Matrix-valued problems reminiscent of optimal transport and their applications to PDEs

Dmitry Vorotnikov

Universidade de Coimbra

IPAM, UCLA, May 2025

Prelude: 'Ballistic' problems

- $\mathfrak{M}$ is a complete, connected Riemannian manifold
- $X : \mathfrak{M} \to T\mathfrak{M}$ is a given vector field on $\mathfrak{M}$
- $x_T \in \mathfrak{M}$ is a given point
- By a **ballistic geodesic problem** we mean finding a curve γ_t that minimizes

$$\min_{\dot{\gamma}_0 = X(\gamma_0),\ \gamma_T = x_T} \int_0^T \langle \dot{\gamma}_t, \dot{\gamma}_t \rangle_{\gamma_t}\, dt \tag{1}$$

- That is, we know the **final position** and the **initial velocity** of an unknown geodesic (mixed boundary condition: inhomogeneous Neumann at $t = 0$, Dirichlet at $t = T$).

- If $X = \operatorname{grad}\Phi$ for some $\Phi : \mathfrak{M} \to \mathbb{R}$, the ballistic problem (1) is formally equivalent to the **Hopf-Lax formula**

$$\min_{x_0 \in \mathfrak{M}} \Phi(x_0) + \frac{1}{2} dist^2(x_0, x_T), \tag{2}$$

  where *dist* is the Riemannian distance on $\mathfrak{M}$
- What matters here is the minimizer x_0 that determines the starting point of the unknown geodesic and not the optimal value
- Formula (2) makes sense if we have no manifold structure but just a metric space, not necessarily connected
- It appears in de Giorgi's minimizing movement scheme — aka JKO scheme — in the infinite-dimensional and/or metric geometric context (in particular, in the context of optimal transport)

- Similar problems are referred to as marginal entropy-transport problems (Liero-Mielke-Savaré, Invent. Math. '18)
- This is also related to the mean-field games in the spirit of Lasry-Lions
- In order to formulate 'ballistic optimal transport' (with quadratic cost), let us start with the quadratic Hamilton-Jacobi equation

$$\partial_t \psi + \frac{1}{2}|\nabla \psi|^2 = 0, (t, x) \in [0, T] \times \Omega$$

- Here, for simplicity, Ω is the periodic box $\mathbb{T}^d$
- Rescaling if needed, we can assume $T = 1$ and $|\Omega| = 1$
- Fix the initial data $\psi(0, x) = \psi_0(x)$

A 'ballistic' problem in the Wasserstein space

- Now set $v = \nabla\psi$ to recast the Cauchy problem in the form

$$\partial_t v + \frac{1}{2}\nabla \operatorname{tr}(v \otimes v) = 0, \ v(0) = \nabla\psi_0 \qquad (3)$$

- Following Brenier (CMP '18), let us consider the problem of finding a weak solution to (3) that would **minimize** the time average of the kinetic energy

$$\frac{1}{2}\iint_{(0,T)\times\Omega} |v|^2(t,x)\,dx\,dt$$

- This problem might not always admit a solution, but leads to a dual variational problem

- The dual problem, at least formally, reads

$$-\int_\Omega \psi_0(x)\rho(0,x)\,dx - \frac{1}{2}\int_0^T \int_\Omega \rho^{-1} q \cdot q\,dx\,dt \to \sup \tag{4}$$

subject to the constraints

$$\partial_t \rho + \operatorname{div} q = 0, \quad d\rho(T) = d\rho_1 := dx, \quad \rho \ge 0 \tag{5}$$

- Define the functional

$$\Psi(\rho) := \int_\Omega \psi_0 d\rho$$

on the 2-Wasserstein space $(\mathcal{P}(\Omega), W_2)$ of probability measures on Ω

- Multiplying by -1, we rewrite (4), (5) in the ballistic/Hopf-Lax form

$$\min_{\rho_0 \in \mathcal{P}(\Omega)} \Psi(\rho_0) + \frac{1}{2} W_2^2(\rho_0, \rho_1) \tag{6}$$

One-dimensional case

- For $d = 1$ (that is, $x \in \mathbb{T}^1$) system (3) is a conservation law, or, more precisely, the inviscid Burgers equation:

$$\partial_t v + \frac{1}{2}\partial_x(v^2) = 0, \ v(0) = \partial_x \psi_0 \tag{7}$$

- Brenier showed that any entropy solution to (7) — possibly discontinuous — can be retrieved from the dual formulation (4), (5).

