A Review of the Adjoint Method: Continuous and Discrete approaches

Oscar Peredo

Computer Applications in Science & Engineering Department,

Barcelona Supercomputing Center

December 20, 2010

Index

Index

Lagrange multipliers

Original problem	Modified problem
$\min\{f(x) : h(x) = 0\}$	$\min\{f(x) : h(x) = \vec{\delta}\}$
x^* solution	$x^* + \Delta x$ solution

How can we estimate the variation in f given the variation $\vec{\delta}$?

Lagrange multipliers

Original problem	Modified problem
$\min\{f(x) : h(x) = 0\}$	$\min\{f(x) : h(x) = \vec{\delta}\}$
x^* solution	$x^* + \Delta x$ solution

How can we estimate the variation in f given the variation $\vec{\delta}$?

Lagrange multipliers

Original problem	Modified problem
$\min\{f(x) : h(x) = 0\}$	$\min\{f(x) : h(x) = \vec{\delta}\}$
x^* solution	$x^* + \Delta x$ solution

How can we estimate the variation in f given the variation $\vec{\delta}$?

Lagrange multipliers

Using Karush-Kuhn-Tucker necessary conditions:

$$\nabla f(x^*) = -\nabla h(x^*)\lambda$$

we have:

$$\begin{aligned}
\Delta f &= f(x^* + \Delta x) - f(x^*) \\
&= \nabla f(x^*)^T \Delta x + o(\|\Delta x\|) \\
&= (-\nabla h(x^*)\lambda)^T \Delta x + o(\|\Delta x\|) \\
&= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|)
\end{aligned}$$

and also:

$$\begin{aligned}
\vec{\delta} &= h(x^* + \Delta x) - h(x^*) \\
&= \nabla h(x^*)^T \Delta x + \vec{o}(\|\Delta x\|)
\end{aligned}$$

Lagrange multipliers

Using Karush-Kuhn-Tucker necessary conditions:

$$\nabla f(x^*) = -\nabla h(x^*)\lambda$$

we have:

$$\begin{aligned} \Delta f &= f(x^* + \Delta x) - f(x^*) \\ &= \nabla f(x^*)^T \Delta x + o(\|\Delta x\|) \\ &= (-\nabla h(x^*)\lambda)^T \Delta x + o(\|\Delta x\|) \\ &= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|) \end{aligned}$$

and also:

$$\begin{aligned} \vec{\delta} &= h(x^* + \Delta x) - h(x^*) \\ &= \nabla h(x^*)^T \Delta x + \vec{o}(\|\Delta x\|) \end{aligned}$$

Lagrange multipliers

Using Karush-Kuhn-Tucker necessary conditions:

$$\nabla f(x^*) = -\nabla h(x^*)\lambda$$

we have:

$$\begin{aligned}
\Delta f &= f(x^* + \Delta x) - f(x^*) \\
&= \nabla f(x^*)^T \Delta x + o(\|\Delta x\|) \\
&= (-\nabla h(x^*)\lambda)^T \Delta x + o(\|\Delta x\|) \\
&= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|)
\end{aligned}$$

and also:

$$\begin{aligned}
\vec{\delta} &= h(x^* + \Delta x) - h(x^*) \\
&= \nabla h(x^*)^T \Delta x + \vec{o}(\|\Delta x\|)
\end{aligned}$$

Lagrange multipliers

Using Karush-Kuhn-Tucker necessary conditions:

$$\nabla f(x^*) = -\nabla h(x^*)\lambda$$

we have:

$$\begin{aligned}
\Delta f &= f(x^* + \Delta x) - f(x^*) \\
&= \nabla f(x^*)^T \Delta x + o(\|\Delta x\|) \\
&= (-\nabla h(x^*)\lambda)^T \Delta x + o(\|\Delta x\|) \\
&= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|)
\end{aligned}$$

and also:

$$\begin{aligned}
\vec{\delta} &= h(x^* + \Delta x) - h(x^*) \\
&= \nabla h(x^*)^T \Delta x + \vec{o}(\|\Delta x\|)
\end{aligned}$$

Lagrange multipliers

Then:

$$\begin{aligned}\Delta f &= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|)\\ &= -\lambda^T(\delta - \vec{o}(\|\Delta x\|)) + o(\|\Delta x\|)\\ &= -\lambda^T \delta + o(\|\Delta\|)\\ &\approx -\lambda^T \delta\end{aligned}$$

Lagrange multipliers

Then:

$$\begin{aligned}
\Delta f &= -\lambda^T \nabla h(x^*)^T \Delta x + o(\|\Delta x\|) \\
&= -\lambda^T (\delta - \vec{o}(\|\Delta x\|)) + o(\|\Delta x\|) \\
&= -\lambda^T \delta + o(\|\Delta\|) \\
&\approx -\lambda^T \delta
\end{aligned}$$

Lagrange multipliers

Formally:

Theorem (Sensitivity with equality constraints)

Let (x^*, λ^*) local minimum and associated Lagrange multiplier. Also assume that the second order sufficiency conditions holds for (x^*, λ^*). Consider the family of problems:

$$\min\{f(x) : h(x) = u\}$$

where u is a parameter.
Then, for all u exists $(x(u), \lambda(u))$ local minimum and associated Lagrange multiplier of the problem $\min\{f(x) : h(x) = u\}$, $x^* = x(0)$, $\lambda^* = \lambda(0)$ and

$$\nabla_u f(x(u)) = -\lambda(u)$$

Lagrange multipliers

Formally:

Theorem (Sensitivity with equality constraints)

