A Numerical Approach of Friedrichs' Systems Under Constraints in Bounded Domains

Clément Mifsud and Bruno Després

Abstract We present here an explicit finite volume scheme on unstructured meshes adapted to first-order hyperbolic systems under constraints in bounded domains. This scheme is based on the work (Coudière, Vila, Villedieu in C R Acad Sci Paris Sér I Math 331:95–100, 2000, [3]) in the unconstrained case and the splitting strategy of Després, Lagoutière, Seguin (Nonlinearity 24:3055–3081, 2011, [4]). We show that this scheme is stable under a Courant–Friedrichs–Lewy condition (and convergent for problems posed in the whole space), and we illustrate the solution constructed by this scheme on the example of the simplified model of perfect plasticity. From the theoretical point of view, the interaction between the constraint and the boundary of the domain in the model of perfect plasticity is encoded by a nonlinear boundary condition. With this numerical approach, we will show that, even if this scheme uses the underlying linear boundary condition, the results are consistent with the nonlinear model (and in particular with the nonlinear boundary condition).

Keywords Finite volume schemes · Friedrichs' systems · Constrained problems

Mathematics Subject Classification 2010 65M08 · 65M12 · 35L50 · 35L60 74C05

1 Introduction

The aim of this article is to examine the numerical approximation of Friedrichs' equations under constraints (posed in the whole space or in bounded domains). To do so, we use a popular method for hyperbolic problems: the method of finite

---

C. Mifsud (✉) · B. Després
Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie – Paris 6, CNRS, UMR 7598, 75005 Paris, France
e-mail: mifsud@ljll.math.upmc.fr

B. Després
Institut Universitaire de France, Paris, France
e-mail: despres@ann.jussieu.fr

C. Klingenberg and M. Westdickenberg (eds.), *Theory, Numerics and Applications of Hyperbolic Problems II*, Springer Proceedings in Mathematics & Statistics 237, https://doi.org/10.1007/978-3-319-91548-7_25

volumes (for a detailed presentation of this method, we refer to [5, 6]). Although there is an important number of schemes that have been developed, the analysis of the convergence and its rate of schemes on unstructured meshes for multidimensional problems (i.e., the domain is a subset of $\mathbb{R}^n$ with $n > 1$, and the solution belongs to $\mathbb{R}^m$ with $m > 1$) are still in its infancy.

However, the article [9] has established a rate of convergence for the RKDG scheme (see [2]), using P0 finite elements in space and the RK1 scheme in time, on unstructured meshes for generic Friedrichs' systems of the following form

$$\begin{cases} \partial_t U + \sum_{j=1}^n \partial_j (A_i U) + BU = f, & \text{in } (0,T) \times \mathbb{R}^n, \\ U(0,x) = U_0(x), & \text{in } \mathbb{R}^n, \end{cases} \tag{1}$$

where $U : (t,x) \in (0,T) \times \mathbb{R}^n \to \mathbb{R}^m$, $A_i : (t,x) \in (0,T) \times \mathbb{R}^n \to \mathbb{M}_{\text{sym}}^{m\times m}$, $B : (t,x) \in (0,T) \times \mathbb{R}^n \to \mathbb{M}^{m\times m}$, $f : (t,x) \in (0,T) \times \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbb{M}^{m\times m}$ (resp. $\mathbb{M}_{\text{sym}}^{m\times m}$) is the space of $m \times m$ (resp. symmetric) matrices with real coefficients. A similar analysis has been performed in the note [3] on bounded domains.

In addition, the study of the convergence of a scheme based on the Rusanov scheme on Cartesian meshes has been performed in [4] for constrained Friedrichs' systems. In fact, to show the existence of a weak solution (in the sense of Definition 1) to the constrained Friedrichs' system

$$\begin{cases} \partial_t U + \sum_{j=1}^n A_j \partial_j U = 0 \quad \text{in } (0,T] \times \mathbb{R}^n; \quad U(0,x) = U^0(x) \quad \text{if } x \in \mathbb{R}^n, \\ U(t,x) \in \mathscr{C} \quad \text{if } (t,x) \in [0,T] \times \mathbb{R}^n, \end{cases} \tag{2}$$

where $\mathscr{C}$ is a fixed closed and convex subset of $\mathbb{R}^m$ (with $0 \in \overset{\circ}{\mathscr{C}}$), the authors construct a numerical solution with a two-step scheme such that a subsequence converges to a weak solution of (2). In this paper, we extend the strategy of [4] to schemes on unstructured meshes and to problems posed in bounded domains.

In Sect. 2, we recall some notations and define our finite volume scheme on unstructured meshes for constrained Friedrichs' systems in bounded domains.

In Sect. 3, we recall some results of [4] on constrained Friedrichs' systems in the whole space and state a convergence result in the whole space on a similar scheme (to the one presented in Sect. 2 on bounded domains). This result tells us that the finite volume scheme on unstructured meshes, based on the work [9], associated with a projection step has the same rate of convergence (in the space $L^2((0,T) \times \mathbb{R}^n; \mathbb{R}^m)$) as in the unconstrained case (obtained in [9]).

