Nonconservative Products

1 Definition

Consider the nonconservative product

$$g(\boldsymbol{q})\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}x},$$

where $g(\boldsymbol{q}) : \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^p$ is continuous, but $\boldsymbol{q}$ is possibly discontinuous. In this case, the product is traditionally not well-defined at the discontinuities of $\boldsymbol{q}$. In order to define this product for discontinuous functions, $\boldsymbol{q}$, it is possible to regularize $\boldsymbol{q}$ with a path ϕ at discontinuities according to the theory laid out by Dal Maso, Le Floch, and Murat. To this end consider Lipschitz continuous paths, $\boldsymbol{\psi} : [0,1] \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^p$, that satisfy the following properties.

1. $\forall \boldsymbol{q}_L, \boldsymbol{q}_R \in \mathbb{R}^p$, $\boldsymbol{\psi}(0, \boldsymbol{q}_L, \boldsymbol{q}_R) = \boldsymbol{q}_L$ and $\boldsymbol{\psi}(1, \boldsymbol{q}_L, \boldsymbol{q}_R) = \boldsymbol{q}_R$
2. $\exists k > 0$, $\forall \boldsymbol{q}_L, \boldsymbol{q}_R \in \mathbb{R}^p$, $\forall s \in [0,1]$, $\left|\frac{\partial \boldsymbol{\psi}}{\partial s}(s, \boldsymbol{q}_L, \boldsymbol{q}_R)\right| \leq k|\boldsymbol{q}_L - \boldsymbol{q}_R|$ elementwise
3. $\exists k > 0$, $\forall \boldsymbol{q}_L, \boldsymbol{q}_R, \boldsymbol{u}_L, \boldsymbol{u}_R \in \mathbb{R}^p$, $\forall s \in [0,1]$, elementwise

$$\left|\frac{\partial \boldsymbol{\psi}}{\partial s}(s, \boldsymbol{q}_L, \boldsymbol{q}_R) - \frac{\partial \boldsymbol{\psi}}{\partial s}(s, \boldsymbol{u}_L, \boldsymbol{u}_R)\right| \leq k(|\boldsymbol{q}_L - \boldsymbol{u}_L| + |\boldsymbol{q}_R - \boldsymbol{u}_R|)$$

Once we have these paths, $\boldsymbol{\psi}$, we can define the nonconservative product.

Let $\boldsymbol{q} : [a,b] \to \mathbb{R}^p$ be a function of bounded variation, let $g : \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^p$ be a continuous function, and let $\boldsymbol{\psi}$ satisfy the properties given above. Then there exists a unique real-valued bounded Borel measure $\boldsymbol{\mu}$ on $[a,b]$ characterized by the two following properties.

1. If q is continuous on a Borel set $B \subset [a,b]$, then

$$\boldsymbol{\mu}(B) = \int_B g(\boldsymbol{q})\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}x}\,\mathrm{d}x$$

2. If q is discontinuous at a point $x_0 \in [a,b]$, then

$$\boldsymbol{\mu}(x_0) = \int_0^1 g(\boldsymbol{\psi}(s; \boldsymbol{q}(x_0^-), \boldsymbol{q}(x_0^+)))\frac{\partial \boldsymbol{\psi}}{\partial s}(s; \boldsymbol{q}(x_0^-), \boldsymbol{q}(x_0^+))\,\mathrm{d}s$$

By definition, this measure $\boldsymbol{\mu}$ is the nonconservative product $g(\boldsymbol{q})\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}x}$ and will be denoted by

$$\boldsymbol{\mu} = \left[g(\boldsymbol{q})\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}x}\right]_{\boldsymbol{\psi}}$$

Note that if there exists a function $\boldsymbol{f}(\boldsymbol{q})$ such that $\boldsymbol{f}'(\boldsymbol{q}) = g(\boldsymbol{q})$, then

$$\int_0^1 g(\boldsymbol{\psi}(s; \boldsymbol{q}(x_0^-), \boldsymbol{q}(x_0^+)))\frac{\partial \boldsymbol{\psi}}{\partial s}(s; \boldsymbol{q}(x_0^-), \boldsymbol{q}(x_0^+))\,\mathrm{d}s = \boldsymbol{f}\big(\boldsymbol{q}(x_0^+)\big) - \boldsymbol{f}\big(\boldsymbol{q}(x_0^-)\big) \tag{1}$$

for any path $\boldsymbol{\psi}$ that satisfies the conditions 1 - 3.