The incompressible Euler system

- The incompressible Euler system reads

$$\begin{aligned} \partial_t v + \operatorname{div}(v \otimes v) + \nabla p &= 0, \\ \operatorname{div} v &= 0, \\ v(0,x) &= v_0(x). \end{aligned}$$

- The unknowns are $v : [0,T] \times \Omega \to \mathbb{R}^d$ and $p : [0,T] \times \Omega \to \mathbb{R}$.
- The (kinetic) energy $\frac{1}{2} \int_\Omega |v|^2(t,x)\, dx$ is formally conserved along the flow.
- Brenier (CMP '18) suggested to search for the solutions of the incompressible Euler that minimize the averaged kinetic energy.

- Up to regularity issues, the dual problem reads

$$-\int_\Omega U_0(x) : G(0,x)\,dx - \frac{1}{2}\int_0^T \int_\Omega G^{-1} q \cdot q\,dx\,dt \longrightarrow \sup \tag{8}$$

subject to the constraints

$$\partial_t G + (\nabla + \nabla^T)(Pq) = 0, \quad G(T) = I. \tag{9}$$

Here P is the Leray-Helmholtz projector, U_0 is any matrix-valued function such that

$$P(\operatorname{div} U_0) = v_0$$

(such U_0 exists if $\operatorname{div} v_0 = 0$ and v_0 has zero mean on the periodic box). Here and below I is the identity matrix of appropriate size.

Interim conclusions

- Ballistic problems can emerge from dual formulation of Cauchy problems for PDEs
- If we extend this framework to systems of PDEs beyond the quadratic HJ case, the associated scalar measure ρ is typically replaced with a non-negative-definite **matrix-valued measure** G.
- Such dual problems are expected to carry information about the physically relavant solutions to the original systems of PDEs

Related work

- Brenier and Moyano (PTRSA '22): dual multi-stream formulation of the Euler-Poisson system
- V. (ARMA '22) developed Brenier's approach by finding structures in nonlinear **quadratic** PDEs that permit to set up similar dual problems (and prove existence, consistency, weak-strong uniqueness, absence of duality gap etc.) in a **systematic** way
- Mirebeau and Stampfli '25: discretization, convergence, and numerical implementation of the dual ballistic formulation of the quadratic porous medium equation and Burgers' equation

Related work

- A similar duality scheme has been (rather independently) proposed by Acharya (QAM '23), and subsequently developed by him and several collaborators (including Zarnescu, Stroffolini, Pego et al.) in a vast series of recent papers
- A notable feature in Acharya's construction of the dual problem is a 'base state' that can be regarded as an 'initial guess' for the solution of the original PDE
- From this point of view, the 'base states' in Brenier's framework are identically zero

In search of a selection principle

- Cauchy problems for systems of nonlinear evolutionary PDEs can have infinitely many weak solutions
- Convex integration (De Lellis, Székelyhidi et al.)
- No universal selection principle for weak solutions
- Many of such problems possess a physically relevant quantity (e.g., a Hamiltonian, energy or entropy) that should be **formally conserved** along the flow, but this **may fail** for weak solutions
- For physically relevant weak solutions, such quantity — hereafter we refer to it as the **total entropy** — is generally expected to remain below or equal to its initial value
- The wording comes from the theory of conservation laws and might be confusing. For example, smooth solutions of the heat equation dissipate the Boltzmann entropy, so this entropy is never conserved. In this talk the 'total entropy' is a formally conserved quantity