In Sect. 4, we show that the scheme presented in Sect. 2 is stable (under a Courant–Friedrichs–Lewy condition) in the space $L^\infty(0,T; L^2(\Omega, \mathbb{R}^m))$.

Then in Sect. 5, we briefly recall the equations of the simplified model of the dynamical perfect plasticity problem (described in [1]) and how this problem is related to the constrained Friedrichs' systems.

Finally, in Sect. 6, we illustrate the solution constructed by this scheme on the example of the simplified model of the dynamical perfect plasticity problem and show that the interaction between the constraint and the boundary condition that

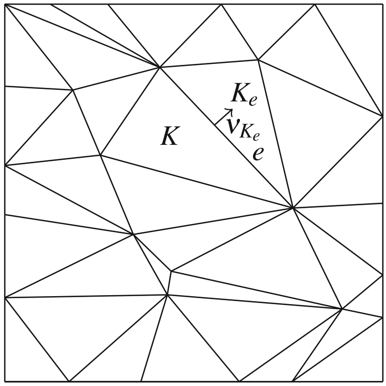


Fig. 1 An unstructured meshes of the square $[0, 1] \times [0, 1]$. Here the polytopes are triangles

has been underlined theoretically by the nonlinear boundary condition can also be observed numerically.

2 Description of the Scheme

In this section, we present the general framework of this work and the scheme we are interested in. Let $\mathcal{T}_h$ be a triangulation of $\Omega \subset \mathbb{R}^n$ (a n-dimensional polytope); i.e., $\mathcal{T}_h = (K_i)_{i\in\mathcal{I}}$, with $\mathcal{I} \subset \mathbb{N}$, is a family of open nonempty convex polytope such that $\cup_{i\in\mathcal{I}} \overline{K_i} = \overline{\Omega}$, for all $i \neq j$, $K_i \cap K_j = \emptyset$ and $h = \sup_{i\in\mathcal{I}} (\mathrm{diam} K_i) < +\infty$. The set of edges of a polytope K is denoted $\mathcal{E}_K$. We introduce the following notations (see also Fig. 1),

$m_K, m_{\partial K}$: $\mathcal{L}^n$-measure of K, $\mathcal{H}^{n-1}$-measure ∂K,
$e \in \mathcal{E}_K$: an edge ($(n-1)$-dimensional polytope) of K with $\mathcal{H}^{n-1}$-measure m_e,
$\mathcal{E}_{K\,\mathrm{i}}, \mathcal{E}_{K\,\mathrm{b}}$: the set of interior edges e of K, the set of boundary edges e of K,
ν_{K_e} : the unit exterior normal of K on the edge e with $\nu_{K_e} = (\nu^1_{K_e}, \nu^2_{K_e}, \ldots, \nu^n_{K_e})$,
K_e : neighboring cell of K with $\overline{K} \cap \overline{K_e} = e$.

We also suppose that the triangulation is regular in the sense that there exists a constant $C_1 > 0$ (independent of the triangulation $\mathcal{T}_h$) such that

$$\forall K \in \mathcal{T}_h, \quad C_1 h^n \leq m_K, \quad \text{and} \quad \forall K \in \mathcal{T}_h, \quad \forall e \in \mathcal{E}_K \quad C_1 h^{n-1} \leq m_e.$$

We want to investigate the numerical approximation (using finite volume schemes) of the following constrained Friedrichs system

$$\begin{cases} \partial_t U + \sum_{i=1}^n A_i \partial_i U = f, \quad \text{on } (0,T)\times\Omega; \quad U(0,x) = U_0(x), \quad \text{on } \Omega, \\ (A_\nu - M_\nu)U = 0, \quad \text{on } (0,T)\times\partial\Omega; \quad U(t,x) \in \mathcal{C}, \quad \text{a.e in } (0,T)\times\Omega \end{cases} \tag{3}$$

where $\mathscr{C} \subset \mathbb{R}^m$ is a closed convex (independent of t and x) with $0 \in \overset{\circ}{\mathscr{C}}$, $A_\nu = \sum_{i=1}^n A_i \nu^i$ with $\nu = (\nu^1, \dots, \nu^n)$ is the unit exterior normal to Ω, and M_ν is a non-negative symmetric matrix that encodes the boundary condition and has to satisfy some algebraic conditions (see [8, Sect. 2.1]).

Remark 1. In particular, due to the hypotheses on A_ν and M_ν, we have

1. For all $k \in \mathbb{R}^m$, there exists a unique triple (k^0, k_-, k_+) such that $k = k^0 + k_- + k_+$ and $k^0 \in \ker A_\nu$, $k^- \in (\ker(A_\nu - M_\nu)) \cap \mathrm{Im} A_\nu$, and $k^+ \in (\ker(A_\nu + M_\nu)) \cap \mathrm{Im} A_\nu$.
2. For all $k, \kappa \in \mathbb{R}^m$, $\langle k | A_\nu \kappa \rangle = \langle k_- | A_\nu \kappa_- \rangle + \langle k_+ | A_\nu \kappa_+ \rangle$.

