

HIDDEN CONVEXITY AND DAFERMOS' PRINCIPLE FOR SOME
DISPERSIVE EQUATIONS

DMITRY VOROTNIKOV

ABSTRACT. We discover an abstract structure behind several nonlinear dispersive equations (including the NLS, NLKG and GKdV equations with generic defocusing power-law nonlinearities) that is reminiscent of hyperbolic conservation laws. The underlying abstract problem admits an “entropy” that is formally conserved. The entropy is determined by a strictly convex function that naturally generates an anisotropic Orlicz space. For such problems, we introduce the dual matrix-valued variational formulation in the spirit of *Y. Brenier. Comm. Math. Phys. (2018) 364(2) 579-605* and prove several results related to existence, uniqueness and consistency on large time intervals. As an application, we show that no subsolution of the problems that fit into our framework is able to dissipate the total entropy earlier or faster than the strong solution on the interval of existence of the latter. This result (we call it Dafermos’ principle) is new even for “isotropic” problems such as the incompressible Euler system.

Keywords: nonlinear PDE, convex duality, entropy, anisotropic Orlicz space, generalized optimal transport, Dafermos’ criterion

MSC [2020] 35D99, 35L90, 37K58, 47A56, 49Q99

1. INTRODUCTION

Cauchy problems for nonlinear evolutionary PDEs can have infinitely many weak solutions, cf. [24, 23, 22, 14]. Many of such problems possess a physically relevant quantity (a Hamiltonian, energy or entropy depending on the context) that should be formally conserved along the flow, but can fail to do it for the weak solutions. There is some consensus that for physically relevant solutions this quantity (for definiteness, hereafter we refer to it as the “total entropy”) cannot exceed its initial value, cf. [21]. No universal selection criterion for weak solutions has been found, but in physics, information theory, chemistry and biology there exist numerous variational principles that can potentially assist. The most famous ones are Prigogine’s principle also known as the minimum entropy production principle [50] that is applicable to open systems and Ziegler’s principle also known as the maximum entropy production principle [58] that is suitable for closed systems. Other notable principles are due to Onsager, Gyarmati, Berthelot, Swenson, Lotka, Enskog, Kohler, Haken, Paltridge, Malkus, Veronis and Jaynes, see [46] for a survey. Such principles can be employed both for derivation of physically relevant systems of PDEs and for selection of physically relevant solutions [31]. In this vein, Dafermos [19, 20] suggested to assume that the physically relevant weak solutions dissipate the total entropy earlier and faster than the irrelevant ones. Typically,

The author thanks Yann Brenier and Iván Moyano for inspiring discussions. The research was partially supported by Centro de Matemática da Universidade de Coimbra (CMUC), funded by the Portuguese Government through FCT/MCTES, DOI 10.54499/UIDB/00324/2020.

this kind of entropy¹ 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.

Appropriateness of Dafermos’ principle was examined for various PDEs, see, e.g., [33, 20, 26, 16, 18]. Dafermos’ criterion has recently been applied [29, 15] at the (wider) level of *subsolutions* to some problems of fluid dynamics. Numerical applications of the criterion have recently been produced in [35, 36, 37].

For systems of hyperbolic conservation laws

$$\partial_t \mathbf{v} = -\partial_x(\mathbf{F}(\mathbf{v})), \quad (1.1)$$

where $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a given flux function, there exist classical selection criteria (Lax’ shock condition, the Kruzhkov-Lax entropy criterion, the vanishing viscosity criterion and some others) that under some assumptions are compatible with Dafermos’ principle, see [19, 41]. For equations of Lagrangian and continuum mechanics, a certain least action admissibility criterion has recently been advocated in [30]. A related least action principle was discussed in [28].

For the incompressible Euler and some related equations, Brenier [11, 12] suggested to search for the solution that minimizes the time average of the kinetic energy. This problem might not always admit a solution, but it generates a dual variational problem that has better convexity features. He found an explicit formula that relates the smooth solutions of the incompressible Euler existing on a small time interval to the solutions of the corresponding dual problem. In [13], this technique was applied to the multi-stream Euler-Poisson system. In [57], we developed Brenier’s approach by finding structures in nonlinear quadratic PDEs that permit to define similar dual problems with decent properties. It was also discussed in [11, 12, 57] that dual problems of this kind have strong analogy with the optimal transportation problems.

A similar duality scheme has been (rather independently) proposed by Acharya and several collaborators, see [1, 2, 52] and the references therein. The recent preprint [3] builds upon both approaches. Among other results, the authors of [3] prove the Γ -convergence of the dual problem for the incompressible Navier-Stokes system to the corresponding dual problem for the incompressible Euler. From the perspective of the analogy with the optimal transport, this result is reminiscent of the Γ -convergence results from [42, 6, 48]. A notable feature of [3] is that their construction of the dual problem involves a “base state” that can be regarded as an “initial guess” for the solution of the primal problem. (From this point of view, the “base states” in [11, 12, 13, 57] are identically zero).

It was clear from the very beginning [11, 1] that the nonlinearity in a PDE does not need to be quadratic in order to implement the construction. However, the majority of rigorous results for this kind of problems has until now been obtained for quadratic nonlinearities. Of course, for many relevant systems the nonlinearity fails to be quadratic, for instance, for multidimensional systems of conservation laws

$$\partial_t \mathbf{v} = -\operatorname{div}(\mathbf{F}(\mathbf{v})), \quad (1.2)$$

¹The wording comes from the theory of conservation laws and might be confusing for some readers. For example, smooth solutions of the heat equation dissipate the Boltzmann entropy, so this entropy is never conserved. In this paper, as we have already agreed, the “total entropy” is a formally conserved quantity.

where \mathbf{F} is a prescribed matrix function, and for various nonlinear dispersive equations. Remarkably, rigorous study of Brenier's duality scheme for multidimensional systems of conservation laws (1.2) that admit a convex entropy, initiated in [11], is still pending.

In this paper, we extend Brenier's approach to systems of PDEs that can be rewritten in the abstract form

$$\partial_t \mathbf{v} = L(\mathbf{F}(\mathbf{v})), \quad (1.3)$$

where the matrix function \mathbf{F} has some convexity and positivity properties to be specified below in Assumptions 3.2 and 3.3, and L is a vector-valued differential operator of any order. To fix the ideas, we will restrict ourselves to unknown vector functions $\mathbf{v}(\mathbf{t}, \mathbf{x})$ with \mathbf{x} varying in the periodic box $\Omega = \mathbb{T}^d$. Note that (1.3) is strongly reminiscent of (1.1) and (1.2). We will also need to assume that (1.3) admits a strictly convex anisotropic *entropy* function $\mathbf{K}(\mathbf{v})$ that is formally conserved along the flow as described in Assumptions 3.1 and 3.7. As we will see, the NLS, NLKG and GKdV equations with generic defocusing power-law nonlinearities can be recast in our abstract form.

As an application of our duality scheme, we rigorously prove the following variant of Dafermos' principle: no subsolution of (1.3) can dissipate the total entropy earlier or faster than the strong solution on the interval of existence of the latter (Theorem 4.3). Our Dafermos' principle is new even for quadratic problems that fit into our framework, including the incompressible Euler equation [11], the ideal incompressible MHD, the multidimensional Camassa-Holm, the Zakharov-Kuznetsov system and many others (various examples can be found in [57]).

Our entropy \mathbf{K} generates an anisotropic Orlicz space, hence our procedure will heavily rely on the theory of anisotropic Orlicz spaces, see, e.g., [17]. In particular, we will prove existence of solutions to the dual problem that belong to an anisotropic Orlicz space. Notably, some authors employed isotropic Orlicz spaces for the study of nonlinear Schrödinger and generalized KdV equations, cf. [32], but we have never heard of any application of anisotropic Orlicz spaces — that appear so naturally in our generic approach — to this sort of PDEs.

Until recently, an unsettling detail of the duality scheme was that the majority of consistency results — that advocate the possibility to view solutions to the dual problems as certain “dual variational solutions” to PDEs that admit the considered duality approach — were only obtained on small time intervals, cf. [11, Theorems 2.3 and 3.1], [57, Theorem 3.8], [13, Theorem 5.1]. In this paper, we are able to guarantee the consistency on large time intervals by introducing appropriate time-dependent weights to the problem. Global in time consistency theorems for the dual formulations of incompressible Euler and Navier-Stokes equations have recently been obtained in [3]. The latter results heavily rely on a suitable choice of the “base state”. Nevertheless, our consistency results seem to be of a different nature, and, furthermore, the two attitudes complement each other to a certain degree, see Remark 4.4 for a thorough comparison.

We believe that it is potentially possible to prove the weak-strong uniqueness property in the sense that existence of a (suitably defined) strong solution to (1.3) implies that the solution to the dual problem is unique in a certain anisotropic Orlicz space. The corresponding result in the quadratic case $\mathbf{F}(\mathbf{v}) = \mathbf{v} \otimes \mathbf{v}$ was established in [57, Section 5]. As the first step towards this conjecture, we will show that the strong solution is always unique. The proof is unusual because it relies on the analysis of a specific “Jeffreys divergence”.

To the best of our knowledge, global in time solvability for the considered dispersive PDEs is merely known under significant restrictions on the exponents in the power laws or on the size of the initial data, cf. [55, 44, 27]. Our Theorem 5.4 and Remark 5.5 provide existence of certain “dual variational solutions” for these problems without such restrictions.

The paper consists of seven sections (including the Introduction) and two appendices. In Section 2, we list some necessary facts about N -functions and anisotropic Orlicz spaces. In Section 3, we describe our abstract formalism. In Section 4, we establish the consistency of the dual problem, and explain how it can be secured on large time intervals. As an application, we prove the already mentioned version of Dafermos' principle. In Section 5, we show that the dual problem always has a solution that belongs to a specific anisotropic Orlicz space. In Section 6, we obtain uniqueness of strong solutions. In Section 7, we demonstrate how the NLS, NLKG and GKdV equations with generic defocusing power-law nonlinearities (as well as scalar conservation laws) fit into our framework. In Appendix A, we prove a technical anisotropic variant of Hardy's inequality (we do not claim much novelty here, but we failed to find it in the existing literature). In Appendix B, we employ an unsophisticated geometric point of view in order to illustrate the analogy of the dual problems in question with the optimal transport theory.

2. ANISOTROPIC ORLICZ SPACES

In this section we collect some basic necessary facts about N -functions and anisotropic Orlicz spaces.

Definition 2.1 (Young function). A convex function $\Phi : \mathbb{R} \rightarrow \mathbb{R}_+$ is called a *Young function* provided Φ is even, $\Phi(s) = 0$ if and only if $s = 0$, and

$$\lim_{s \rightarrow 0} \frac{\Phi(s)}{s} = 0, \quad \lim_{s \rightarrow +\infty} \frac{\Phi(s)}{s} = +\infty.$$

Note that the Legendre transform Φ^* of a Young function is well defined and is also a Young function.

Definition 2.2 (N -function). A convex function $\mathbf{K} : \mathbb{R}^n \rightarrow \mathbb{R}_+$ is called an *N -function* provided \mathbf{K} is even, and there exist two Young functions Φ_- and Φ_+ such that

$$\Phi_-(|y|) \leq \mathbf{K}(y) \leq \Phi_+(|y|), \quad y \in \mathbb{R}^n.$$

It follows from the definition that if \mathbf{K} is an N -function, then

- (i) $\mathbf{K}(y) = 0$ if and only if $y = 0$;
- (ii) The Legendre transform \mathbf{K}^* is well-defined and is also an N -function.

Definition 2.3 (Δ_2 -condition). An N -function $\mathbf{K} : \mathbb{R}^n \rightarrow \mathbb{R}_+$ satisfies the Δ_2 -condition provided²

$$\mathbf{K}(2y) \lesssim \mathbf{K}(y) + 1, \quad y \in \mathbb{R}^n.$$

Assumption 2.4. *Henceforth in this section, we fix an N -function \mathbf{K} . We always assume that both \mathbf{K} and \mathbf{K}^* satisfy the Δ_2 -condition, and all the facts are claimed under this assumption.*

Let Q be either an interval, or a flat torus (periodic box), or the Cartesian product of an interval and a torus, equipped with the Lebesgue measure $d\mathbf{x}$.

Definition 2.5 (Modular and anisotropic Orlicz class). The integral

$$\int_Q \mathbf{K}(f(\mathbf{x})) d\mathbf{x}$$

is called the *modular* of a measurable function $f : Q \rightarrow \mathbb{R}^n$. Denote by $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ the set of functions f whose modular is finite.

²Hereafter, we write $b \lesssim a$ for two scalar expressions a and b to indicate that $b \leq Ca$ with a uniform constant C .

Note that $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ is a linear space, cf. [17, Lemma 3.3.1], and it is called an *anisotropic Orlicz space*. For $\mathbf{K}(\mathbf{v}) = \frac{|\mathbf{v}|^p}{p}$, $p > 1$, the spaces $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ are classical Lebesgue spaces $L^p(Q; \mathbb{R}^n)$.

There are two different ways to define a relevant norm on $L_{\mathbf{K}}(Q; \mathbb{R}^n)$.

Definition 2.6 (Luxemburg and Orlicz norms). The *Luxemburg norm* of $f \in L_{\mathbf{K}}(Q; \mathbb{R}^n)$ is defined as

$$\|f\|_{L_{\mathbf{K}}(Q)} := \inf \left\{ \lambda > 0 : \int_Q \mathbf{K} \left(\frac{f(x)}{\lambda} \right) dx \leq 1 \right\}.$$

The *Orlicz norm* of $f \in L_{\mathbf{K}}(Q; \mathbb{R}^n)$ is defined as

$$\|f\|_{\|L_{\mathbf{K}}(Q)} := \sup \left\{ \int_Q f(x) \cdot g(x) dx : \int_Q \mathbf{K}^*(g(x)) dx \leq 1 \right\}.$$

Both norms are indeed norms. Moreover, $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ equipped with any of these norms is a separable reflexive Banach space, see [17]. Both norms are equivalent with uniform constants:

$$\|f\|_{L_{\mathbf{K}}(Q)} \leq \|f\|_{\|L_{\mathbf{K}}(Q)} \leq 2\|f\|_{L_{\mathbf{K}}(Q)}. \quad (2.1)$$

Furthermore, the strong topology associated with both norms coincides with the *modular topology* [17], i.e.,

$$\|f_k\|_{L_{\mathbf{K}}(Q)} \rightarrow 0 \Leftrightarrow \int_Q \mathbf{K}(f_k(x)) dx \rightarrow 0 \quad (2.2)$$

as $k \rightarrow +\infty$.

Proposition 2.7 (Hölder inequality [17]). For every $f \in L_{\mathbf{K}}(Q; \mathbb{R}^n)$ and $g \in L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$ one has

$$\left| \int_Q f(x) \cdot g(x) dx \right| \leq 2\|f\|_{L_{\mathbf{K}}(Q)} \|g\|_{L_{\mathbf{K}^*}(Q)}.$$

Another useful estimate, [17, Lemma 3.1.14], is

$$\|f\|_{L_{\mathbf{K}}(Q)} \leq \max \left\{ 1, \int_Q \mathbf{K}(f(x)) dx \right\}. \quad (2.3)$$

Proposition 2.8 (Duality [17]). Bounded linear functionals on $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ can be represented in the form

$$f \mapsto \int_Q f(x) \cdot g(x) dx$$

for some $g \in L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$. Consequently, $L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$ is the dual space of $L_{\mathbf{K}}(Q; \mathbb{R}^n)$.

We will need the following (refined) embedding of anisotropic Orlicz spaces.

Proposition 2.9. Let $\mathbf{K}, \check{\mathbf{K}}$ be two N -functions. Assume that there exists $r > 0$ such that $\mathbf{K}(\mathbf{v}) \leq \check{\mathbf{K}}(\mathbf{v})$ for $|\mathbf{v}| \geq r$. Then for any $f \in L_{\check{\mathbf{K}}}(Q; \mathbb{R}^n)$ we have $f \in L_{\mathbf{K}}(Q; \mathbb{R}^n)$ and

$$\|f\|_{L_{\mathbf{K}}(Q)} \leq C_r \|f\|_{L_{\check{\mathbf{K}}}(Q)}$$

where C_r merely depends on r and Q .

The proof of Proposition 2.9 is omitted because the classical proof for the isotropic case, cf. [39, Theorem 13.3], is applicable here as well, mutatis mutandis.