Let (x^*, λ^*) *local minimum and associated Lagrange multiplier. Also assume that the second order sufficiency conditions holds for* (x^*, λ^*). *Consider the family of problems:*

$$\min\{f(x) : h(x) = u\}$$

where u is a parameter.
Then, for all u exists $(x(u), \lambda(u))$ *local minimum and associated Lagrange multiplier of the problem* $\min\{f(x) : h(x) = u\}$, $x^* = x(0)$, $\lambda^* = \lambda(0)$ *and*

$$\nabla_u f(x(u)) = -\lambda(u)$$

Lagrange multipliers

Formally:

Theorem (Sensitivity with equality constraints)

Let (x^*, λ^*) local minimum and associated Lagrange multiplier. Also assume that the second order sufficiency conditions holds for (x^*, λ^*). Consider the family of problems:

$$\min\{f(x) : h(x) = u\}$$

where u is a parameter.
Then, for all u exists $(x(u), \lambda(u))$ local minimum and associated Lagrange multiplier of the problem $\min\{f(x) : h(x) = u\}$, $x^* = x(0)$, $\lambda^* = \lambda(0)$ and

$$\nabla_u f(x(u)) = -\lambda(u)$$

Index

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \tag{1}$$

  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

- **Central idea**:
  - [illegible]ing derivatives of $J(\gamma, s)$ with respect to design variables [illegible]

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \qquad (1)$$

  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

- **Central idea**:
  - [illegible]ing derivatives of $J(\gamma, s)$ with respect to design variables [illegible]

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \qquad (1)$$

  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

- **Central idea**:
  - ...ing derivatives of $J(\gamma, s)$ with respect to design variables

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \quad (1)$$

  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

- **Central idea**:
  - [illegible]ing derivatives of $J(\gamma, s)$ with respect to design variables
  - [illegible]

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \tag{1}$$

  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

- **Central idea**:
  - Computing derivatives of $J(\gamma, s)$ with respect to design variables using Lagrange multipliers (adjoint variables).

Overview

- **Objective**:
  - We want to solve optimal control problems associated with numerical simulations governed by a partial differential equation:

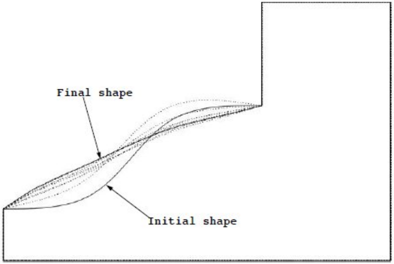
$$\min\{J(\gamma, s) : C(\gamma, s) = 0, (\gamma, s) \in D \times S\} \tag{1}$$

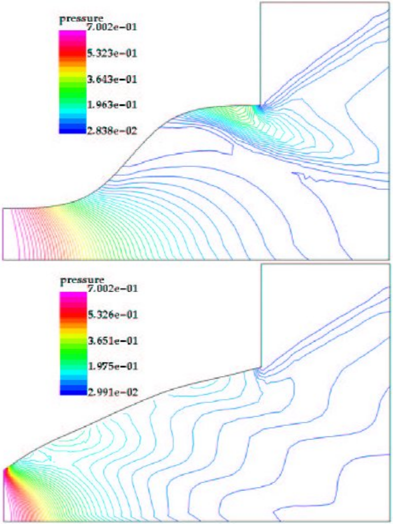
  - Using gradient-based optimization methods for design variables γ:

$$\gamma_{k+1} = \gamma_k - \alpha_k \frac{\partial J}{\partial \gamma}(\gamma_k)$$

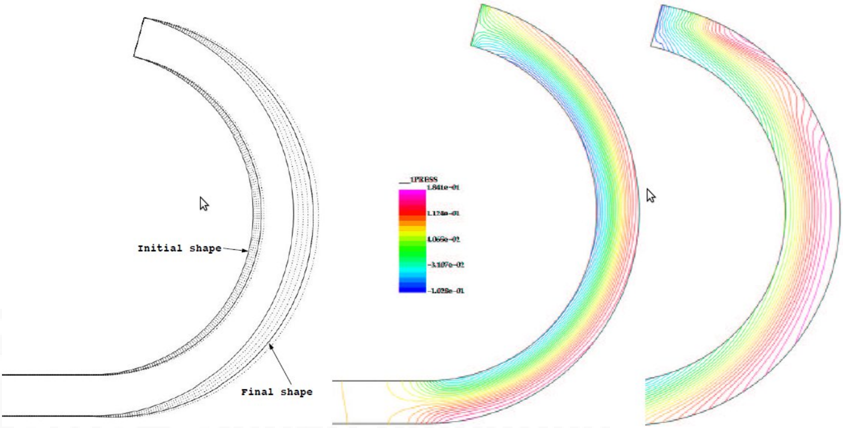
- **Central idea**:
  - Computing derivatives of $J(\gamma, s)$ with respect to design variables using Lagrange multipliers (adjoint variables).

Overview: examples (Nozzle)

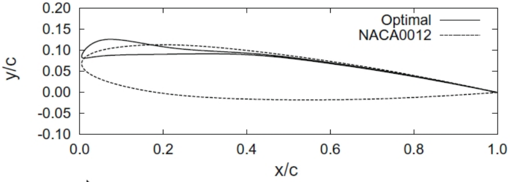


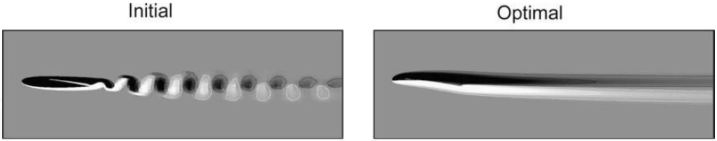


Overview: examples (Air duct)



Overview: examples (Wing airfoil)





Overview: approaches

- **Continuous approach**:
  - First differentiate, then discretize.
- **Discrete approach**:
  - First discretize, then differentiate.