Extremal principles in natural sciences

- Prigogine's principle (aka minimum entropy production principle, 1945, thermodynamics and beyond, open systems)
- Ziegler's principle (aka maximum entropy production principle, 1963, originally for nonequilibrium thermodynamics, closed systems)
- Related principles in physics, information theory, chemistry and biology: Onsager, Gyarmati, Berthelot, Swenson, Lotka, Enskog, Kohler, Haken, Paltridge, Malkus, Veronis, Jaynes et al.
- Such principles can be employed both for derivation of physically relevant systems of PDEs and for selection of physically relevant solutions

A more mathematical approach: Dafermos' principle

- Dafermos' principle (1973): physically relevant solutions dissipate the total entropy earlier and faster than the irrelevant ones (cf. Ziegler's principle)
- Typically, the total entropy is conserved for smooth solutions but can dissipate for weak solutions due to shocks etc.
- It is important that the criterion acts locally near the 'bifurcation' moment of time, i.e., before a certain moment the total entropies of a 'good' and a 'bad' solution are equal (this stage is optional and the discrepancy can already occur at the initial moment of time), but shortly after that moment the total entropy of a 'bad' solution becomes larger than the total entropy of a 'good' solution; however, it is permitted that, as time elapses, the total entropy of a 'bad' solution returns to being smaller than or equal to the one of a 'good' solution

Dafermos' principle

- Appropriateness of Dafermos' principle was examined for various PDEs: Dafermos, Hsiao, Feireisl, Chiodaroli, Kreml, Cieślak, Jamróz et al.
- Numerical applications: Klein (2023, 2024)
- For systems of conservation laws, there exist other selection principles due to Kruzhkov, Lax, Liu et al.
- Dafermos' principle complies with the above criteria for scalar conservation laws ($n = 1$) and for several other models.
- Notably, Dafermos' principle needs just one physically relevant entropy (that is formally conserved) and its applicabilty is not restricted to conservation laws or first-order PDEs.

Back to the duality scheme: non-quadratic examples

- It was clear from the very beginning (Brenier, Acharya) that the nonlinearity in a PDE does not need to be quadratic in order to implement the duality construction
- However, the majority of rigorous results for this kind of problems has until recently been obtained for quadratic nonlinearities
- Of course, for many relevant systems the nonlinearity **fails to be quadratic**
- This is the case for **systems of conservation laws** and for various **nonlinear dispersive equations**, including the **defocusing NLS** equation

$$\mathrm{i}\partial_t \Psi = -\Delta\Psi + f(|\Psi|^2)\Psi$$

- The unknown is $\Psi : [0, T] \times \Omega \to \mathbb{C}$. Here $f : \mathbb{R} \to \mathbb{R}$ is a given increasing nonlinear function

Representative non-quadratic examples

- The following famous system of conservation laws describes motion of compressible barotropic fluids:

$$\partial_t q + \operatorname{div}\left(\frac{q \otimes q}{\rho}\right) + \nabla(P(\rho)) = 0,$$
$$\partial_t \rho + \operatorname{div} q = 0,$$
$$q(0) = q_0, \quad \rho(0) = \rho_0.$$

- The unknowns are $(q, \rho) : [0, T] \times \Omega \to \mathbb{R}^d \times \mathbb{R}_+$
- Here $P : \mathbb{R}_+ \to \mathbb{R}_+$ is the pressure function (which is often assumed to be convex)