1.1 Higher Dimensions

In higher dimensions the paths, $\boldsymbol{\psi}$ must also have the property that

4. $\boldsymbol{\psi}(s, \boldsymbol{q}_L, \boldsymbol{q}_R) = \boldsymbol{\psi}(1 - s, \boldsymbol{q}_L, \boldsymbol{q}_R)$

Then the following Thereom can be given in spacetime Let $\boldsymbol{q} : \Omega \to \mathbb{R}^m$ be a bounded function of bounded variation defined on an open subset Ω of $\mathbb{R}^{n+1}$ and $\boldsymbol{t} : \mathbb{R}^m \to \mathbb{R}^m$ be a locally bounded Borel function. Then there existsa unique family of real-valued bounded Borel measures μ_i on Ω, $i = 1, 2, \ldots, m$ such that

1. if B is a continuous Borel subset of Ω, then

$$\mu_i(B) = \int_B t_{ik}(q)\boldsymbol{q}_{x_k} \, \mathrm{d}\lambda$$

where λ is the Borel measure;

2. if B is a discontinuous subset of Ω of approximate jump, then

$$\mu_i(B) = \int_B \int_0^1 t_{ik}(\boldsymbol{\psi}(s, \boldsymbol{q}^L, \boldsymbol{q}^R)) \frac{\partial \boldsymbol{\psi}}{\partial s}(s, \boldsymbol{q}^L, \boldsymbol{q}^R) \, \mathrm{d}s \boldsymbol{n}_k^L \, \mathrm{d}H^n$$

with $\boldsymbol{q}^L$ and $\boldsymbol{q}^R$ the left and right traces at the discontinuity, where H^n is the n-dimensional Hausdorf measure and where we choose $\boldsymbol{n}^L$ the outward normal with respect to the left state,

3. if B is an irregular Borel subset of Ω, then $\mu_i(B) = 0$

This is given in Rhebergen without proof, but I haven't found any outside original references for this Theorem. Mostly this reflects the one dimensional theorem, but I don't understand the appearance of the outward facing normal. It appears to have been an arbitrary choice between $\boldsymbol{n}^L$ and $\boldsymbol{n}^R$. Also I am not sure why it is necessary at all.

2 Weak Solutions

A function $\boldsymbol{q}$ of bounded variation is a weak solution to

$$\boldsymbol{q}_t + g(\boldsymbol{q})\boldsymbol{q}_x = 0 \tag{2}$$

if

$$\boldsymbol{q}_t + [g(\boldsymbol{q})\boldsymbol{q}_x]_\phi = 0 \tag{3}$$

as a bounded Borel measure on $\mathbb{R} \times \mathbb{R}_+$. This is equivalent to finding $\boldsymbol{q}$ that satisfies,

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}} v_t(t, x)\boldsymbol{q}(t, x) \, \mathrm{d}x \, \mathrm{d}t + \int_{\mathbb{R}_+} \int_{\mathbb{R}} v(t, \cdot)[g(\boldsymbol{q}(t, \cdot))\boldsymbol{q}_x(t, \cdot)]_\psi \, \mathrm{d}t = \mathbf{0} \tag{4}$$

for all functions $v \in C_0^\infty(\mathbb{R}_t \times \mathbb{R})$.

3 DG Weak Formulation

3.1 Rhebergen Weak Formulation

Find $\boldsymbol{q} \in V_h$ such that for all $\boldsymbol{v} \in V_h$,

$$\sum_j \left(\int_{K_j} \boldsymbol{v}^T \boldsymbol{q}_t - \boldsymbol{v}_x^T \boldsymbol{f}(\boldsymbol{q}) + \boldsymbol{v}^T g(\boldsymbol{q})\boldsymbol{q}_x \, \mathrm{d}x \right) + \sum_S \left(\int_S (\boldsymbol{v}^L - \boldsymbol{v}^R)^T \hat{\boldsymbol{P}}^{nc} \, \mathrm{d}S \right) \tag{5}$$