Proposition 2.10 (Density of smooth functions). Smooth functions are dense in $L_{\mathbf{K}}(Q; \mathbb{R}^n)$.

Proof. This fact is well known, so we just give a sketch of the proof. It follows from [17, Lemma 2.3.16] that there exist $r > 0$ and $\rho > 1$ such that $\mathbf{K}(\mathbf{v}) \leq \frac{|\mathbf{v}|^\rho}{\rho}$ for $|\mathbf{v}| \geq r$. By Proposition 2.9, $L^\rho(Q) \subset L_{\mathbf{K}}(Q)$ and $\|\cdot\|_{L_{\mathbf{K}}(Q)} \leq C_r \|\cdot\|_{L^\rho(Q)}$. Simple functions are dense in $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ by [17, Corollary 3.4.12]. On the other hand, in any L^ρ -neighborhood (and hence in any $L_{\mathbf{K}}$ -neighborhood) of a simple function there is a smooth function. \square

We will also need the following variant of Krasnoselskii's theorem about continuity of Nemytskii operators.

Proposition 2.11. *Let $\eta : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a continuous function such that the Nemytskii operator*

$$\mathcal{N} : L_{\mathbf{K}}(Q; \mathbb{R}^n) \rightarrow L^1(Q; \mathbb{R}^m), \quad \mathcal{N}(\mathbf{w})(\mathbf{x}) = \eta(\mathbf{w}(\mathbf{x}))$$

is well-defined. Then \mathcal{N} is continuous.

We omit the proof because in the original proof of Krasnoselskii, see [38, Chapter 1, Theorem 2.1], it is possible to substitute the anisotropic Orlicz source $L_{\mathbf{K}}$ for the Lebesgue source L^ρ without destroying the validity of the argument³.

We finish this section with a statement about mollification. We restrict ourselves to the case when Q is a d -dimensional periodic box. Let ρ be a smooth compactly supported probability density on \mathbb{R}^d . Set $\rho_k(\mathbf{x}) := k\rho(k\mathbf{x})$, $k \in \mathbb{N}$. For any fixed $f \in L_{\mathbf{K}}(Q; \mathbb{R}^n)$, the convolution $\rho_k * f$ is well-defined, smooth, and belongs to $L_{\mathbf{K}}(Q; \mathbb{R}^n)$.

Proposition 2.12. *We have $\rho_k * f \rightarrow f$ strongly in $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ as $k \rightarrow +\infty$.*

Proof. We give just a brief sketch of the proof. It follows from the density of smooth functions that the “translation” map

$$\mathbf{x} \mapsto f(\cdot - \mathbf{x})$$

is well-defined and continuous from \mathbb{R}^d to $L_{\mathbf{K}}(Q; \mathbb{R}^n)$. We conclude by mimicking the “isotropic” Donaldson-Trudinger argument from [25, Lemma 2.1]. \square

3. THE ABSTRACT SETTING

Let us start with fixing some basic notation. In the what follows, Ω is the periodic box \mathbb{T}^d , $d \in \mathbb{N}$, equipped with the Lebesgue measure $d\mathbf{x}$. For two measurable functions $u, v : \Omega \rightarrow \mathbb{R}^m$, $m \in \mathbb{N}$, we will use the shortcut

$$(u, v) := \int_{\Omega} u(\mathbf{x}) \cdot v(\mathbf{x}) d\mathbf{x}.$$

We will also use the notations $\mathbb{R}^{N \times N}$, $\mathbb{R}_s^{N \times N}$ and $\mathbb{R}_+^{N \times N}$ for the spaces of $N \times N$ matrices, symmetric matrices and nonnegative-definite matrices, resp. For $A, B \in \mathbb{R}_s^{N \times N}$, we write $A \succeq B$ when $A - B \in \mathbb{R}_+^{N \times N}$ (the Loewner order [45]; we stress that the notation $A \geq B$ is reserved for a less restrictive order that will introduced in Assumption 3.3 below), and $A : B$ for the scalar product of A and B generated by the Frobenius norm $|\cdot|$. The symbol I will stand for the identity matrix of a relevant size.

As anticipated in the Introduction, we are interested in the problem

$$\partial_t \mathbf{v} = L(\mathbf{F}(\mathbf{v})), \quad \mathbf{v}(0, \cdot) = \mathbf{v}_0. \tag{3.1}$$

³The case of anisotropic targets is more tricky and lies beyond this discussion; for instance, the proof for isotropic Orlicz sources and targets that can be found in [39] does not work in the anisotropic case.

Here $\mathbf{v}_0 : \Omega \rightarrow \mathbb{R}^n$, $n \in \mathbb{N}$, is the initial datum, $\mathbf{v} : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is an unknown vector function,

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}_S^{N \times N}$$

is a prescribed C^2 -smooth matrix function with some convexity and positivity properties to be specified below, and L is a vector-valued differential operator⁴ with constant coefficients

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

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

Assumption 3.1 (Entropy). *Define the “entropy” function for our problem by*

$$\mathbf{K} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{K}(\mathbf{v}) := \frac{1}{2} \text{Tr}(\mathbf{F}(\mathbf{v}) - \mathbf{F}(\mathbf{0})).$$

Assume that \mathbf{K} is strictly convex (but not necessarily uniformly convex). Assume also that \mathbf{K} is an N -function, and both \mathbf{K} and its Legendre transform \mathbf{K}^ satisfy the Δ_2 -condition.*

To any given $\mathbf{v} \in \mathbb{R}^n$ we associate the “sharp” vector

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

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

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

is C^1 -smooth and injective. Moreover,

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

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

It follows from [17, Lemma 2.3.16] that there exist $p, q > 1$ and $r > 0$ such that

$$\frac{\mathbf{v} \cdot \mathbf{v}^\#}{\mathbf{K}(\mathbf{v})} \leq p, \quad \frac{\mathbf{v}^\# \cdot \mathbf{v}}{\mathbf{K}^*(\mathbf{v}^\#)} \geq q$$

for $|\mathbf{v}| \geq r$. Consequently,

$$\mathbf{K}^*(\mathbf{v}^\#) \lesssim \mathbf{K}(\mathbf{v}) + 1, \quad \mathbf{v} \in \mathbb{R}^n. \quad (3.2)$$

Similarly,

$$\mathbf{K}(\mathbf{v}) \lesssim \mathbf{K}^*(\mathbf{v}^\#) + 1, \quad \mathbf{v} \in \mathbb{R}^n. \quad (3.3)$$

Assumption 3.2 (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 [40, 9], namely, $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}_S^{N \times N}$ 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, cf. Remark 7.1 and Appendix B, so we will merely assume that \mathbf{F} is Λ -convex in the following more relaxed sense. Let $\mathbf{u} : \Omega \rightarrow \mathbb{R}^n$ 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 \mathbf{u} and containing I such that $L^*(\mathbf{u}(\mathbf{x})) \in \Lambda$, $\mathbf{x} \in \Omega$. We say that \mathbf{F} is*

⁴As a matter of fact, L can be a more general (and possibly nonlocal) linear operator; it is easy to see from the considerations below that it is enough to assume that L commutes with the mollification $L(\rho_k * M) = \rho_k * L(M)$ and is a closed linear operator with a dense domain $D(L)$ in a certain space.

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.

Assumption 3.3 (Λ -order). We will also require some positivity of the matrix function \mathbf{F} . For this purpose, we will use the less restrictive Λ -order instead of the Loewner order. Namely, denote

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

We will assume that

$$\mathbf{F}(\mathbb{R}^n) \subset \mathbb{R}_\Lambda^{N \times N}. \quad (3.4)$$

For $A, B \in \mathbb{R}_S^{N \times N}$, we write $B \leq A$ when $A - B \in \mathbb{R}_\Lambda^{N \times N}$. Obviously, $B \leq A$ implies $B \leq A$, and the converse is true only if $\Lambda = \mathbb{R}_S^{N \times N}$. Employing some simple algebra, it is easy to derive from $\mathbf{F} \geq 0$ that

$$|\pi_\Lambda \mathbf{F}(\mathbf{v})| \lesssim \mathbf{K}(\mathbf{v}) + 1, \quad (3.5)$$

where $\pi_\Lambda : \mathbb{R}_S^{N \times N} \rightarrow \Lambda$ denotes the orthogonal projection w.r.t. the Frobenius product. We will also need to assume that pointwise control similar to (3.5) holds at the level of partial derivatives

$$\mathbf{F}_l(\mathbf{v}) := \mathbf{F}_{v_l}(\mathbf{v}), \quad l = 1, \dots, n.$$

More exactly, we assume that

$$|\pi_\Lambda \partial_l \mathbf{F}(\mathbf{v})| \lesssim |\nabla \mathbf{K}(\mathbf{v})| + 1. \quad (3.6)$$

Remark 3.4. In some examples, including the incompressible Euler equation, cf. [11, 57], Λ coincides with $\mathbb{R}_S^{N \times N}$, hence in this case the Λ -order and the Λ -convexity coincide with their Loewner counterparts. For the dispersive equations considered in Section 7, the subspace Λ is significantly smaller than $\mathbb{R}_S^{N \times N}$.

Remark 3.5. The “flux” $\mathbf{F}(\mathbf{v})$ can be viewed as an “entropy matrix” that very vaguely resembles the entropy matrix from [51]. Notably, the definition of convexity of matrix-valued functions employed in [51] is very restrictive, it is much stronger than the Loewner convexity.

Remark 3.6. The quadratic flux $\mathbf{F}(\mathbf{v}) = \mathbf{v} \otimes \mathbf{v}$ and the corresponding entropy $\mathbf{K}(\mathbf{v}) = \frac{1}{2}|\mathbf{v}|^2$ obviously satisfy Assumptions 3.1, 3.2 and 3.3 for any subspace $\Lambda \subset \mathbb{R}_S^{n \times n}$. Note that in this case $N = n$ and $\mathbf{v}^\# = \mathbf{v}$. This applies to the incompressible Euler and other quadratic examples mentioned in the Introduction.

Assumption 3.7 (Conservativity). We will focus on the situation when L satisfies the “formal conservativity condition”

$$(\mathbf{F}(\mathbf{v}), L^*(\mathbf{v}^\#)) = 0 \quad (3.7)$$

provided $\mathbf{v}^\# : \Omega \rightarrow \mathbb{R}^n$ is a smooth function (remember that $\mathbf{v} = \nabla \mathbf{K}^*(\mathbf{v}^\#)$).

Problem (3.1) admits the following natural weak formulation:

$$\int_0^T [(\mathbf{v}, \partial_t \alpha) + (\mathbf{F}(\mathbf{v}), L^* \alpha)] dt + (\mathbf{v}_0, \alpha(0, \cdot)) = 0 \quad (3.8)$$

for all smooth vector fields $\alpha : [0, T] \times \Omega \rightarrow \mathbb{R}^n$, $\alpha(T, \cdot) \equiv 0$.

Remark 3.8 (Formulation in terms of the sharp variable). In the conservative case (3.1) admits a formally equivalent formulation in terms of the “sharp” unknown variable $\mathbf{v}^\#$:

$$\partial_t(\mathbf{v}^\#)_l + L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\nabla \mathbf{K}^*(\mathbf{v}^\#)) = 0, \quad \mathbf{v}^\#(0, \cdot) = \mathbf{v}_0^\# := (\mathbf{v}_0)^\#, \quad l = 1, \dots, n. \quad (3.9)$$

Indeed, let $\mathbf{a} : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ be a smooth test function, $s \in \mathbb{R}$ be a real number, and \mathbf{v} be a solution to (3.1). Then due to the conservativity

$$(\mathbf{F}(\mathbf{v} + s\mathbf{a}), L^*(\nabla \mathbf{K}(\mathbf{v} + s\mathbf{a}))) = 0. \quad (3.10)$$

Taking the derivative w.r.t. s at $s = 0$ yields

$$\sum_{l=1}^n [(\partial_l \mathbf{F}(\mathbf{v}) \mathbf{a}_l, L^*(\mathbf{v}^\#)) + (L(\mathbf{F}(\mathbf{v})), \partial_{v_l}(\nabla \mathbf{K}(\mathbf{v})) \mathbf{a}_l)] = 0. \quad (3.11)$$

Using (3.1), we recast (3.11) in the form

$$\sum_{l=1}^n (\partial_l \mathbf{F}(\mathbf{v}) : L^*(\mathbf{v}^\#), \mathbf{a}_l) + \sum_{l,m=1}^n (\partial_t v_m, \partial_{lm} \mathbf{K}(\mathbf{v}) \mathbf{a}_l) = 0. \quad (3.12)$$

By the chain rule, $\partial_t(\mathbf{v}^\#)_l = \partial_t(\partial_l \mathbf{K}(\mathbf{v})) = \sum_{m=1}^n \partial_{lm} \mathbf{K}(\mathbf{v}) \partial_t v_m$. Thus (3.12) becomes

$$\sum_{l=1}^n (\partial_l \mathbf{F}(\mathbf{v}) : L^*(\mathbf{v}^\#) + \partial_t(\mathbf{v}^\#)_l, \mathbf{a}_l) = 0, \quad (3.13)$$

which immediately implies (3.9) due to arbitrariness of \mathbf{a} and the relation $\mathbf{v} = \nabla \mathbf{K}^*(\mathbf{v}^\#)$.

Let us now rewrite problem (3.8) in terms of the test functions $\mathbf{B} := L^* \mathbf{a}$ and $\mathbf{E} := \partial_t \mathbf{a}$ (note that the conservativity (3.7) is not needed at this stage). We first observe that

$$(\mathbf{v}_0, \mathbf{a}(0)) = - \int_0^T (\mathbf{v}_0, \partial_t \mathbf{a}) dt = - \int_0^T (\mathbf{v}_0, \mathbf{E}) dt. \quad (3.14)$$

The link between \mathbf{B} and \mathbf{E} can alternatively be described by the conditions

$$\partial_t \mathbf{B} = L^* \mathbf{E}, \quad \mathbf{B}(T, \cdot) \equiv 0. \quad (3.15)$$

Hence, (3.8) becomes

$$\int_0^T [(\mathbf{v} - \mathbf{v}_0, \mathbf{E}) + (\mathbf{F}(\mathbf{v}), \mathbf{B})] dt = 0 \quad (3.16)$$

for all smooth vector fields $\mathbf{B} : [0, T] \times \Omega \rightarrow \mathbb{R}_s^{N \times N}$, $\mathbf{E} : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ satisfying the constraints (3.15).

Observe that (3.15) can be rewritten in the following weak form

$$\int_0^T [(\mathbf{B}, \partial_t \Psi) + (\mathbf{E}, L\Psi)] dt = 0 \quad (3.17)$$

for all smooth matrix fields $\Psi : [0, T] \times \Omega \rightarrow \mathbb{R}_s^{N \times N}$, $\Psi(0) = 0$.

Motivated by the discussion above, we adopt the following definitions, where we tacitly assume $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$.

Definition 3.9 (Weak solutions). A function

$$\mathbf{v} \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n) \quad (3.18)$$

is a *weak solution* to (3.1) if it satisfies (3.16) for all pairs

$$(E, B) \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}_s^{N \times N}) \quad (3.19)$$

meeting the constraint (3.17).

Note that if \mathbf{v} is a weak solution, then $\mathbf{F}(\mathbf{v}) \in L^1((0, T) \times \Omega; \mathbb{R}_s^{N \times N})$ due to (3.5), so (3.16) makes sense.

Definition 3.10 (Subsolutions). A pair of functions

$$(\mathbf{v}, M) \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n) \times L^1((0, T) \times \Omega; \mathbb{R}_s^{N \times N}), \quad \mathbf{F}(\mathbf{v}) \leq M, \quad (3.20)$$

is a *subsolution* to (3.1) if it satisfies

$$\int_0^T [(\mathbf{v} - \mathbf{v}_0, E) + (M, B)] dt = 0 \quad (3.21)$$

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 (3.17).

The second inequality in (3.20) is understood in the sense of the Λ -order a.e. in $(0, T) \times \Omega$. Obviously, if \mathbf{v} is a weak solution, then $(\mathbf{v}, \mathbf{F}(\mathbf{v}))$ is a subsolution. Accordingly, the subsolution *entropy* function that complies with Assumption 3.1 is

$$M \mapsto \frac{1}{2} \text{Tr}(M - \mathbf{F}(\mathbf{0})), \quad M \in \mathbb{R}_\Lambda^{N \times N}.$$

Observe that for a weak solution \mathbf{v} the modular (that we will call the *total entropy*)

$$K(t) := \int_\Omega \mathbf{K}(\mathbf{v}(t, \mathbf{x})) d\mathbf{x}. \quad (3.22)$$

belongs to $L^1(0, T)$ and thus is finite for almost all times. Similarly, the total entropy

$$\tilde{K}(t) := \frac{1}{2} \text{Tr} \int_\Omega [M(t, \mathbf{x}) - \mathbf{F}(\mathbf{0})] d\mathbf{x} = \frac{1}{2} (M(t, \cdot) - \mathbf{F}(\mathbf{0}), I). \quad (3.23)$$

of a subsolution also belongs to $L^1(0, T)$.