A general framework

Our aim is to provide a systematic framework for capturing various aspects of the duality scheme across a broad class of systems of PDEs, including non-quadratic ones. For this purpose, we consider the following abstract problem

$$\partial_t v = L(\mathbf{F}(v)),\ v(0,\cdot) = v_0.$$

Here $v_0 : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$, is the initial datum, $v : [0,T] \times \Omega \to \mathbb{R}^n$ is an unknown vector function,

$$\mathbf{F} : \mathbb{R}^n \to \mathbb{R}_s^{N\times N}$$

is a prescribed C^2-smooth matrix function with some convexity and positivity properties to be discussed below (if time permits), and L is a vector-valued differential operator with constant coefficients

$$L(\Xi)_i = \sum_{j=0}^{\nu} \sum_{l,m=1}^{N} \sum_{|\alpha|=j} b_{ilm\alpha} \partial_\alpha \Xi_{lm},\ i = 1,\dots,n,$$

where α is a generic multiindex and $\Xi(x) \in \mathbb{R}_s^{N\times N}$.

Entropy

Define the 'entropy' function for our problem by

$$\mathbf{K} : \mathbb{R}^n \to \mathbb{R}, \ \mathbf{K}(v) := \frac{1}{2}\mathrm{tr}(\mathbf{F}(v) - \mathbf{F}(0)).$$

Assume that $\mathbf{K}$ is strictly convex (but not necessarily uniformly convex). We will also require some extra technical assumptions in order to work with the *anisotropic Orlicz space* $L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. (Roughly speaking, that Orlicz space consists of functions $v : \Omega \to \mathbb{R}^n$ such that $\mathbf{K}(v(x))$ is integrable on Ω.)

‘Legendre map’

To any given $v \in \mathbb{R}^n$ we associate the ‘sharp’ vector

$$v^{\#} := \nabla \mathbf{K}(v),$$

where the gradient ∇ is taken w.r.t. v. Note that $0^{\#} = 0$ because $\nabla \mathbf{K}(0) = 0$. Since $\mathbf{K}$ is strictly convex, the map

$$\nabla \mathbf{K} : v \mapsto v^{\#}$$

is C^1-smooth and injective. Moreover,

$$v = \nabla \mathbf{K}^{*}(v^{\#})$$

for any $v^{\#} \in \nabla \mathbf{K}(\mathbb{R}^n) = \mathbb{R}^n$.

Λ-convexity

We will require some convexity of the matrix function $\mathbf{F}$. The classical notion of convexity of matrix-valued functions is the Loewner convexity, namely, $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}_s^{N\times N}$ is Loewner convex if $\mathbf{F} : P$ is a convex function for any matrix $P \in \mathbb{R}_+^{N\times N}$. This is too restrictive for our purposes, so we will merely assume that $\mathbf{F}$ is Λ*-convex* in the following more relaxed sense. Let $u : \Omega \to \mathbb{R}^n$ be an arbitrary smooth function, and let L^* be the differential operator adjoint to L. Let $\Lambda \subset \mathbb{R}_s^{N\times N}$ be the smallest linear subspace of $\mathbb{R}_s^{N\times N}$ independent of u and containing I such that $L^*(u(x)) \in \Lambda$, $x \in \Omega$. We say that $\mathbf{F}$ is Λ-convex if $\mathbf{F} : P$ is a convex function for any $P \in \Lambda \cap \mathbb{R}_+^{N\times N}$. Of course, any Loewner convex matrix function is Λ-convex, but not vice versa. Notably, it follows from the Λ-convexity of $\mathbf{F}$ that $\mathbf{K}$ is convex.

Λ-order

We will also require some positivity of the matrix function **F**. For this purpose, we will use the less restrictive Λ-order instead of the Loewner order. Namely, denote

$$\mathbb{R}_\Lambda^{N\times N} = \{\Xi \in \mathbb{R}_s^{N\times N} | \, \Xi : P \geq 0, \; \forall P \in \Lambda \cap \mathbb{R}_+^{N\times N}\}.$$

We will assume that

$$\mathbf{F}(\mathbb{R}^n) \subset \mathbb{R}_\Lambda^{N\times N}.$$

For $A, B \in \mathbb{R}_s^{N\times N}$, we write $B \leq A$ when $A - B \in \mathbb{R}_\Lambda^{N\times N}$.

