k \left[\frac{\partial \boldsymbol{t}_0}{\partial \boldsymbol{u}} \right]_{r_l^k}$$

Solid mechanics with AD

Solid mechanics with AD

Minimum free energy principle is

$$\delta\Psi = 0 \qquad \text{with} \qquad \Psi = \int_{V_0} (\phi - b_0)\,\mathrm{d}V_0 - \int_{S_0} t_0\,\mathrm{d}S_0$$

Solid mechanics with AD

Minimum free energy principle is

$$\delta\Psi = 0 \qquad \text{with} \qquad \Psi = \int_{V_0} (\phi - b_0)\, \mathrm{d}V_0 - \int_{S_0} t_0\, \mathrm{d}S_0$$

we can use a FE discretization for evaluating Ψ

$$\Psi(\boldsymbol{u}) = \sum_{k=1}^{N_{BE}} \sum_{i=1}^{N_{BW}^k} w_I^k \, [\phi + b_0]_{r_I^k} + \sum_{k=1}^{N_{SE}} \sum_{i=1}^{N_{SW}^k} v_I^k \, [t_0]_{r_I^k}$$

Solid mechanics with AD

Minimum free energy principle is

$$\delta\Psi = 0 \qquad \text{with} \qquad \Psi = \int_{V_0} (\phi - b_0)\, \mathrm{d}V_0 - \int_{S_0} t_0 \,\mathrm{d}S_0$$

we can use a FE discretization for evaluating Ψ

$$\Psi(\boldsymbol{u}) = \sum_{k=1}^{N_{BE}} \sum_{i=1}^{N_{BW}^k} w_l^k \, [\phi + b_0]_{r_l^k} + \sum_{k=1}^{N_{SE}} \sum_{i=1}^{N_{SW}^k} v_l^k \, [t_0]_{r_l^k}$$

equilibrium is given by

$$\delta\Psi = \frac{\partial \Psi}{\partial \boldsymbol{u}} \cdot \delta\boldsymbol{u} = 0 \quad \forall\, \delta\boldsymbol{u} \qquad \Longleftrightarrow \qquad \boldsymbol{r} = \frac{\partial \Psi}{\partial \boldsymbol{u}} = 0$$

Solid mechanics with AD

Minimum free energy principle is

$$\delta\Psi = 0 \qquad \text{with} \qquad \Psi = \int_{V_0} (\phi - b_0)\, \mathrm{d}V_0 - \int_{S_0} t_0 \, \mathrm{d}S_0$$

we can use a FE discretization for evaluating Ψ

$$\Psi(\boldsymbol{u}) = \sum_{k=1}^{N_{BE}} \sum_{i=1}^{N_{BW}^k} w_l^k \left[\phi + b_0\right]_{r_l^k} + \sum_{k=1}^{N_{SE}} \sum_{i=1}^{N_{SW}^k} v_l^k \, \left[t_0\right]_{r_l^k}$$

equilibrium is given by

$$\delta\Psi = \frac{\partial \Psi}{\partial \boldsymbol{u}} \cdot \delta\boldsymbol{u} = 0 \quad \forall\, \delta\boldsymbol{u} \qquad \Longleftrightarrow \qquad \boldsymbol{r} = \frac{\partial \Psi}{\partial \boldsymbol{u}} = 0$$

the tangent stiffness matrix is

$$\boldsymbol{K}_t = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{u}} = \frac{\partial^2 \Psi}{\partial \boldsymbol{u} \partial \boldsymbol{u}}$$

Boundary conditions and constraints

We use AD with Laplace Multipliers

$$L\left(\boldsymbol{u},\boldsymbol{\lambda}\right)=\Psi\left(\boldsymbol{u}\right)-\boldsymbol{\lambda}\cdot\boldsymbol{g}\left(\boldsymbol{u}\right)$$

Boundary conditions and constraints

We use AD with Laplace Multipliers

$$L(\boldsymbol{u}, \boldsymbol{\lambda}) = \Psi(\boldsymbol{u}) - \boldsymbol{\lambda} \cdot \boldsymbol{g}(\boldsymbol{u})$$

equilibrium is

$$\nabla L = \mathbf{0}$$

with:

