Numerical simulation for wave equations using continuous and discontinuous Galerkin method

1 Some basic concepts for numerical simulation

Let us consider a partial difference equation in time and space $\mathcal{D}u = 0$, and a numerical method $\tilde{\mathcal{D}}_h U = 0$ to solve the equation, in which h denotes the size of the mesh.

Definition (Consistency)**.** Let u^* denote the accurate solution of equation $\mathcal{D}u = 0$, then we define $\tilde{\mathcal{D}}_h u^*$ as the *local truncation error* of the numerical method. If

$$\lim_{h\to 0} ||\tilde{\mathcal{D}}_h u^*|| = 0$$

, then we say that the numerical method is *consistent*. Furthermore, if

$$||\tilde{\mathcal{D}}_h u^*|| \sim \mathcal{O}(h^p)$$

as $h \to 0$, we say that the method is *of order p*.

Definition (Convergence)**.** Let the accurate solution of the equation be u^* and the numerical solution be U_h, we say that the method is *convergent* in norm $||\cdot||$ if

$$\lim_{h\to 0} ||u^*(\cdot, t) - U_h(\cdot, t)|| = 0, \qquad \text{for } \forall t \geq 0.$$

Definition (Stablity)**.** Suppose the change in initial value δU^0 result in the change in numerical solution δU_h, we say that the method is *stable* if for $\forall T \geq 0$, $\exists C_s$ and $h_0 > 0$ s.t.

$$\frac{||\delta U_h(\cdot, t)||}{||\delta U^0||} \leq C_s, \qquad \text{for } \forall k < k_0.$$

The following theorem states that for linear equation, it is easy to formulate a convergent method: just construct a consistent scheme and guarantee that the scheme is stable while decreasing the mesh size.

Theorem (Lax equivalence theorem)**.** *For a linear consistent difference method of a linear equation, it is convergent if and only if it is stable.*

Example. The 1-d advection equation

$$\begin{aligned} &\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \\ &u(x, 0) = f(x). \end{aligned}$$

The accurate solution is the shape f propagating in the $+x$ direction with velocity 1:

$$u(x, t) = f(x - t).$$

we use the following two explicit schemes to solve it (both consistent), the first one is an upwind scheme:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{U_i^n - U_{i-1}^n}{\Delta x} = 0 \tag{1}$$

for time step $t = n\Delta t$ and spatial point $x = i\Delta x$, and a downwind scheme:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{U_{i+1}^n - U_i^n}{\Delta x} = 0. \tag{2}$$

We find that the upwind scheme is stable when the CFL condition $\Delta t < \Delta x$ is satisfied, and thus converges to the right solution, but the result exhibits dissipation feature (please keep this in mind, we will come back to this very important feature later!). The downwind scheme is unconditionally instable, and thus does not converge. We will talk about the reason later.

Example. Unfortunately nonlinear equations are far more complicated. Now let us consider the Burgers' equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$
$$u(x,0) = \begin{cases} 1.2, & x < 0 \\ 0.4, & x \geq 0. \end{cases}$$

The first scheme we try is an upwind approximation to the spatial derivative:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + U_i^n \frac{U_i^n - U_{i-1}^n}{\Delta x} = 0 \tag{3}$$

for time step $t = n\Delta t$ and spatial point $x = i\Delta x$. It can be verified that this scheme is consistent, we can tune the time step to achieve stability, but the numerical solution converges to a wrong solution (compared with the known accurate solution to Burgers' equation). This example shows the difficulty to deal with nonlinear problems: we construct a consistent scheme and the numerical result stably converges to a solution, but this solution may be wrong!

The key to tackle this problem turns out to be construct a conservative method. The Burgers' equation can be reformulated into

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} = 0,$$

then we can write a scheme in conservation form:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{(U_i^n)^2 - (U_{i-1}^n)^2}{2\Delta x} = 0 \tag{4}$$

Definition (Conservation form)**.** If a numerical method is in the form

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{F(U_{j-p}^n, \ldots, U_{j+q}^n) - F(U_{j-p-1}^n, \ldots, U_{j+q-1}^n)}{\Delta x} = 0,$$

we say that the method is in *conservation form. F* is called *numerical flux function.*

The following theorem states that a conservative method converges to the right solution provided that it converges:

Theorem (Lax-Wendroff)**.** *If the numerical solution u_h is computed with a consistent and conservative method, and u_h converges to a function u as $h \to 0$, then u is a weak solution of the conservation law.*

2 A brief introduction to continuous Galerkin and spectral element method

2.1 Elastodynamic equation

From now on we focus on a particular type of wave equation – the elastodynamic equation:

$$\rho(\boldsymbol{x})\ddot{\boldsymbol{u}}(\boldsymbol{x},t) - \nabla\cdot(\boldsymbol{\Lambda}(x):\nabla\boldsymbol{u}(\boldsymbol{x},t)) = \boldsymbol{f}(\boldsymbol{x},t) \tag{EQ}$$

for $(\boldsymbol{x},t)\in\Omega\times I$ in which $I=[0,+\infty)$, with initial condition:

$$\boldsymbol{u}(\boldsymbol{x},0)=\dot{\boldsymbol{u}}(\boldsymbol{x},0)=0 \qquad \text{for } \forall\boldsymbol{x}\in\Omega \tag{IC}$$

and boundary condition of Dirichlet type

$$\boldsymbol{u}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}\in\partial\Omega}=0 \qquad \text{for } \forall t\in I \tag{BC1}$$

or Neumann type

$$\boldsymbol{n}\cdot(\boldsymbol{\Lambda}(x):\nabla\boldsymbol{u}(\boldsymbol{x},t))\Big|_{\boldsymbol{x}\in\partial\Omega}=0 \qquad \text{for } \forall t\in I. \tag{BC2}$$

ρ is density, $\boldsymbol{\Lambda}$ is the stiffness tensor of 4-order, $\boldsymbol{f}$ is body-force per volume and Ω is the domain to solve the equation. This is the form we are familiar with, but a very crucial problem for this form is that some physically reasonable situation is excluded. In order to be represented in this form, $\boldsymbol{u}$ must have a well-defined second-order spatial derivative, i.e., each component of $\boldsymbol{u}$ must be C^2. However, some wave propagation problem does not give rise to displacement in C^2 but still has physical meanings, as shown in the following example.

Example. Consider 1-d wave equation

$$\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x^2}=0$$

with initial condition

$$u(x,0)=f(x), \qquad \frac{\partial u(x,0)}{\partial t}=0.$$

The solution is the shape of f propagating forward and backward:

$$u(x,t)=\frac{1}{2}\left(f(x-t)+f(x+t)\right).$$

We can choose f to be a function that is not C^2, e.g., a tent function, the problem still physically makes sense, but the second-order derivative in (EQ) is not well-defined.