The equations of (3) have to be understood in a weak sense (see Definition 1 for the case $\Omega = \mathbb{R}^n$ and Sect. 5 for the general case). To approximate the solutions of this kind of problem, we first forget about the constraint and use a finite volume scheme (explicit in time) based on the note [3]. More precisely, we use a piecewise constant approximation of U, denoted by V_h, such that

$$\forall (t, x) \in [t^p, t^{p+1}) \times K, \qquad V_h(t,x) = v_K^p, \quad \text{with } v_K^0 = \frac{1}{m_K} \int_K U_0(x)\, \mathrm{d}x,$$

where $0 = t^0 < t^1 < \cdots < t^{N+1} = T$ $(t^{p+1} - t^p = \Delta t)$, and in a first step, we construct

$$\frac{m_K}{\Delta t} \left(v_K^{p+1,*} - v_K^p \right) + \sum_{e \in \mathscr{E}_K} g_{K_e} m_e = f_K^p := \frac{1}{m_K \Delta t} \int_{t^p}^{t^{p+1}} \int_K f(t,x)\, \mathrm{d}x\, \mathrm{d}t,$$

where $A_{K_e} = \sum_{i=1}^n A_i \nu_{K_e}^i$ and we define the interior fluxes ($e \cap \partial\Omega = \emptyset$),

$$g_{K_e} = \underbrace{(A_{K_e})^+ v_K^p}_{\text{Outcoming flow from } K \text{ to } K_e} + \underbrace{(A_{K_e})^- v_{K_e}^p}_{\text{Incoming flow in } K \text{ from } K_e}, \tag{4}$$

where we denote $(A_{K_e})^-$ (resp. $(A_{K_e})^+$) the negative (resp. positive) part of A_{K_e}, and the (centered) boundary fluxes,

$$g_{K_e} = \frac{A_{K_e} + M_{K_e}}{2} v_K^p, \tag{5}$$

with $M_{K_e} = M_{\nu_{K_e}}$ a matrix satisfying the conditions of [8, Sect. 2.1] (see also Remark 1). In order to take account of the constraint, we simplify project on each cell K the value $v_K^{p+1,*}$ onto the set $\mathscr{C}$. Hence, the second step is

$$v_K^{p+1} = P_{\mathscr{C}} \left(v_K^{p+1,*} \right).$$

where $P_{\mathscr{C}}$ is the projection onto $\mathscr{C}$. It leads us to the following scheme for $U_0 \in L^2(\mathbb{R}^n; \mathscr{C})$,

$$\begin{cases} \forall K \in \mathscr{T}_h, \quad v_K^0 = \frac{1}{m_K} \int_K U_0(x)\,\mathrm{d}x, \\ \forall K \in \mathscr{T}_h, \forall 0 \leq p \leq N,\ v_K^{p+1,*} = v_K^p - \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_K} g_{K_e} m_e + \Delta t f_K^p, \\ \forall K \in \mathscr{T}_h, \forall 0 \leq p \leq N, \quad v_K^{p+1} = P_{\mathscr{C}}\left(v_K^{p+1,*}\right). \end{cases} \tag{6}$$

Thanks to the following discrete Green formula

$$\sum_{e \in \mathscr{E}_K} A_{K_e} m_e = 0 \quad \Leftrightarrow \quad \sum_{e \in \mathscr{E}_{K\mathrm{b}}} A_{K_e} m_e + \sum_{e \in \mathscr{E}_{K\mathrm{i}}} (A_{K_e})^+ m_e = \sum_{e \in \mathscr{E}_{K\mathrm{i}}} -(A_{K_e})^- m_e, \tag{7}$$

one can rewrite the first step of the scheme (6) in a nonconservative form

$$\frac{v_K^{p+1,*} - v_K^p}{\Delta t} = \sum_{e \in \mathscr{E}_{K\mathrm{i}}} \frac{m_e}{m_K} (A_{K_e})^- (v_K^p - v_{K_e}^p) - \sum_{e \in \mathscr{E}_{K\mathrm{b}}} \frac{m_e}{m_K} \frac{M_{K_e} - A_{K_e}}{2} v_K^p + f_K^p. \tag{8}$$

Remark 2. We denote by $\langle ; \rangle$ the canonical scalar product of $\mathbb{R}^m$ and $|.|$ the associated norm. By abuse of notation, we also use the notation $|.|$ for the (matrix) operator norm associated with the canonical norm of $\mathbb{R}^m$.

Remark 3. When $\Omega = \mathbb{R}^n$, one can use the scheme (6) to approximate the solution of the problem (2). In that case, all the sums over $\mathscr{E}_{K\mathrm{b}}$ are empty sums.