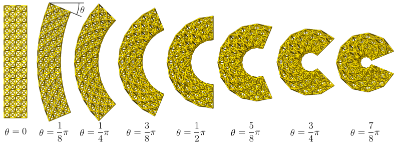
$$\nabla L = \begin{bmatrix} \frac{\partial \Psi}{\partial \boldsymbol{u}} - \boldsymbol{\lambda} \cdot \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{u}} \\ -\boldsymbol{g} \end{bmatrix} \qquad \nabla^2 L = \begin{bmatrix} \frac{\partial^2 \Psi}{\partial \boldsymbol{u} \partial \boldsymbol{u}} - \boldsymbol{\lambda} \cdot \frac{\partial^2 \boldsymbol{g}}{\partial \boldsymbol{u} \partial \boldsymbol{u}} & -\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{u}}^T \\ -\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{u}} & \mathbf{0} \end{bmatrix}$$

a few more remarks

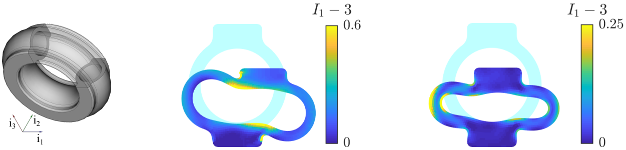
- General simplification in implementing classical FE
- Remarkable reduction in code writing
- Free energy is always a scalar
- Simplifies the implementation of sophisticated material models and boundary conditions
- FE implementation trough AD does not require a stress tensor
- AD on the free energy is a radical approach to FE automation

Examples

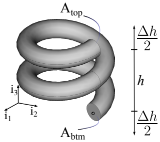
- Truss structure

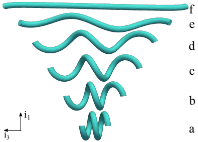


- Axisymmetric problem



- Three dimensional continuum





The rod element

The strain energy for the rod element is

$$\phi^{\text{rod}} = \frac{E_s A l_o}{2} \left(\frac{l}{l_0} - 1 \right)^2 \qquad \text{with} \qquad \begin{aligned} l_0 &= \| \boldsymbol{r}_2 - \boldsymbol{r}_1 \| \\ l &= \| \boldsymbol{r}_2 + \boldsymbol{u}_2 - (\boldsymbol{r}_1 + \boldsymbol{u}_1) \|. \end{aligned}$$

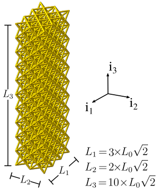
The strain energy of a truss structure is $\Phi^{\text{truss}} = \sum_i \phi_i^{\text{rod}}$

The rod element

The strain energy for the rod element is

$$\phi^{\text{rod}} = \frac{E_s A l_o}{2} \left(\frac{l}{l_0} - 1 \right)^2 \qquad \text{with} \qquad \begin{aligned} l_0 &= \| \boldsymbol{r}_2 - \boldsymbol{r}_1 \| \\ l &= \| \boldsymbol{r}_2 + \boldsymbol{u}_2 - (\boldsymbol{r}_1 + \boldsymbol{u}_1) \|. \end{aligned}$$

The strain energy of a truss structure is $\Phi^{\text{truss}} = \sum_i \phi_i^{\text{rod}}$