Remark 3.11. For the incompressible Euler equation, in view of Remark 3.4 and some observations made in [21, 57], Definition 3.10 is equivalent to the conventional definition of a subsolution that, among other applications, is used in the theory of convex integration [23, 22].

Our primal problem is to search for the weak solutions to (3.1) that minimize a suitable weighted time integral of the total entropy⁵ (3.22).

More exactly, fix a positive scalar function $\zeta(\mathbf{t})$ that is bounded away from 0 and ∞ on $[0, T]$. Let $\mathcal{H}(t) := \int_t^T \zeta(s) ds$. The idea is to search for the weak solutions that minimize $\int_0^T \zeta(t) K(t) dt$.

⁵See also Remark 3.18.

Remark 3.12 (“Rough” Dafermos’ criterion). A typical weight is $\zeta(t) := \exp(-\gamma t)$ with large $\gamma > 0$, cf. Remark 4.2. In this case our selection criterion can be viewed as a “rough” Dafermos’ principle, because we prioritize the solutions that dissipate the total entropy as fast and as early as possible and that at later times do not admit dramatic exponential growth (of order $\exp(\gamma t)$) of the total entropy.

We will now employ the “sharp” formulation (3.9) in order to define *strong solutions* to (3.1).

Definition 3.13 (Strong solutions). Assume that $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. A function \mathbf{v} satisfying (3.18) and

$$\partial_t(\zeta \mathbf{v}^\#) \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n), \quad \zeta L^*(\mathbf{v}^\#) \in L^\infty((0, T) \times \Omega; \mathbb{R}^{N \times N}) \quad (3.24)$$

is a *strong solution* to (3.1) if it is a weak solution and

$$\partial_t(\mathbf{v}^\#)_l + L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\nabla \mathbf{K}^*(\mathbf{v}^\#)) = 0, \quad \mathbf{v}^\#(0, \cdot) = \mathbf{v}_0^\# \quad (3.25)$$

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

Remark 3.14. Let us explain why (3.25) makes sense for \mathbf{v} from the regularity class (3.18), (3.24). Firstly, (3.2) and (3.18) imply $\mathbf{v}^\# \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$. The first equality in (3.25) is equivalent to

$$\zeta(\mathbf{v}^\#)_l + \partial_t(\zeta \mathbf{v}^\#)_l + \zeta L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\nabla \mathbf{K}^*(\mathbf{v}^\#)) = 0 \quad (3.26)$$

a.e. in $(0, T) \times \Omega$. By (3.6), $\pi_\wedge \partial_l \mathbf{F}(\mathbf{v})$ is pointwise controlled by $\nabla \mathbf{K}(\mathbf{v}) = \mathbf{v}^\#$, therefore all members of (3.26) belong to $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$. We moreover claim that

$$\mathbf{v}^\# \in C([0, T]; L_{\mathbf{K}^*}(\Omega; \mathbb{R}^n)) \quad (3.27)$$

and that all members in the second equality in (3.25) belong to $L_{\mathbf{K}^*}(\Omega; \mathbb{R}^n)$. Indeed, since $\mathbf{K}^*(\mathbf{v}_0^\#) \lesssim \mathbf{K}(\mathbf{v}_0) + 1$, we have $\mathbf{v}_0^\# \in L_{\mathbf{K}^*}(\Omega; \mathbb{R}^n)$. But

$$\partial_t \mathbf{v}^\# = \frac{1}{\zeta} \partial_t(\zeta \mathbf{v}^\#) + \frac{\zeta'}{\zeta} \mathbf{v}^\# \in L_{\mathbf{K}^*, \text{loc}}([0, T) \times \Omega; \mathbb{R}^n) \subset L^1_{\text{loc}}([0, T); L^1(\Omega)),$$

whence

$$\mathbf{v}^\# \in AC_{\text{loc}}([0, T); L^1(\Omega)).$$

Leveraging convexity of \mathbf{K}^* , for every $t \in [0, T)$ and small $h > 0$ we obtain

$$\begin{aligned} \int_{\Omega} \mathbf{K}^*(\mathbf{v}^\#(t+h, \mathbf{x}) - \mathbf{v}^\#(t, \mathbf{x})) d\mathbf{x} &= \int_{\Omega} \mathbf{K}^* \left(\frac{1}{h} \int_t^{t+h} h \partial_t \mathbf{v}^\#(s, \mathbf{x}) ds \right) d\mathbf{x} \\ &\leq \int_{\Omega} \frac{1}{h} \int_t^{t+h} \mathbf{K}^*(h \partial_t \mathbf{v}^\#(s, \mathbf{x})) ds d\mathbf{x} \\ &= \int_{\Omega} \frac{1}{h} \int_t^{t+h} \mathbf{K}^*(h \partial_t \mathbf{v}^\#(s, \mathbf{x}) + (1-h)0) ds d\mathbf{x} \\ &\leq \int_{\Omega} \frac{1}{h} \int_t^{t+h} (h \mathbf{K}^*(\partial_t \mathbf{v}^\#(s, \mathbf{x})) + (1-h)\mathbf{K}^*(0)) ds d\mathbf{x} \\ &= \int_{\Omega} \int_t^{t+h} \mathbf{K}^*(\partial_t \mathbf{v}^\#(s, \mathbf{x})) ds d\mathbf{x} \rightarrow 0 \quad (3.28) \end{aligned}$$

as $h \rightarrow 0$ due to absolute continuity of the Lebesgue integral. This secures the right-continuity of the map $\mathbf{v}^\# : [0, T] \rightarrow L_{\mathbf{K}^*}(\Omega)$; the left-continuity is proved in a similar fashion.

Remark 3.15. Definition 3.13 is actually independent of the weight function χ . Indeed, let \mathbf{v} be a strong solution, χ_1 be another positive scalar function bounded away from 0 and ∞ , and $\chi_1 = \int_t^T \chi_1(s) ds$. Obviously, $\chi(t) \lesssim \chi_1(t) \lesssim \chi(t)$. Hence,

$$\chi_1 L^*(\mathbf{v}^\#) \in L^\infty((0, T) \times \Omega; \mathbb{R}_s^{N \times N}),$$

and

$$\partial_t(\chi_1 \mathbf{v}^\#) = -\chi_1 \mathbf{v}^\# + \chi_1 \partial_t(\mathbf{v}^\#) = -\chi_1 \mathbf{v}^\# + \frac{\chi_1}{\chi} \chi \mathbf{v}^\# + \frac{\chi_1}{\chi} \partial_t(\chi \mathbf{v}^\#).$$

Since $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$ is a linear space invariant w.r.t. multiplication by bounded functions,

$$\partial_t(\chi_1 \mathbf{v}^\#) \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n).$$

The following lemma shows that the strong solutions soar above the rest of the weak solutions.

Lemma 3.16. *If (3.7) is assumed, then for any strong solution \mathbf{v} the total entropy is continuous in time for $t \in [0, T]$ and is conserved, i.e., $K(t) = K(0) = \int_\Omega \mathbf{K}(\mathbf{v}_0) dx$.*

Proof. Let $\varphi(t)$ be an arbitrary smooth compactly supported function on $(0, T)$. Observe that $\frac{\varphi}{\chi}$ is bounded, Lipschitz and compactly supported on $(0, T)$. Let us prove that

$$\int_0^T (\mathbf{K}(\mathbf{v}), \partial_t \varphi) dt = 0. \quad (3.29)$$

We first claim that

$$\int_0^T (\mathbf{F}(\mathbf{v}), L^*(\varphi \mathbf{v}^\#)) dt = 0 \quad (3.30)$$

and

$$\int_0^T [(\mathbf{v} - \mathbf{v}_0, \partial_t(\varphi \mathbf{v}^\#)) - (\mathbf{K}(\nabla \mathbf{K}^*(\mathbf{v}^\#)), \partial_t \varphi)] dt = 0. \quad (3.31)$$

Taking (3.30) and (3.31) for granted, and employing (3.16) with

$$B = L^*(\varphi \mathbf{v}^\#) = \frac{\varphi}{\chi} \chi L^*(\mathbf{v}^\#) \in L^\infty((0, T) \times \Omega; \mathbb{R}^n),$$

$$E = \partial_t(\varphi \mathbf{v}^\#) = \partial_t\left(\frac{\varphi}{\chi}\right) \chi \mathbf{v}^\# + \partial_t(\chi \mathbf{v}^\#) \frac{\varphi}{\chi} \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n),$$

we infer

$$\int_0^T (\mathbf{v} - \mathbf{v}_0, \partial_t(\varphi \mathbf{v}^\#)) dt = 0$$

and, consequently,

$$\int_0^T (\mathbf{K}(\nabla \mathbf{K}^*(\mathbf{v}^\#)), \partial_t \varphi) dt = 0.$$

We have proved (3.29), i.e., that the total entropy is conserved in the sense of distributions. In order to prove that it is conserved in the classical sense, we just need to establish the continuity

of the total entropy w.r.t. time. By Remark 3.14, $\mathbf{v}^\# \in C([0, T]; L_{\mathbf{K}^*}(\Omega; \mathbb{R}^n))$. Proposition 2.11 and (3.3) imply that the Nemytskii operator

$$\mathcal{N}_1 : w \mapsto \mathbf{K}(\nabla \mathbf{K}^*(w))$$

is continuous from $L_{\mathbf{K}^*}(\Omega; \mathbb{R}^n)$ to $L^1(\Omega)$. Consequently, $\mathbf{K}(\mathbf{v}) \in C([0, T]; L^1(\Omega))$, whence $K(t) \in C([0, T])$.

It remains to prove (3.30) and (3.31). Without loss of generality (continuing by zero outside $[0, T]$) we may assume that $\mathbf{v}^\#$ and φ are defined on the torus $Q := 2T\mathbb{T}^1 \times \Omega$. Consider the mollifications $\mathbf{v}_k^\# := \rho_k * \mathbf{v}^\#$, where ρ_k are as in Section 2.

We now prove (3.31). Due to the presence of the cut-off function φ , the behaviour of $\mathbf{v}^\#$ for small $T-t$ does not affect the validity of (3.31), and thus without loss of generality we may assume that

$$\partial_t \mathbf{v}^\# = \frac{1}{\mathcal{L}} \partial_t (\mathcal{L} \mathbf{v}^\#) + \frac{\mathcal{L}}{\mathcal{L}} \mathbf{v}^\# \in L_{\mathbf{K}^*}(Q; \mathbb{R}^n).$$

Then

$$\mathbf{v}_k^\# \rightarrow \mathbf{v}^\#, \quad \partial_t \mathbf{v}_k^\# \rightarrow \partial_t \mathbf{v}^\#, \quad \partial_t (\varphi \mathbf{v}_k^\#) \rightarrow \partial_t (\varphi \mathbf{v}^\#) \quad (3.32)$$

strongly in $L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$ as $k \rightarrow +\infty$. Set $\mathbf{v}_k := \nabla \mathbf{K}^*(\mathbf{v}_k^\#)$. Observe now that

$$\int_0^T \left[(\mathbf{v}_k, \partial_t (\varphi \mathbf{v}_k^\#)) - (\mathbf{K}(\nabla \mathbf{K}^*(\mathbf{v}_k^\#)), \partial_t \varphi) \right] dt = 0. \quad (3.33)$$

Indeed, integration by parts gives

$$\begin{aligned} & \int_0^T \left[(\mathbf{v}_k, \partial_t (\varphi \mathbf{v}_k^\#)) - (\mathbf{K}(\mathbf{v}_k), \partial_t \varphi) \right] dt \\ &= - \int_0^T \left[(\partial_t \mathbf{v}_k, \varphi \mathbf{v}_k^\#) - (\nabla \mathbf{K}(\mathbf{v}_k) \cdot \partial_t \mathbf{v}_k, \varphi) \right] dt \\ &= - \int_0^T \left[(\partial_t \mathbf{v}_k, \varphi \mathbf{v}_k^\#) - (\mathbf{v}_k^\# \cdot \partial_t \mathbf{v}_k, \varphi) \right] dt = 0. \end{aligned} \quad (3.34)$$

By (3.3),

$$\int_Q \mathbf{K}(\nabla \mathbf{K}^*(w)) \lesssim \int_Q \mathbf{K}^*(w) + 1. \quad (3.35)$$

Proposition 2.11 implies that the Nemytskii operator

$$\mathcal{N}_1 : w \mapsto \mathbf{K}(\nabla \mathbf{K}^*(w))$$

is continuous from $L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$ to $L^1(Q)$. Hence, we can pass to the limit in the second term of (3.33).

On the other hand, it follows from (3.32) and (3.35) that the sequence $\mathbf{v}_k = \nabla \mathbf{K}^*(\mathbf{v}_k^\#)$ is bounded in $L_{\mathbf{K}}(Q; \mathbb{R}^n)$ and a.e. converges to \mathbf{v} . Passing to a subsequence if necessary, we infer $\mathbf{v}_k \rightarrow \mathbf{v}$ weakly in $L_{\mathbf{K}}(Q; \mathbb{R}^n)$. We can now legitimately pass to the limit in the first term of (3.33).

Moreover,

$$\int_0^T (\mathbf{v}_0, \partial_t (\varphi \mathbf{v}_k^\#)) dt = 0 \quad (3.36)$$

since φ is compactly supported in $(0, T)$ and \mathbf{v}_0 does not depend on t . Since $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n) \subset L_{\mathbf{K}}(Q; \mathbb{R}^n)$, we can pass to the limit in (3.36) as well.

As a result, (3.33) and (3.36) imply (3.31).

Let us prove (3.30). Due to the presence of φ , the behaviour of $\mathbf{v}^\#$ for small $T - t$ is again not relevant, and we may assume that

$$L^*(\mathbf{v}^\#) \in L^\infty(Q; \mathbb{R}_s^{N \times N}).$$

Passing to a subsequence if necessary, we get

$$L^*(\mathbf{v}_k^\#) \rightarrow L^*(\mathbf{v}^\#) \quad (3.37)$$

weakly- $*$ in $L^\infty(Q; \mathbb{R}_s^{N \times N})$ as $k \rightarrow +\infty$. On the other hand, (3.7) yields

$$\int_0^T (\mathbf{F}(\nabla \mathbf{K}^*(\mathbf{v}_k^\#)), L^*(\varphi \mathbf{v}_k^\#)) dt = \int_0^T (\mathbf{F}(\nabla \mathbf{K}^*(\mathbf{v}_k^\#)), L^*(\mathbf{v}_k^\#)) \varphi(t) dt = 0. \quad (3.38)$$

Estimates (3.5), (3.35) and Proposition 2.11 imply that the Nemytskii operator

$$\mathcal{N}_2 : \mathbf{w} \mapsto \mathbf{F}(\nabla \mathbf{K}^*(\mathbf{w}))$$

is continuous from $L_{\mathbf{K}^*}(Q; \mathbb{R}^n)$ to $L^1(Q; \mathbb{R}_s^{N \times N})$. Hence,

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

strongly in $L^1(Q; \mathbb{R}_s^{N \times N})$, and passing to the limit in (3.38) we obtain (3.30). \square

The suggested approach of selecting the weak solutions that minimize $\int_0^T \varkappa(t) K(t) dt$ — or, equivalently, $\frac{1}{2} \int_{(0, T) \times \Omega} \varkappa(t) \operatorname{Tr} \mathbf{F}(t) dx dt$ — can be implemented via the saddle-point problem

$$\mathcal{I}(\mathbf{v}_0, T) = \inf_{\mathbf{v}} \sup_{E, B: (3.17)} \int_0^T \left[(\mathbf{v} - \mathbf{v}_0, E) + \frac{1}{2} (\mathbf{F}(\mathbf{v}), \varkappa I + 2B) \right] dt. \quad (3.39)$$

The infimum in (3.39) is taken over all $\mathbf{v} \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)$, and the supremum is taken over all pairs (E, B) satisfying (3.19) and the linear constraint (3.17).