3 Previous Results on Constrained Friedrichs' Systems in the Whole Space

The aim of this section is to recall the definition of weak solutions to Friedrichs' systems under convex constraints in the whole space and to state some numerical results about these systems. We consider the following Cauchy problem: find $U : [0, T] \times \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\begin{cases} \partial_t U + \sum_{j=1}^n A_j \partial_j U = 0 \quad \text{in } (0, T] \times \mathbb{R}^n; \quad U(0, x) = U^0(x) \quad \text{if } x \in \mathbb{R}^n, \\ U(t, x) \in \mathscr{C} \quad \text{if } (t, x) \in [0, T] \times \mathbb{R}^n, \end{cases} \tag{9}$$

where $\mathscr{C}$ is a fixed (*i.e.,* independent of the time and space variables) nonempty closed and convex subset of $\mathbb{R}^m$ containing 0 in its interior, the matrices A_j are $m \times m$ symmetric matrices independent of time and space, and $T > 0$. This type of nonlinear hyperbolic problems has been introduced in [4] where a notion of weak solutions to problem (9) has been defined.

Definition 1. Let $U^0 \in L^2(\mathbb{R}^n, \mathscr{C})$ and $T > 0$. A function $U \in L^2([0,T] \times \mathbb{R}^n, \mathscr{C})$ is a weak constrained solution of (9) if we have for all $\kappa \in \mathscr{C}$ and $\phi \in \mathscr{C}_c^\infty([0,T[\times \mathbb{R}^n)$ with $\phi \geq 0$,

$$\int_0^T \int_{\mathbb{R}^n} \Big(|U-\kappa|^2 \partial_t \phi + \sum_{j=1}^n \big\langle U-\kappa;\, A_j(U-\kappa)\big\rangle\, \partial_j \phi \Big)\, \mathrm{d}x\, \mathrm{d}t \\ + \int_{\mathbb{R}^n} |U^0(x)-\kappa|^2 \phi(0,x)\, \mathrm{d}x \geq 0. \tag{10}$$

We recall here the existence and uniqueness result of [4].

Theorem 1. *Assume that $U^0 \in L^2(\mathbb{R}^n, \mathscr{C})$. There exists a unique weak constrained solution $U \in L^2([0,T] \times \mathbb{R}^n, \mathscr{C})$ to (9) in the sense of Definition 1.*

The existence of a solution has been obtained in [4] thanks to a finite volume scheme on Cartesian grids. At each time step, the scheme first let the solution evolve according to the Rusanov scheme without taking care about the constraint. Then, on each mesh they project the solution onto the set of constraints.

Thanks to this splitting strategy and to a compactness argument (which relies on the fact that the mesh is Cartesian), they show that the numerical solution admits a convergent subsequence and they prove that the limit of this subsequence has to be a solution of (9) in the sense of Definition 1.

In this paper, we use this splitting strategy for schemes defined on unstructured meshes. One can show that the scheme (6) (see Remark 3) enjoys the same rate of convergence as in the unconstrained case (for the complete proof, see [7]).

Theorem 2. *Let $U \in H^1((0,T) \times \mathbb{R}^n; \mathscr{C})$ be a dissipative solution associated with the initial condition $U_0 \in H^1(\mathbb{R}^n; \mathscr{C})$. Let V_h be the solution constructed from U_0 thanks to the scheme (6) (see Remark 3). Then we have,*

$$\|U - V_h\|_{L^2((0,T)\times\mathbb{R}^n;\mathbb{R}^m)} \leq C\sqrt{h},$$

for some constant C depending on ε, n, T, U_0 and the matrices A_i.

4 Stability in Time of Schemes

Once we know that the strategy of [4] combined with the scheme, analyzed in [9], leads to a convergent scheme (on unstructured meshes) for constrained Friedrichs' systems in $(0,T) \times \mathbb{R}^n$, one can analyze this splitting strategy on bounded domains (i.e., for Problem (3)). In this section, we prove that the scheme (6) enjoys a stability property under a Courant–Friedrichs–Lewy condition. For simplicity, we decide to derive this stability property in the case where the source term is null. In that case, the $L^2(\mathbb{R}^n)$-norm of the solution does not increase in time.

Proposition 1. *Suppose that the following CFL condition holds:*

$$\max\left(\sup_{K,e\in\mathscr{E}_K}\frac{\Delta t m_{\partial K}}{m_K}\left|(A_{K_e})^-\right|,\sup_{K,e\in\mathscr{E}_{K\,b}}\frac{\Delta t m_{\partial K}}{m_K}\left|(M_{K_e}-A_{K_e})/2\right|\right)\le 1, \quad (11)$$

the scheme (6) *is stable; i.e., the approximate solution* V_h *satisfies (here* $f\equiv 0$*)*

$$\forall t\in[0,T],\qquad \|V_h(t,\cdot)\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\le\|U_0\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\,.$$

Proof. From the nonconservative form (8), we have

$$v_K^{p+1,*}=\sum_{e\in\mathscr{E}_K}\frac{m_e}{m_{\partial K}}v_K^{p+1,*}(e),$$

where we set

$$v_K^{p+1,*}(e)=\begin{cases}v_K^p+\frac{\Delta t m_{\partial K}}{m_K}(A_{K_e})^-(v_K^p-v_{K_e}^p), & \text{if } e\in\mathscr{E}_{K\,\mathrm{i}},\\ v_K^p-\frac{\Delta t m_{\partial K}}{m_K}\frac{M_{K_e}-A_{K_e}}{2}v_K^p, & \text{if } e\in\mathscr{E}_{K\,\mathrm{b}}.\end{cases}$$