Boundary conditions are

$$(X_2 + u_2) \cos(\theta) + (X_3 + u_3) \sin(\theta) - \frac{1}{2} L_3 = 0$$

$$(X_2 + u_2) \cos(\theta) - (X_3 + u_3) \sin(\theta) + \frac{1}{2} L_3 = 0$$

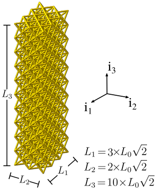
$$\sum u_1 = 0$$

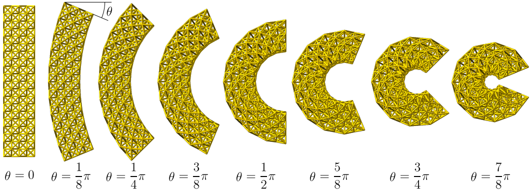
The rod element

The strain energy for the rod element is

$$\phi^{\text{rod}} = \frac{E_s A l_o}{2}\left(\frac{l}{l_0} - 1\right)^2 \qquad \text{with} \qquad \begin{aligned} l_0 &= \|\boldsymbol{r}_2 - \boldsymbol{r}_1\| \\ l &= \|\boldsymbol{r}_2 + \boldsymbol{u}_2 - (\boldsymbol{r}_1 + \boldsymbol{u}_1)\|. \end{aligned}$$

The strain energy of a truss structure is $\Phi^{\text{truss}} = \sum_i \phi_i^{\text{rod}}$





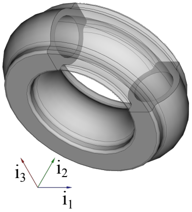
Boundary conditions are

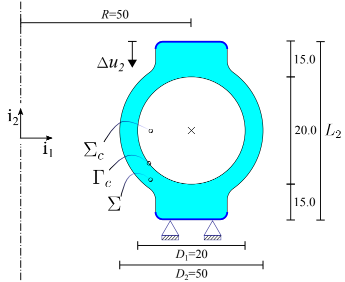
$$(X_2 + u_2)\cos(\theta) + (X_3 + u_3)\sin(\theta) - \frac{1}{2}L_3 = 0$$

$$(X_2 + u_2)\cos(\theta) - (X_3 + u_3)\sin(\theta) + \frac{1}{2}L_3 = 0$$

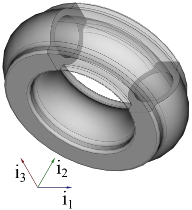
$$\sum u_1 = 0$$

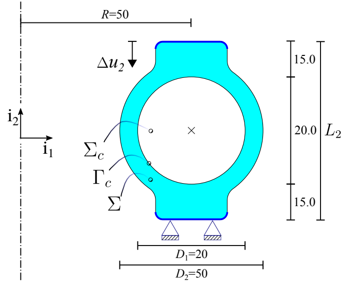
Cylindrical symmetry problem with volume constraint





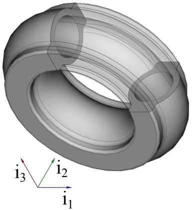
Cylindrical symmetry problem with volume constraint

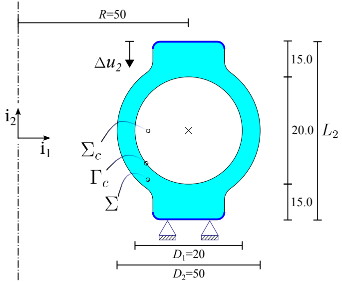




$$\Phi^{\text{cyl}} = 2\pi \int_{\Sigma} \phi\left(\boldsymbol{F}^{\text{cyl}}\right) X_1 \mathrm{d}\Sigma \quad \text{with} \quad \boldsymbol{F}^{\text{cyl}} = \begin{bmatrix} 1+u_{1,1} & u_{1,2} & 0 \\ u_{2,1} & 1+u_{2,2} & 0 \\ 0 & 0 & 1+\frac{u_1}{X_1} \end{bmatrix}$$

Cylindrical symmetry problem with volume constraint



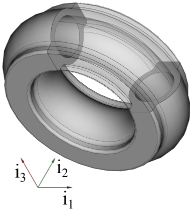


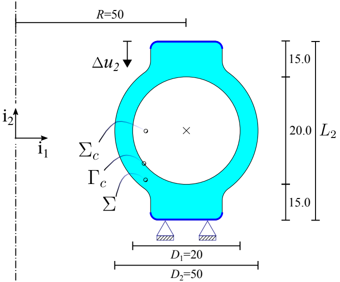
$$\Phi^{\mathrm{cyl}} = 2\pi \int_{\Sigma} \phi\left(\boldsymbol{F}^{\mathrm{cyl}}\right)\, X_1 \mathrm{d}\Sigma \qquad \text{with} \qquad \boldsymbol{F}^{\mathrm{cyl}} = \begin{bmatrix} 1+u_{1,1} & u_{1,2} & 0 \\ u_{2,1} & 1+u_{2,2} & 0 \\ 0 & 0 & 1+\frac{u_1}{X_1} \end{bmatrix}$$

