Gradient-based shape optimization

Dr.-Ing. Andreas Apostolatos

April 16, 2020

1 Problem placement

Given is a *Boundary Value Problem* (BVP) which governs a primal unkown field $\mathbf{u} \in \boldsymbol{V}$ and which in residual form writes,

$$\mathbf{R}(\mathbf{u}) = \mathbf{0} \; . \tag{1}$$

Exemplarily, in case of linear elasticity the aforementioned residual form is,

$$\begin{aligned}
\mathbf{R}(\mathbf{u}) &= \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} + \mathbf{b} && \text{in } \Omega \; , && \text{(2a)}\\
\boldsymbol{\sigma} &= \boldsymbol{C} : \boldsymbol{\varepsilon} && \text{in } \Omega \; , && \text{(2b)}\\
\boldsymbol{\varepsilon} &= \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}^{\mathrm{T}}\mathbf{u}\right) && \text{in } \Omega \; , && \text{(2c)}\\
\mathbf{u} &= \bar{\mathbf{u}} && \text{on } \Gamma_{\mathrm{d}} \; , && \text{(2d)}\\
\boldsymbol{\sigma} \cdot \mathbf{n} &= \bar{\mathbf{t}} && \text{on } \Gamma_{\mathrm{n}} \; . && \text{(2e)}
\end{aligned}$$

Moreover, given is an objective function $\mathcal{J}(\mathbf{s}, \mathbf{u})$ where $\mathbf{s} \in \boldsymbol{\mathcal{S}}$ represents a continuous design variable and $\mathbf{u}$ the primal solution of problem in Eq. (2). The task is to minimize $\mathcal{J}$ with respect to the design variables $\mathbf{s}$ subject to the residual equation in Eq. (1), namely,

$$\min_{\bar{\mathbf{s}} \in \boldsymbol{\mathcal{S}}} \mathcal{J}(\bar{\mathbf{s}}, \mathbf{u}) \text{ subject to } \mathbf{R}(\mathbf{s}, \mathbf{u}) = \mathbf{0} \text{ in } \Omega(\mathbf{s}) \; . \tag{3}$$

Problem in Eq. (3) can be formulated as the minimization of the following augmented by Lagrange Multipliers potential,

$$\mathcal{L}(\mathbf{s}, \mathbf{u}, \boldsymbol{\lambda}) = \mathcal{J}(\mathbf{s}, \mathbf{u}) + \int_{\Omega(\mathbf{s})} \boldsymbol{\lambda} \cdot \mathbf{R}(\mathbf{s}, \mathbf{u}) \, \mathrm{d}\Omega \; . \tag{4}$$

The stationary point of potential in Eq. (4) is sought for which,

$$\delta\mathcal{L} = \delta\mathcal{J} + \int_{\Omega(\mathbf{s})} \delta\boldsymbol{\lambda} \cdot \mathbf{R} + \boldsymbol{\lambda} \frac{\partial \mathbf{R}}{\partial \mathbf{s}} \cdot \delta\mathbf{s} + \frac{\partial \mathbf{R}}{\partial \mathbf{u}} \cdot \delta\mathbf{u} \, \mathrm{d}\Omega \; . \tag{5}$$

Taking advantage of the fact that $\mathbf{R}(\mathbf{s}, \mathbf{u}) = \mathbf{0}$ in $\Omega(\mathbf{s})$, Eq. (5) simplifies to,

$$\delta\mathcal{L} = \delta\mathcal{J} + \int_{\Omega(\mathbf{s})} \boldsymbol{\lambda} \cdot \frac{\partial \mathbf{R}}{\partial \mathbf{s}} \cdot \delta\mathbf{s} + \frac{\partial \mathbf{R}}{\partial \mathbf{u}} \cdot \delta\mathbf{u} \, \mathrm{d}\Omega \; . \tag{6}$$

2 Shape optimization in linear elasticity

Herein assumed is that the residual $\mathbf{R}$ is given by the discrete problem meaning,

$$\mathbf{R}(\hat{\mathbf{s}}, \hat{\mathbf{u}}) = \mathbf{K}(\hat{\mathbf{s}})\hat{\mathbf{u}} - \mathbf{F} \; , \tag{7}$$

where $\hat{\mathbf{s}}$ and $\hat{\mathbf{u}}$ stand for the discrete design variables $\hat{\mathbf{s}} = [\hat{s}_0 \; \cdots \; \hat{s}_i \; \cdots \; \hat{s}_n]$ and the discrete primal solution field $\hat{\mathbf{u}} = [\hat{u}_0 \; \cdots \; \hat{u}_i \; \cdots \; \hat{u}_n]$, n being the number of *Degrees of Freedom* (DOFs). Moreover concerning the variables appearing in Eq. (7),

$$\mathbf{K}(\hat{\mathbf{s}}) = \int_{\Omega_{\mathrm{h}}(\hat{\mathbf{s}})} \mathbf{B}(\hat{\mathbf{s}})^{\mathrm{T}} \mathbf{C} \mathbf{B}(\hat{\mathbf{s}}) \, \mathrm{d}\Omega \; , \tag{8a}$$