Overview: approaches

- **Continuous approach**:
  - First differentiate, then discretize.
- **Discrete approach**:
  - First discretize, then differentiate.

Overview: approaches

- **Continuous approach**:
  - First differentiate, then discretize.
- **Discrete approach**:
  - First discretize, then differentiate.

Index

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, ...
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, . . .
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, . . .
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, . . .
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, ...
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, ...
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

- **Main task**: formulate the adjoint system.
- **Analytical tools required**:
  - Differential calculus (in Banach spaces for some cases), mathematical identities (Green), integration by parts, geometrical constraints, ...
- **Computational tools required**:
  - PDE solver.
  - Parameterization scheme for design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Continuous approach

Summary:

- Differentiate the continuous original system with respect to the design variables.
- Formulate the adjoint system.
- Solve the adjoint system (discrete).
- Use the adjoint solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Example: Wave equation with homogeneous b.c.

- State equations:

$$
\begin{aligned}
\underbrace{\rho(x)\partial_t^2 u - \nabla\cdot(\mu(x)\nabla u)}_{Lu} &= g, & \Omega\times[0,T] \qquad (2a)\\
u &= 0, & \partial\Omega_1\times[0,T] \qquad (2b)\\
\nabla u\cdot n &= 0, & \partial\Omega_2\times[0,T] \qquad (2c)\\
u(x,0)=\partial_t u(x,0) &= 0, & \Omega\times\{0\} \qquad (2d)
\end{aligned}
$$

- Cost functional:

$$
J(u) = \int_\Omega\int_0^T \underbrace{\frac{1}{2}[u(x,t)-u_0(x,t)]^2\delta(x-\xi)\,dtdV}_{f(u)} \qquad (3)
$$

- Variables:

Example: Wave equation with homogeneous b.c.

- State equations:

$$\underbrace{\rho(x)\partial_t^2 u - \nabla\cdot(\mu(x)\nabla u)}_{Lu} = g, \quad \Omega\times[0,T] \tag{2a}$$

$$u = 0, \quad \partial\Omega_1\times[0,T] \tag{2b}$$

$$\nabla u\cdot n = 0, \quad \partial\Omega_2\times[0,T] \tag{2c}$$

$$u(x,0) = \partial_t u(x,0) = 0, \quad \Omega\times\{0\} \tag{2d}$$

- Cost functional:

$$J(u) = \int_\Omega\int_0^T \underbrace{\frac{1}{2}[u(x,t)-u_0(x,t)]^2\delta(x-\xi)}_{f(u)}\,dtdV \tag{3}$$

- Design variables:

$$(\rho,\mu)$$

Example: Wave equation with homogeneous b.c.

- State equations:

$$
\begin{aligned}
\underbrace{\rho(x)\partial_t^2 u - \nabla \cdot (\mu(x)\nabla u)}_{Lu} &= g, && \Omega \times [0, T] && (2a)\\
u &= 0, && \partial\Omega_1 \times [0, T] && (2b)\\
\nabla u \cdot n &= 0, && \partial\Omega_2 \times [0, T] && (2c)\\
u(x,0) = \partial_t u(x,0) &= 0, && \Omega \times \{0\} && (2d)
\end{aligned}
$$

- Cost functional:

$$
J(u) = \int_\Omega \int_0^T \underbrace{\frac{1}{2}[u(x,t) - u_0(x,t)]^2 \delta(x - \xi)}_{f(u)} \, dtdV \qquad (3)
$$

- Design variables:

$$
(\rho, \mu)
$$

Example: Wave equation with homogeneous b.c.

- State equations:

$$\begin{aligned}
\underbrace{\rho(x)\partial_t^2 u - \nabla\cdot(\mu(x)\nabla u)}_{Lu} &= g, & \Omega\times[0,T] && \text{(2a)}\\
u &= 0, & \partial\Omega_1\times[0,T] && \text{(2b)}\\
\nabla u\cdot n &= 0, & \partial\Omega_2\times[0,T] && \text{(2c)}\\
u(x,0)=\partial_t u(x,0) &= 0, & \Omega\times\{0\} && \text{(2d)}
\end{aligned}$$

- Cost functional:

$$J(u) = \int_\Omega\int_0^T \underbrace{\frac{1}{2}[u(x,t)-u_0(x,t)]^2\delta(x-\xi)}_{f(u)}\,dtdV \qquad (3)$$

- Design variables:

$$(\rho,\mu)$$

Example: Wave equation with homogeneous b.c.

- Adjoint equations:

$$
\begin{aligned}
\underbrace{\rho(x)\partial_t^2\lambda - \nabla\cdot(\mu(x)\nabla\lambda)}_{L^*\lambda} &= -D_u f(u), && \Omega\times[0,T] && \text{(4a)}\\
\lambda &= 0, && \partial\Omega_1\times[0,T] && \text{(4b)}\\
\nabla\lambda\cdot n &= 0, && \partial\Omega_2\times[0,T] && \text{(4c)}\\
\lambda(x,T)=\partial_t\lambda(x,T) &= 0, && \Omega\times\{0\} && \text{(4d)}
\end{aligned}
$$

- Differentiated cost functional:

$$
D_{(\rho,\mu)}J(u) = \int_\Omega\int_0^T \mu'\nabla u\cdot\nabla\lambda\, dt dV - \int_\Omega\int_0^T \rho'\partial_t u\partial_t\lambda\, dt dV \qquad (5)
$$

Example: Wave equation with homogeneous b.c.