2.2 Weak derivative and Sobolev space

Definition (Weak derivative)**.** Let $\Omega\subset\mathbb{R}^n$ be an open set, if u and v satisfy

$$\int_\Omega v\phi \mathrm{d}\boldsymbol{x}=-\int_\Omega u\frac{\partial\phi}{\partial x_i}\mathrm{d}\boldsymbol{x} \qquad \text{for } \forall\phi\in C_0^\infty(\Omega),$$

then we say that v is a weak derivative of u. We denote

$$v=\frac{\partial u}{\partial x_i}.$$

Remark. If u is differentiable in classical sense, then it is weakly differentiable, and its classical derivative is a weak derivative. Conversely, this is not true. A weakly differentiable function does not even have classical derivative. See the example below.

Example. Piecewise polynomial which forms a globally continuous function is weakly differentiable. Let $\Omega = \bigcup_e \Omega_e$, $u\Big|_{\Omega_e} = p_e \in \mathbb{P}(\Omega_e)$, and $u \in C(\Omega)$, then u is weakly differentiable, and

$$\frac{\partial u}{\partial x_i}\bigg|_{\Omega_e} = \frac{\partial p_e}{\partial x_i}.$$

Note that the global continuity is necessary for the function to be weakly differentiable.

Definition (Sobolev space)**.** Let open set $\Omega \subset \mathbb{R}^n$. The Sobolev space $H^1(\Omega)$ is defined by

$$H^1(\Omega) = \left\{ u \,\middle|\, \left(\int_\Omega |u|^2 + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right|^2 \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{2}} < \infty \right\}.$$

$H_0^1(\Omega)$ is defined by

$$H_0^1(\Omega) = \left\{ u \,\middle|\, u \in H^1(\Omega), u\Big|_{\partial\Omega} = 0 \right\}.$$

Moreover, H^1 and H_0^1 are Hilbert spaces with inner product

$$\langle u, v \rangle = \int_\Omega uv + \sum_{i=1}^n \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \mathrm{d}\boldsymbol{x},$$

and the norm induced by inner product is given by

$$||u|| = \sqrt{\langle u, u \rangle} = \left(\int_\Omega |u|^2 + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right|^2 \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{2}}.$$

Example. The space of piecewise polynomial with global continuity is a subspace of Sobolev space:

$$\left\{ u \,\middle|\, u\Big|_{\Omega_e} = p_e \in \mathbb{P}(\Omega_e), u \in C(\Omega) \right\} \subset H_1(\Omega)$$

in which $\Omega = \bigcup_e \Omega_e$.

2.3 Strong and weak form of the elastodynamic equation

Definition (Strong form)**.** The *strong form* of the elastodynamic equation (EQ) is

$$\begin{array}{ll} \text{Find} & \boldsymbol{u}(\boldsymbol{x}, t) \in \left(C^2(\bar{\Omega})\right)^n \\ \text{s.t.} & \text{equation (EQ) holds for } \forall t \in I, \\ & \text{initial condition (IC) holds for } \forall \boldsymbol{x} \in \Omega \text{ and} \\ & \text{boundary condition (BC1) or (BC2) holds for } \forall t \in I. \end{array} \tag{*}$$

The solution to the strong form (*) is the *classical solution*.

Definition (Weak form)**.** The *Weak form* of the elastodynamic equation (EQ) is

$$\begin{aligned} &\text{Find} \quad \boldsymbol{u}(\boldsymbol{x},t) \in V^n \\ &\text{s.t.} \quad \int_\Omega \rho(\boldsymbol{x})\ddot{\boldsymbol{u}}(\boldsymbol{x},t)\cdot\boldsymbol{v}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + a\left(\boldsymbol{u}(\boldsymbol{x},t),\boldsymbol{v}(\boldsymbol{x})\right) = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t)\cdot\boldsymbol{v}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \qquad (**) \\ &\qquad \text{holds for } \forall\, t \in I \text{ and } \forall\, \boldsymbol{v}(\boldsymbol{x}) \in V^n, \end{aligned}$$

in which $V^n = \left(H_0^1(\Omega)\right)^n$ for Dirichlet boundary condition and $V^n = \left(H^1(\Omega)\right)^n$ for Neumann boundary condition, and the symmetric bilinear operator $a(\cdot,\cdot)$ is defined by

$$a\left(\boldsymbol{u}(\boldsymbol{x}),\boldsymbol{v}(\boldsymbol{x})\right) = \int_\Omega \nabla\boldsymbol{u}(\boldsymbol{x}) : \boldsymbol{\Lambda}(\boldsymbol{x}) : \nabla\boldsymbol{v}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.$$

Function $\boldsymbol{v}$ is usually referred to as *test function.* Solution of weak form (**) is *weak solution.*

Theorem (Equivalence of classical and weak solution)**.** *Suppose* $\boldsymbol{f} \in \left(C(\bar{\Omega})\right)^n$*, if* $\boldsymbol{u} \in \left(C^2(\bar{\Omega})\right)^n$ *for* $\forall\, t \in I$ *is classical solution, then it is a weak solution. Conversely, if a weak solution* $\boldsymbol{u}$ *happens to be in* $\left(C^2(\bar{\Omega})\right)^n$*, then it is classical solution.*

3 Galerkin method and spectral-element discretization

Definition (Galerkin method)**.** Let V_h^n be a finite-dimensional subspace of V^n, then the *Galerkin method* to solve (**) is

$$\begin{aligned} &\text{Find} \quad \boldsymbol{u}_h(\boldsymbol{x},t) \in V_h^n \\ &\text{s.t.} \quad \int_\Omega \rho(\boldsymbol{x})\ddot{\boldsymbol{u}}_h(\boldsymbol{x},t)\cdot\boldsymbol{v}_h(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + a\left(\boldsymbol{u}_h(\boldsymbol{x},t),\boldsymbol{v}_h(\boldsymbol{x})\right) = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t)\cdot\boldsymbol{v}_h(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \qquad (***) \\ &\qquad \text{holds for } \forall\, t \in I \text{ and } \forall\, \boldsymbol{v}_h(\boldsymbol{x}) \in V_h^n. \end{aligned}$$

Remark. The Galerkin method is just a finite-dimensional version of the weak form, it has the following important features: the space we use to approximate the solution is the same as that to approximate the test function, and both are subspace of V^n.

To solve this semi-discrete form, we choose a set of basis $\{\boldsymbol{\phi}_i(\boldsymbol{x}), i = 1,2,\ldots,N\}$, in which N is the degree of freedom. Then we can write

$$\boldsymbol{u}_h(\boldsymbol{x},t) = \sum_{i=1}^{N} U_i(t)\boldsymbol{\phi}_i(\boldsymbol{x}).$$

We let the test function to be

$$\boldsymbol{v}_h(\boldsymbol{x}) = \boldsymbol{\phi}_j(\boldsymbol{x}), \qquad j = 1,2,\ldots,N.$$

Then the Galerkin method can be written as

$$\sum_{i=1}^{N} \ddot{U}_i(t)\int_\Omega \rho(\boldsymbol{x})\boldsymbol{\phi}_i(\boldsymbol{x})\cdot\boldsymbol{\phi}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \sum_{i=1}^{N} U_i(t)a\left(\boldsymbol{\phi}_i(\boldsymbol{x}),\boldsymbol{\phi}_i(\boldsymbol{x})\right) = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t)\cdot\boldsymbol{\phi}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad j = 1,2,\ldots,N.$$