The dual problem is

$$\mathcal{J}(\mathbf{v}_0, T) = \sup_{E, B: (3.17)} \inf_{\mathbf{v}} \int_0^T \left[(\mathbf{v} - \mathbf{v}_0, E) + \frac{1}{2} (\mathbf{F}(\mathbf{v}), \varkappa I + 2B) \right] dt, \quad (3.40)$$

where \mathbf{v}, E, B are varying in the same function spaces as above.

It is also relevant to consider the corresponding problems for the subsolutions:

$$\tilde{\mathcal{I}}(\mathbf{v}_0, T) = \inf_{\mathbf{v}, M: \mathbf{F}(\mathbf{v}) \leq M} \sup_{E, B: (3.17)} \int_0^T \left[(\mathbf{v} - \mathbf{v}_0, E) + \frac{1}{2} (M, \varkappa I + 2B) \right] dt. \quad (3.41)$$

The infimum in (3.39) is taken over all $\mathbf{v} \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)$ and $M \in L^1((0, T) \times \Omega; \mathbb{R}_\Lambda^{N \times N})$, and the supremum is taken over all pairs (E, B) satisfying (3.19) and the linear constraint (3.17). The problem dual to (3.41) is

$$\tilde{\mathcal{J}}(\mathbf{v}_0, T) = \sup_{E, B: (3.17)} \inf_{\mathbf{v}, M: \mathbf{F}(\mathbf{v}) \leq M} \int_0^T \left[(\mathbf{v} - \mathbf{v}_0, E) + \frac{1}{2} (M, \varkappa I + 2B) \right] dt, \quad (3.42)$$

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

Remark 3.17 (Simple observations about the optimal values). Denote

$$K_0 := \frac{1}{2} \int_{\Omega} \operatorname{Tr} \mathbf{F}(\mathbf{v}_0) dx = K(0) + \frac{1}{2} \int_{\Omega} \operatorname{Tr} \mathbf{F}(0) dx. \quad (3.43)$$

Since $\inf \sup \geq \sup \inf$, one has

$$\mathcal{I}(\mathbf{v}_0, T) \geq \tilde{\mathcal{I}}(\mathbf{v}_0, T) \geq \tilde{\mathcal{J}}(\mathbf{v}_0, T), \quad \mathcal{I}(\mathbf{v}_0, T) \geq \mathcal{J}(\mathbf{v}_0, T) \geq \tilde{\mathcal{J}}(\mathbf{v}_0, T).$$

Note that the \sup in (3.39) is always $+\infty$ if \mathbf{v} is not a weak solution, whence $\mathcal{I}(\mathbf{v}_0, T) = +\infty$ if there are no weak solutions. On the other hand, if (3.7) holds, and there exists a strong solution \mathbf{v} , then the corresponding \sup in (3.39) is equal to

$$\begin{aligned} \frac{1}{2} \int_0^T (\mathbf{F}(\mathbf{v}), \varkappa I) dt &= \int_0^T \varkappa \left[K(t) + \frac{1}{2} \int_{\Omega} \operatorname{Tr} \mathbf{F}(0) dx \right] dt \\ &= \int_0^T \varkappa \left[K(0) + \frac{1}{2} \int_{\Omega} \operatorname{Tr} \mathbf{F}(0) dx \right] dt = \int_0^T \varkappa K_0 dt = \varkappa(0) K_0, \end{aligned}$$

which yields $\mathcal{I}(\mathbf{v}_0, T) \leq \varkappa(0) K_0$. Finally, testing by $(E, B) = (0, 0)$ we see that

$$\begin{aligned} \tilde{\mathcal{J}}(\mathbf{v}_0, T) &\geq \frac{1}{2} \inf_{\mathbf{v}, M: \mathbf{F}(\mathbf{v}) \leq M} \int_0^T (M, \varkappa I) dt \geq \\ &\inf_{\mathbf{v}, M: \mathbf{F}(\mathbf{v}) \leq M} \int_{(0, T) \times \Omega} \varkappa \left(\mathbf{K}(\mathbf{v}) + \frac{1}{2} \operatorname{Tr} \mathbf{F}(0) \right) dx dt \geq \frac{\varkappa(0) |\Omega| \operatorname{Tr} \mathbf{F}(0)}{2} \geq 0. \end{aligned}$$

It is easy to see (taking into account that $M + A \geq M$ for any $A \geq 0$) that if $\varkappa I + 2B$ is not non-negative-definite on a set of positive Lebesgue measure in $(0, T) \times \Omega$, then the \inf in (3.42) equals $-\infty$. Hence, any solution to (3.42) necessarily satisfies

$$\varkappa I + 2B \geq 0 \text{ a.e. in } (0, T) \times \Omega. \quad (3.44)$$

Consider the nonlinear functional

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

defined by the formula

$$\mathcal{K}(E, B) = \inf_{\mathbf{F}(z) \leq M} \int_0^T \left[(z, E) + \frac{1}{2} (M, \varkappa I + 2B) \right] dt, \quad (3.45)$$

where the infimum is taken over all pairs $(z, M) \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n) \times L^1((0, T) \times \Omega; \mathbb{R}_s^{N \times N})$.

Then (3.42) is equivalent to

$$\tilde{\mathcal{J}}(\mathbf{v}_0, T) = \sup_{E, B: (3.17), (3.44)} \int_0^T -(\mathbf{v}_0, E) dt + \mathcal{K}(E, B), \quad (3.46)$$

the supremum is taken over all pairs (E, B) belonging to the class (3.19).

Remark 3.18 (A variant of our theory based on the relative entropy). Inspired by [3], is possible to develop a variant of our theory for which the primal problem consists in minimizing a suitable time integral of the *relative entropy* (also known as the Bregman divergence)

$$\mathcal{D}(t) := \int_{\Omega} [\mathbf{K}(\mathbf{v}(t, \mathbf{x})) - \mathbf{K}(\tilde{\mathbf{v}}(t, \mathbf{x})) - \tilde{\mathbf{v}}^{\#}(t, \mathbf{x}) \cdot (\mathbf{v}(t, \mathbf{x}) - \tilde{\mathbf{v}}(t, \mathbf{x}))] d\mathbf{x}. \quad (3.47)$$

Here

$$\tilde{\mathbf{v}} : [0, T] \times \Omega \rightarrow \mathbb{R}^n$$

is a fixed “base state”. (Since $\mathbf{K}(\mathbf{0}) = \mathbf{0}$ and $\mathbf{0}^{\#} = \mathbf{0}$, the setting that we have adopted in this paper corresponds to $\tilde{\mathbf{v}} \equiv \mathbf{0}$.) The implementation of this idea lies beyond the scope of this paper, but it is very plausible that an existence theorem (similar to Theorem 5.4) and a global in time consistency result similar to [3, Theorem 3.6], cf. Remark 4.4, should hold for the corresponding dual problem. An interesting task would be to obtain consistency results in which $\tilde{\mathbf{v}}$ would neither be identically zero (as in Theorem 4.1) nor coincide with the strong solution (as in [3, Theorem 3.6]).

4. CONSISTENCY AND DAFERMOS’ PRINCIPLE

The following theorem shows that a strong solution determines a solution to the optimization problem (3.46), and vice versa. This advocates the possibility to view the maximizers of (3.46) as “dual variational solutions” to (3.1), cf. [11, 57, 3], see also Remark 5.5.

Theorem 4.1 (Consistency). *Let $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. Let \mathbf{v} be a strong solution to (3.1) satisfying*

$$\mathcal{I} \geq -2\mathcal{L}(t)L^*(\mathbf{v}^{\#}) \text{ a.e. in } (0, T) \times \Omega. \quad (4.1)$$

Then $\mathcal{I}(\mathbf{v}_0, T) = \tilde{\mathcal{J}}(\mathbf{v}_0, T) = \tilde{\mathcal{I}}(\mathbf{v}_0, T) = \mathcal{J}(\mathbf{v}_0, T) = \mathcal{L}(0)\mathcal{K}_0$. The pair $(\mathbf{E}_+, \mathbf{B}_+)$ defined by

$$\mathbf{B}_+ = L^* \mathbf{a}, \mathbf{E}_+ = \partial_t \mathbf{a}, \quad (4.2)$$

where

$$\mathbf{a} = \mathcal{L}\mathbf{v}^{\#}, \quad (4.3)$$

belongs to the class (3.19) and maximizes (3.46). Moreover, one can invert these formulas and express \mathbf{v} in terms of \mathbf{E}_+ as follows

$$\mathbf{v}(t, \mathbf{x}) = \nabla \mathbf{K}^* \left(\frac{1}{\mathcal{L}(t)} \int_t^T (-\mathbf{E}_+)(s, \mathbf{x}) ds \right), \quad t < T. \quad (4.4)$$

Proof. By construction, the pair $(\mathbf{E}_+, \mathbf{B}_+)$ belongs to $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n) \times L^{\infty}((0, T) \times \Omega; \mathbb{R}_s^{N \times N})$. It follows from $\mathcal{L}(T) = \mathbf{0}$ that this pair verifies (3.17). Moreover, (4.1) implies (3.44) for \mathbf{B}_+ . We now claim that

$$\frac{1}{2}(\mathcal{I} + 2\mathbf{B}_+) : \partial_t \mathbf{F}(\mathbf{v}) + (\mathbf{E}_+)_t = 0. \quad (4.5)$$

Indeed, using (3.9) we compute

$$\begin{aligned}
& \frac{1}{2}(\varepsilon I + 2B_+) : \partial_l \mathbf{F}(\mathbf{v}) + (E_+)_l \\
&= \frac{1}{2}(\varepsilon I + 2\mathcal{L}L^*(\mathbf{v}^\#)) : \partial_l \mathbf{F}(\mathbf{v}) + (-\varepsilon(\mathbf{v}^\#)_l + \mathcal{L}\partial_t(\mathbf{v}^\#)_l) \\
&= \varepsilon \partial_l \mathbf{K}(\mathbf{v}) + \mathcal{L}L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\mathbf{v}) + (-\varepsilon(\mathbf{v}^\#)_l + \mathcal{L}\partial_t(\mathbf{v}^\#)_l) \\
&= \mathcal{L}(L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\mathbf{v}) + \partial_t(\mathbf{v}^\#)_l) = 0. \quad (4.6)
\end{aligned}$$

On the other hand, since \mathbf{v} is in particular a weak solution, it satisfies (3.16) with test functions (E_+, B_+) . Thus we have

$$\int_0^T [(\mathbf{v} - \mathbf{v}_0, E_+) + (\mathbf{F}(\mathbf{v}), B_+)] dt = 0. \quad (4.7)$$

Hence, by (4.5),

$$\int_0^T [-(\mathbf{v}_0, E_+) + (\mathbf{F}(\mathbf{v}), B_+)] dt = \frac{1}{2} \sum_{l=1}^n \int_0^T (\mathbf{v}_l(\varepsilon I + 2B_+), \partial_l \mathbf{F}(\mathbf{v})) dt. \quad (4.8)$$

Employing (4.8) and Lemma 3.16, we obtain

$$\begin{aligned}
& \frac{1}{2} \int_0^T (\mathbf{F}(\mathbf{v}), \varepsilon I + 2B_+) dt - \frac{1}{2} \sum_{l=1}^n \int_0^T (\mathbf{v}_l \partial_l \mathbf{F}(\mathbf{v}), \varepsilon I + 2B_+) dt \\
&= \int_0^T \left[(\mathbf{v}_0, E_+) + \varepsilon(t) \left(\mathbf{K}(\mathbf{v}(t)) + \frac{1}{2} \text{Tr} \mathbf{F}(0), 1 \right) \right] dt \\
&= \int_0^T \left[(\mathbf{v}_0, E_+) + \varepsilon(t) \left(\mathbf{K}(\mathbf{v}_0) + \frac{1}{2} \text{Tr} \mathbf{F}(0), 1 \right) \right] dt \\
&= \int_0^T (\mathbf{v}_0, E_+) dt + \varepsilon(0) K_0. \quad (4.9)
\end{aligned}$$

By Remark 3.17, we have $\mathcal{I}(\mathbf{v}_0, T) \leq \varepsilon(0) K_0$. Hence, it suffices to show that

$$\int_0^T -(\mathbf{v}_0, E_+) dt + \varepsilon(0) K_0 \geq \varepsilon(0) K_0, \quad (4.10)$$

so that there is no duality gap.

The Λ -convexity of \mathbf{F} implies that the function $\mathbf{y} \mapsto \frac{1}{2} \mathbf{F}(\mathbf{y}) : (\varepsilon I + 2B_+(t, \mathbf{x}))$ is convex w.r.t. $\mathbf{y} \in \mathbb{R}^n$ for a.e. fixed pair of parameters $(t, \mathbf{x}) \in (0, T) \times \Omega$. Now (4.5), (4.9) and (3.44) for B_+

yield

$$\begin{aligned}
& \int_0^T -(\mathbf{v}_0, E_+) dt + \mathcal{K}(E_+, B_+) \\
&= \int_0^T -(\mathbf{v}_0, E_+) dt \\
&+ \inf_{\mathbf{F}(z) \leq M} \int_0^T \left[-\frac{1}{2} \sum_{l=1}^n (z_l \partial_l \mathbf{F}(\mathbf{v}), \varkappa I + 2B_+) + \frac{1}{2} (M, \varkappa I + 2B_+) \right] dt \\
&\geq \int_0^T -(\mathbf{v}_0, E_+) dt \\
&+ \inf_{z \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)} \int_0^T \left[-\frac{1}{2} \sum_{l=1}^n (z_l \partial_l \mathbf{F}(\mathbf{v}), \varkappa I + 2B_+) + \frac{1}{2} (\mathbf{F}(z), \varkappa I + 2B_+) \right] dt \\
&= \int_0^T -(\mathbf{v}_0, E_+) dt \\
&+ \int_0^T \left[-\frac{1}{2} \sum_{l=1}^n (\mathbf{v}_l \partial_l \mathbf{F}(\mathbf{v}), \varkappa I + 2B_+) + \frac{1}{2} (\mathbf{F}(\mathbf{v}), \varkappa I + 2B_+) \right] dt \\
&= \varkappa(\mathbf{0})K_0.
\end{aligned}$$

The penultimate equality follows from the well-known fact that, given a convex function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$, the function $\mathbf{y} \mapsto \Phi(\mathbf{y}) - \mathbf{y} \cdot \nabla \Phi(\boldsymbol{\xi})$ attains its minimum at $\mathbf{y} = \boldsymbol{\xi}$.

Finally, since $\varkappa(T) = \mathbf{0}$, we deduce from (4.2) and (4.3) that

$$\int_t^T E_+(s, \mathbf{x}) ds = \int_t^T \partial_s (\varkappa(s) \mathbf{v}^\#(s, \mathbf{x})) ds = -\varkappa(t) \mathbf{v}^\#(t, \mathbf{x}), \quad (4.11)$$

which yields (4.4). \square

Remark 4.2 (Limitation (4.1) can be overcome by adapting the weight \varkappa). At first glance, condition (4.1) indicates that the interval for the which the consistency holds can be smaller than the interval $[0, T)$ on which the strong solution⁶ exists. However, the weight \varkappa can be selected to guarantee the consistency on any interval $[0, T_1]$, $T_1 < T$. Indeed, by Remark 3.15 with $\varkappa_1 \equiv \mathbf{1}$, for any strong solution on $[0, T)$ we have that $(T-t)L^*(\mathbf{v}^\#(t, \mathbf{x}))$ is essentially bounded. Therefore, for any $T_1 < T$ there exists $\gamma > \mathbf{0}$ such that

$$\gamma I \geq -2L^*(\mathbf{v}^\#) \quad (4.12)$$

a.e. on $(0, T_1) \times \Omega$. It is easy to see from Definition 3.13 that \mathbf{v} (more accurately, the restriction $\mathbf{v}|_{[0, T_1] \times \Omega}$) is a strong solution on $[0, T_1]$. Consider the weight $\varkappa(t) := \exp(-\gamma t)$ and let $\varkappa(t) := \int_t^{T_1} \varkappa(s) ds \geq 0$ (note that we are now working on the interval $[0, T_1]$). Observe that

$$\gamma \varkappa \leq \varkappa.$$

⁶We say “the strong solution” because strong solutions are unique, see Theorem 6.1.

Because of (4.12), this implies that (4.1) holds on $(0, T_1) \times \Omega$.

As an application of Theorem 4.1 we establish a variant of Dafermos' principle for (3.1) affirming that a strong solution dissipates the total entropy not slower and not later than any other subsolution.