Quadratic nonlinearities are not ruled out

The quadratic matrix function

$$\mathbf{F}(v) = v \otimes v$$

and the corresponding entropy

$$\mathbf{K}(v) = \frac{1}{2}|v|^2$$

obviously satisfy our assumptions for any subspace $\Lambda \subset \mathbb{R}_s^{n\times n}$. Note that in this case $N = n$ and $v^{\#} = v$. This applies to the incompressible Euler, ideal incompressible MHD, the Muskat problem, Camassa-Holm, equations of ideal convection, incompressible isotropic elastic fluids and many other examples.

Conservativity

We will focus on the situation when L satisfies the "formal conservativity condition"

$$(\mathbf{F}(v), L^*(v^\#)) = 0 \tag{10}$$

provided $v^\# : \Omega \to \mathbb{R}^n$ is a smooth. Here and below we use the shortcut

$$(U, V) := \int_\Omega U(x) : V(x)\, dx.$$

This **formally** yields that the total entropy

$$K(t) := \int_\Omega \mathbf{K}(v(t))\, dx$$

is conserved along the solutions to our problem $\partial_t v = L(\mathbf{F}(v))$.

Example: defocusing NLS

We will discuss the NLS (but similar results are true for the NLKG and a nonlinear "defocusing" variant of KdV):

$$\mathrm{i}\partial_t \Psi = -\Delta\Psi + |\Psi|^{2q}\Psi,\ \Psi(0) = \Psi_0,$$

where $q \geq \frac{1}{2}$ is a given constant. The unknown is $\Psi : [0, T] \times \Omega \to \mathbb{C}$.
We first change the variable $\psi := \Psi e^{-\mathrm{i}t}$ to rewrite this in the form

$$\mathrm{i}\partial_t \psi = -\Delta\psi + |\psi|^{2q}\psi + \psi,\ \psi(0) = \Psi_0.$$

We now let $a = \mathfrak{Re}\,\psi, b = \mathfrak{Im}\,\psi, \delta = \nabla a, \beta = \nabla b$. Then the system becomes

$$\begin{gathered}
\partial_t a = -\operatorname{div}\beta + (a^2 + b^2)^q b + b, \\
\partial_t b = \operatorname{div}\delta - (a^2 + b^2)^q a - a, \\
\partial_t \delta = -\Delta\beta + \nabla((a^2 + b^2)^q b + b), \\
\partial_t \beta = \Delta\delta - \nabla((a^2 + b^2)^q a - a).
\end{gathered}$$

We let

$$n = 2d+2, N = 2d+4,$$
$$v = (a, b, \delta, \beta),$$
$$\bar{v} = (1, a, b, \delta, \beta, (a^2+b^2)^q),$$

$$\underbrace{\mathbf{K}(v) = \frac{1}{2}\left(|\delta|^2 + |\beta|^2 + a^2 + b^2 + \frac{1}{q+1}(a^2+b^2)^{q+1}\right),}_{\text{anisotropic entropy}}$$

$$\mathbf{F}(v) = \epsilon\bar{v}\otimes\bar{v} + \left(\frac{2}{N}\mathbf{K}(v)+1\right)e_1\otimes e_1 + \left(-\epsilon a^2 + \frac{2}{N}\mathbf{K}(v)\right)e_2\otimes e_2$$
$$+\left(-\epsilon b^2 + \frac{2}{N}\mathbf{K}(v)\right)e_3\otimes e_3 + \sum_{m=1}^{d}\left(-\epsilon|\delta_m|^2 + \frac{2}{N}\mathbf{K}(v)\right)e_{3+m}\otimes e_{3+m}$$
$$+\sum_{m=1}^{d}\left(-\epsilon|\beta_m|^2 + \frac{2}{N}\mathbf{K}(v)\right)e_{3+d+m}\otimes e_{3+d+m} + \left(-\epsilon(a^2+b^2)^{2q} + \frac{2}{N}\mathbf{K}(v)\right)e_N\otimes e_N,$$

where $\epsilon = \epsilon(q,d) > 0$ is fixed but sufficiently small.