We introduce vectors $\underline{\mathsf{U}}$ and $\underline{\mathsf{F}}$, matrices $\underline{\underline{\mathsf{M}}}$ and $\underline{\underline{\mathsf{K}}}$ whose components are defined as

$$(\underline{\mathsf{U}}(t))_i = U_i(t),$$

$$(\underline{\mathsf{F}}(t))_i = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t)\cdot\boldsymbol{\phi}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

$$(\underline{\underline{\mathsf{M}}})_{ij} = \int_\Omega \rho(\boldsymbol{x})\boldsymbol{\phi}_i(\boldsymbol{x}) \cdot \boldsymbol{\phi}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$
$$(\underline{\underline{\mathsf{K}}})_{ij} = a\left(\boldsymbol{\phi}_i(\boldsymbol{x}), \boldsymbol{\phi}_i(\boldsymbol{x})\right),$$

then Galerkin method can be written as a linear system

$$\underline{\underline{\mathsf{M}}}\ \underline{\ddot{\mathsf{U}}}(t) + \underline{\underline{\mathsf{K}}}\ \underline{\mathsf{U}}(t) = \underline{\mathsf{F}}(t).$$

This linear system is solvable since $\underline{\underline{\mathsf{M}}}$ is positive definite.

For simplicity we denote

$$B(\boldsymbol{u}(\boldsymbol{x},t), \boldsymbol{v}(\boldsymbol{x})) = \int_\Omega \rho(\boldsymbol{x})\ddot{\boldsymbol{u}}(\boldsymbol{x},t) \cdot \boldsymbol{v}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + a\left(\boldsymbol{u}(\boldsymbol{x},t), \boldsymbol{v}(\boldsymbol{x})\right).$$

Definition (Energy)**.** For a function $\boldsymbol{u}(\boldsymbol{x},t) \in V^n$ for $\forall\, t \in I$, its *energy* at time t is defined as

$$E\{\boldsymbol{u}\}(t) = \frac{1}{2}\int_\Omega \rho(\boldsymbol{x})\dot{\boldsymbol{u}}(\boldsymbol{x},t) \cdot \dot{\boldsymbol{u}}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} + \frac{1}{2}a\left(\boldsymbol{u}(\boldsymbol{x},t), \boldsymbol{u}(\boldsymbol{x},t)\right).$$

Theorem (Energy conservation of Galerkin method)**.** *If $\boldsymbol{u}(\boldsymbol{x},t) \in V^n$ is a weak solution, i.e., a solution to (**), then*

$$\frac{\mathrm{d}E\{\boldsymbol{u}\}(t)}{\mathrm{d}t} = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t) \cdot \dot{\boldsymbol{u}}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}. \tag{5}$$

Similarly, if $\boldsymbol{u}_h(\boldsymbol{x},t) \in V_h^n$ is a Galerkin approximation, i.e., a solution to (††), then

$$\frac{\mathrm{d}E\{\boldsymbol{u}_h\}(t)}{\mathrm{d}t} = \int_\Omega \boldsymbol{f}(\boldsymbol{x},t) \cdot \dot{\boldsymbol{u}}_h(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}. \tag{6}$$

Proof. Obviously for $\boldsymbol{u}(\boldsymbol{x},t) \in V^n$,

$$\frac{\mathrm{d}E\{\boldsymbol{u}\}(t)}{\mathrm{d}t} = B(\boldsymbol{u}(\boldsymbol{x},t), \dot{\boldsymbol{u}}(\boldsymbol{x},\boldsymbol{t})).$$

If $\boldsymbol{u}$ is weak solution then let $\boldsymbol{v}(\boldsymbol{x}) = \dot{\boldsymbol{u}}(\boldsymbol{x},t)$ in (**) at time t, and we can get (5). Similarly, if $\boldsymbol{u}_h$ is weak solution then let $\boldsymbol{v}_h(\boldsymbol{x}) = \dot{\boldsymbol{u}}_h(\boldsymbol{x},t)$ in (††) at time t, and we can get (6). □

If the domain of interest Ω is discretized in to elements $\Omega = \bigcup_e \Omega_e$, and in (††) we select

$$V_h^n = \left\{ \boldsymbol{u} \,\middle|\, \boldsymbol{u}\Big|_{\Omega_e} = \left(\mathbb{P}_m(\Omega_e)\right)^n, \boldsymbol{u} \in \left(C(\bar{\Omega})\right)^n \right\} \tag{SE}$$

in which $\mathbb{P}_m(\Omega_e)$ denote the space of polynomial of order at most m on Ω_e, the resulting method is *spectral element method* (SEM).

Theorem (Convergence of spectral element method)**.** *If the accurate solution to (**) is $\boldsymbol{u}^*(\boldsymbol{x},t)$, V_h^n is selected as (*SE*) in (††), then*

$$||\boldsymbol{u}^*(\cdot,t) - \boldsymbol{u}_h(\cdot,t)|| \sim O(h^{m+1})$$

as $h \to 0$.

This theorem indicates that with the same degree of freedom, it is a better idea to increase the order of approximation space than to decrease the mesh size.

Remark. To conclude the introduction to continuous Galerkin and spectral element method, let us summarize the features of spectral element method:

- SEM can achieve arbitrarily high order in space. This makes SEM very advantageous for accurately computing wave field.
- SEM is conservative in energy. This means that the numerical wave field neither dissipate nor explode in time (provided that the time integration is accurate), which is very preferable in simulating wave propagation.
- The mass matrix of SEM is global. This requires the computation of the inverse of a very large matrix, which can be problematic for parallel computation (can be solved by GLL interpolation).
- In order for the weak derivatives to be well-defined, the global continuity is required for spectral element approximation. This unfortunately makes the simulation of discontinuities in the wave field difficult.
- Although the continuity of displacement is required, the continuity of traction is never explicitly imposed.

The high-order feature in continuous areas and the conservation in energy makes SEM very ideal for simulating waves in the far-field. However, it can be troublesome in dealing with near-field simulation in which discontinuity must be taken into account.

4 General formulation of discontinuous Galerkin method

4.1 Velocity-strain form of elastodynamic equation

In order to permit discontinuities, we further lower the order of derivation in the equation. We denote $\boldsymbol{v}(\boldsymbol{x},t) = \dot{\boldsymbol{u}}(\boldsymbol{x},t)$ to be the velocity, and the 2-order tensor $\boldsymbol{\varepsilon} = \frac{1}{2}\left(\nabla \boldsymbol{u}(\boldsymbol{x},t) + (\nabla \boldsymbol{u}(\boldsymbol{x},t))^T\right)$ to be the strain.