Theorem 4.3 (Dafermos' principle). *Let $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. Let \mathbf{v} be a strong solution to (3.1) on the interval $[0, T]$ with total entropy $K(t) = K(0) = \int_{\Omega} \mathbf{K}(\mathbf{v}_0) dx$, and let (u, M) be a subsolution on $[0, T]$ with total entropy $\tilde{K}(t) = \frac{1}{2}(M(t, \cdot) - \mathbf{F}(0), I)$. Then for any $0 \leq t_0 < t_1 \leq T$ it cannot simultaneously be that $\tilde{K}(t) \leq K(t)$ for a.a. $t \in (0, t_1)$ and $\tilde{K}(t) < K(t)$ for a.a. $t \in (t_0, t_1)$. In particular, it is impossible that $\tilde{K}(t) < K(t)$ for a.a. $t \in (0, \epsilon)$, $\epsilon > 0$.*

Proof. We just prove the first claim, to obtain the second one it suffices to put $t_0 = 0$ and $t_1 = \epsilon$.

Assume that a subsolution (u, M) satisfies $\tilde{K}(t) \leq K(0)$ for a.a. $t \in (0, t_1)$ and $\tilde{K}(t) < K(0)$ for a.a. $t \in (t_0, t_1)$ for some $0 \leq t_0 < t_1 \leq T$. W.l.o.g. $t_1 < T$. Fix any $T_1 \in (t_1, T)$. Let γ be a sufficiently large number that satisfies (4.12) a.e. on $(0, T_1) \times \Omega$ and other lower bounds to be determined below. As in Remark 4.2, we set $\varsigma(t) := \exp(-\gamma t)$ and $\varkappa(t) := \int_t^{T_1} \varsigma(s) ds$, and infer that (4.1) holds on $(0, T_1) \times \Omega$.

We now claim that

$$\int_{t_0}^{T_1} \varsigma(t)(\tilde{K}(t) - K(t)) dt < 0, \quad (4.13)$$

which implies that

$$\int_0^{T_1} \varsigma(t)(\tilde{K}(t) - K(t)) dt < 0. \quad (4.14)$$

Since \mathbf{v} is a strong solution, its total entropy $K(t) = K(0)$, so (4.14) is equivalent to

$$\int_0^{T_1} \varsigma(t)\tilde{K}(t) dt < \int_0^{T_1} \varsigma(t)K(0) dt = \varkappa(0)K(0).$$

Hence,

$$\frac{1}{2} \int_0^{T_1} \varsigma(t)(M(t, \cdot), I) dt < \frac{1}{2} \int_0^{T_1} \varsigma(t)(\mathbf{F}(0), I) dt + \varkappa(0)K(0) = \varkappa(0)K_0. \quad (4.15)$$

It follows from Theorem 4.1 that

$$\tilde{I}(\mathbf{v}_0, T_1) = \varkappa(0)K_0. \quad (4.16)$$

It is easy to conclude from Definition 3.10 that the restriction $(u, M)|_{(0, T_1) \times \Omega}$ is a subsolution on $(0, T_1)$. Thus in order to get a contradiction it is enough to put together (3.41) with $T = T_1$, (4.15) and (4.16).

It remains to prove (4.13). Let

$$\epsilon := \int_{t_1}^{T_1} \varsigma(t)(\tilde{K}(t) - K(t)) dt.$$

Then our claim (4.13) becomes

$$\int_{t_0}^{t_1} \varsigma(t)(\tilde{K}(t) - K(t)) dt < -\epsilon. \quad (4.17)$$

If $\epsilon \leq 0$, (4.17) is obvious, so let us assume $\epsilon > 0$. Observe now that

$$\epsilon \leq \varsigma(t_1) \int_0^{T_1} |\tilde{K}(t) - K(t)| dt \leq C_\epsilon \varsigma(t_1),$$

where C_ϵ does not depend on γ . Fix any point $t_2 \in (t_0, t_1)$. Let γ be so large that

$$\exp(\gamma(t_1 - t_2)) \int_{t_0}^{t_2} (K(t) - \tilde{K}(t)) dt \geq C_\epsilon.$$

Then we conclude that

$$\begin{aligned} \int_{t_0}^{t_1} \varsigma(t)(\tilde{K}(t) - K(t)) dt &< \int_{t_0}^{t_2} \varsigma(t)(\tilde{K}(t) - K(t)) dt \\ &\leq \varsigma(t_2) \int_{t_0}^{t_2} (\tilde{K}(t) - K(t)) dt = \exp(-\gamma t_2) \int_{t_0}^{t_2} (\tilde{K}(t) - K(t)) dt \\ &\leq -\exp(-\gamma t_1) C_\epsilon \leq -\epsilon. \end{aligned}$$

□

Remark 4.4 (Comparison with [3]). The results of this section are new even for the quadratic case $\mathbf{F}(\mathbf{v}) = \mathbf{v} \otimes \mathbf{v}$, and are consequently applicable to the incompressible Euler equation, cf. [11, 57]. Another global in time consistency result for the dual formulation of the incompressible Euler has recently been obtained in [3, Theorem 3.6]. Nevertheless, our results (Theorem 4.1 and Remark 4.2) seem to be of a different nature, and, furthermore, the two attitudes complement each other to a certain degree. Indeed, [3, Theorem 3.6] states that if the strong solution \mathbf{v} to the incompressible Euler coincides with the “base state” $\tilde{\mathbf{v}}$, then the solution of the dual problem is identically zero (in the variables used in [3]). Thus the information about the strong solution is contained not in the solution of the dual problem but in the “base state” only. Moreover, the proof of [3, Theorem 3.6] ignores the formal conservativity of the problem (and therefore can be extended to the Navier-Stokes). In contrast, our “base state” is identically zero (see Remark 3.18), and the information about the strong solution is contained in the solution $(\mathbf{B}_+, \mathbf{E}_+)$ of the dual problem. This information can be retrieved by formula (4.4) that strongly relies on the formal conservativity (3.7).

5. SOLVABILITY OF THE DUAL PROBLEM

In this section we prove solvability of the dual problem (3.46) under an additional technical assumption on the operator L .

Definition 5.1 (Strong trace condition). The operator L is said to satisfy the strong trace condition if there exists a uniform constant c such that for any $\zeta \in D(L^*)$ such that the eigenvalues of the matrix $-L^* \zeta(\mathbf{x})$ are uniformly bounded from above by a constant k for a.e. $\mathbf{x} \in \Omega$, the eigenvalues of the matrix $L^* \zeta(\mathbf{x})$ are also uniformly bounded from above a.e. in Ω by ck .

Remark 5.2. The strong trace condition is particularly satisfied provided

$$L(qI) = 0 \tag{5.1}$$

for any smooth scalar function $q(\mathbf{x})$. Indeed, it suffices to observe that in this case the trace of the matrix function $L^* \zeta$ vanishes almost everywhere because

$$(\mathrm{Tr}(L^* \zeta), q) = (L^* \zeta, qI) = (\zeta, L(qI)) = 0.$$

Remark 5.3. Although the strong trace condition is not very restrictive and holds in many situations, cf. Section 7, see also [57], it fails for the system of conservation laws (1.2) with a symmetric flux matrix \mathbf{F} . Indeed, the adjoint of $-\mathrm{div}$ is the symmetric part of the Jacobian matrix, and there is no way to control its eigenvalues from above if they are bounded from below. However, in many cases it is possible to recast the problem in a form that admits the strong trace condition. In Section 7 we will show how to implement this for the scalar conservation laws.

Theorem 5.4 (Existence in anisotropic Orlicz spaces). *Assume that L satisfies the strong trace condition. Then for any $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$ there exists a maximizer (E, B) to (3.46) in the class (3.19), and*

$$0 \leq \frac{\varkappa(0)|\Omega| \mathrm{Tr} \mathbf{F}(0)}{2} \leq \tilde{\mathcal{J}}(\mathbf{v}_0, T) < +\infty.$$

Proof. It suffices to consider the pairs (E, B) that meet the restrictions (3.17), (3.44). Testing (3.46) with $E = 0$, $B = 0$ as in Remark 3.17, we see that

$$\tilde{\mathcal{J}}(\mathbf{v}_0, T) \geq \frac{\varkappa(0)|\Omega| \mathrm{Tr} \mathbf{F}(0)}{2} \geq 0.$$

Let (E_m, B_m) be a maximizing sequence. Since $0 \leq \tilde{\mathcal{J}}(\mathbf{v}_0, T)$, without loss of generality we may assume that

$$0 \leq - \int_0^T (\mathbf{v}_0, E_m) dt + \mathcal{K}(E_m, B_m). \quad (5.2)$$

The eigenvalues of $-B_m$ are uniformly bounded from above because $\varkappa I + 2B_m \geq 0$. Since the strong trace condition is assumed, a uniform L^∞ bound on B_m follows directly from Definition 5.1. In other words, $\varkappa I + 2B_m \preceq \varkappa c I$ with some constant $c > 0$ a.e. in $(0, T) \times \Omega$. By the definition of \mathcal{K} in (3.45), we have

$$\begin{aligned} \mathcal{K}(E_m, B_m) &\leq \inf_{\mathbf{F}(z) \leq M} \int_0^T \left[(z, E_m) + \frac{\varkappa c}{2} (M, I) \right] dt \\ &= \inf_{z \in L_{\mathbf{K}}} \int_0^T \left[(z, E_m) + \left(\varkappa c, \mathbf{K}(z) + \frac{1}{2} \mathrm{Tr} \mathbf{F}(0) \right) \right] dt \\ &= - \int_0^T \left(\varkappa c, \mathbf{K}^* \left(-\frac{E_m}{\varkappa c} \right) \right) dt + \frac{c \varkappa(0) |\Omega| \mathrm{Tr} \mathbf{F}(0)}{2}. \end{aligned} \quad (5.3)$$

Employing the Fenchel–Young inequality we infer that

$$\begin{aligned} \int_0^T \left(\varkappa c, \mathbf{K}^* \left(-\frac{E_m}{\varkappa c} \right) \right) dt &\leq - \int_0^T (\nu_0, E_m) dt + \frac{c\mathfrak{z}(0)|\Omega| \operatorname{Tr} \mathbf{F}(0)}{2} \\ &= \int_0^T \frac{\varkappa c}{2} \left(2\nu_0, -\frac{E_m}{\varkappa c} \right) dt + \frac{c\mathfrak{z}(0)|\Omega| \operatorname{Tr} \mathbf{F}(0)}{2} \\ &\leq \int_0^T \frac{1}{2} (\mathbf{K}(2\nu_0), \varkappa c) dt + \int_0^T \frac{1}{2} \left(\mathbf{K}^* \left(-\frac{E_m}{\varkappa c} \right), \varkappa c \right) dt + \frac{c\mathfrak{z}(0)|\Omega| \operatorname{Tr} \mathbf{F}(0)}{2}, \end{aligned} \quad (5.4)$$

whence

$$\int_0^T \frac{1}{2} \left(\mathbf{K}^* \left(-\frac{E_m}{\varkappa c} \right), \varkappa c \right) dt \leq \mathfrak{z}(0)C, \quad (5.5)$$

where C depends only on the L^1 -norm of $\mathbf{K}(2\nu_0)$ that is finite. This together with (2.3) yields a uniform $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$ -bound on E_m . It follows from (5.3), (5.4) and (5.5) that the right-hand side of (5.2) is uniformly bounded, whence $\tilde{\mathcal{J}}(\nu_0, T) < +\infty$. The functional \mathcal{K} is concave and upper semicontinuous on $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n) \times L^\infty((0, T) \times \Omega; \mathbb{R}_s^{n \times n})$ as an infimum of affine continuous functionals, cf. (3.45). The functional $\int_0^T (\nu_0, \cdot) dt$ is a linear bounded functional on $L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$. Consequently, every weak- $*$ accumulation point of (E_m, B_m) is a maximizer of (3.46). Note that the constraints (3.17), (3.44) are preserved by the limit. \square

Remark 5.5 (Back to the original problem). Let (E, B) be any maximizer of (3.46) satisfying (3.19). Mimicking (4.4), we can define a “generalized solution” to (3.1) by setting

$$\mathbf{v}^\#(t, \mathbf{x}) := -\frac{1}{\mathfrak{z}(t)} \int_t^T E(s, \mathbf{x}) ds, \quad \mathbf{v}(t, \mathbf{x}) := \nabla \mathbf{K}^*(\mathbf{v}^\#(t, \mathbf{x})). \quad (5.6)$$

This object \mathbf{v} automatically belongs to the same regularity class as the strong solutions. Indeed,

$$\partial_t[\mathfrak{z}\mathbf{v}^\#] = E \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n).$$

On the other hand, since

$$\frac{1}{\mathfrak{z}(t)} \lesssim \frac{1}{T-t},$$

our anisotropic Hardy inequality (see Appendix A) implies that $\mathbf{v}^\# \in L_{\mathbf{K}^*}((0, T) \times \Omega; \mathbb{R}^n)$, whence $\mathbf{v} \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)$ by (3.3). Finally, $\mathbf{v}^\#(t) \in D(L^*)$ for a.a. $t \in (0, T)$, and

$$\mathfrak{z}L^*(\mathbf{v}^\#) = B \in L^\infty((0, T) \times \Omega; \mathbb{R}_s^{n \times n}). \quad (5.7)$$

Indeed, let $\psi : (0, T) \times \Omega \rightarrow \mathbb{R}_s^{N \times N}$ be an arbitrary smooth compactly supported matrix field, and set $\Psi(t, \cdot) := \int_0^t \psi(\tau, \cdot) d\tau$; integrating by parts and using (3.17), we deduce that

$$\begin{aligned} \int_0^T (B(t, \cdot), \psi(t, \cdot)) dt &= - \int_0^T (E(t, \cdot), L\Psi(t, \cdot)) dt = \\ &= - \int_0^T \left(\int_t^T E(\tau, \cdot) d\tau, L\Psi(t, \cdot) \right) dt = \int_0^T (\mathfrak{z}(t)L^*\mathbf{v}^\#(t, \cdot), \psi(t, \cdot)) dt. \end{aligned}$$

However, at this level of generality, \mathbf{v} is not necessarily a strong solution.

6. UNIQUENESS OF STRONG SOLUTIONS

We believe that it is possible to prove the weak-strong uniqueness property in the sense that existence of a strong solution (as in Definition 3.13) implies that the corresponding solution (4.2) to the dual problem (3.42) is unique in the class (3.19). The corresponding result in the quadratic case $\mathbf{F}(\mathbf{v}) = \mathbf{v} \otimes \mathbf{v}$ was established in [57, Section 5]. As the first step towards this conjecture, we show here that a strong solution is always unique. The proof heavily relies on Λ -convexity of \mathbf{F} .

Theorem 6.1 (Uniqueness of strong solutions). *Let u, \mathbf{v} be two strong solutions to (3.1) with the same initial datum $\mathbf{v}_0 \in L_{\mathbf{K}}(\Omega; \mathbb{R}^n)$. Then, for every $t \in [0, T)$, $u(t, \cdot) = \mathbf{v}(t, \cdot)$ a.e. in Ω .*

Proof. To fix the ideas, we will assume that u, \mathbf{v} are regular enough to perform the manipulations below. (The proof in the general case follows exactly the same strategy with some tedious and rather standard technicalities: in the (comparatively easy) quadratic case $\mathbf{F}(\mathbf{v}) = \mathbf{v} \otimes \mathbf{v}$ the rigorous implementation can be found in [57, Lemma 5.2]). In order to avoid heavy notation, we will write $u(t)$ and $\mathbf{v}(t)$, and often even u and \mathbf{v} , instead of $u(t, \cdot)$ and $\mathbf{v}(t, \cdot)$.

The key idea is to look at the evolution of the ‘‘Jeffreys divergence’’⁷

$$\zeta(t) := (u(t) - \mathbf{v}(t), u^\#(t) - \mathbf{v}^\#(t)) = (u(t) - \mathbf{v}(t), \nabla \mathbf{K}(u)(t) - \nabla \mathbf{K}(\mathbf{v})(t)).$$

By strict convexity of \mathbf{K} , $\zeta(t) \geq 0$, and $\zeta(t) = 0$ if and only if $u(t) = \mathbf{v}(t)$. Obviously, $\zeta(0) = 0$, and it is enough to show that

$$\zeta(t) = 0, \quad t \in [0, T_1] \tag{6.1}$$

for any $T_1 < T$.