To represent the NLS in our absract form, define the operator L as follows:

$$L\begin{pmatrix} a_{11} & a_{12} & a_{13} & A_{14} & A_{15} & a_{16} \\ a_{12} & a_{22} & a_{23} & A_{24} & A_{25} & a_{26} \\ a_{13} & a_{23} & a_{33} & A_{34} & A_{35} & a_{36} \\ A_{14}^\top & A_{24}^\top & A_{34}^\top & A_{44} & A_{45} & A_{46} \\ A_{15}^\top & A_{25}^\top & A_{35}^\top & A_{45}^\top & A_{55} & A_{56} \\ a_{16} & a_{26} & a_{36} & A_{46}^\top & A_{56}^\top & a_{66} \end{pmatrix} = \epsilon^{-1}\begin{pmatrix} -\operatorname{div} A_{15} + a_{36} + a_{13} \\ \operatorname{div} A_{14} - a_{26} - a_{12} \\ -\Delta A_{15} + \nabla(a_{36} + a_{13}) \\ \Delta A_{14} - \nabla(a_{26} + a_{12}) \end{pmatrix}.$$

It can easily be shown that Λ consists of the elements of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & A_{14} & A_{15} & 0 \\ a_{12} & a_{11} & 0 & 0 & 0 & a_{26} \\ a_{13} & 0 & a_{11} & 0 & 0 & a_{36} \\ A_{14}^\top & 0^\top & 0^\top & a_{11} I & 0 & 0 \\ A_{15}^\top & 0 & 0 & 0 & a_{11} I & 0 \\ 0 & a_{26} & a_{36} & 0 & 0 & a_{11} \end{pmatrix}.$$

Moreover,

$$v^{\#} = (y, z, \delta, \beta) = (a + (a^2 + b^2)^q a, b + (a^2 + b^2)^q b, \delta, \beta),$$

and all our assumptions, including the formal conservativity, can be verified.

Weak solutions

Definition

A function $v \in L_{\mathbf{K}}((0,T)\times\Omega;\mathbb{R}^n)$ is a *weak solution* to the abstract problem

$$\partial_t v = L(\mathbf{F}(v)),\ v(0,\cdot) = v_0. \tag{11}$$

if

$$\int_0^T [(v - v_0, E) + (\mathbf{F}(v), B)]\, dt = 0$$

for all pairs $(E,B) \in L_{\mathbf{K}^*}((0,T)\times\Omega;\mathbb{R}^n)\times L^\infty((0,T)\times\Omega;\mathbb{R}_s^{N\times N})$ meeting the constraint

$$\int_0^T [(B, \partial_t\Psi) + (E, L\Psi)]\, dt = 0 \tag{12}$$

for all sufficiently smooth vector fields $\Psi : [0,T] \to X_s^{N\times N}$, $\Psi(0) = 0$.

Definition

A pair of functions

$$(v, M) \in L_{\mathbf{K}}((0,T)\times\Omega;\mathbb{R}^n)\times L^1((0,T)\times\Omega;\mathbb{R}_s^{N\times N}),\ \mathbf{F}(v)\leq M, \tag{13}$$

is a *subsolution* to (11) if it satisfies

$$\int_0^T [(v - v_0, E) + (M, B)]\,dt = 0$$

for all pairs

$$(E, B) \in L_{\mathbf{K}^*}((0,T)\times\Omega;\mathbb{R}^n)\times L^\infty((0,T)\times\Omega;\mathbb{R}_s^{N\times N})$$

meeting the constraint (12).

The corresponding entropy of a subsolution is of course $\frac{1}{2}\operatorname{tr}(M - \mathbf{F}(0))$.