$$+ \sum_S \left(\int_S \frac{1}{2} (\boldsymbol{v}^R + \boldsymbol{v}^L)^T \int_0^1 g\big(\boldsymbol{\psi}(\tau, \boldsymbol{q}^L, \boldsymbol{q}^R)\big) \frac{\partial \boldsymbol{\psi}}{\partial \tau}(\tau, \boldsymbol{q}^L, \boldsymbol{q}^R) \, \mathrm{d}\tau \, \mathrm{d}S \right) \tag{6}$$

where $\hat{\boldsymbol{P}}^{nc}$ is the nonconservative numerical flux, if symmetrical wave speeds are assumed, then the Rusanov or Local Lax Friedrichs flux can be used, otherwise the nonconservative product will affect the numerical flux.

3.2 Standard Hyperbolic Conservation Law DG Formulation

Let $\{K_j\}$ be a mesh of the domain $[a, b]$. Also denote the DG space as

$$V_h = \left\{ v \in L^1([a,b]) \middle| \, v|_{K_j} \in \mathbb{P}^M(K_j) \right\}$$

Consider the hyperbolic conservation law given below with the corresponding classical and semi discrete weak solutions.

$$\boldsymbol{q}_t + \boldsymbol{f}(\boldsymbol{q})_x = 0 \tag{7}$$

$$\int_a^b v\boldsymbol{q}_t - v_x \boldsymbol{f}(\boldsymbol{q}) \, \mathrm{d}x = 0 \tag{8}$$

The DG formulation requires finding $\boldsymbol{q}_h \in V_h$ for all $v_h \in V_h$ such that

$$\int_a^b v_h \boldsymbol{q}_{h,t} + v_h \boldsymbol{f}(\boldsymbol{q}_h)_x \, \mathrm{d}x = 0 \tag{9}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} + v_h \boldsymbol{f}(\boldsymbol{q}_h)_x \, \mathrm{d}x \right) = 0 \tag{10}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(\int_{K_j} v_h \boldsymbol{f}(\boldsymbol{q}_h)_x \, \mathrm{d}x \right) = 0 \tag{11}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(\hat{v}_h \hat{f}(\boldsymbol{q}_h) \Big|_{x_{j-1/2}}^{x_{j+1/2}} - \int_{K_j} v_{h,x} \boldsymbol{f}(\boldsymbol{q}_h) \, \mathrm{d}x \right) = 0 \tag{12}$$

Usually the value of $\hat{v}_h$ is the interior value of the test function on the element integral that is being integrated by parts. That is

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(v_h\left(x_{j+1/2}^-\right) \hat{f}_{j+1/2} - v_h\left(x_{j-1/2}^+\right) \hat{f}_{j-1/2} - \int_{K_j} v_{h,x} \boldsymbol{f}(\boldsymbol{q}_h) \, \mathrm{d}x \right) = 0 \tag{13}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_{I_j} \left(\left(v_h(x^-) - v_h(x^+) \right) \hat{f} \right) - \sum_j \left(\int_{K_j} v_{h,x} \boldsymbol{f}(\boldsymbol{q}_h) \, \mathrm{d}x \right) = 0 \tag{14}$$

Using these values for the test functions at the interfaces and then grouping the interfaces together reveals jump terms in the test functions at the interfaces.

I see that if the value of the test functions had a single value at the interfaces like the numerical fluxes, $\hat{f}$, then when combining the values at each interface the terms would cancel out. Does choosing the interior value for the test functions just make sure that those interface terms don't cancel out, or what is the theoretical reason for the values of the test functions at the interfaces.

3.3 Pure Nonconservative DG Formulation

Consider the 1D nonconservative equation shown below,

$$\boldsymbol{q}_t + g(\boldsymbol{q})\boldsymbol{q}_x = \boldsymbol{0} \qquad x \in [a, b], 0 < t < T$$

Now the semi discrete DG formulation for this problem becomes finding $\boldsymbol{q}_h \in V_h$ for all $v_h \in V_h$ that satisfies

$$\int_a^b v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x + \int_a^b v_h [g(\boldsymbol{q}_h)\boldsymbol{q}_{h,x}]_\psi \, \mathrm{d}x = \boldsymbol{0} \tag{15}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(\int_{K_j} v_h g(\boldsymbol{q}_h) \boldsymbol{q}_{h,x} \, \mathrm{d}x \right) + \sum_I \left(\hat{v}_h \int_0^1 g(\psi(s, \boldsymbol{q}_h^L, \boldsymbol{q}_h^R)) \frac{\partial \psi}{\partial s}(s, \boldsymbol{q}_h^L, \boldsymbol{q}_h^R) \, \mathrm{d}s \right) = 0 \tag{16}$$

Consider the case where there exists a function $\boldsymbol{f}(\boldsymbol{q}_h)$ such that $\boldsymbol{f}'(\boldsymbol{q}_h) = g(\boldsymbol{q})$.