- Adjoint equations:

$$
\begin{aligned}
\underbrace{\rho(x)\partial_t^2\lambda - \nabla\cdot(\mu(x)\nabla\lambda)}_{L^*\lambda} &= -D_u f(u), && \Omega\times[0,T] && \text{(4a)}\\
\lambda &= 0, && \partial\Omega_1\times[0,T] && \text{(4b)}\\
\nabla\lambda\cdot n &= 0, && \partial\Omega_2\times[0,T] && \text{(4c)}\\
\lambda(x,T)=\partial_t\lambda(x,T) &= 0, && \Omega\times\{0\} && \text{(4d)}
\end{aligned}
$$

- Differentiated cost functional:

$$
D_{(\rho,\mu)}J(u) = \int_\Omega\int_0^T \mu'\nabla u\cdot\nabla\lambda dtdV - \int_\Omega\int_0^T \rho'\partial_t u\partial_t\lambda dtdV \qquad (5)
$$

Example: Wave equation with homogeneous b.c.

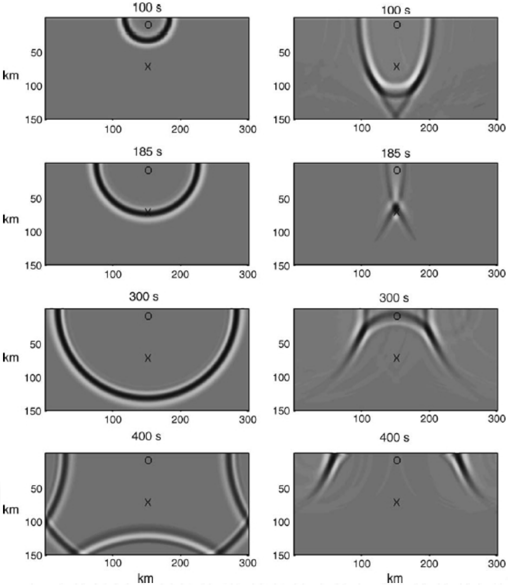
- Adjoint equations:

$$\begin{aligned}
\underbrace{\rho(x)\partial_t^2\lambda - \nabla\cdot(\mu(x)\nabla\lambda)}_{L^*\lambda} &= -D_u f(u), && \Omega\times[0,T] && \text{(4a)}\\
\lambda &= 0, && \partial\Omega_1\times[0,T] && \text{(4b)}\\
\nabla\lambda\cdot n &= 0, && \partial\Omega_2\times[0,T] && \text{(4c)}\\
\lambda(x,T)=\partial_t\lambda(x,T) &= 0, && \Omega\times\{0\} && \text{(4d)}
\end{aligned}$$

- Differentiated cost functional:

$$D_{(\rho,\mu)}J(u) = \int_\Omega\int_0^T \mu'\nabla u\cdot\nabla\lambda dtdV - \int_\Omega\int_0^T \rho'\partial_t u\partial_t\lambda dtdV \tag{5}$$

Example: Wave equation with homogeneous b.c.



Continuous approach: reviewed applications

- Aerodynamics
  - Transonic small disturbances
  - 2D Euler equations using conformal mapping
  - 3D Euler equations
- Meteorology
  - Parameter estimation
  - Athmosphere-Ocean interaction
- Seismology
  - Wave equation
  - veform tomography
  - c anisotropic waves

Continuous approach: reviewed applications

- Aerodynamics
  - Transonic small disturbances
  - 2D Euler equations using conformal mapping
  - 3D Euler equations
- Meteorology
  - Parameter estimation
  - Athmosphere-Ocean interaction
- Seismology
  - Wave equation
  - Waveform tomography
  - ...c anisotropic waves

Continuous approach: reviewed applications

- Aerodynamics
  - Transonic small disturbances
  - 2D Euler equations using conformal mapping
  - 3D Euler equations
- Meteorology
  - Parameter estimation
  - Athmosphere-Ocean interaction
- Seismology
  - Wave equation
  - …veform tomography
  - …c anisotropic waves

Continuous approach: reviewed applications

- Aerodynamics
  - Transonic small disturbances
  - 2D Euler equations using conformal mapping
  - 3D Euler equations
- Meteorology
  - Parameter estimation
  - Athmosphere-Ocean interaction
- Seismology
  - Wave equation
  - Waveform tomography
  - Seismic anisotropic waves

Index

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

- **Main task**: computarized calculation of derivatives and resolution of KKT system.
- **Analytical tools required**:
  - None.
- **Computational tools required**:
  - Automatic differentiation.
  - PDE solver.
  - Parameterization scheme for design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Discrete approach

Summary:

- Solve the original system (discrete).
- Differentiate the discrete original system with respect to the design variables.
- Solve KKT system and obtain adjoint solution.
- Use the adjoint solution and the original solution to calculate derivatives of the cost function $J(\gamma, s)$ with respect to design variables.