Strong solutions

Remarkably, the formal conservativity implies that we can (at least formally) change the variables and rewrite our abstract problem (11) in terms of $v^\#$:

$$\partial_t(v^\#)_l + L^*(v^\#) : \partial_l \mathbf{F}(\nabla \mathbf{K}^*(v^\#)) = 0, \quad v^\#(0,\cdot) = v_0^\# \tag{14}$$

Definition

Assume that $v_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. A function v is a *strong solution* to our abstract problem (11) if it is a weak solution,

$$\partial_t((T-t)v^\#) \in L_{\mathbf{K}^*}((0,T)\times\Omega; \mathbb{R}^n),$$
$$(T-t)L^*(v^\#) \in L^\infty((0,T)\times\Omega; \mathbb{R}_s^{N\times N}),$$

and (14) holds a.e. in $(0,T)\times\Omega$ and in Ω, resp.

Strong solutions

- We can prove that the strong solutions are **unique and conserve the total entropy** $K(t)$
- The proof employs the funny 'Jeffreys divergence'

$$\mathcal{J}(t) := (u(t) - v(t), u^{\#}(t) - v^{\#}(t)).$$

- Both conservativity and uniqueness are not valid for weak solutions, let alone for subsolutions

Primal problem

Fix a positive **weight** function $\mathfrak{h}(t)$ bounded away from 0 and ∞ on $[0, T]$ (a typical choice is $\mathfrak{h}(t) = e^{-\gamma t}$ with large γ). The idea is to search for the weak solution that **minimizes**

$$\int_0^T \mathfrak{h}(t) K(t)\, dt.$$

This leads to the saddle-point problem

$$\mathcal{I}(v_0, T) = \inf_{v} \sup_{E,B:\,(12)} \int_0^T \left[(v - v_0, E) + \frac{1}{2}(\mathbf{F}(v), \mathfrak{h} I + 2B) \right] dt. \tag{15}$$

The infimum in (15) is taken over all $v \in L_{\mathbf{K}}((0,T) \times \Omega; \mathbb{R}^n)$, and the supremum is taken over all pairs

$$(E, B) \in L_{\mathbf{K}^*}((0,T) \times \Omega; \mathbb{R}^n) \times L^\infty((0,T) \times \Omega; \mathbb{R}_s^{N\times N})$$

meeting the linear constraint (12).

Duality

It is actually relevant to minimize the weighted total entropy along the larger class of subsolutions:

$$\tilde{\mathcal{I}}(v_0, T) = \inf_{v,M:\mathbf{F}(v)\leq M} \sup_{E,B:\,(12)} \int_0^T \left[(v - v_0, E) + \frac{1}{2}(M, \mathfrak{h}I + 2B) \right] dt. \qquad (16)$$

The infimum in (15) is taken over all $v \in L_{\mathbf{K}}((0,T)\times\Omega;\mathbb{R}^n)$ and $M \in L^1((0,T)\times\Omega;\mathbb{R}^{N\times N}_{\Lambda})$, and the supremum is taken over all pairs (E, B) satisfying the same restrictions as above. The problem **dual** to (16) is

$$\tilde{\mathcal{J}}(v_0, T) = \sup_{\underbrace{E,B:\,(12)}_{\text{'transport' constraint}}} \inf_{v,M:\mathbf{F}(v)\leq M} \int_0^T \left[(v - v_0, E) + \frac{1}{2}(M, \underbrace{\mathfrak{h}I + 2B}_{\text{matricial 'density'}}) \right] dt,$$

where v, M, E, B are varying in the same function spaces as above.