Observe that we have for all $e\in\mathscr{E}_{K\,\mathrm{i}}$, since $(A_{K_e})^-\in\mathbb{M}_{\mathrm{sym}}^{m\times m}$,

$$\begin{aligned}|v_K^{p,*}(e)|^2=|v_K^p|^2&-\tfrac{\Delta t m_{\partial K}}{m_K}\left(-\left\langle v_K^p;(A_{K_e})^-v_K^p\right\rangle+\left\langle v_{K_e}^p;(A_{K_e})^-v_{K_e}^p\right\rangle\right)\\&+\tfrac{\Delta t m_{\partial K}}{m_K}\left\langle v_K^p-v_{K_e}^p;\left(\mathrm{Id}+\tfrac{\Delta t m_{\partial K}}{m_K}(A_{K_e})^-\right)(A_{K_e})^-(v_K^p-v_{K_e}^p)\right\rangle\end{aligned}$$

Using the CFL condition, we obtain that

$$\forall y\in\mathbb{R}^m,\left\langle\left(\mathrm{Id}+\frac{\Delta t m_{\partial K}}{m_K}(A_{K_e})^-\right)y;y\right\rangle\ge 0. \quad (12)$$

In particular, if we apply (12) to $y=\left(-(A_{K_e})^-\right)^{1/2}(v_K^p-v_{K_e}^p)$, it yields

$$|v_K^{p,*}(e)|^2\le|v_K^p|^2-\frac{\Delta t m_{\partial K}}{m_K}\left(-\left\langle v_K^p;(A_{K_e})^-v_K^p\right\rangle+\left\langle v_{K_e}^p;(A_{K_e})^-v_{K_e}^p\right\rangle\right). \quad (13)$$

Now, if $e\in\mathscr{E}_{K\,\mathrm{b}}$, we have, again since A_{K_e} and M_{K_e} belong to $\mathbb{M}_{\mathrm{sym}}^{m\times m}$,

$$\begin{aligned}|v_K^{p+1,*}(e)|^2=|v_K^p|^2&-\frac{\Delta t m_{\partial K}}{m_K}\left\langle v_K^p;\frac{M_{K_e}-A_{K_e}}{2}v_K^p\right\rangle\\&-\frac{\Delta t m_{\partial K}}{m_K}\left\langle\frac{M_{K_e}-A_{K_e}}{2}\left(\mathrm{Id}-\frac{\Delta t m_{\partial K}}{m_K}\left(\frac{M_{K_e}-A_{K_e}}{2}\right)\right)v_K^p;v_K^p\right\rangle. \quad (14)\end{aligned}$$

Similarly, the CFL condition (11) implies that for all $y \in \mathbb{R}^m$, we have

$$\left\langle \mathrm{Id} - \frac{\Delta t m_{\partial K}}{m_K} \left(\frac{M_{K_e} - A_{K_e}}{2} \right) y; y \right\rangle \geq 0,$$

and algebraic manipulations (see Remark 1) tell us that

$$\left\langle \frac{M_{K_e} - A_{K_e}}{2} \left(\mathrm{Id} - \frac{\Delta t m_{\partial K}}{m_K} \left(\frac{M_{K_e} - A_{K_e}}{2} \right) \right) v_K^p; v_K^p \right\rangle$$
$$= \left\langle \left(\mathrm{Id} - \frac{\Delta t m_{\partial K}}{m_K} \left(\frac{M_{K_e} - A_{K_e}}{2} \right) \right) M_{K_e}^{1/2} (v_K^p)_+; M_{K_e}^{1/2} (v_K^p)_+ \right\rangle \geq 0,$$

which implies that (14) becomes

$$|v_K^{p+1,*}(e)|^2 \leq |v_K^p|^2 - \frac{\Delta t m_{\partial K}}{m_K} \left\langle v_K^p; \frac{M_{K_e} - A_{K_e}}{2} v_K^p \right\rangle.$$

Using convexity, it yields

$$|v_K^{p+1,*}|^2 \leq |v_K^p|^2 - \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_{K\mathrm{i}}} \left(-\langle v_K^p; (A_{K_e})^- v_K^p \rangle + \langle v_{K_e}^p; (A_{K_e})^- v_{K_e}^p \rangle \right) m_e$$
$$- \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_{K\mathrm{b}}} \left\langle v_K^p; \frac{M_{K_e} - A_{K_e}}{2} v_K^p \right\rangle m_e.$$

Furthermore, if we use the relation (7), we obtain

$$|v_K^{p+1,*}|^2 \leq |v_K^p|^2 - \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_{K\mathrm{i}}} \left(\rangle v_K^p; (A_{K_e})^+ v_K^p \rangle + \langle v_{K_e}^p; (A_{K_e})^- v_{K_e}^p \rangle \right) m_e$$
$$- \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_{K\mathrm{b}}} \left\langle v_K^p; \frac{A_{K_e} + M_{K_e}}{2} v_K^p \right\rangle m_e. \tag{15}$$