Example: Data assimilation in Geophysics

- State equations:

$$\frac{d\theta(t)}{dt} = \frac{1}{h(t)} C_T V (1+K)(T(t)-\theta(t)) \quad (6a)$$

$$\frac{dh(t)}{dt} = K C_T V \frac{T(t)-\theta(t)}{\sigma(t)} + W \quad (6b)$$

$$\frac{d\sigma(t)}{dt} = -\frac{d\theta(t)}{dt} + \gamma \frac{dh(t)}{dt} \quad (6c)$$

$$(\theta(0), h(0), \sigma(0)) = (\theta_0, h_0, \sigma_0) \quad (6d)$$

- Cost functional:

$$J(\theta, h, \sigma) = \sum_{i=0}^{n} \left[\frac{(\theta_i - \tilde{\theta}_i)^2}{\delta_\theta^2} + \frac{(h_i - \tilde{h}_i)^2}{\delta_h^2} + \frac{(\sigma_i - \tilde{\sigma}_i)^2}{\delta_\sigma^2} \right] \quad (7)$$

$$\frac{(\bar{\theta} - \theta_{clim})^2}{\delta^2_{\theta_{clim}}} + \frac{(\bar{H} - H_{clim})^2}{\delta^2_{H_{clim}}} \quad (8)$$

- Design variables:

$$\gamma = (\theta_0, h_0, \sigma_0, K, W)$$

Example: Data assimilation in Geophysics

- State equations:

$$\frac{d\theta(t)}{dt} = \frac{1}{h(t)} C_T V(1+K)(T(t)-\theta(t)) \tag{6a}$$

$$\frac{dh(t)}{dt} = K C_T V \frac{T(t)-\theta(t)}{\sigma(t)} + W \tag{6b}$$

$$\frac{d\sigma(t)}{dt} = -\frac{d\theta(t)}{dt} + \gamma \frac{dh(t)}{dt} \tag{6c}$$

$$(\theta(0), h(0), \sigma(0)) = (\theta_0, h_0, \sigma_0) \tag{6d}$$

- Cost functional:

$$J(\theta, h, \sigma) = \sum_{i=0}^{n} \left[\frac{(\theta_i - \tilde{\theta}_i)^2}{\delta_\theta^2} + \frac{(h_i - \tilde{h}_i)^2}{\delta_h^2} + \frac{(\sigma_i - \tilde{\sigma}_i)^2}{\delta_\sigma^2} \right] \tag{7}$$

$$+ \frac{(\bar{\theta} - \theta_{clim})^2}{\delta_{\theta_{clim}}^2} + \frac{(\bar{H} - H_{clim})^2}{\delta_{H_{clim}}^2} \tag{8}$$

- Design variables:

$$\gamma = (\theta_0, h_0, \sigma_0, K, W)$$

Example: Data assimilation in Geophysics

- State equations:

$$\frac{d\theta(t)}{dt} = \frac{1}{h(t)} C_T V(1+K)(T(t)-\theta(t)) \tag{6a}$$

$$\frac{dh(t)}{dt} = K C_T V \frac{T(t)-\theta(t)}{\sigma(t)} + W \tag{6b}$$

$$\frac{d\sigma(t)}{dt} = -\frac{d\theta(t)}{dt} + \gamma \frac{dh(t)}{dt} \tag{6c}$$

$$(\theta(0), h(0), \sigma(0)) = (\theta_0, h_0, \sigma_0) \tag{6d}$$

- Cost functional:

$$J(\theta, h, \sigma) = \sum_{i=0}^{n} \left[\frac{(\theta_i - \tilde{\theta}_i)^2}{\delta_\theta^2} + \frac{(h_i - \tilde{h}_i)^2}{\delta_h^2} + \frac{(\sigma_i - \tilde{\sigma}_i)^2}{\delta_\sigma^2} \right] \tag{7}$$

$$+ \frac{(\bar{\theta} - \theta_{clim})^2}{\delta^2_{\theta_{clim}}} + \frac{(\bar{H} - H_{clim})^2}{\delta^2_{H_{clim}}} \tag{8}$$

- Design variables:

$$\gamma = (\theta_0, h_0, \sigma_0, K, W)$$

Example: Data assimilation in Geophysics

- State equations:

$$\frac{d\theta(t)}{dt} = \frac{1}{h(t)} C_T V (1+K)(T(t)-\theta(t)) \tag{6a}$$

$$\frac{dh(t)}{dt} = K C_T V \frac{T(t)-\theta(t)}{\sigma(t)} + W \tag{6b}$$

$$\frac{d\sigma(t)}{dt} = -\frac{d\theta(t)}{dt} + \gamma \frac{dh(t)}{dt} \tag{6c}$$

$$(\theta(0), h(0), \sigma(0)) = (\theta_0, h_0, \sigma_0) \tag{6d}$$

- Cost functional:

$$J(\theta, h, \sigma) = \sum_{i=0}^{n} \left[\frac{(\theta_i - \tilde{\theta}_i)^2}{\delta_\theta^2} + \frac{(h_i - \tilde{h}_i)^2}{\delta_h^2} + \frac{(\sigma_i - \tilde{\sigma}_i)^2}{\delta_\sigma^2} \right] \tag{7}$$

$$+ \frac{(\bar{\theta} - \theta_{clim})^2}{\delta^2_{\theta_{clim}}} + \frac{(\bar{H} - H_{clim})^2}{\delta^2_{H_{clim}}} \tag{8}$$

- Design variables:

$$\gamma = (\theta_0, h_0, \sigma_0, K, W)$$

Example: Data assimilation in Geophysics

- Discretization (Euler scheme):

$$
\begin{aligned}
\theta_{i+1} &= \theta_i + \frac{\tau}{h_i} C_T V(1+K)(T_i - \theta_i) \\
h_{i+1} &= h_i + \tau K C_T V \frac{t_i - \theta_i}{\sigma_i} + \tau W \\
\sigma_{i+1} &= \sigma_i - \frac{\tau}{h_i} C_T V(1+K)(T_i - \theta_i) + \gamma\tau K C_T V \frac{T_i - \theta_i}{\sigma_i} + \gamma\tau W
\end{aligned}
$$