The momentum conservation law is

$$\rho(\boldsymbol{x})\dot{\boldsymbol{v}}(\boldsymbol{x},t) - \nabla \cdot (\boldsymbol{\Lambda}(\boldsymbol{x}) : \boldsymbol{\varepsilon}(\boldsymbol{x},t)) = \boldsymbol{f}(\boldsymbol{x},t).$$

We define the tensor operator $\boldsymbol{\mathcal{F}}_1$ such that

$$\boldsymbol{\mathcal{F}}_1(\boldsymbol{x})\boldsymbol{\varepsilon} = -\boldsymbol{\Lambda}(\boldsymbol{x}) : \boldsymbol{\varepsilon},$$

Then the momentum conservation law can be written as

$$\rho(\boldsymbol{x})\dot{\boldsymbol{v}}(\boldsymbol{x},t) + \nabla \cdot (\boldsymbol{\mathcal{F}}_1\boldsymbol{\varepsilon}(\boldsymbol{x},t)) = \boldsymbol{f}(\boldsymbol{x},t).$$

Introduce a tensor operator $\boldsymbol{\mathcal{F}}_2$ such that

$$\boldsymbol{a} \cdot \boldsymbol{\mathcal{F}}_2\boldsymbol{u} = -\frac{1}{2}\left(\boldsymbol{a}\boldsymbol{u} + (\boldsymbol{a}\boldsymbol{u})^T\right),$$

holds for every vector $\boldsymbol{a}$, then the constitutional law can be written as

$$\dot{\boldsymbol{\varepsilon}}(\boldsymbol{x},t) + \nabla \cdot (\boldsymbol{\mathcal{F}}_2\boldsymbol{v}(\boldsymbol{x},t)) = 0.$$

Denote

$$\boldsymbol{Q}(\boldsymbol{x}) = \begin{bmatrix} \rho(\boldsymbol{x})\boldsymbol{I} & 0 \\ 0 & \boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{q}(\boldsymbol{x},t) = \begin{bmatrix} \boldsymbol{v}(\boldsymbol{x},t) \\ \boldsymbol{\varepsilon}(\boldsymbol{x},t) \end{bmatrix}, \quad \boldsymbol{\mathcal{F}}(\boldsymbol{x}) = \begin{bmatrix} 0 & \boldsymbol{\mathcal{F}}_1(\boldsymbol{x}) \\ \boldsymbol{\mathcal{F}}_2 & 0 \end{bmatrix}, \quad \boldsymbol{g}(\boldsymbol{x},t) = \begin{bmatrix} \boldsymbol{f}(\boldsymbol{x},t) \\ 0 \end{bmatrix},$$

then the equation of the system can be written as

$$\boldsymbol{Q}(\boldsymbol{x})\dot{\boldsymbol{q}}(\boldsymbol{x},t)+\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t))=\boldsymbol{g}(\boldsymbol{x},t) \tag{$\dagger$}$$

Equation (†) is the velocity-strain form of the elastodynamic equation. Note that (†) is a conservation-law equation: $\boldsymbol{Q}(x)\boldsymbol{q}(\boldsymbol{x},t)$ is the capacity, its change is the net effect of flux into the area $\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)$ and source $\boldsymbol{g}(\boldsymbol{x},t)$.

With this 1-order PDE, we are ready to construct a numerical method that permits discontinuity and preserves high-order accuracy. We define a bilinear form as the multiplication of two states $\boldsymbol{q}=\begin{bmatrix}\boldsymbol{v}\\ \boldsymbol{\varepsilon}\end{bmatrix}$ and $\boldsymbol{p}=\begin{bmatrix}\boldsymbol{w}\\ \boldsymbol{\psi}\end{bmatrix}$

$$\mathcal{E}(\boldsymbol{q},\boldsymbol{p})=\boldsymbol{v}\cdot\boldsymbol{w}+\boldsymbol{\varepsilon}:\boldsymbol{\Lambda}(\boldsymbol{x}):\boldsymbol{\psi}.$$

Multiply (†) with test function $\boldsymbol{p}$ and integrate on Ω_e, we get

$$\int_{\Omega_e}\mathcal{E}\left(\boldsymbol{Q}(\boldsymbol{x})\dot{\boldsymbol{q}}(\boldsymbol{x},t),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}+\int_{\Omega_e}\mathcal{E}\left(\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}=\int_{\Omega_e}\mathcal{E}\left(\boldsymbol{g}(\boldsymbol{x},t),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}.$$

Using integration-by-parts, we have

$$\begin{aligned}
&\int_{\Omega_e}\mathcal{E}\left(\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\\
=&-\int_{\Omega_e}\boldsymbol{w}(\boldsymbol{x},t)\cdot(\nabla\cdot(\boldsymbol{\Lambda}(\boldsymbol{x}):\boldsymbol{\varepsilon}(\boldsymbol{x},t)))\,\mathrm{d}\boldsymbol{x}-\int_{\Omega_e}\boldsymbol{\psi}(\boldsymbol{x},t):\boldsymbol{\Lambda}(\boldsymbol{x}):\nabla\boldsymbol{v}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\
=&\int_{\partial\Omega_e}\mathcal{E}\left(\hat{\boldsymbol{n}}\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}+\int_{\Omega_e}\nabla\boldsymbol{w}(\boldsymbol{x},t):\boldsymbol{\Lambda}(\boldsymbol{x}):\boldsymbol{\varepsilon}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\
&\quad+\int_{\Omega_e}(\nabla\cdot(\boldsymbol{\Lambda}(\boldsymbol{x}):\boldsymbol{\psi}(\boldsymbol{x},t)))\cdot\boldsymbol{v}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}\\
=&\int_{\partial\Omega_e}\mathcal{E}\left(\hat{\boldsymbol{n}}\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}-\int_{\Omega_e}\mathcal{E}\left(\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x},t)),\boldsymbol{q}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

Insert in the above equation,

$$\begin{aligned}
\int_{\Omega_e}\mathcal{E}\left(\boldsymbol{Q}(\boldsymbol{x})\dot{\boldsymbol{q}}(\boldsymbol{x},t),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}+&\int_{\partial\Omega_e}\mathcal{E}\left(\hat{\boldsymbol{n}}\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\\
&-\int_{\Omega_e}\mathcal{E}\left(\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x},t)),\boldsymbol{q}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}=\int_{\Omega_e}\mathcal{E}\left(\boldsymbol{g}(\boldsymbol{x},t),\boldsymbol{p}(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

To guarantee the continuity of flow on the contact between two elements, we substitute the flux term $\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{q}(\boldsymbol{x},t)$ on the boundary of the element with a numerical flux $\boldsymbol{\mathcal{F}}^*(\boldsymbol{x})\boldsymbol{q}^*(\boldsymbol{x},t)$, depending on the media parameters and states on both sides of the contact. We then obtain the discontinuous Galerkin method of the velocity-strain form (†)