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(\int_{K_j} v_h \boldsymbol{f}(\boldsymbol{q}_h)_x \, \mathrm{d}x \right) + \sum_I \left(\hat{v}_h \int_0^1 f'(\psi(s, \boldsymbol{q}_h^L, \boldsymbol{q}_h^R)) \frac{\partial \psi}{\partial s}(s, \boldsymbol{q}_h^L, \boldsymbol{q}_h^R) \, \mathrm{d}s \right) = 0 \tag{17}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) + \sum_j \left(\int_{K_j} v_h \boldsymbol{f}(\boldsymbol{q}_h)_x \, \mathrm{d}x \right) + \sum_I \left(\hat{v}_h \left(\boldsymbol{f}\left(\boldsymbol{q}_h^R\right) - \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \right) \right) = 0 \tag{18}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) - \sum_j \left(\int_{K_j} v_{h,x} \boldsymbol{f}(\boldsymbol{q}_h) \, \mathrm{d}x \right) + \sum_I \left(v_h^L \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) - v_h^R \boldsymbol{f}\left(\boldsymbol{q}_h^R\right) \right) + \sum_I \left(\hat{v}_h \left(\boldsymbol{f}\left(\boldsymbol{q}_h^R\right) - \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \right) \right) = 0 \tag{19}$$

$$\sum_j \left(\int_{K_j} v_h \boldsymbol{q}_{h,t} \, \mathrm{d}x \right) - \sum_j \left(\int_{K_j} v_{h,x} \boldsymbol{f}(\boldsymbol{q}_h) \, \mathrm{d}x \right) + \sum_I \left(\left(\hat{v} - v_h^R \right) \boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \left(v_h^L - \hat{v} \right) \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \right) = 0 \tag{20}$$

Now we want to choose $\hat{v}$ such that this is equivalent to the traditional DG formulation. However we don't have any numerical flux terms so instead we want the interface terms to look like $\left(v_h^L - v_h^R\right) \frac{1}{2} \left(\boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \right)$. At least this is what Rhebergen does and then replaces the flux average with the numerical flux.

$$\left(\hat{v} - v_h^R \right) \boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \left(v_h^L - \hat{v} \right) \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) = \left(v_h^L - v_h^R \right) \frac{1}{2} \left(\boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \right) \tag{21}$$

$$\left(\hat{v} - v_h^R \right) \boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \left(v_h^L - \hat{v} \right) \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) = \frac{1}{2} \left(v_h^L - v_h^R \right) \boldsymbol{f}\left(\boldsymbol{q}_h^R\right) + \frac{1}{2} \left(v_h^L - v_h^R \right) \boldsymbol{f}\left(\boldsymbol{q}_h^L\right) \tag{22}$$

$$\left(\hat{v} - v_h^R \right) = \frac{1}{2} \left(v_h^L - v_h^R \right) \tag{23}$$

$$\hat{v} = \frac{1}{2} \left(v_h^L + v_h^R \right) \tag{24}$$

$$\left(v_h^L - \hat{v} \right) = \frac{1}{2} \left(v_h^L - v_h^R \right) \tag{25}$$

$$-\hat{v} = \frac{1}{2} \left(-v_h^L - v_h^R \right) \tag{26}$$

$$\hat{v} = \frac{1}{2} \left(v_h^L + v_h^R \right) \tag{27}$$

$$\tag{28}$$

We see that the appropriate numerical flux for the test function when multiplying the nonconservative product at the interface should be the average value. This agrees with the results given in Rhebergen. My one question about this is the swap from the average value of f to the numerical flux of f. I am tempted to use the numerical flux of f when integrating by parts, but then in order to agree with the traditional method $\hat{v}$ should be zero.