Remark that, thanks to Remark 1, we have for all $e \in \mathscr{E}_{K\mathrm{b}}$

$$\left\langle v_K^p; \frac{A_{K_e} + M_{K_e}}{2} v_K^p \right\rangle = \langle (v_K^p)_-; M_{K_e} (v_K^p)_- \rangle \geq 0.$$

Consequently, from (15) and since for all $y \in \mathbb{R}^m$, $|P_{\mathscr{C}}(y)| \leq |y|$, we obtain

$$|v_K^{p+1}|^2 \leq |v_K^p|^2 - \frac{\Delta t}{m_K} \sum_{e \in \mathscr{E}_{K\mathrm{i}}} \left(\langle v_K^p; (A_{K_e})^+ v_K^p \rangle + \langle v_{K_e}^p; (A_{K_e})^- v_{K_e}^p \rangle \right) m_e. \tag{16}$$

Then, we remark

$$\sum_{K\in\mathscr{T}_h}\sum_{e\in\mathscr{E}_{K\mathrm{i}}}\left(\left\langle v_K^p;(A_{K_e})^+v_K^p\right\rangle+\left\langle v_{K_e}^p;(A_{K_e})^-v_{K_e}^p\right\rangle\right)m_e=0.$$

Consequently, summing the inequality (16) over $K\in\mathscr{T}_h$ and from $p=0$ to $q-1$, where $t\in[0,T]$ and q an integer such that $t\in[t^q,t^{q+1})$ (or $q=N+1$ if $t=T$), leads to the stability property.

5 The Simplified Model of the Dynamical Perfect Plasticity

Let us briefly recall the equations of this model and the two points of view that one can use to describe its (theoretical) solution. First, the equations, derived from the physics of solids (see [1, Sects. 3.1 and 3.2]), of this simplified model of dynamical perfect plasticity are

$$\begin{cases}\partial_t v-\mathrm{div}\sigma=f, \qquad \nabla v=\partial_t\sigma+\partial_t p,\\ |\sigma|\le 1, \quad \text{and } \langle\sigma;\partial_t p\rangle=|\partial_t p|.\end{cases} \tag{17}$$

where $v:\Omega\times[0,T]\to\mathbb{R}$ is the velocity of the material, $\sigma:\Omega\times[0,T]\to\mathbb{R}^2$ the Cauchy stress tensor, and $p:\Omega\times[0,T]\to\mathbb{R}^2$ the plastic deformation tensor and Ω is a open bounded subset of $\mathbb{R}^2$. The tensor σ is constrained to stay in the unit closed Euclidean ball of $\mathbb{R}^2$, denoted $\overline{B}$. To these equations, we add initial and boundary conditions. The boundary condition, that comes from the hyperbolic point of view, is the following nonlinear one

$$\langle\sigma;\nu\rangle+T(v)=0, \quad \text{on } (0,T)\times\partial\Omega, \tag{18}$$

where $T(z)=\min(-1,\max(z,1))$. It shows a threshold on the velocity (due to the constraint) in the boundary condition. We also need an initial condition

$$(v,\sigma)(t=0)=(v_0,\sigma_0) \tag{19}$$

that has to satisfy two hypotheses

$$\langle\sigma_0;\nu\rangle+v_0=0 \quad \mathscr{H}^1 \text{ on } \partial\Omega, \tag{20}$$

$$|\sigma_0|\le 1 \text{ a.e. in } \Omega. \tag{21}$$

The first condition asserts that the initial condition has to satisfy the hyperbolic boundary condition that one could use in the unconstrained case, and the second condition states that the initial condition satisfies the constraint. In fact, one can show (see [1, Proposition 7.1]) that the solution of this simplified model satisfies the following inequality for all $(k, \tau) \in \mathbb{R} \times \overline{B}$ and all $\varphi \in W^{1,\infty}(\mathbb{R} \times \mathbb{R}^2)$ (with $\varphi \geq 0$ and compactly supported in $\mathbb{R} \times \mathbb{R}^2$)

$$\int_0^T \int_\Omega \left((v-k)^2 + |\sigma - \tau|^2\right) \partial_t \varphi \, \mathrm{d}x \, \mathrm{d}t + \int_\Omega \left((v_0-k)^2 + |\sigma_0 - \tau|^2\right) \varphi(0) \, \mathrm{d}x \\ - 2 \int_0^T \int_\Omega (\sigma - \tau) \cdot \nabla \varphi (v-k) \, \mathrm{d}x \, \mathrm{d}t + 2 \int_0^T \int_\Omega f(v-k) \varphi \, \mathrm{d}x \, \mathrm{d}t \\ + 2 \int_0^T \int_{\partial\Omega} (\sigma \cdot \nu - \tau \cdot \nu)(T(v) - k) \varphi \, \mathrm{d}\mathscr{H}^{n-1} \, \mathrm{d}t \geq 0. \quad (22)$$