$$\mathbf{F} = \mathbf{N}(\mathbf{X}_0)^{\mathrm{T}} \mathbf{f} \; . \tag{8b}$$

$\Omega_{\mathrm{h}}(\hat{\mathbf{s}})$ being the discretized domain of the problem. In the discrete case, $\hat{\mathbf{s}}$ stands for the coordinates of the nodes in the provided mesh $\Omega_{\mathrm{h}}(\hat{\mathbf{s}})$. In Eq. (8), $\mathbf{B}(\hat{\mathbf{s}})$ stands for the B-operator matrix of the problem in Eq. (2), $\mathbf{N}^{\mathrm{T}}(\mathbf{X}_0)$ for the matrix of the basis functions corresponding to the discretization evaluated at $\mathbf{X}_0$ and $\mathbf{f}$ the constant load which is applied at $\mathbf{X}_0$. Accordingly, the objective function is defined by the strain-energy meaning that,

$$\mathcal{J}(\hat{\mathbf{s}}, \hat{\mathbf{u}}) = -\frac{1}{2} \hat{\mathbf{u}}^{\mathrm{T}} \mathbf{K}(\hat{\mathbf{s}}) \hat{\mathbf{u}} = -\frac{1}{2} \hat{\mathbf{u}}^{\mathrm{T}} \mathbf{F} \; , \tag{9}$$

which has to be minimized so that the strain-energy is maximized. In this way, the augmented potential in Eq. (4) becomes,

$$\mathcal{L}(\hat{\mathbf{s}}, \hat{\mathbf{u}}, \bar{\boldsymbol{\lambda}}) = \mathcal{J}(\hat{\mathbf{s}}, \hat{\mathbf{u}}) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}} \mathbf{R}(\hat{\mathbf{s}}, \hat{\mathbf{u}}) \; , \tag{10}$$

where $\bar{\boldsymbol{\lambda}}$ stands for the vector of the discrete ajoint variables (Lagrange Multipliers) in this case. The first variation of the augmented potential in Eq. (6) becomes,

$$\delta \mathcal{L} = \left(\frac{\partial \mathcal{J}}{\partial \hat{\mathbf{s}}} + \bar{\boldsymbol{\lambda}}^{\mathrm{T}} \frac{\partial \mathbf{R}}{\partial \hat{\mathbf{s}}} \right) \delta \hat{\mathbf{s}} + \left(\frac{\partial \mathcal{J}}{\partial \hat{\mathbf{u}}} + \bar{\boldsymbol{\lambda}}^{\mathrm{T}} \frac{\partial \mathbf{R}}{\partial \hat{\mathbf{u}}} \right) \delta \hat{\mathbf{u}} \; . \tag{11}$$

To simplify the problem, the discrete vector of the Lagrange Multipliers $\bar{\boldsymbol{\lambda}}$ is chosen such that the following condition holds,

$$\frac{\partial \mathcal{J}}{\partial \hat{\mathbf{u}}} + \bar{\boldsymbol{\lambda}}^{\mathrm{T}} \frac{\partial \mathbf{R}}{\partial \hat{\mathbf{u}}} = \mathbf{0} \; , \tag{12}$$

since the vector of the Lagrange Multipliers $\bar{\boldsymbol{\lambda}}$ can be freely selected. Eq. (12) can be alternatively written as,

$$\left(\frac{\partial \mathbf{R}}{\partial \hat{\mathbf{u}}} \right)^{\mathrm{T}} \hat{\boldsymbol{\lambda}} = - \left(\frac{\partial \mathcal{J}}{\partial \hat{\mathbf{u}}} \right)^{\mathrm{T}} \; . \tag{13}$$

For linear elastic problems, relation in Eq. (7) results into,

$$\left(\frac{\partial \mathbf{R}}{\partial \hat{\mathbf{u}}} \right)^{\mathrm{T}} = \mathbf{K}(\hat{\mathbf{s}})^{\mathrm{T}} = \mathbf{K}(\hat{\mathbf{s}}) \; , \tag{14}$$

and by substituting Eq. (14) into Eq. (13) given that $\mathbf{F} = \mathbf{K}(\hat{\mathbf{s}})\hat{\mathbf{u}}$ from Eq. (7) and given that $\partial\mathcal{J}/\partial\hat{\mathbf{u}} = -1/2\mathbf{F}$ from Eq. (9), Eq. (13) results in,

$$\bar{\boldsymbol{\lambda}} = \frac{1}{2}\hat{\mathbf{u}} \ , \tag{15}$$

concerning the computation of the discrete adjoint variables $\bar{\boldsymbol{\lambda}}$. Substituting Eq. (15) into Eq. (11) it results in,

$$\delta\mathcal{L} = \left(\frac{\partial\mathcal{J}}{\partial\hat{\mathbf{s}}} + \frac{1}{2}\hat{\mathbf{u}}^{\mathrm{T}}\frac{\partial\mathbf{R}}{\partial\hat{\mathbf{s}}}\right)\delta\hat{\mathbf{s}} \ . \tag{16}$$