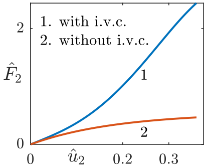
$$L^{\mathrm{cyl}}\left(u_k, \lambda_c\right) = \Phi^{\mathrm{cyl}}(u_k) - \lambda_c\left[V_c\left(u_k\right) - V_{c_0}\right] \qquad \text{with}$$

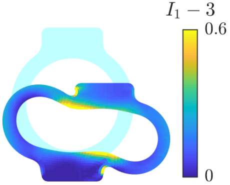
$$V_c = 2\pi \int_{\Sigma_c} x_1\, \mathrm{d}\Sigma = \pi \int_{\Gamma_c} x_1\, x_2\, \mathrm{d}x_1$$

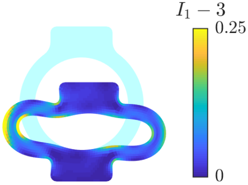
Cylindrical symmetry problem with volume constraint





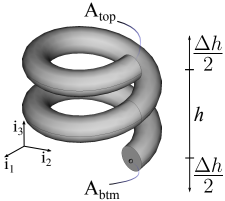




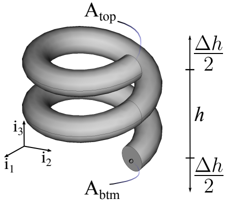


Three-dimensional continuum problem

Three-dimensional continuum problem



Three-dimensional continuum problem

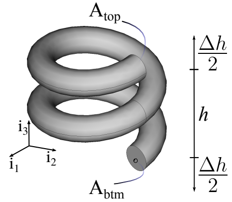


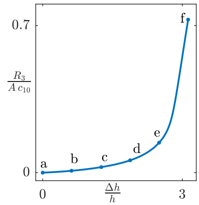
$$\int_{A_{\text{btm}}} u_1 \,\mathrm{d}A = 0 \qquad \int_{A_{\text{top}}} u_1 \,\mathrm{d}A = 0$$

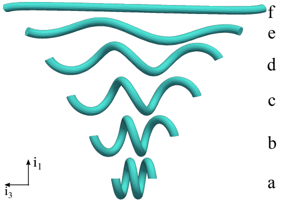
$$\int_{A_{\text{btm}}} u_2 \,\mathrm{d}A = 0 \qquad \int_{A_{\text{top}}} u_2 \,\mathrm{d}A = 0$$

$$\int_{A_{\text{btm}}} u_3 \,\mathrm{d}A = -\frac{\Delta h}{2} \qquad \int_{A_{\text{top}}} u_3 \,\mathrm{d}A = \frac{\Delta h}{2}$$

Three-dimensional continuum problem







$$\int_{A_{btm}} u_1 \,\mathrm{d}A = 0 \qquad \int_{A_{top}} u_1 \,\mathrm{d}A = 0$$

$$\int_{A_{btm}} u_2 \,\mathrm{d}A = 0 \qquad \int_{A_{top}} u_2 \,\mathrm{d}A = 0$$

$$\int_{A_{btm}} u_3 \,\mathrm{d}A = -\frac{\Delta h}{2} \qquad \int_{A_{top}} u_3 \,\mathrm{d}A = \frac{\Delta h}{2}$$

one final remark

- In the context of FE for solid mechanics, AD works.

Forward Mode Automatic Differentiation with an application to solid mechanics

Andrea Vigliotti [1]   Ferdinando Auricchio [2]

[1]Italian Aerospace Research Centre

[2]University of Pavia

21/01/2020