Using (3.1) and (3.9), we compute

$$\begin{aligned} \zeta'(t) &= (\partial_t u(t) - \partial_t \mathbf{v}(t), u^\#(t) - \mathbf{v}^\#(t)) + (u(t) - \mathbf{v}(t), \partial_t u^\#(t) - \partial_t \mathbf{v}^\#(t)) \\ &= (\mathbf{F}(u) - \mathbf{F}(\mathbf{v}), L^*(u^\#) - L^*(\mathbf{v}^\#)) \\ &\quad + \sum_{l=1}^n (u_l - \mathbf{v}_l, L^*(\mathbf{v}^\#) : \partial_l \mathbf{F}(\mathbf{v}) - L^*(u^\#) : \partial_l \mathbf{F}(u)) \\ &= \left(-L^*(u^\#), \mathbf{F}(\mathbf{v}) - \mathbf{F}(u) - \sum_{l=1}^n (\mathbf{v}_l - u_l) : \partial_l \mathbf{F}(u) \right) \\ &\quad + \left(-L^*(\mathbf{v}^\#), \mathbf{F}(u) - \mathbf{F}(\mathbf{v}) - \sum_{l=1}^n (u_l - \mathbf{v}_l) : \partial_l \mathbf{F}(\mathbf{v}) \right). \end{aligned} \tag{6.2}$$

It follows from Definition 3.13, cf. (3.24), that $L^*(u^\#)$ and $L^*(\mathbf{v}^\#)$ are essentially bounded on $[0, T_1] \times \Omega$. Thus there exists a constant $\gamma > 0$ such that $\gamma I - L^*(u^\#)$ and $\gamma I - L^*(\mathbf{v}^\#)$ are positive-definite matrix functions a.e. on $[0, T_1] \times \Omega$, and for some uniform $c > 0$ we have

⁷The classical Jeffreys divergence corresponds to the case $n = 1$ and $\mathbf{K}(\mathbf{v}) = \mathbf{v} \log \mathbf{v}$, which is however ruled out by our assumptions.

$-L^*(u^\#) - L^*(v^\#) \leq cI$. Consequently,

$$\begin{aligned} \zeta'(t) = & \left(\gamma I - L^*(u^\#), \mathbf{F}(v) - \mathbf{F}(u) - \sum_{l=1}^n (v_l - u_l) : \partial_l \mathbf{F}(u) \right) \\ & + \left(\gamma I - L^*(v^\#), \mathbf{F}(u) - \mathbf{F}(v) - \sum_{l=1}^n (u_l - v_l) : \partial_l \mathbf{F}(v) \right) \\ & - \gamma \sum_{l=1}^n (u_l - v_l, \text{Tr}[\partial_l \mathbf{F}(u) - \partial_l \mathbf{F}(v)]). \end{aligned} \quad (6.3)$$

Due to the Λ -convexity of \mathbf{F} , the functions $y \mapsto (\gamma I - L^*(u^\#(t, \mathbf{x}))) : \mathbf{F}(y)$ and $y \mapsto (\gamma I - L^*(v^\#(t, \mathbf{x}))) : \mathbf{F}(y)$ are convex w.r.t. $y \in \mathbb{R}^n$ for a.e. fixed pair of parameters $(t, \mathbf{x}) \in [0, T_1] \times \Omega$. On the other hand, it is easy to see that for any convex function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ we have

$$0 \leq \Phi(y) - \Phi(\xi) - (y - \xi) \cdot \nabla \Phi(\xi) \leq (y - \xi) \cdot (\nabla \Phi(y) - \nabla \Phi(\xi)), \quad y, \xi \in \mathbb{R}^n. \quad (6.4)$$

Leveraging (6.4), we derive from (6.3) that

$$\begin{aligned} \zeta'(t) \leq & \left(\gamma I - L^*(u^\#), \sum_{l=1}^n (v_l - u_l) : (\partial_l \mathbf{F}(v) - \partial_l \mathbf{F}(u)) \right) \\ & + \left(\gamma I - L^*(v^\#), \sum_{l=1}^n (u_l - v_l) : (\partial_l \mathbf{F}(u) - \partial_l \mathbf{F}(v)) \right) \\ & - 2\gamma \sum_{l=1}^n (u_l - v_l, \partial_l \mathbf{K}(u) - \partial_l \mathbf{K}(v)) \\ \leq & \left((2\gamma + c)I, \sum_{l=1}^n (u_l - v_l) : (\partial_l \mathbf{F}(u) - \partial_l \mathbf{F}(v)) \right) - 2\gamma (u - v, u^\# - v^\#) \\ = & (4\gamma + 2c) \sum_{l=1}^n (u_l - v_l, \partial_l \mathbf{K}(u) - \partial_l \mathbf{K}(v)) - 2\gamma \zeta(t) = 2(c + \gamma)\zeta(t), \end{aligned}$$

and (6.1) follows by Grönwall's inequality. \square

7. APPLICATIONS TO PDES

In this section present three relatively simple examples of dispersive equations to which our abstract theory is germane (NLS, NLKG and GKdV equations with generic defocusing power-law nonlinearities). We however believe that it is possible to go much beyond these examples by considering PDEs of higher order (as, e.g., in [41, 10]), with nonlocal terms (e.g., of Benjamin-Ono or Hartree type) or more general nonlinearities. In all these examples, we will be able to verify the assumptions of Section 3 and of Theorem 5.4. Consequently, the theorems of Sections 4, 5 and 6 are fully applicable here.

To the best of our knowledge, global in time solvability for these PDEs is merely known under significant restrictions on the exponents in the power laws or on the size of the initial data, cf. [55, 44, 27]. Theorem 5.4 and Remark 5.5 provide existence of certain “dual variational solutions” for these problems without such restrictions.

As a warm-up exercise, we show how our theory can be applied to scalar conservation laws.

7.1. Scalar conservation laws. In this example we further restrict ourselves to $\Omega = \mathbb{T}^1$. Consider the scalar conservation law

$$\partial_t v = -\partial_x(\mathbf{K}(v)), \quad (7.1)$$

where $\mathbf{K} : \mathbb{R} \rightarrow \mathbb{R}$ is as in Assumption 3.1. Note that there is a tiny ‘‘anisotropy’’ because \mathbf{K} does not need to be even. As anticipated in Remark 5.3, the operator $-\partial_x$ does not satisfy the strong trace condition, so Theorem 5.4 is not directly applicable. In order to overcome this obstacle, we let

$$n = 1, N = 2,$$

$$\mathbf{F}(v) = \left(\begin{array}{c|c} \mathbf{K}(v) & \mathbf{K}(v) \\ \hline \mathbf{K}(v) & \mathbf{K}(v) \end{array} \right).$$

Obviously, $\mathbf{K} = \frac{1}{2} \text{Tr}(\mathbf{F} - \mathbf{F}(0))$ and $v^\# = \mathbf{K}'(v)$.

We now define the operator L by the formula

$$L \left(\begin{array}{c|c} a_{11} & a_{12} \\ \hline a_{12} & a_{22} \end{array} \right) = -\partial_x a_{12}.$$

Then we can rewrite (7.1) in the abstract form (3.1) with L and \mathbf{F} that we have introduced.

Condition (5.1) is obviously true, so the strong trace condition holds. Moreover,

$$L^*(\kappa) = \frac{1}{2} \left(\begin{array}{c|c} 0 & \partial_x \kappa \\ \hline \partial_x \kappa & 0 \end{array} \right).$$

Let us check the conservativity condition (3.7). Remembering that we are now working on \mathbb{T}^1 , we easily see that

$$(\mathbf{F}(v), L^*(v^\#)) = (\mathbf{K}(v), \partial_x(\mathbf{K}'(v))) = -(\partial_x v, (\mathbf{K}'(v))^2) = 0.$$

It is clear that $\mathbf{F} \succeq 0$ and \mathbf{F} is Loewner convex. The subspace Λ consists of the elements of the form

$$\left(\begin{array}{c|c} a_{11} & a_{12} \\ \hline a_{12} & a_{11} \end{array} \right).$$

The validity of Assumptions 3.2 and 3.3 is now obvious.

7.2. GKdV. We keep assuming $\Omega = \mathbb{T}^1$. Consider the ‘‘defocusing’’ generalized Korteweg-de Vries equation [53, 27, 47] (GKdV)⁸

$$\partial_t u + \partial_{\xi\xi\xi} u = |u|^\alpha \partial_\xi u, \quad u(0) = u_0. \quad (7.2)$$

The unknown is $u : [0, T] \times \Omega \rightarrow \mathbb{R}$, and $\alpha \geq 1$ is a prescribed constant (not necessarily integer). Performing a simple change of variable $x := \xi - t$ we rewrite this equation as follows:

$$\partial_t u + \partial_{xxx} u = (|u|^\alpha + 1) \partial_x u, \quad u(0) = u_0. \quad (7.3)$$

⁸To simplify the presentation we consider the real-valued equation, but the same approach works for the complex-valued case.

Letting $w := \partial_x u$ we can rewrite the problem in the following way:

$$\partial_t u = -\partial_{xx} w + \frac{1}{\alpha+1} \partial_x (|u|^\alpha u) + \partial_x u, \quad (7.4)$$

$$\partial_t w = -\partial_{xxx} w + \frac{1}{\alpha+1} \partial_{xx} (|u|^\alpha u) + \partial_{xx} u, \quad (7.5)$$

$$u(0) = u_0, \quad w(0) = w_0. \quad (7.6)$$

Rigorously speaking, this is an extended system, but what is important is that it still possesses an ‘‘entropy’’ that is formally conserved even if we do not assume the compatibility condition $w_0 = \partial_x u_0$.

We let

$$n = 2, N = 4,$$

$$v = (u, w),$$

$$\tilde{v} = (1, u, w, |u|^\alpha),$$

$$\mathbf{K}(v) = \frac{1}{2}(u^2 + w^2) + \frac{|u|^{\alpha+2}}{(\alpha+2)(\alpha+1)},$$

$$\begin{aligned} \mathbf{F}(v) = \epsilon \tilde{v} \otimes \tilde{v} + \left(\frac{1}{2} \mathbf{K}(v) + 1 \right) e_1 \otimes e_1 + \left(-\epsilon u^2 + \frac{1}{2} \mathbf{K}(v) \right) e_2 \otimes e_2 \\ + \left(-\epsilon w^2 + \frac{1}{2} \mathbf{K}(v) \right) e_3 \otimes e_3 + \left(-\epsilon |u|^{2\alpha} + \frac{1}{2} \mathbf{K}(v) \right) e_4 \otimes e_4, \end{aligned}$$

where small $\epsilon > 0$ will be selected later. It is straightforward to check that $\mathbf{K} = \frac{1}{2} \text{Tr}(\mathbf{F} - \mathbf{F}(0))$ and that \mathbf{K} fits into the framework of Assumption 3.1. Furthermore,

$$v^\# = (z, w) = (u + (\alpha+1)^{-1} |u|^\alpha u, w);$$

observe that the second component coincides with w .

Define the operator L by the formula

$$L \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{pmatrix} = \epsilon^{-1} \begin{pmatrix} -\partial_{xx} a_{13} + \frac{1}{\alpha+1} \partial_x a_{24} + \partial_x a_{12} \\ -\partial_{xxx} a_{13} + \frac{1}{\alpha+1} \partial_{xx} a_{24} + \partial_{xx} a_{12} \end{pmatrix}.$$

Then (7.4)–(7.5) can be written in the abstract form (3.1).

Condition (5.1) is obviously true, so the strong trace condition holds. Moreover,

$$L^* \begin{pmatrix} \kappa \\ \theta \end{pmatrix} = \frac{\epsilon^{-1}}{2} \times \begin{pmatrix} 0 & -\partial_x \kappa + \partial_{xx} \theta & -\partial_{xx} \kappa + \partial_{xxx} \theta & 0 \\ -\partial_x \kappa + \partial_{xx} \theta & 0 & 0 & -\frac{1}{\alpha+1} \partial_x \kappa + \frac{1}{\alpha+1} \partial_{xx} \theta \\ -\partial_{xx} \kappa + \partial_{xxx} \theta & 0 & 0 & 0 \\ 0 & -\frac{1}{\alpha+1} \partial_x \kappa + \frac{1}{\alpha+1} \partial_{xx} \theta & 0 & 0 \end{pmatrix}.$$

Let us check the conservativity condition (3.7). We compute, integrating by parts where needed,

$$\begin{aligned}
(\mathbf{F}(\mathbf{v}), L^*(\mathbf{v}^\#)) &= \\
&= (u, -\partial_x Z + \partial_{xx} W) + (w, -\partial_{xx} Z + \partial_{xxx} W) + \left(|u|^\alpha u, -\frac{1}{\alpha+1} \partial_x Z + \frac{1}{\alpha+1} \partial_{xx} W \right) \\
&= (u, -\partial_x Z + \partial_{xx} W) + (\partial_{xx} W, -Z + \partial_x W) + (Z - u, -\partial_x Z + \partial_{xx} W) \\
&= (\partial_{xx} W, \partial_x W) + (Z, -\partial_x Z) = 0.
\end{aligned}$$

Observe that Λ consists of the elements of the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{12} & a_{11} & 0 & a_{24} \\ a_{13} & 0 & a_{11} & 0 \\ 0 & a_{24} & 0 & a_{11} \end{pmatrix}$$

Let us check that $\mathbf{F} \geq 0$ and that \mathbf{F} is Λ -convex. Fix any matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & 0 \\ p_{12} & p_{11} & 0 & p_{24} \\ p_{13} & 0 & p_{11} & 0 \\ 0 & p_{24} & 0 & p_{11} \end{pmatrix} \geq 0.$$

Then

$$\begin{aligned}
\mathbf{F}(\mathbf{v}) : P &= \epsilon p_{11} + \epsilon u^2 p_{11} + \epsilon w^2 p_{11} + \epsilon |u|^{2\alpha} p_{11} \\
&\quad + 2\epsilon u p_{12} + 2\epsilon w p_{13} + 2\epsilon |u|^\alpha u p_{24} \\
&\quad + \frac{1}{2} \mathbf{K}(\mathbf{v}) p_{11} + p_{11} + \left(-\epsilon u^2 + \frac{1}{2} \mathbf{K}(\mathbf{v}) \right) p_{11} \\
&\quad + \left(-\epsilon w^2 + \frac{1}{2} \mathbf{K}(\mathbf{v}) \right) p_{11} + \left(-\epsilon |u|^{2\alpha} + \frac{1}{2} \mathbf{K}(\mathbf{v}) \right) p_{11} \\
&= \left[\epsilon + 1 + u^2 + w^2 + 2 \frac{|u|^{\alpha+2}}{(\alpha+2)(\alpha+1)} \right] p_{11} + 2\epsilon u p_{12} + 2\epsilon w p_{13} + 2\epsilon |u|^\alpha u p_{24} \quad (7.7)
\end{aligned}$$

Since $P \geq 0$,

$$\max\{|p_{12}|, |p_{13}|, |p_{24}|\} \leq p_{11}. \quad (7.8)$$

Thus the function in (7.7) is non-negative provided ϵ is sufficiently small. The Hessian of the function in (7.7) w.r.t. $\mathbf{v} = (u, w)$ is

$$\begin{pmatrix} 2p_{11} + 2|u|^\alpha p_{11} + 2\epsilon\alpha(\alpha+1)|u|^{\alpha-2} u p_{24} & 0 \\ 0 & 2p_{11} \end{pmatrix}.$$

This matrix is always non-negative-definite provided ϵ is less than or equal to a certain constant that depends only on α .

We now we need to compute the partial derivatives of the non-diagonal components of $\pi_\Lambda \mathbf{F}$ w.r.t. the components of \mathbf{v} :

$$\begin{aligned}
\partial_u \mathbf{F}_{12} &= \epsilon, \\
\partial_u \mathbf{F}_{13} &= 0, \\
\partial_u \mathbf{F}_{24} &= \epsilon(\alpha+1)|u|^\alpha \lesssim |u|^{\alpha+1} + 1 \lesssim |\mathbf{v}^\#| + 1,
\end{aligned}$$

$$\begin{aligned}\partial_w \mathbf{F}_{12} &= 0, \\ \partial_w \mathbf{F}_{13} &= \epsilon, \\ \partial_w \mathbf{F}_{24} &= 0.\end{aligned}$$

This immediately implies (3.6).