- KKT equations:

$$
\begin{aligned}
\nabla L &= 0 \\
&\Downarrow \\
Ax &= d
\end{aligned}
$$

Example: Data assimilation in Geophysics

- Discretization (Euler scheme):

$$
\begin{aligned}
\theta_{i+1} &= \theta_i + \frac{\tau}{h_i} C_T V (1+K)(T_i - \theta_i) \\
h_{i+1} &= h_i + \tau K C_T V \frac{t_i - \theta_i}{\sigma_i} + \tau W \\
\sigma_{i+1} &= \sigma_i - \frac{\tau}{h_i} C_T V (1+K)(T_i - \theta_i) + \gamma\tau K C_T V \frac{T_i - \theta_i}{\sigma_i} + \gamma\tau W
\end{aligned}
$$

- KKT equations:

$$
\begin{aligned}
\nabla L &= 0 \\
&\Downarrow \\
A\lambda &= d
\end{aligned}
$$

Example: Data assimilation in Geophysics

- Discretization (Euler scheme):

$$
\begin{aligned}
\theta_{i+1} &= \theta_i + \frac{\tau}{h_i} C_T V (1+K)(T_i - \theta_i) \\
h_{i+1} &= h_i + \tau K C_T V \frac{t_i - \theta_i}{\sigma_i} + \tau W \\
\sigma_{i+1} &= \sigma_i - \frac{\tau}{h_i} C_T V (1+K)(T_i - \theta_i) + \gamma \tau K C_T V \frac{T_i - \theta_i}{\sigma_i} + \gamma \tau W
\end{aligned}
$$

- KKT equations:

$$
\begin{aligned}
\nabla L &= 0 \\
&\Downarrow \\
A\lambda &= d
\end{aligned}
$$

Discrete approach: applications

- Meteorology
  - Data assimilation
  - Athmosphere-Ocean interaction
- Aerodynamics
  - Shape parameterization
  - Derivatives with respect to node positions
- Finance
  - Calculation of price sensitivities (*Greeks*)

Discrete approach: applications

- Meteorology
  - Data assimilation
  - Athmosphere-Ocean interaction
- Aerodynamics
  - Shape parameterization
  - Derivatives with respect to node positions
- Finance
  - Calculation of price sensitivities (*Greeks*)

Discrete approach: applications

- Meteorology
  - Data assimilation
  - Athmosphere-Ocean interaction
- Aerodynamics
  - Shape parameterization
  - Derivatives with respect to node positions
- Finance
  - Calculation of price sensitivities (*Greeks*)

Discrete approach: applications

- Meteorology
  - Data assimilation
  - Athmosphere-Ocean interaction
- Aerodynamics
  - Shape parameterization
  - Derivatives with respect to node positions
- Finance
  - Calculation of price sensitivities (*Greeks*)

Index

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
    - Physical meaning of adjoint variables is clearer.
    - Each problem has its own adjoint (some are very complex).
    - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
    - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
    - More clean and easy to understand.
    - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Conclusions

- Continuous approach:
  - Physical meaning of adjoint variables is clearer.
  - Each problem has its own adjoint (some are very complex).
  - Adjoint solver and Original solver are the same in many cases (reusable code).
- Discrete approach:
  - Automatic differentiation $\Rightarrow$ Exact gradient calculation.
  - More clean and easy to understand.
  - More computational resources involved.

Comments

- Powerful tool to calculate derivatives in complex systems.
- Not all problems can use this method (a study of the design variables must be done).
- Gradient-free methods (simulated annealing or genetic algorithms) are an alternative in some cases.

Comments

- Powerful tool to calculate derivatives in complex systems.
- Not all problems can use this method (a study of the design variables must be done).
- Gradient-free methods (simulated annealing or genetic algorithms) are an alternative in some cases.

Comments

- Powerful tool to calculate derivatives in complex systems.
- Not all problems can use this method (a study of the design variables must be done).
- Gradient-free methods (simulated annealing or genetic algorithms) are an alternative in some cases.

Thanks for your attention!

Bibliography

H. Brezis.
Analyse fonctionnelle théorie et applications.
Masson, 1992.

L. D. Evans.
Partial Differential Equations.
AMS, 2010.

A. Fichtner, H.-P. Bunge, and H. Igel.
The adjoint method in seismology: I. theory.
Physics of The Earth and Planetary Interiors, 157(1-2):86 – 104, 2006.

R. Fiedler.
Computational Aspects of the Adjoint Method.
PhD thesis, University of Oklahoma, 1998.

M. B. Giles and N. A. Pierce.
Adjoint equations in cfd: duality, boundary conditions and solution behaviour.
AIAA Paper 97-1850, 1997.

L. Hascoët, M. Vázquez, B. Koobus, and A. Dervieux.
A framework for adjoint-based shape design and error control.
CFD Journal, 16(4):454–464, 2008.

A. Jameson.
Aerodynamic design via control theory.
Journal of Scientific Computing, 3:233–260, September 1988.

O. Pironneau.
Optimal Shape Design for Elliptic Systems.
Springer-Verlag, New York, 1984.

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$\begin{aligned} D_\gamma Lu &= D_\rho Lu + D_\mu Lu \\ &= 0 \qquad (D_\gamma g = 0) \end{aligned}$$

$$\begin{aligned} D_\rho Lu &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\ &\quad - \nabla \cdot (\mu \nabla (D_\rho u)) \end{aligned}$$

$$\begin{aligned} D_\mu Lu &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\ &\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u)) \end{aligned}$$

$$\rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \tag{11}$$

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$
\begin{aligned}
D_\gamma L u &= D_\rho L u + D_\mu L u \\
&= 0 \qquad (D_\gamma g = 0)
\end{aligned}
$$

$$
\begin{aligned}
D_\rho L u &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\
&\quad - \nabla \cdot (\mu \nabla (D_\rho u))
\end{aligned}
$$

$$
\begin{aligned}
D_\mu L u &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\
&\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u))
\end{aligned}
$$

$$
\rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \quad (11)
$$

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$\begin{aligned} D_\gamma L u &= D_\rho L u + D_\mu L u \\ &= 0 \qquad (D_\gamma g = 0) \end{aligned}$$

$$\begin{aligned} D_\rho L u &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\ &\quad - \nabla \cdot (\mu \nabla (D_\rho u)) \end{aligned}$$

$$\begin{aligned} D_\mu L u &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\ &\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u)) \end{aligned}$$