Consistency

Theorem

Denote $\mathfrak{H}(t) := \int_t^T \mathfrak{h}(s)\,ds$. *Let* v *be a strong solution to* (11) *satisfying*

$$\mathfrak{h}I \geq -2\mathfrak{H}(t)L^*(v^{\#}) \text{ a.e. in } (0,T)\times\Omega. \quad (17)$$

Then $\mathcal{I}(v_0,T) = \tilde{\mathcal{J}}(v_0,T) = \tilde{\mathcal{I}}(v_0,T)$. *The pair* (E_+,B_+) *defined by*

$$B_+ = L^* a,\ E_+ = \partial_t a,$$

where

$$a = \mathfrak{H}v^{\#},$$

is a maximizer for the dual problem. Moreover, one can invert these formulas and express v *in terms of* E_+ *as follows*

$$v(t,x) = \nabla \mathbf{K}^*\left(\frac{1}{\mathfrak{H}(t)}\int_t^T (-E_+)(s,x)\,ds\right), \quad t < T. \quad (18)$$

Limitation (17) can be overcome by adapting the weight $\mathfrak{h}$

At first glance, condition (17) indicates that the interval for the which the consistency holds can be smaller than the interval $[0, T)$ on which the strong solution exists. However, we can guarantee the consistency on any interval $[0, T_1]$, $T_1 < T$, by setting $\mathfrak{h}(t) := \exp(-\gamma t)$ with sufficiently large γ.

Application: a reinterpretation of Dafermos' principle

Theorem

*Let $v_0 \in L_{\mathbf{K}}(\Omega;\mathbb{R}^n)$. Let v be a strong solution to (11) on the interval $[0,T]$ with total entropy $K(t) = K(0) = \int_\Omega \mathbf{K}(v_0)\,dx$, and let (u,M) be a weak solution on $[0,T]$ with total entropy $\tilde{K}(t)$. Then for any $0 \le t_0 < t_1 \le T$ it cannot simultaneously be that $\tilde{K}(t) \le K(t)$ for a.a. $t \in (0,t_1)$ and $\tilde{K}(t) < K(t)$ for a.a. $t \in (t_0,t_1)$. Moreover, the same result is true if u is **merely a subsolution**. In particular, it is impossible that $\tilde{K}(t) < K(t)$ for a.a. $t \in (0,\epsilon)$, $\epsilon > 0$.*

Quadratic case and comparison with [Acharya et al.]

These results are new even for the quadratic case $\mathbf{F}(v) = v \otimes v$, and are consequently applicable to the incompressible Euler. Another global in time consistency result for the dual formulation of the incompressible Euler has recently been obtained by Acharya, Stroffolini and Zarnescu. Nevertheless, our results seem to be of a different nature, and, furthermore, the two attitudes complement each other to a certain degree. Indeed, they prove that if the strong solution v to the incompressible Euler coincides with the 'base state' $\tilde{v}$, then the solution of the dual problem is identically zero. Thus, in their case the information about the strong solution is contained not in the solution of the dual problem but in the 'base state' only. Moreover, their proof ignores the formal conservativity of the problem (and therefore can be extended to the Navier-Stokes). In contrast, our 'base state' is zero, and the information about the strong solution is contained in the solution of the dual problem. This information can be retrieved by formula (18) that strongly relies on the formal conservativity.

Existence in the Orlicz setting

Theorem

Under a minor technical assumption, for any $v_0 \in L_{\mathbf{K}}(\Omega;\mathbb{R}^n)$ there exists a maximizer (E,B) for the dual problem, and

$$0 \le \frac{\mathfrak{H}(0)|\Omega| \operatorname{tr} \mathbf{F}(0)}{2} \le \tilde{\mathcal{J}}(v_0, T) < +\infty.$$

In this setting formula

$$v^{\#}(t,x) := -\frac{1}{\mathfrak{H}(t)} \int_t^T E(s,x)\, ds, \;\; v(t,x) := \nabla \mathbf{K}^*(v^{\#}(t,x))$$

provides an object that lies in the same class as the strong solutions and can be viewed as a generalized solution to the original problem (cf. Brenier's shock-free substitutes of entropy solutions to Burgers' equation).

Rmk. We also expect weak-strong uniqueness for the dual problem, but so far we have only managed to prove it for quadratic nonlinearities.