Remark 7.1 (Lack of Loewner convexity). It is interesting to observe the full Loewner convexity is missing for the matrix function \mathbf{F} above. Indeed, let

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \succeq 0.$$

Then, for any $\alpha \geq 1$ and any fixed $\epsilon > 0$, the function

$$(u, w) \mapsto \mathbf{F}(v) : P = \mathbf{K}(v) + 2\epsilon|u|^\alpha w = \frac{1}{2}(u^2 + w^2) + \frac{|u|^{\alpha+2}}{(\alpha+2)(\alpha+1)} + 2\epsilon|u|^\alpha w$$

is obviously not convex on \mathbb{R}^2 . Similar counterexamples can be constructed for the systems discussed in the sequel.

Remark 7.2 (Sharp formulation of GKdV). For any $\alpha \geq 1$, using the obtained expressions of the partial derivatives of \mathbf{F} , it is elementary to see that the “sharp” problem (3.9) for GKdV reads

$$\begin{aligned}\partial_t z + (1 + \Xi(|z|))(-\partial_x z + \partial_{xx} w) &= 0, \\ \partial_t w - \partial_{xx} z + \partial_{xxx} w &= 0,\end{aligned}$$

where Ξ is the inverse of the function $s \mapsto s^{1/\alpha}(1 + (\alpha+1)^{-1}s)$.

7.3. Defocusing NLS. Let $\Omega = \mathbb{T}^d$. Consider the nonlinear Schrödinger equation with a generic defocusing power-law nonlinearity

$$i\partial_t \Psi = -\Delta \Psi + |\Psi|^{2q} \Psi, \quad \Psi(0) = \Psi_0, \quad (7.9)$$

where $q \geq \frac{1}{2}$ is a given constant. The unknown is $\Psi : [0, T] \times \Omega \rightarrow \mathbb{C}$.

We first change the variable $\psi := \Psi e^{-it}$ to rewrite this in the form

$$i\partial_t \psi = -\Delta \psi + |\psi|^{2q} \psi + \psi, \quad \psi(0) = \Psi_0. \quad (7.10)$$

We now let $a = \Re \psi$, $b = \Im \psi$, $f = \nabla a$, $\theta = \nabla b$. Then the system becomes

$$\partial_t a = -\operatorname{div} \theta + (a^2 + b^2)^q b + b, \quad (7.11)$$

$$\partial_t b = \operatorname{div} f - (a^2 + b^2)^q a - a, \quad (7.12)$$

$$\partial_t f = -\Delta \theta + \nabla((a^2 + b^2)^q b + b), \quad (7.13)$$

$$\partial_t \theta = \Delta f - \nabla((a^2 + b^2)^q a - a). \quad (7.14)$$

As in the previous example, this is an extended system, but we will see that it still possesses an “entropy” that is formally conserved.

We let

$$\begin{aligned}n &= 2d + 2, N = 2d + 4, \\ v &= (a, b, f, \theta),\end{aligned}$$

$$\begin{aligned}\tilde{v} &= (1, a, b, f, \theta, (a^2 + b^2)^q), \\ \mathbf{K}(v) &= \frac{1}{2} \left(f^2 + \theta^2 + a^2 + b^2 + \frac{1}{q+1} (a^2 + b^2)^{q+1} \right), \\ \mathbf{F}(v) &= \epsilon \tilde{v} \otimes \tilde{v} + \left(\frac{2}{N} \mathbf{K}(v) + 1 \right) e_1 \otimes e_1 + \left(-\epsilon a^2 + \frac{2}{N} \mathbf{K}(v) \right) e_2 \otimes e_2 \\ &\quad + \left(-\epsilon b^2 + \frac{2}{N} \mathbf{K}(v) \right) e_3 \otimes e_3 + \sum_{m=1}^d \left(-\epsilon |f_m|^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{3+m} \otimes e_{3+m} \\ &\quad + \sum_{m=1}^d \left(-\epsilon |\theta_m|^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{3+d+m} \otimes e_{3+d+m} + \left(-\epsilon (a^2 + b^2)^{2q} + \frac{2}{N} \mathbf{K}(v) \right) e_N \otimes e_N,\end{aligned}$$

where small $\epsilon > 0$ will be selected later. It is straightforward to check that $\mathbf{K} = \frac{1}{2} \text{Tr}(\mathbf{F} - \mathbf{F}(0))$ and that

$$v^\# = (y, z, f, \theta) = (a + (a^2 + b^2)^q a, b + (a^2 + b^2)^q b, f, \theta).$$

Define the operator L by the formula

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

Then (7.11)–(7.14) can be rewritten in the abstract form (3.1).

Condition (5.1) is obviously true, so the strong trace condition holds. Moreover,

$$L^* \begin{pmatrix} \kappa \\ \theta \\ \eta \\ \xi \end{pmatrix} = \frac{\epsilon^{-1}}{2} \times \begin{pmatrix} 0 & -\theta + \text{div} \xi & \kappa - \text{div} \eta & -\nabla \theta + \Delta \xi & \nabla \kappa - \Delta \eta & 0 \\ -\theta + \text{div} \xi & 0 & 0 & 0 & 0 & -\theta + \text{div} \xi \\ \kappa - \text{div} \eta & 0 & 0 & 0 & 0 & \kappa - \text{div} \eta \\ (-\nabla \theta + \Delta \xi)^\top & 0 & 0 & 0 & 0 & 0 \\ (\nabla \kappa - \Delta \eta)^\top & 0 & 0 & 0 & 0 & 0 \\ 0 & -\theta + \text{div} \xi & \kappa - \text{div} \eta & 0 & 0 & 0 \end{pmatrix}.$$

Let us check the conservativity condition (3.7). We compute, integrating by parts where needed,

$$\begin{aligned}(\mathbf{F}(v), L^*(v^\#)) &= (a, -z + \text{div} \theta) + (b, y - \text{div} f) + (f, -\nabla z + \Delta \theta) + (\theta, \nabla y - \Delta f) \\ &\quad + (a(a^2 + b^2)^q, -z + \text{div} \theta) + (b(a^2 + b^2)^q, y - \text{div} f) \\ &= (y, -z + \text{div} \theta) + (z, y - \text{div} f) + (\text{div} f, z) - (\nabla f, \nabla \theta) - (\text{div} \theta, y) + (\nabla \theta, \nabla f) = 0.\end{aligned}$$

Observe now that Λ consists of the elements of the form

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

Let us show that $\mathbf{F} \geq 0$ and that \mathbf{F} is Λ -convex. Fix

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & P_{14} & P_{15} & 0 \\ p_{12} & p_{11} & 0 & 0 & 0 & p_{26} \\ p_{13} & 0 & p_{11} & 0 & 0 & p_{36} \\ P_{14}^\top & 0^\top & 0^\top & p_{11}I & 0 & 0 \\ P_{15}^\top & 0 & 0 & 0 & p_{11}I & 0 \\ 0 & p_{26} & p_{36} & 0 & 0 & p_{11} \end{pmatrix} \succeq 0.$$

Then, in the same spirit as for the GKdV, we compute

$$\begin{aligned} \mathbf{F}(\mathbf{v}) : P = & \left[\epsilon + 1 + f^2 + \theta^2 + a^2 + b^2 + \frac{1}{q+1}(a^2 + b^2)^{q+1} \right] p_{11} \\ & + 2\epsilon a p_{12} + 2\epsilon b p_{13} + 2\epsilon P_{14} f + 2\epsilon P_{15} \theta \\ & + 2\epsilon (a^2 + b^2)^q a p_{26} + 2\epsilon (a^2 + b^2)^q b p_{36}. \end{aligned} \quad (7.15)$$

Since $P \succeq 0$, it is tedious but elementary to check that the function in (7.15) and its Hessian w.r.t. $\mathbf{v} = (a, b, f, \theta)$ are non-negative and non-negative-definite, resp., provided ϵ is controlled by a constant that depends only on $q \geq \frac{1}{2}$.

It remains to estimate the partial derivatives of the non-diagonal components of $\pi_\Lambda \mathbf{F}$ w.r.t. the components of \mathbf{v} . The only potentially problematic terms (because of the presence of a nonlinearity) are

$$\begin{aligned} \partial_a \mathbf{F}_{26} &= \epsilon (a^2 + b^2)^q + 2\epsilon q a^2 (a^2 + b^2)^{q-1}, \\ \partial_b \mathbf{F}_{26} &= 2\epsilon q a b (a^2 + b^2)^{q-1}, \\ \partial_a \mathbf{F}_{36} &= 2\epsilon q a b (a^2 + b^2)^{q-1}, \\ \partial_b \mathbf{F}_{36} &= \epsilon (a^2 + b^2)^q + 2\epsilon q b^2 (a^2 + b^2)^{q-1}. \end{aligned}$$

All of them are

$$\lesssim (a^2 + b^2)^q \lesssim (a^2 + b^2)^{\frac{2q+1}{2}} + 1 \lesssim |\mathbf{v}^\#| + 1,$$

which yields validity of assumption (3.6).

8. COMPLEX NLKG

Let $\Omega = \mathbb{T}^d$. Consider the complex Klein-Gordon equation with a generic (defocusing) power-law nonlinearity

$$\partial_{tt}\psi - \Delta\psi + |\psi|^{2q}\psi + \psi = 0, \quad \psi(0) = \psi_0. \quad (8.1)$$

Here $q \geq \frac{1}{2}$ is a given constant. The unknown is $\psi : [0, T] \times \Omega \rightarrow \mathbb{C}$.

We let $a = \Re\psi$, $b = \Im\psi$, $f = \nabla a$, $\theta = \nabla b$, $u = \partial_t a$, $v = \partial_t b$. Then the system can be rewritten in the following form:

$$\partial_t a = u, \quad (8.2)$$

$$\partial_t b = v, \quad (8.3)$$

$$\partial_t f = \nabla u, \quad (8.4)$$

$$\partial_t \theta = \nabla v, \quad (8.5)$$

$$\partial_t u = \operatorname{div} f - (a^2 + b^2)^q a - a, \quad (8.6)$$

$$\partial_t v = \operatorname{div} \theta - (a^2 + b^2)^q b - b. \quad (8.7)$$

As above, this is an extended system, but what matters is that it is formally conservative.

We let

$$n = 2d + 4, N = 2d + 6, \quad v = (a, b, f, \theta, u, v),$$

$$\tilde{v} = (1, a, b, f, \theta, u, v, (a^2 + b^2)^q),$$

$$\mathbf{K}(v) = \frac{1}{2} \left(u^2 + v^2 + f^2 + \theta^2 + a^2 + b^2 + \frac{1}{q+1} (a^2 + b^2)^{q+1} \right),$$

$$\begin{aligned} \mathbf{F}(v) = & \epsilon \tilde{v} \otimes \tilde{v} + \left(\frac{2}{N} \mathbf{K}(v) + 1 \right) e_1 \otimes e_1 + \left(-\epsilon a^2 + \frac{2}{N} \mathbf{K}(v) \right) e_2 \otimes e_2 \\ & + \left(-\epsilon b^2 + \frac{2}{N} \mathbf{K}(v) \right) e_3 \otimes e_3 + \sum_{m=1}^d \left(-\epsilon |f_m|^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{3+m} \otimes e_{3+m} \\ & + \sum_{m=1}^d \left(-\epsilon |\theta_m|^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{3+d+m} \otimes e_{3+d+m} \\ & + \left(-\epsilon u^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{N-2} \otimes e_{N-2} \\ & + \left(-\epsilon v^2 + \frac{2}{N} \mathbf{K}(v) \right) e_{N-1} \otimes e_{N-1} \\ & + \left(-\epsilon (a^2 + b^2)^{2q} + \frac{2}{N} \mathbf{K}(v) \right) e_N \otimes e_N, \end{aligned}$$

where small $\epsilon > 0$ will be selected later. It is straightforward to check that $\mathbf{K} = \frac{1}{2} \operatorname{Tr}(\mathbf{F} - \mathbf{F}(0))$. We have

$$v^\# = (y, z, f, \theta, u, v) = (a + a(a^2 + b^2)^q, b + b(a^2 + b^2)^q, f, \theta, u, v).$$

We now set

$$L \begin{pmatrix} a_{11} & a_{12} & a_{13} & A_{14} & A_{15} & a_{16} & a_{17} & a_{18} \\ a_{12} & a_{22} & a_{23} & A_{24} & A_{25} & a_{26} & a_{27} & a_{28} \\ a_{13} & a_{23} & a_{33} & A_{34} & A_{35} & a_{36} & a_{37} & a_{38} \\ A_{14}^\top & A_{24}^\top & A_{34}^\top & A_{44} & A_{45} & A_{46} & A_{47} & A_{48} \\ A_{15}^\top & A_{25}^\top & A_{35}^\top & A_{45}^\top & A_{55} & A_{56} & A_{57} & A_{58} \\ a_{16} & a_{26} & a_{36} & A_{46}^\top & A_{56}^\top & a_{66} & a_{67} & a_{68} \\ a_{17} & a_{27} & a_{37} & A_{47}^\top & A_{57}^\top & a_{67} & a_{77} & a_{78} \\ a_{18} & a_{28} & a_{38} & A_{48}^\top & A_{58}^\top & a_{68} & a_{78} & a_{88} \end{pmatrix} = \epsilon^{-1} \begin{pmatrix} a_{16} \\ a_{17} \\ \nabla a_{16} \\ \nabla a_{17} \\ \operatorname{div} A_{14} - (a_{28} + a_{12}) \\ \operatorname{div} A_{15} - (a_{38} + a_{13}) \end{pmatrix}.$$

Then (7.11)–(7.14) can be rewritten in the abstract form (3.1).

It is clear that

$$L^* \begin{pmatrix} \kappa \\ \theta \\ \eta \\ \xi \\ \chi \\ \omega \end{pmatrix} = \frac{\epsilon^{-1}}{2} \begin{pmatrix} 0 & -\chi & -\omega & -\nabla\chi & -\nabla\omega & \kappa - \operatorname{div}\eta & \theta - \operatorname{div}\xi & 0 \\ -\chi & 0 & 0 & 0 & 0 & 0 & 0 & -\chi \\ -\omega & 0 & 0 & 0 & 0 & 0 & 0 & -\omega \\ -\nabla\chi^\top & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\nabla\omega^\top & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \kappa - \operatorname{div}\eta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta - \operatorname{div}\xi & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\chi & -\omega & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The validity of Assumptions 3.2 and 3.3 and of condition (5.1) can be shown in a way similar to the NLS case. Let us just check Assumption 3.7. We compute, integrating by parts where needed,

$$\begin{aligned} (\mathbf{F}(v), L^*(v^\#)) &= (a, -u) + (b, -v) + (f, -\nabla u) + (\theta, -\nabla v) + (u, y - \operatorname{div} f) \\ &\quad + (v, z - \operatorname{div} \theta) + (a(a^2 + b^2)^q, -u) + (b(a^2 + b^2)^q, -v) \\ &= (y, -u) + (z, -v) + (\operatorname{div} f, u) + (\operatorname{div} \theta, v) + (\nabla \theta, \nabla f) \\ &\quad + (u, y - \operatorname{div} f) + (v, z - \operatorname{div} \theta) = 0. \end{aligned}$$

APPENDIX A. AN ANISOTROPIC HARDY INEQUALITY

Proposition A.1. *Let \mathbf{K} be a C^1 -smooth N -function such that \mathbf{K} and \mathbf{K}^* satisfy the Δ_2 -condition. Given $f \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)$, define the function*

$$F(t, x) := \frac{1}{t} \int_0^t f(s, x) ds.$$

Then

$$F \in L_{\mathbf{K}}((0, T) \times \Omega; \mathbb{R}^n)$$

and

$$\|F\|_{L_{\mathbf{K}}((0, T) \times \Omega)} \lesssim \|f\|_{L_{\mathbf{K}}((0, T) \times \Omega)}. \quad (\text{A.1})$$

Proof. It follows from [17, Lemma 2.3.16] that there exist $r > 0$ and $p > 1$ such that

$$\frac{v \cdot \nabla \mathbf{K}(v)}{\mathbf{K}(v)} \geq p \quad (\text{A.2})$$

for $|v| \geq r$.

We first observe that

$$\mathbf{K}(sv) \leq \mathbf{K}(v)s^p, s \in (0, 1), v \in \mathbb{R}^n, |sv| \geq r. \quad (\text{A.3})$$

Indeed, fix $s_0 \in (0, 1)$ with $|s_0 v| \geq r$, and consider the function

$$\phi(s) = \log \mathbf{K}(sv) - \log(\mathbf{K}(v)s^p), s \in [s_0, 1].$$

Then

$$\phi'(s) = s^{-1} \frac{sv \cdot \nabla \mathbf{K}(sv)}{\mathbf{K}(sv)} - s^{-1}p \geq 0.$$

Since $\phi(1) = 0$, we infer $\mathbf{K}(s_0 v) \leq \mathbf{K}(v)s_0^p$.