Definition (Discontinuous Galerkin method)**.** Let V_h be a finite-dimensional space, then the *discontinuous Galerkin method* to solve (†) is

$$\begin{aligned}
&\text{Find}\quad \boldsymbol{q}_h(\boldsymbol{x},t)=\begin{bmatrix}\boldsymbol{v}_h(\boldsymbol{x},t)\\ \boldsymbol{\varepsilon}_h(\boldsymbol{x},t)\end{bmatrix},\boldsymbol{v}_h(\boldsymbol{x},t)\in V_h^n,\boldsymbol{\varepsilon}_h(\boldsymbol{x},t)\in V_h^{n\times n,\mathrm{sym}}\\
&\text{s.t.}\quad \int_{\Omega_e}\mathcal{E}\left(\boldsymbol{Q}(\boldsymbol{x})\dot{\boldsymbol{q}}_h(\boldsymbol{x},t),\boldsymbol{p}_h(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}+\int_{\partial\Omega_e}\mathcal{E}\left(\hat{\boldsymbol{n}}\cdot(\boldsymbol{\mathcal{F}}^*(\boldsymbol{x})\boldsymbol{q}_h^*(\boldsymbol{x},t)),\boldsymbol{p}_h(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}\\
&\qquad-\int_{\Omega_e}\mathcal{E}\left(\nabla\cdot(\boldsymbol{\mathcal{F}}(\boldsymbol{x})\boldsymbol{p}_h(\boldsymbol{x},t)),\boldsymbol{q}_h(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}=\int_{\Omega_e}\mathcal{E}\left(\boldsymbol{g}(\boldsymbol{x},t),\boldsymbol{p}_h(\boldsymbol{x},t)\right)\mathrm{d}\boldsymbol{x}.\\
&\qquad\text{holds for }\forall\, t\in I\text{ and }\forall\,\boldsymbol{p}_h(\boldsymbol{x},t)=\begin{bmatrix}\boldsymbol{w}_h(\boldsymbol{x},t)\\ \boldsymbol{\psi}_h(\boldsymbol{x},t)\end{bmatrix},\boldsymbol{w}_h(\boldsymbol{x},t)\in V_h^n,\boldsymbol{\psi}_h(\boldsymbol{x},t)\in V_h^{n\times n,\mathrm{sym}}.
\end{aligned} \tag{$\dagger\dagger$}$$

5 Discontinuous Galerkin method for one-dimensional wave equation

In order to show the basic concepts of the discontinuous Galerkin method, we start from the one-dimensional case. The 1-d elastodynamic equation is

$$\begin{bmatrix} \rho(x) & 0 \\ 0 & 1 \end{bmatrix} \frac{\partial}{\partial t} \begin{bmatrix} v(x,t) \\ \varepsilon(x,t) \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} 0 & -\mu(x) \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v(x,t) \\ \varepsilon(x,t) \end{bmatrix} = \begin{bmatrix} f(x,t) \\ 0 \end{bmatrix}. \tag{7}$$

Let

$$\underline{\underline{\mathsf{Q}}}(x) = \begin{bmatrix} \rho(x) & 0 \\ 0 & 1 \end{bmatrix}, \quad \underline{\mathsf{q}}(x,t) = \begin{bmatrix} v(x,t) \\ \varepsilon(x,t) \end{bmatrix}, \quad \underline{\underline{\mathsf{A}}}(x) = \begin{bmatrix} 0 & -\mu(x) \\ -1 & 0 \end{bmatrix}, \quad \underline{\mathsf{f}} = \begin{bmatrix} f(x,t) \\ 0 \end{bmatrix},$$

then the equation can be written as

$$\frac{\partial\Big(\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\Big)}{\partial t} + \frac{\partial\Big(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\Big)}{\partial x} = \underline{\mathsf{f}}(x,t). \tag{8}$$

Note that equation (8) is in a typical form of a conservation law: the change of physical quantity $\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)$ in a domain equals to the amount of flux $\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)$ that flows into the domain plus the amount of source $\underline{\mathsf{f}}(x,t)$ in the domain. In this section we discuss how to solve this equation numerically.

5.1 Finite volume method of conservation-law equation

Suppose the entire domain of interest is discretized into elements, the i^{th} element is represented as $\Omega_i = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$. Integrate equation (8) on element i, we obtain

$$\frac{\partial}{\partial t}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\mathrm{d}x + \Big(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\Big)\Big|_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{f}}(x,t)\mathrm{d}x.$$

Now it is natural to approximate everything with piecewise constant to construct a numerical method that is of 1-order accuracy. Let

$$\underline{\underline{\mathsf{Q}}}(x)\Big|_{\Omega_i} \approx \underline{\underline{\mathsf{Q}}}_i, \quad \underline{\mathsf{q}}(x,t)\big|_{\Omega_i} \approx \underline{\mathsf{q}}_i(t), \quad \underline{\underline{\mathsf{A}}}(x)\big|_{\Omega_i} \approx \underline{\underline{\mathsf{A}}}_i, \quad \underline{\mathsf{f}}(x,t)|_{\Omega_i} \approx \underline{\mathsf{f}}_i(t),$$

then

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\mathrm{d}x = |\Omega_i|\ \underline{\underline{\mathsf{Q}}}_i\ \underline{\mathsf{q}}_i(t), \quad \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{f}}(x,t)\mathrm{d}x = |\Omega_i|\ \underline{\mathsf{f}}_i(t).$$

However, the approximation of the flux $\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)$ at the element boundary remains undefined. If we can find an appropriate approximation for the flux

$$\Big(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\Big)\Big|_{x_{i-\frac{1}{2}}} \approx \underline{\mathsf{F}}_{i-\frac{1}{2}}(t), \quad \Big(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\Big)\Big|_{x_{i+\frac{1}{2}}} \approx \underline{\mathsf{F}}_{i+\frac{1}{2}}(t), \tag{9}$$

then we can formulate a numerical method

$$\underline{\underline{\mathsf{Q}}}_i \frac{\partial \underline{\mathsf{q}}_i(t)}{\partial t} + \frac{\underline{\mathsf{F}}_{i+\frac{1}{2}}(t) - \underline{\mathsf{F}}_{i-\frac{1}{2}}(t)}{|\Omega_i|} = \underline{\mathsf{f}}_i(t). \tag{10}$$

Numerical method (10) is called *finite volume method*. The flux approximation 9 is called *numerical flux*. In the next subsection we discuss how to choose a reasonable numerical flux.

5.2 Constant- and variable-coefficient linear Riemann problem

In order to choose a reasonable approximation for the flux, we want to investigate how the discontinuity evolves with time. Since we are only interested in a very small time interval and the information propagates at a finite speed, we can study a much simpler system, the system without source that has the identical discontinuity but is constant elsewhere, instead.

Definition (Riemann problem)**.** The initial-value problem defined on $(-\infty, +\infty)$

$$\begin{aligned} &\frac{\partial\Big(\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\Big)}{\partial t} + \frac{\partial\Big(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\Big)}{\partial x} = 0 \\ \text{in which}\quad &\underline{\underline{\mathsf{Q}}}(x) = \begin{cases} \underline{\underline{\mathsf{Q}}}_l, & x<0 \\ \underline{\underline{\mathsf{Q}}}_r, & x>0 \end{cases}, \quad \underline{\underline{\mathsf{A}}}(x) = \begin{cases} \underline{\underline{\mathsf{A}}}_l, & x<0 \\ \underline{\underline{\mathsf{A}}}_r, & x>0 \end{cases} \\ \text{with initial condition}\quad &\underline{\mathsf{q}}(x,0) = \begin{cases} \underline{\mathsf{q}}_l, & x<0 \\ \underline{\mathsf{q}}_r, & x>0 \end{cases} \end{aligned} \tag{11}$$

is called the *Riemann problem*.