It follows:

$$D_\gamma L u = \rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \quad (11)$$

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$\begin{aligned} D_\gamma Lu &= D_\rho Lu + D_\mu Lu \\ &= 0 \qquad (D_\gamma g = 0) \end{aligned}$$

$$\begin{aligned} D_\rho Lu &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\ &\quad - \nabla \cdot (\mu \nabla (D_\rho u)) \end{aligned}$$

$$\begin{aligned} D_\mu Lu &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\ &\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u)) \end{aligned}$$

It follows:

$$D_\gamma Lu = \rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \quad (11)$$

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$\begin{aligned} D_\gamma Lu &= D_\rho Lu + D_\mu Lu \\ &= 0 \qquad (D_\gamma g = 0) \end{aligned}$$

$$\begin{aligned} D_\rho Lu &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\ &\quad - \nabla \cdot (\mu \nabla (D_\rho u)) \end{aligned}$$

$$\begin{aligned} D_\mu Lu &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\ &\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u)) \end{aligned}$$

It follows:

$$D_\gamma Lu = \rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \quad (11)$$

Continuous approach: calculating the adjoint

First step: differentiate equation (2a) with respect to $\gamma = (\rho, \mu)$:

$$
\begin{aligned}
D_\gamma L u &= D_\rho L u + D_\mu L u \\
&= 0 \qquad (D_\gamma g = 0)
\end{aligned}
$$

$$
\begin{aligned}
D_\rho L u &= \rho' \frac{\partial^2 u}{\partial t^2} + \rho \frac{\partial^2}{\partial t^2}(D_\rho u) \\
&\quad - \nabla \cdot (\mu \nabla (D_\rho u))
\end{aligned}
$$

$$
\begin{aligned}
D_\mu L u &= \rho \frac{\partial^2}{\partial t^2}(D_\mu u) \\
&\quad - \nabla \cdot (\mu' \nabla u) - \nabla \cdot (\mu \nabla (D_\mu u))
\end{aligned}
$$

It follows:

$$
D_\gamma L u = \rho' \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\mu' \nabla u) + \rho \frac{\partial^2}{\partial t^2}(D_\gamma u) - \nabla \cdot (\mu \nabla (D_\gamma u)) \tag{11}
$$

Continuous approach: calculating the adjoint

Second step: multiply (11) by a test function v and integrate over $\Omega \times [0, T]$.

$$
\begin{aligned}
\int_\Omega \int_0^T v D_\gamma(Lu)\,dtdV = & \\
& \int_\Omega \int_0^T v\rho(x)' \frac{\partial^2 u}{\partial t^2} dtdV && (12)\\
& - \int_\Omega \int_0^T v\nabla \cdot (\mu'(x)\nabla u) dtdV && (13)\\
& + \int_\Omega \int_0^T v\rho(x) \frac{\partial^2}{\partial t^2}(D_\gamma u) dtdV && (14)\\
& - \int_\Omega \int_0^T v\nabla \cdot (\mu(x)\nabla (D_\gamma u)) dtdV && (15)
\end{aligned}
$$

Continuous approach: calculating the adjoint

Second step: multiply (11) by a test function v and integrate over $\Omega \times [0, T]$.

$$\int_\Omega \int_0^T v D_\gamma(Lu) dt dV =$$

$$\int_\Omega \int_0^T v \rho(x)' \frac{\partial^2 u}{\partial t^2} dt dV \tag{12}$$

$$- \int_\Omega \int_0^T v \nabla \cdot (\mu'(x) \nabla u) dt dV \tag{13}$$

$$+ \int_\Omega \int_0^T v \rho(x) \frac{\partial^2}{\partial t^2} (D_\gamma u) dt dV \tag{14}$$

$$- \int_\Omega \int_0^T v \nabla \cdot (\mu(x) \nabla (D_\gamma u)) dt dV \tag{15}$$

Continuous approach: calculating the adjoint

Second step: multiply (11) by a test function v and integrate over $\Omega \times [0, T]$.

$$
\begin{aligned}
\int_\Omega \int_0^T v D_\gamma (Lu) dtdV = & \\
& \int_\Omega \int_0^T v \rho(x)' \frac{\partial^2 u}{\partial t^2} dtdV \quad (12) \\
& - \int_\Omega \int_0^T v \nabla \cdot (\mu'(x) \nabla u) dtdV \quad (13) \\
& + \int_\Omega \int_0^T v \rho(x) \frac{\partial^2}{\partial t^2} (D_\gamma u) dtdV \quad (14) \\
& - \int_\Omega \int_0^T v \nabla \cdot (\mu(x) \nabla (D_\gamma u)) dtdV \quad (15)
\end{aligned}
$$

Continuous approach: calculating the adjoint

Second step: multiply (11) by a test function v and integrate over $\Omega \times [0, T]$.