Thanks to (18) and algebraic manipulations, one has

$$(\sigma \cdot \nu - \tau \cdot \nu)(T(v) - k) \\ = \frac{1}{4} \left((k + \tau \cdot \nu)^2 - (T(v) - k - (\sigma \cdot \nu - \tau \cdot \nu))^2\right) \geq \frac{1}{4} (k + \tau \cdot \nu)^2 , \quad (23)$$

Equation (23) allows us to rewrite (22), using the hyperbolic variable $U =^t (v, \sigma)$ as

$$\int_0^T \int_\Omega |U - \kappa|^2 \, \partial_t \varphi + \sum_{i=1}^{2} \langle U - \kappa ; A_i (U - \kappa) \rangle \, \partial_i \varphi + 2 \langle F ; U - \kappa \rangle \, \varphi \, \mathrm{d}x \, \mathrm{d}t \\ + \int_\Omega |U_0 - \kappa|^2 \, \varphi(t=0) \, \mathrm{d}x + \int_0^T \int_{\partial\Omega} \langle \kappa_+ ; M_\nu \kappa_+ \rangle \, \varphi \, \mathrm{d}\mathscr{H}^{n-1}(x) \, \mathrm{d}t \geq 0, \quad (24)$$

where $F =^t (f, 0, 0)$, $U_0 =^t (v_0, \sigma_0)$, $\kappa =^t (k, \tau)$

$$A_1 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad M_\nu = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\nu^1)^2 & \nu^1 \nu^2 \\ 0 & \nu^1 \nu^2 & (\nu^2)^2 \end{pmatrix}, \quad (25)$$

and κ_+ stands for the projection onto $(\ker(A_\nu + M_\nu)) \cap \mathrm{Im} A_\nu$. The fact that Eq. (24) is satisfied for all κ and all φ is the definition of a solution to Problem (3) (see also [8]). In addition, when the solution U is in $W^{1,\infty}([0, T]; L^2(\Omega; \mathscr{C})$, one can show (see [1, Sect. 7]) that Eqs. (17)–(19) are equivalent to this definition of a weak constrained solution to Problem (3).

6 A Numerical Test on the Simplified Model of the Dynamical Perfect Plasticity

Now that this mechanical problem has been put into the hyperbolic framework (3), the simplified model of dynamical perfect plasticity can be approached thanks to the scheme described in Sect. 2. One important point to notice first is that this scheme does not include a special treatment at the boundary to model the nonlinear boundary condition (18). Indeed, we only take into account the constraint thanks to a projection step on every mesh and the first step of this scheme uses the linear boundary condition

$$(A_\nu - M_\nu)U = 0 \qquad \Leftrightarrow \qquad \langle\sigma;\nu\rangle + v = 0. \tag{26}$$

Our goal now is to test numerically the interactions between the boundary condition and the constraint for this particular hyperbolic system under constraint and to see if the nonlinear boundary condition is obtained with this scheme. The major point that allows us to bring to light these facts is the velocity threshold overrun in the boundary condition (18). To observe this overrun, we present here one test case (for more test cases, see [7, Sect. 4.4]).

The test is based on the following formal motivation: We want to observe large velocities near the boundary. But if we look at the equation of motion

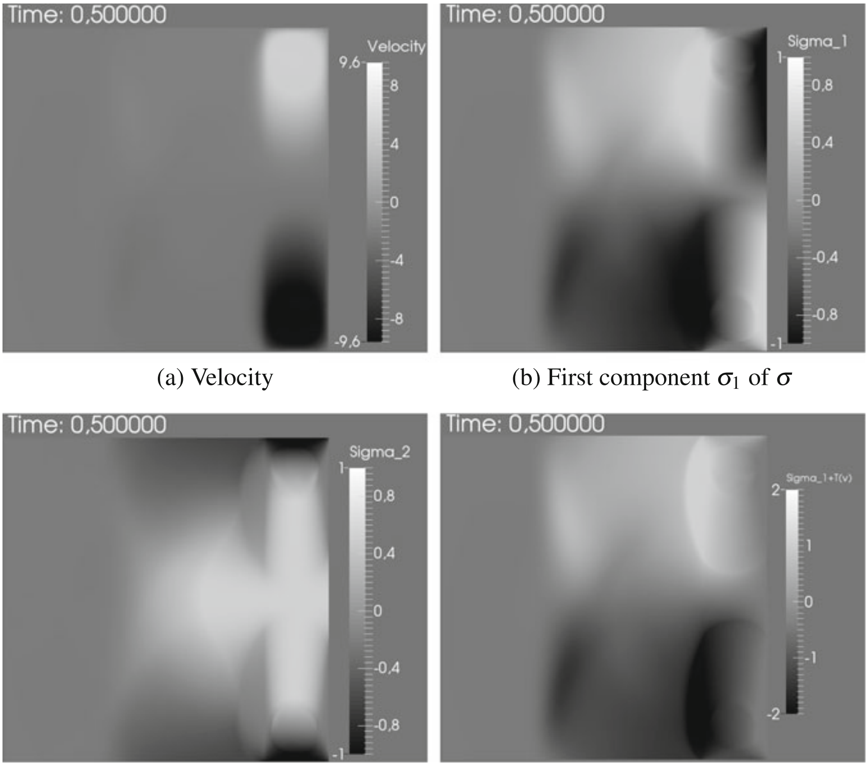
$$\partial_t v - \mathrm{div}\sigma = f,$$

we see that if f is positive (for example) near the boundary (for each time), then the velocity is going to increase over time near the boundary. Hence, we present a test case when the source term f varies from -50 to 50 near the boundary and is equal to zero elsewhere.