Using Eq. (7) it can be seen that,

$$\frac{\partial\mathbf{R}}{\partial\hat{\mathbf{s}}} = \frac{\partial\mathbf{K}}{\partial\hat{\mathbf{s}}}\hat{\mathbf{u}} \ , \tag{17}$$

provided that the load vector $\mathbf{F}$ does not depend on the design variables. Substituting Eq. (17) into Eq. (16) results in,

$$\delta\mathcal{L} = \overbrace{\left(\frac{\partial\mathcal{J}}{\partial\hat{\mathbf{s}}} + \frac{1}{2}\hat{\mathbf{u}}^{\mathrm{T}}\frac{\partial\mathbf{K}}{\partial\hat{\mathbf{s}}}\hat{\mathbf{u}}\right)}^{\text{sensitivity field } \mathbf{S}}\delta\hat{\mathbf{s}} \ . \tag{18}$$

Since the objective function can be defined either way in Eq. (9) using the last equality one has that $\partial\mathcal{J}/\partial\hat{\mathbf{s}} = -\hat{\mathbf{u}}^{\mathrm{T}}\partial\mathbf{F}/\partial\hat{\mathbf{s}} = \mathbf{0}$ and thus the sensitivity vector is given by,

$$\mathbf{S} = \frac{1}{2}\hat{\mathbf{u}}^{\mathrm{T}}\frac{\partial\mathbf{K}}{\partial\hat{\mathbf{s}}}\hat{\mathbf{u}} \ . \tag{19}$$

3 Sensitivity analysis

Concerning the computation of the sensitivities defined in Eq. (19) one can either compute them analytically using relation in Eq. (8a) or evaluate them using a finite differencing scheme. Herein the second option is preferred since it is more general and does not require lengthy derivations. Within the finite differencing approach the derivative $\partial\mathbf{K}/\partial\hat{\mathbf{s}}$ in Eq. (19) is approximated as,

$$\frac{\partial\mathbf{K}}{\partial\hat{s}_i} = \frac{1}{d_i}\left(\mathbf{K}(s_i + d_i) - \mathbf{K}(s_i)\right) \ , \tag{20}$$

for each design variable s_i in $\mathbf{s}$ where d_i stands for the perturbation of design variable s_i. Note that $d_i \neq d_j$ for $i \neq j$ with $i, j = 1, \ldots, n$ and those perturbation are chosen based on the corresponding element size. However, in this work the perturbations are all chosen the same, that is $d_1 = \ldots = d_n = d$ and this value may be scaled with a small number $\epsilon \approx 1e-4$.

4 Gradient-based shape optimization

Within the gradient-based shape optimization a set of optimization iterations is formulated within which we are moving in the negative gradient direction (which is in this case is the sensitivity vector) in the design space $\boldsymbol{\mathcal{S}}$. In this was, at each optimization iteration the design variables are updated as,

$$\mathbf{s}^{\hat{i}} = \mathbf{s}^{\hat{i}} - \alpha^{\hat{i}} \mathbf{S}^{\hat{i}} \,, \tag{21}$$

where $\alpha^{\hat{i}}$ is a small number in the neighbourhood of $\epsilon = 1e-3$ which weights the sensitivities $\mathbf{S}^{\hat{i}}$ at the $\hat{i}$-th optimization iteration. Typically $\alpha^{\hat{i}}$ adapts within the optimization iterations but herein we chose it constant, that is $\alpha^{\hat{i}} = \alpha$ for all $\hat{i}$. The algorithm for the aforentioned gradient-based shape optimization writes,

Algorithm 1: Gradient-based shape optimization

Data: Define perturbation d, sensitivity weight α, maximum number of iterations $n_{\max}$ and a tolerance ϵ

Result: Shape optimal design

Initialization $\hat{i} = 1$;

while $\hat{i} \leq n_{max}$ **do**

  Compute the unperturbed stiffness matrix $\mathbf{K}(\hat{\mathbf{s}}^{\hat{i}})$ using Eq.(8a);

  Solve for the primal field $\mathbf{u}^{\hat{i}}$ using Eq. (7);

  Compute the ojective function $\mathcal{J}$ using Eq. (9);

  if $|\mathcal{J}| \leq \epsilon$ **then**

    break; // Convergence has been achieved

  end

  for *all design variables* $\hat{s}_i^{\hat{i}}$ **do**

    Compute the perturbed stiffness matrix $\mathbf{K}(\hat{\mathbf{s}}^{\hat{i}} + d_i^{\hat{i}})$ with Eq. (8a);

    Compute the derivative $\partial \mathbf{K} / \partial \hat{s}_i^{\hat{i}}$ using Eq. (20);

    Compute the corresponding sensitivity S_i using Eq. (19);

    Recover the unperturbed stiffness matrix $\mathbf{K}(\hat{\mathbf{s}}^{\hat{i}})$;

  end

  Update the design variables using Eq. (21);

  $\hat{i} = \hat{i} + 1$;

end