$$
\begin{aligned}
\int_\Omega \int_0^T v D_\gamma (Lu) dtdV \quad = & \\
& \int_\Omega \int_0^T v \rho(x)' \frac{\partial^2 u}{\partial t^2} dtdV \qquad (12) \\
& - \int_\Omega \int_0^T v \nabla \cdot (\mu'(x) \nabla u) dtdV \qquad (13) \\
& + \int_\Omega \int_0^T v \rho(x) \frac{\partial^2}{\partial t^2} (D_\gamma u) dtdV \qquad (14) \\
& - \int_\Omega \int_0^T v \nabla \cdot (\mu(x) \nabla (D_\gamma u)) dtdV \qquad (15)
\end{aligned}
$$

Continuous approach: calculating the adjoint

Applying integration by parts to each term and fixing boundary and initial conditions to the test function v, gives us:

$$\begin{aligned}
\int_\Omega \int_0^T D_\gamma u \cdot (\rho \frac{\partial^2 v}{\partial t^2} - \nabla \cdot (\mu \nabla v)) dtdV \\
- \int_\Omega \int_0^T \rho' \frac{\partial u}{\partial t} \frac{\partial v}{\partial t} dtdV \\
+ \int_\Omega \int_0^T \mu' \nabla u \cdot \nabla v dtdV &= 0 \\
v &= 0, \qquad \partial\Omega_1 \times [0, T] \\
\frac{\partial v}{\partial n} &= 0, \qquad \partial\Omega_2 \times [0, T] \\
v(x, T) = \frac{\partial v}{\partial t}(x, T) &= 0, \qquad \Omega \times T
\end{aligned}$$

Continuous approach: calculating the adjoint

Applying integration by parts to each term and fixing boundary and initial conditions to the test function v, gives us:

$$
\begin{aligned}
\int_\Omega \int_0^T D_\gamma u \cdot (\rho \frac{\partial^2 v}{\partial t^2} - \nabla \cdot (\mu \nabla v)) dtdV & & \\
- \int_\Omega \int_0^T \rho' \frac{\partial u}{\partial t} \frac{\partial v}{\partial t} dtdV & & \\
+ \int_\Omega \int_0^T \mu' \nabla u \cdot \nabla v dtdV &= 0 & \\
v &= 0, & \partial\Omega_1 \times [0, T] \\
\frac{\partial v}{\partial n} &= 0, & \partial\Omega_2 \times [0, T] \\
v(x, T) = \frac{\partial v}{\partial t}(x, T) &= 0, & \Omega \times T
\end{aligned}
$$

Discrete approach: calculating λ

Relative to the problem (1), we define the Lagrangian as:

$$L(\gamma, s) = J(\gamma, s) - \lambda^T C(\gamma, s) \tag{16}$$

If appropriate conditions are met (continuity and differentiability), we can discretize $J(\gamma, s)$ and $C(\gamma, s)$, and write the Karush-Kuhn-Tucker conditions for the discretized problem:

$$\begin{aligned} \nabla L(\gamma, s) &= 0 \qquad (17) \\ C(\gamma, s) &= 0 \qquad (18) \end{aligned}$$

or equivalently:

$$\begin{aligned} \nabla_\gamma J(\gamma, s) - (\lambda_\gamma)^T \nabla_\gamma C(\gamma, s) &= 0 \qquad (19) \\ \nabla_s J(\gamma, s) - (\lambda_s)^T \nabla_s C(\gamma, s) &= 0 \qquad (20) \\ C(\gamma, s) &= 0 \qquad (21) \end{aligned}$$

Discrete approach: calculating λ

Let's consider that the state variables s depend on the design variables γ, ie, $s(\gamma) : D \subset \mathbb{R}^{d_\gamma} \rightarrow \mathbb{R}^{d_s}$, then, the cost function J can be written:

$$j(\gamma) = J(\gamma, s(\gamma)) \tag{22}$$

and the derivative of the state equation with respect of γ is:

$$\underbrace{\nabla_\gamma C(\gamma, s(\gamma))}_{d_C \times d_\gamma} + \underbrace{\nabla_s C(\gamma, s(\gamma))}_{d_C \times d_s} \underbrace{\nabla_\gamma s(\gamma)}_{d_s \times d_\gamma} = 0 \tag{23}$$

If $d_C = d_s$ holds (number of state equations equal to the number of state variables) and assuming that $\nabla_s C(\gamma, s(\gamma))$ is invertible, it follows that:

$$\nabla_\gamma s(\gamma) = -\left(\nabla_s C(\gamma, s(\gamma))\right)^{-1} \nabla_\gamma C(\gamma, s(\gamma)) \tag{24}$$

Discrete approach: calculating λ

Derivating j we get:

$$D_\gamma j(\gamma) \quad = \quad \nabla_\gamma J(\gamma, s(\gamma)) + \nabla_s J(\gamma, s(\gamma)) \nabla_\gamma s(\gamma) \qquad (25)$$

and replacing (24) into (25), and using the equation (20) from the KKT system in order to define $\nabla_s J$, it follows:

$$\begin{aligned} D_\gamma j(\gamma) &= \nabla_\gamma J(\gamma, s(\gamma)) - \nabla_s J(\gamma, s(\gamma)) \left(\nabla_s C(\gamma, s(\gamma))\right)^{-1} \nabla_\gamma C(\gamma, s(\gamma)) \\ &= \nabla_\gamma J(\gamma, s(\gamma)) - \lambda_s^T \nabla_s C(\gamma, s(\gamma)) \left(\nabla_s C(\gamma, s(\gamma))\right)^{-1} \nabla_\gamma C(\gamma, s(\gamma) \\ &= \nabla_\gamma J(\gamma, s(\gamma)) - \lambda_s^T \nabla_\gamma C(\gamma, s(\gamma)) \end{aligned} \qquad (26)$$

[GP97],[Pir84],[Bre92],[Eva10],[Jam88],[HVKD08], [Fie98],[FBI06]