This test allows us to obtain large velocity near the boundary (i.e., $|v| \gg 1$ near $\partial\Omega$) and to bring to light that the nonlinear boundary is taken into account by our scheme. For this test case, we use the following data

- Spatial domain : $\Omega = [0, 1] \times [0, 1]$. Our mesh is regular and contains 80000 triangles.
- Final time : $T = 1$. We use 800 time steps, and consequently, the CFL condition (11) is approximately equal to 0.71.
- Initial data : In this test, we use data that touch the boundary $x = 1$. The initial velocity v_0 is null outside the open ball B_1 of radius 0.3 and center $(1, 0.5)$, and v_0 is equal to -1 on the open ball B_2 of radius 0.25 and center $(1, 0.5)$. In the strip between these two balls, we join these two constants using a $\mathscr{C}^1$ connection. It is important to notice that $-1 \leq v_0 \leq 0$. In order to satisfy the (linear) boundary condition at $x = 1$, the first component of σ is equal to $-v_0$. The second component of σ is null on Ω. Consequently, we have $v_0 + \langle\sigma;\nu\rangle = 0$ on $\partial\Omega$. Remark also that the initial data belong to the convex set of constraints.
- The term source f is equal to $100y - 50$ for all $t \in [0, T]$, for all $y \in [0, 1]$ and $x > 0.8$ and to 0 elsewhere.

We decide to highlight the interaction between the constraint and the boundary at time $t = 0.5$ in Fig. 2. In this figure, we display the velocity (top left of the figure), the first component, denoted σ_1 in the following, of σ (top right), the second component (bottom left), denoted σ_2, and the term $\sigma_1 + T(v)$ (which is involved in the boundary condition at $x = 1$: $\sigma_1 + T(v) = 0$).



(a) Velocity

(b) First component σ_1 of σ

(c) Second component σ_2 of σ

(d) $\sigma_1 + T(v)$ (*i.e.* the boundary term on the right of the domain)

Fig. 2 Test case at time $t = 0.5$

We observe that the introduction of our term source in the strip $[0.8, 1] \times [0, 1]$ allows us to get a large velocity (i.e., $|v| \gg 1$) near the boundary $x = 1$ (see Fig. 2a). The theoretical boundary condition implies that in this situation we should see that $\sigma_1 = -1$ at the upper end of the boundary $x = 1$ (and, consequently, $\sigma_2 = 0$ due to the constraint) and $\sigma_1 = 1$ at the lower end of the boundary $x = 1$ (and $\sigma_2 = 0$ due to the constraint). Numerically, the scheme produces a solution that matches the mathematical model (see Fig. 2b, c). Consequently, the nonlinear boundary condition is satisfied by the numerical approximation (see Fig. 2d) despite the fact that we have not implemented any particular treatment at the boundary to get this nonlinear boundary condition. This fact may be seen as a first validation of our scheme.

References

1. J.-F. Babadjian, C. Mifsud, Hyperbolic structure for a simplified model of dynamical perfect plasticity. Arch. Ration. Mech. Anal. **223**(2), 761–815 (2017)
2. B. Cockburn, C.-W. Shu, TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. Math. Comp. **52**(186), 411–435 (1989)
3. Y. Coudière, J.-P. Vila, P. Villedieu, Convergence d'un schéma volumes finis explicite en temps pour les systèmes hyperboliques linéaires symétriques en domaines bornés. C. R. Acad. Sci. Paris Sér. I Math. **331**(1), 95–100 (2000)
4. B. Després, F. Lagoutière, N. Seguin, Weak solutions to Friedrichs systems with convex constraints. Nonlinearity **24**(11), 3055–3081 (2011)
5. R. Eymard, T. Gallouët, R. Herbin, Finite volume methods, in *Handbook of Numerical Analysis*, vol. VII (North-Holland, Amsterdam, 2000), pp. 713–1020
6. R.J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics (Cambridge University Press, Cambridge, 2002)
7. C. Mifsud, Variational and hyperbolic methods applied to constrained mechanical systems. Ph.D. thesis, Université Pierre et Marie Curie (2016)
8. C. Mifsud, B. Després, N. Seguin, Dissipative formulation of initial boundary value problems for Friedrichs' systems. Commun. Partial Differ. Equ. **41**(1), 51–78 (2016)
9. J.-P. Vila, P. Villedieu, Convergence of an explicit finite volume scheme for first order symmetric systems. Numer. Math. **94**(3), 573–602 (2003)