To try to solve this variable-coefficient Riemann problem for linear system, we start with the constant-coefficient Riemann problem form linear equation

$$\begin{aligned} &Q\frac{\partial q(x,t)}{\partial t} + A\frac{\partial q(x,t)}{\partial x} = 0 \\ \text{with initial condition}\quad &q(x,0) = \begin{cases} q_l, & x<0 \\ q_r, & x>0 \end{cases}. \end{aligned} \tag{12}$$

Note that Q, A, q are scaler quantities. This is the 1-d advection equation we have talked about. The evolution of this system is quite clear,

$$q(x,t) = \begin{cases} q_l, & x < \frac{A}{Q}t \\ q_r, & x > \frac{A}{Q}t \end{cases}.$$

The solution means that discontinuity simply propagates at speed $\frac{A}{Q}$.

Let us move forward to linear system that has m equations, but keep the coefficients constant

$$\begin{aligned} &\underline{\underline{\mathsf{Q}}}\frac{\partial \underline{\mathsf{q}}(x,t)}{\partial t} + \underline{\underline{\mathsf{A}}}\frac{\partial \underline{\mathsf{q}}(x,t)}{\partial x} = 0 \\ \text{with initial condition}\quad &\underline{\mathsf{q}}(x,0) = \begin{cases} \underline{\mathsf{q}}_l, & x<0 \\ \underline{\mathsf{q}}_r, & x>0 \end{cases}. \end{aligned} \tag{13}$$

Assume that the m-by-m matrix $\underline{\underline{\mathsf{Q}}}^{-1}\underline{\underline{\mathsf{A}}}$ has m linearly independent eigenvectors

$$\underline{\underline{\mathsf{A}}}\ \underline{\mathsf{y}}_k = \lambda_k \underline{\underline{\mathsf{Q}}}\ \underline{\mathsf{y}}_k.$$

$\{\underline{\mathsf{y}}_k, k = 1, \ldots, m\}$ is a basis of $\mathbb{R}^m$, thus $\underline{\mathsf{q}}$ can be uniquely represented as the linear combination of $\underline{\mathsf{y}}_k$

$$\underline{\mathsf{q}}(x,t) = \sum_{k=1}^{m} \alpha_k(x,t)\underline{\mathsf{y}}_k.$$

Insert this linear representation into (13), we obtain for $\forall\, k = 1, \ldots m$

$$\begin{aligned} &\frac{\partial \alpha_k(x,t)}{\partial t} + \lambda_k\frac{\partial \alpha_k(x,t)}{\partial x} = 0 \\ \text{with initial condition}\quad &\alpha_k(x,0) = \begin{cases} \alpha_{k,l}, & x<0 \\ \alpha_{k,r}, & x>0 \end{cases}, \end{aligned} \tag{14}$$

in which $\alpha_{k,l}$ and $\alpha_{k,r}$ are coefficient that express $\underline{\mathsf{q}_l}$ and $\underline{\mathsf{q}_r}$ with linear combination of $\underline{\mathsf{y}_k}$

$$\underline{\mathsf{q}_l} = \sum_{k=1}^{m} \alpha_{k,l} \underline{\mathsf{y}_k}, \quad \underline{\mathsf{q}_r} = \sum_{k=1}^{m} \alpha_{k,r} \underline{\mathsf{y}_k}.$$

Note that equation (14) is in the form of (12), thus we can directly write the solution

$$\alpha_k(x,t) = \begin{cases} \alpha_{k,l}, & x > \lambda_k t \\ \alpha_{k,r}, & x < \lambda_k t \end{cases}.$$

This means that the **discontinuities only occur at line $x = \lambda_k t$, and the jump must be an eigenvector of $\underline{\underline{\mathsf{Q}}}^{-1}\underline{\underline{\mathsf{A}}}$ with eigenvalue λ_k**

$$\underline{\underline{\mathsf{A}}} \left(\underline{\mathsf{q}}(\lambda_k t + 0^+, t) - \underline{\mathsf{q}}(\lambda_k t - 0^+, t) \right) = \lambda_k \underline{\underline{\mathsf{Q}}} \left(\underline{\mathsf{q}}(\lambda_k t + 0^+, t) - \underline{\mathsf{q}}(\lambda_k t - 0^+, t) \right).$$

This relation is referred to as *Rankine-Hugoniot jump condition*.

Note that the R-H condition is not confined to the constant-coefficient linear system, it holds in variable-coefficient and even nonlinear case. Now let us solve the Riemann problem for variable-coefficient linear system (11) using the R-H condition. According to the R-H condition, the speed of the propagation of the discontinuity must be an eigenvalue, and the jump at this discontinuity must be an eigenvector corresponding to this eigenvalue. Suppose that the eigenvalues and eigenvectors of $\underline{\underline{\mathsf{Q_l}}}^{-1}\underline{\underline{\mathsf{A_l}}}$ are λ_k^l and $\underline{\mathsf{y}_k^l}$, those of $\underline{\underline{\mathsf{Q_r}}}^{-1}\underline{\underline{\mathsf{A_r}}}$ are λ_k^r and $\underline{\mathsf{y}_k^r}$. Then the solution at the left side of the boundary can be written as

$$\underline{\mathsf{q}}(0 - 0^+, t) = \underline{\mathsf{q}_l} + \sum_{\lambda_k^l < 0} \alpha_k^l \underline{\mathsf{y}_k^l}, \tag{RH1}$$

and that at the right side of the boundary can be written as

$$\underline{\mathsf{q}}(0 + 0^+, t) = \underline{\mathsf{q}_r} + \sum_{\lambda_k^r < 0} \alpha_k^r \underline{\mathsf{y}_k^r}. \tag{RH2}$$

Note that (RH1) and (RH2) holds as long as $t > 0$. The coefficients in (RH1) and (RH2) are determined by the conservation condition across the boundary

$$\underline{\underline{\mathsf{A_l}}}\ \underline{\mathsf{q}}(0 - 0^+, t) = \underline{\underline{\mathsf{A_r}}}\ \underline{\mathsf{q}}(0 + 0^+, t) \tag{CONS}$$

After we solve for the coefficients α_k^l and α_k^r, we obtain the solution for the Riemann problem (11). Now we can let the flux of the solution of the Riemann problem be the numerical flux

$$\underline{\mathsf{F}_{i-\frac{1}{2}}}(t) = \underline{\underline{\mathsf{A}_{i-1}}}\ \underline{\mathsf{q}}(0 - 0^+, 0^+) = \underline{\underline{\mathsf{A}_i}}\ \underline{\mathsf{q}}(0 + 0^+, 0^+)$$

in which $\underline{\mathsf{q}}(x,t)$ is the solution of the Riemann problem

$$\begin{aligned} &\frac{\partial\left(\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\right)}{\partial t} + \frac{\partial\left(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\right)}{\partial x} = 0 \\ \text{in which} \quad &\underline{\underline{\mathsf{Q}}}(x) = \begin{cases} \underline{\underline{\mathsf{Q}_{i-1}}}, & x < 0 \\ \underline{\underline{\mathsf{Q}_i}}, & x > 0 \end{cases}, \quad \underline{\underline{\mathsf{A}}}(x) = \begin{cases} \underline{\underline{\mathsf{A}_{i-1}}}, & x < 0 \\ \underline{\underline{\mathsf{A}_i}}, & x > 0 \end{cases} \\ \text{with initial condition} \quad &\underline{\mathsf{q}}(x,0) = \begin{cases} \underline{\mathsf{q}_{i-1}}(t), & x < 0 \\ \underline{\mathsf{q}_i}(t), & x > 0 \end{cases}. \end{aligned} \tag{15}$$

Such way to determine the numerical flux is usually referred to as *Godunov flux*.

5.3 Godunov flux for 1-d elastodynamic equation

Let us use the method in the last subsection to compute the Godunov flux for equation (7).

$$\underline{\underline{\mathsf{Q}}}_i = \begin{bmatrix} \rho_i & 0 \\ 0 & 1 \end{bmatrix}, \qquad \underline{\underline{\mathsf{A}}}_i = \begin{bmatrix} 0 & -\mu_i \\ -1 & 0 \end{bmatrix}$$

The eigenvalues and eigenvectors of $\underline{\underline{\mathsf{Q}}}_i{}^{-1}\underline{\underline{\mathsf{A}}}_i$ are

$$\lambda_1^i = -\sqrt{\frac{\mu_i}{\rho_i}}, \qquad \underline{\mathsf{y}}_1^i = \begin{bmatrix} \sqrt{\frac{\mu_i}{\rho_i}} \\ 1 \end{bmatrix};$$
$$\lambda_2^i = \sqrt{\frac{\mu_i}{\rho_i}}, \qquad \underline{\mathsf{y}}_2^i = \begin{bmatrix} -\sqrt{\frac{\mu_i}{\rho_i}} \\ 1 \end{bmatrix}.$$

The solution to the Riemann problem (15) is

$$v(x,t) = \begin{cases} v_{i-1}(t), & x < -c_{i-1}t \\ \dfrac{(\mu_i c_{i-1} v_i(t) + \mu_{i-1} c_i v_{i-1}(t)) - c_{i-1}c_i\left(\mu_{i-1}\varepsilon_{i-1}(t) - \mu_i \varepsilon_i(t)\right)}{\mu_i c_{i-1} + \mu_{i-1} c_i}, & -c_{i-1}t < x < c_i t \\ v_i(t), & c_i t < x \end{cases}$$

$$\varepsilon(x,t) = \begin{cases} \varepsilon_{i-1}(t), & x < -c_{i-1}t \\ \mu_i \dfrac{(v_i(t) - v_{i-1}(t)) + (c_{i-1}\varepsilon_{i-1}(t) + c_i\varepsilon_i(t))}{\mu_i c_{i-1} + \mu_{i-1} c_i}, & -c_{i-1}t < x < 0 \\ \mu_{i-1} \dfrac{(v_i(t) - v_{i-1}(t)) + (c_{i-1}\varepsilon_{i-1}(t) + c_i\varepsilon_i(t))}{\mu_i c_{i-1} + \mu_{i-1} c_i}, & 0 < x < c_i t \\ \varepsilon_i(t), & c_i t < x \end{cases}$$

The Godunov flux at the interface $i - \frac{1}{2}$ is

$$\underline{\mathsf{F}}_{i-\frac{1}{2}}(t) = \begin{bmatrix} -\mu_{i-1}\varepsilon(0-0^+, 0^+) \\ -v(0-0^+, 0^+) \end{bmatrix} = \begin{bmatrix} -\mu_i \varepsilon(0+0^+, 0^+) \\ -v(0+0^+, 0^+) \end{bmatrix}$$
$$= \begin{bmatrix} -\mu_{i-1}\mu_i \dfrac{(v_i(t) - v_{i-1}(t)) + (c_{i-1}\varepsilon_{i-1}(t) + c_i\varepsilon_i(t))}{\mu_i c_{i-1} + \mu_{i-1} c_i} \\ -\dfrac{(\mu_i c_{i-1} v_i(t) + \mu_{i-1} c_i v_{i-1}(t)) - c_{i-1}c_i\left(\mu_{i-1}\varepsilon_{i-1}(t) - \mu_i \varepsilon_i(t)\right)}{\mu_i c_{i-1} + \mu_{i-1} c_i} \end{bmatrix}$$

Now we have obtained everything to formulate a finite-volume method. However, the finite-volume method is an 1-order accurate method. If we want to have a high-order method, we must change the formulation. Consider the technique we used in the Galerkin method, one obvious way is to multiply the test function before integration over the element. This leads to the discontinuous Galerkin method.

5.4 Formulation of discontinuous Galerkin method for 1-d elastodynamic equation

Let a test vector function be

$$\underline{\mathsf{p}}(x) = \begin{bmatrix} w(x) \\ \psi(x) \end{bmatrix},$$

and we introduce a matrix

$$\underline{\underline{\tilde{\mathsf{A}}}}(x) = \begin{bmatrix} 1 & 0 \\ 0 & \mu(x) \end{bmatrix}.$$

Multiply (8) with $\left(\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\mathsf{p}}(x)\right)^T$ on the left and integrate over element Ω_i, we get

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\underline{\mathsf{Q}}}(x)\dot{\underline{\mathsf{q}}}(x,t)\mathrm{d}x + \left(\underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\right)\Big|_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}$$
$$-\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial}{\partial x}\left(\underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\right)\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\mathrm{d}x = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\mathsf{f}}(x,t)\mathrm{d}x.$$

To determine the flux at the boundary $i-\frac{1}{2}$, we can solve the Riemann problem

$$\frac{\partial\left(\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\right)}{\partial t} + \frac{\partial\left(\underline{\underline{\mathsf{A}}}(x)\underline{\mathsf{q}}(x,t)\right)}{\partial x} = 0$$
$$\text{in which}\quad \underline{\underline{\mathsf{Q}}}(x) = \begin{cases} \underline{\underline{\mathsf{Q}}}(x^-_{i-\frac{1}{2}}), & x<0 \\ \underline{\underline{\mathsf{Q}}}(x^+_{i-\frac{1}{2}}), & x>0 \end{cases}, \quad \underline{\underline{\mathsf{A}}}(x) = \begin{cases} \underline{\underline{\mathsf{A}}}(x^-_{i-\frac{1}{2}}), & x<0 \\ \underline{\underline{\mathsf{A}}}(x^+_{i-\frac{1}{2}}), & x>0 \end{cases} \tag{16}$$
$$\text{with initial condition}\quad \underline{\mathsf{q}}(x,0) = \begin{cases} \underline{\mathsf{q}}(x^-_{i-\frac{1}{2}},t), & x<0 \\ \underline{\mathsf{q}}(x^+_{i-\frac{1}{2}},t), & x>0 \end{cases},$$

and obtain the Godunov flux

$$\underline{\mathsf{F}_{i-\frac{1}{2}}}(t) = \begin{bmatrix} -\mu^-_{i-\frac{1}{2}}\mu^+_{i-\frac{1}{2}}\dfrac{\left(v^+_{i-\frac{1}{2}}(t)-v^-_{i-\frac{1}{2}}(t)\right)+\left(c^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t)+c^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}}+\mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \\ -\dfrac{\left(\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}}v^+_{i-\frac{1}{2}}(t)+\mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}v^-_{i-\frac{1}{2}}(t)\right)-c^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}\left(\mu^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t)-\mu^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}}+\mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \end{bmatrix} \tag{GF}$$

in which

$$\mu^\pm_{i-\frac{1}{2}} = \mu(x^\pm_{i-\frac{1}{2}}), \quad v^\pm_{i-\frac{1}{2}}(t) = v(x^\pm_{i-\frac{1}{2}},t), \quad \varepsilon^\pm_{i-\frac{1}{2}}(t) = \varepsilon(x^\pm_{i-\frac{1}{2}},t).$$

Replace the flux term at boundary with the Godunov flux, we obtain

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\underline{\mathsf{Q}}}(x)\dot{\underline{\mathsf{q}}}(x,t)\mathrm{d}x + \underline{\mathsf{p}}^T(x^-_{i+\frac{1}{2}})\underline{\underline{\tilde{\mathsf{A}}}}(x^-_{i+\frac{1}{2}})\underline{\mathsf{F}_{i+\frac{1}{2}}}(t) - \underline{\mathsf{p}}^T(x^+_{i-\frac{1}{2}})\underline{\underline{\tilde{\mathsf{A}}}}(x^+_{i-\frac{1}{2}})\underline{\mathsf{F}_{i-\frac{1}{2}}}(t)$$
$$-\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\partial}{\partial x}\left(\underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\right)\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\mathrm{d}x = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{p}}^T(x)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\mathsf{f}}(x,t)\mathrm{d}x. \tag{dG}$$

The superscripts $^+$ and $^-$ denote evaluation at the positive and negative sides of the boundary. Thus the discontinuous Galerkin method of equation (7) can be defined as

$$\text{Find}\quad \underline{\mathsf{q}}(x,t) = \begin{bmatrix} v(x,t) \\ \varepsilon(x,t) \end{bmatrix}, v(x,t), \varepsilon(x,t) \in V_h$$
$$\text{s.t.}\quad \text{(dG) holds for}\forall\ i,\ \forall\ t\in I \text{ and } \forall\ \underline{\mathsf{p}}(x) = \begin{bmatrix} w(x) \\ \psi(x) \end{bmatrix}, w(x), \psi(x) \in V_h,$$

in which the numerical flux $\underline{\mathsf{F}_{i-\frac{1}{2}}}(t)$ is given by the Godunov flux (GF), and

$$V_h = \left\{ u \,\middle|\, u\Big|_{\Omega_i} \in \mathbb{P}_m(\Omega_i),\ \forall\ i \right\}.$$

5.5 Properties of discontinuous Galerkin method

Let us define the energy of in the setting of the conservation-law equation. We first define the energy in element i

$$\begin{aligned} E_i\{\underline{\mathsf{q}}\}(t) &= \frac{1}{2}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \rho(x)v^2(x,t) + \mu(x)\varepsilon^2(x,t)\mathrm{d}x \\ &= \frac{1}{2}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \underline{\mathsf{q}}^T(x,t)\underline{\underline{\tilde{\mathsf{A}}}}(x)\underline{\underline{\mathsf{Q}}}(x)\underline{\mathsf{q}}(x,t)\mathrm{d}x, \end{aligned}$$

and the total energy is

$$E\{\underline{\mathsf{q}}\}(t) = \sum_i E_i\{\underline{\mathsf{q}}\}(t).$$

We can obtain a very important feature of the discontinuous Galerkin method which implies the dissipative characteristics in energy

$$\frac{\mathrm{d}E\{\underline{\mathsf{q}}\}(t)}{\mathrm{d}t} \le \int_\Omega f(x,t)v(x,t)\mathrm{d}x.$$

The order of accuracy of the discontinuous Galerkin method is also $m+1$ provided that the real solution is smooth:

Proposition (Accuracy of discontinuous Galerkin method)**.** $\underline{\mathsf{q}}^*(x,t)$ *is the real solution and* $\underline{\mathsf{q}}(x,t)$ *is the solution of the discontinuous Galerkin method (*dG*). If* $\underline{\mathsf{q}}^*$ *is smooth, then*

$$||\underline{\mathsf{q}}(\cdot,t) - \underline{\mathsf{q}}^*(\cdot,t)|| \sim O(h^{m+1})$$

as the mesh size $h \to 0$.

5.6 Imposing dislocation in discontinuous Galerkin method

In simulation of seismic waves, we usually need to compute waves excited by a prescribed dislocation. This can be easily included in the discontinuous Galerkin method by simply modifying the numerical flux. If we want to impose a dislocation at boundary $i-\frac{1}{2}$, we can solve the Riemann problem with

$$\begin{cases} v(0+0^+,0^+) - v(0-0^+,0^+) = v_d(t) \\ \mu(x^+_{i-\frac{1}{2}})\varepsilon(0+0^+,0^+) = \mu(x^-_{i-\frac{1}{2}})\varepsilon(0+0^-,0^+) \end{cases}.$$

The resulting numerical flux is

$$\underline{\mathsf{F}^-_{i-\frac{1}{2}}}(t) = \begin{bmatrix} -\mu^-_{i-\frac{1}{2}}\mu^+_{i-\frac{1}{2}}\dfrac{\left(v^+_{i-\frac{1}{2}}(t) - v^-_{i-\frac{1}{2}}(t) - v_d(t)\right) + \left(c^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t) + c^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}} + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \\ -\dfrac{\left(\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}}\left(v^+_{i-\frac{1}{2}}(t) - v_d(t)\right) + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}v^-_{i-\frac{1}{2}}(t)\right) - c^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}\left(\mu^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t) - \mu^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}} + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \end{bmatrix},$$

$$\underline{\mathsf{F}^+_{i-\frac{1}{2}}}(t) = \begin{bmatrix} -\mu^-_{i-\frac{1}{2}}\mu^+_{i-\frac{1}{2}}\dfrac{\left(v^+_{i-\frac{1}{2}}(t) - v^-_{i-\frac{1}{2}}(t) - v_d(t)\right) + \left(c^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t) + c^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}} + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \\ -\dfrac{\left(\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}}v^+_{i-\frac{1}{2}}(t) + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}\left(v^-_{i-\frac{1}{2}}(t) + v_d(t)\right)\right) - c^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}\left(\mu^-_{i-\frac{1}{2}}\varepsilon^-_{i-\frac{1}{2}}(t) - \mu^+_{i-\frac{1}{2}}\varepsilon^+_{i-\frac{1}{2}}(t)\right)}{\mu^+_{i-\frac{1}{2}}c^-_{i-\frac{1}{2}} + \mu^-_{i-\frac{1}{2}}c^+_{i-\frac{1}{2}}} \end{bmatrix}.$$