It follows from (A.3) that

$$\mathbf{K}(v) \leq s\mathbf{K}(vs^{-1/p}), s \in (0, 1), |v| \geq r. \quad (\text{A.4})$$

Denote $\mathbf{K}_s(v) := s\mathbf{K}(vs^{-1/p})$ and $f_s(t, x) := f(st, x)$, $s \in (0, 1)$. Then

$$F(t, x) = \int_0^1 f_s(t, x) ds \quad (\text{A.5})$$

and

$$\mathbf{K}(v) \leq \mathbf{K}_s(v), |v| \geq r. \quad (\text{A.6})$$

We now claim that

$$\|f_s\|_{L_{\mathbf{K}_s}((0, T) \times \Omega)} \leq s^{-1/p} \|f\|_{L_{\mathbf{K}}((0, T) \times \Omega)}. \quad (\text{A.7})$$

Indeed,

$$\begin{aligned} \|f_s\|_{L_{\mathbf{K}_s}((0, T) \times \Omega)} &= \inf \left\{ \lambda > 0 : \int_{(0, T) \times \Omega} \mathbf{K}_s \left(\frac{f(ts, x)}{\lambda} \right) dx dt \leq 1 \right\} \\ &= \inf \left\{ \lambda > 0 : \int_{(0, sT) \times \Omega} \mathbf{K}_s \left(\frac{f(t, x)}{\lambda} \right) dx dt \leq s \right\} \\ &\leq \inf \left\{ \lambda > 0 : \int_{(0, T) \times \Omega} \mathbf{K}_s \left(\frac{f(t, x)}{\lambda} \right) dx dt \leq s \right\} \\ &= \inf \left\{ \lambda > 0 : \int_{(0, T) \times \Omega} \mathbf{K} \left(\frac{f(t, x)}{\lambda s^{1/p}} \right) dx dt \leq 1 \right\} \\ &= s^{-1/p} \|f\|_{L_{\mathbf{K}}((0, T) \times \Omega)}. \end{aligned}$$

Employing convexity of the Luxemburg norm, Proposition 2.9, (A.5) and (A.6), and integrating (A.7) w.r.t. $s \in (0, 1)$, we conclude that

$$\begin{aligned} \|F\|_{L_{\mathbf{K}}((0, T) \times \Omega)} &= \left\| \int_0^1 f_s ds \right\|_{L_{\mathbf{K}}((0, T) \times \Omega)} \\ &\leq \int_0^1 \|f_s\|_{L_{\mathbf{K}}((0, T) \times \Omega)} ds \leq C_r \int_0^1 \|f_s\|_{L_{\mathbf{K}_s}((0, T) \times \Omega)} ds \leq \frac{pC_r}{p-1} \|f\|_{L_{\mathbf{K}}((0, T) \times \Omega)}. \end{aligned}$$

□

APPENDIX B. BALLISTIC OPTIMAL TRANSPORT

Let us briefly describe the link of our dual problem (3.46) with the optimal transportation problems. We employ a geometric intuition that did not explicitly appear in the previous works.

Let us start from the following heuristic setting. Let \mathfrak{M} be a complete, connected Riemannian manifold. Let $X : \mathfrak{M} \rightarrow T\mathfrak{M}$ be a given vector field on \mathfrak{M} , and let $x_T \in \mathfrak{M}$ be a prescribed point. By a *ballistic*⁹ geodesic problem we mean finding the curve¹⁰ $\gamma_t \subset \mathfrak{M}$ that minimizes

$$\min_{\dot{\gamma}_0=X(\gamma_0), \gamma_T=x_T} \frac{1}{2} \int_0^T \langle \dot{\gamma}_t, \dot{\gamma}_t \rangle_{\gamma_t} dt, \quad (\text{B.1})$$

i.e., we know the final position and the initial direction of an unknown geodesic.

Assume that the vector field X that determines the direction of the “howitzer” at every point of \mathfrak{M} is a potential field, i.e., $X = \text{grad } \Phi$ for some $\Phi : \mathfrak{M} \rightarrow \mathbb{R}$. Then (B.1) is equivalent to the Hopf-Lax formula

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

where *dist* is the Riemannian distance on \mathfrak{M} (here we are merely interested in the minimizer x_0 that determines the starting point of the unknown geodesic and not in the optimal value itself).

Formula (B.2) makes sense even if we have no manifold structure and \mathfrak{M} is merely a metric space, not necessarily connected. This is particularly relevant for the metric spaces with “Riemannian flavour”, see [49, p. 6] for a list of examples of such spaces. Problem (B.2) indeed appears in infinite-dimensional and metric-geometric contexts, including the context of optimal transport. The generalization to $p > 1$, i.e.,

$$\min_{x_0 \in \mathfrak{M}} \Phi(x_0) + \frac{1}{p} \text{dist}^p(x_0, x_T) \quad (\text{B.3})$$

was studied in [5] in a metric context. Problems of this kind are called marginal entropy-transport problems in [43]. They are crucial ingredients of de Giorgi’s minimizing movement scheme, also known as the JKO scheme [34, 4].

Consider now the quadratic Hamilton-Jacobi equation

$$\partial_t \psi + \frac{1}{2} |\nabla \psi|^2 = 0, \quad \psi(0, x) = \psi_0(x), \quad (t, x) \in [0, T] \times \Omega.$$

Setting $v = \nabla \psi$, we rewrite it in the form

$$\partial_t v + \frac{1}{2} \nabla \text{Tr}(v \otimes v) = 0, \quad v(0) = \nabla \psi_0. \quad (\text{B.4})$$

Note that (B.4) is a particular case of our abstract equation (3.1) with $\mathbf{F}(v) = v \otimes v$, $L = -\frac{1}{2} \nabla \text{Tr}$.

Let $\varkappa \equiv \mathbf{1}$. As explained in [57, Section 2.2], the dual problem (3.46) for the quadratic Hamilton-Jacobi equation may be rewritten (this is formal but can be made rigorous) as

$$- \int_{\Omega} \psi_0 \rho(0) dx - \frac{1}{2} \int_0^T \int_{\Omega} \rho^{-1} |q|^2 dx dt \rightarrow \sup \quad (\text{B.5})$$

⁹The wording is borrowed from [7].

¹⁰The subscript t is not a time derivative but just the value at time t .

subject to the constraints

$$\partial_t \rho + \operatorname{div} q = 0, \quad \rho(T) = 1, \quad \rho \geq 0. \quad (\text{B.6})$$

Rescaling if needed, we can assume $T = 1$ and $|\Omega| = 1$. Defining the functional $\Psi(\rho) := \int_{\Omega} \psi_0 d\rho$ on the Wasserstein space $\mathcal{P}(\Omega)$ of probability measures on Ω , multiplying by -1 and leveraging the Benamou-Brenier formula [56, 8], we can recast (B.5), (B.6) in the Hopf-Lax form

$$\min_{\rho_0 \in \mathcal{P}(\Omega)} \Psi(\rho_0) + \frac{1}{2} W_2^2(\rho_0, \rho_1). \quad (\text{B.7})$$

Here W_2 is the quadratic Wasserstein distance [56], and $d\rho_1$ is the Lebesgue measure dx on Ω . Hence, in this very particular but instructive case, the dual problem (3.46) can be viewed as a ballistic problem on the Wasserstein space.

The observations above can be generalized to the case of the ρ -Wasserstein space (and very likely to the Orlicz-Wasserstein spaces [54, 5]). The generating Hamilton-Jacobi equation is

$$\partial_t v + \frac{\rho - 1}{\rho} \nabla \operatorname{Tr} \mathbf{F}(v) = 0$$

with

$$\mathbf{F}(v) = |v|^{\frac{2-\rho}{\rho-1}} v \otimes v. \quad (\text{B.8})$$

We omit the implementation but call attention to the details that the matrix function \mathbf{F} in (B.8) is Λ -convex but not Loewner convex, and that the resulting Hopf-Lax formula is similar to (B.3) because the ρ -Wasserstein space is not ‘Riemannian-like’ unless $\rho = 2$.

REFERENCES

- [1] A. Acharya. Variational principles for nonlinear PDE systems via duality. *Quart. Appl. Math.*, 81(1):127–140, 2023.
- [2] A. Acharya and A. N. Sengupta. Action principles for dissipative, non-holonomic Newtonian mechanics. *Proc. A.*, 480(2293):Paper No. 20240113, 21, 2024.
- [3] A. Acharya, B. Stroffolini, and A. Zarnescu. Variational dual solutions for incompressible fluids. *arXiv preprint arXiv:2409.04911*, 2024.
- [4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [5] L. Ambrosio, N. Gigli, and G. Savaré. Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces. *Revista Matemática Iberoamericana*, 29(3):969–996, 2013.
- [6] A. Baradat and L. Monsaingeon. Small noise limit and convexity for generalized incompressible flows, Schrödinger problems, and optimal transport. *Arch. Ration. Mech. Anal.*, 235(2):1357–1403, 2020.
- [7] A. Barton and N. Ghoussoub. Dynamic and stochastic propagation of the Brenier optimal mass transport. *European Journal of Applied Mathematics*, 30(6):1264–1299, 2019.
- [8] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [9] J. M. Borwein and J. D. Vanderwerff. *Convex functions: constructions, characterizations and counterexamples*, volume 109 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.
- [10] T. Boulenger and E. Lenzmann. Blowup for biharmonic NLS. *Ann. Sci. Éc. Norm. Supér. (4)*, 50(3):503–544, 2017.
- [11] Y. Brenier. The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem. *Comm. Math. Phys.*, 364(2):579–605, 2018.
- [12] Y. Brenier. Examples of hidden convexity in nonlinear pdes. *Preprint*, 2020.
- [13] Y. Brenier and I. Moyano. Relaxed solutions for incompressible inviscid flows: a variational and gravitational approximation to the initial value problem. *Philos. Trans. Roy. Soc. A*, 380(2219):Paper No. 20210078, 12, 2022.

- [14] A. C. Bronzi, M. C. Lopes Filho, and H. J. Nussenzveig Lopes. Wild solutions for 2d incompressible ideal flow with passive tracer. *Communications in Mathematical Sciences*, 13(5):1333–1343, 2015.
- [15] Á. Castro, D. Faraco, and B. Gebhard. Entropy solutions to macroscopic ipm. *arXiv preprint arXiv:2309.03637*, 2023.
- [16] E. Chiodaroli and O. Kreml. On the energy dissipation rate of solutions to the compressible isentropic Euler system. *Arch. Ration. Mech. Anal.*, 214(3):1019–1049, 2014.
- [17] I. Chlebicka, P. Gwiazda, A. Świerczewska-Gwiazda, and A. Wróblewska-Kamińska. *Partial differential equations in anisotropic Musielak-Orlicz spaces*. Springer Monographs in Mathematics. Springer, Cham, 2021.
- [18] T. Cieślak and G. Jamróz. Maximal dissipation in Hunter-Saxton equation for bounded energy initial data. *Adv. Math.*, 290:590–613, 2016.
- [19] C. M. Dafermos. The entropy rate admissibility criterion for solutions of hyperbolic conservation laws. *J. Differential Equations*, 14:202–212, 1973.
- [20] C. M. Dafermos. Maximal dissipation in equations of evolution. *J. Differential Equations*, 252(1):567–587, 2012.
- [21] S. Daneri and L. Székelyhidi, Jr. Non-uniqueness and h-principle for Hölder-continuous weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.*, 224(2):471–514, 2017.
- [22] C. De Lellis and L. Székelyhidi, Jr. On admissibility criteria for weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.*, 195(1):225–260, 2010.
- [23] C. de Lellis and L. Székelyhidi Jr. The Euler equations as a differential inclusion. *Annals of mathematics*, 170(3):1417–1436, 2009.
- [24] R. J. DiPerna. Uniqueness of solutions to hyperbolic conservation laws. *Indiana Univ. Math. J.*, 28(1):137–188, 1979.
- [25] T. K. Donaldson and N. S. Trudinger. Orlicz-Sobolev spaces and imbedding theorems. *J. Functional Analysis*, 8:52–75, 1971.
- [26] E. Feireisl. Maximal dissipation and well-posedness for the compressible Euler system. *J. Math. Fluid Mech.*, 16(3):447–461, 2014.
- [27] I. Friedman, O. Riaño, S. Roudenko, D. Son, and K. Yang. Well-posedness and dynamics of solutions to the generalized KdV with low power nonlinearity. *Nonlinearity*, 36(1):584–635, 2023.
- [28] B. Gebhard, J. Hirsch, and J. J. Kolumbán. On a degenerate elliptic problem arising in the least action principle for Rayleigh-Taylor subsolutions. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 41(6):1527–1594, 2024.
- [29] B. Gebhard and J. J. Kolumbán. Relaxation of the Boussinesq system and applications to the Rayleigh-Taylor instability. *NoDEA Nonlinear Differential Equations Appl.*, 29(1):Paper No. 7, 38, 2022.
- [30] H. Gimperlein, M. Grinfeld, R. J. Knops, and M. Slemrod. The least action admissibility principle. *arXiv preprint arXiv:2409.07191*, 2024.
- [31] J. Glimm, D. Lazarev, and G.-Q. G. Chen. Maximum entropy production as a necessary admissibility condition for the fluid Navier-Stokes and Euler equations. *SN Applied Sciences*, 2:1–9, 2020.
- [32] F. Höfer and N. A. Nikov. On growth of Sobolev norms for periodic nonlinear Schrödinger and generalised Korteweg-de Vries equations under critical Gibbs dynamics. *arXiv preprint arXiv:2412.08630*, 2024.
- [33] L. Hsiao. The entropy rate admissibility criterion in gas dynamics. *J. Differential Equations*, 38(2):226–238, 1980.
- [34] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [35] S.-C. Klein. Using the Dafermos entropy rate criterion in numerical schemes. *BIT*, 62(4):1673–1701, 2022.
- [36] S.-C. Klein. Stabilizing discontinuous Galerkin methods using Dafermos’ entropy rate criterion: I—One-dimensional conservation laws. *J. Sci. Comput.*, 95(2):Paper No. 55, 37, 2023.
- [37] S.-C. Klein. Stabilizing discontinuous Galerkin methods using Dafermos’ entropy rate criterion: II—Systems of conservation laws and entropy inequality predictors. *J. Sci. Comput.*, 100(2):Paper No. 42, 49, 2024.
- [38] M. A. Krasnoselskii. *Topological methods in the theory of nonlinear integral equations*. The Macmillan Company, New York, 1964.
- [39] M. A. Krasnoselskii and J. B. Rutickii. *Convex functions and Orlicz spaces*. P. Noordhoff Ltd., Groningen, 1961.
- [40] F. Kraus. Über konvexe Matrixfunktionen. *Math. Z.*, 41(1):18–42, 1936.
- [41] P. Krejčí and I. Straškraba. A uniqueness criterion for the Riemann problem. *Hiroshima Math. J.*, 27(2):307–346, 1997.
- [42] C. Léonard. From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.*, 262(4):1879–1920, 2012.

- [43] M. Liero, A. Mielke, and G. Savaré. Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. *Invent. Math.*, 211(3):969–1117, 2018.
- [44] F. Linares and G. Ponce. *Introduction to nonlinear dispersive equations*. Springer, 2014.
- [45] K. Löwner. Über monotone Matrixfunktionen. *Math. Z.*, 38(1):177–216, 1934.
- [46] L. M. Martyushev and V. D. Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Phys. Rep.*, 426(1):1–45, 2006.
- [47] S. Masaki and J.-i. Segata. Existence of a minimal non-scattering solution to the mass-subcritical generalized Korteweg–de Vries equation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 35(2):283–326, 2018.
- [48] L. Monsaingeon, L. Tamanini, and D. Vorotnikov. The dynamical Schrödinger problem in abstract metric spaces. *Adv. Math.*, 426:Paper No. 109100, 2023.
- [49] M. Muratori and G. Savaré. Gradient flows and Evolution Variational Inequalities in metric spaces. I: Structural properties. *Journal of Functional Analysis*, 278(4):108347, 2020.
- [50] I. Prigogine. *Introduction to thermodynamics of irreversible processes*. Interscience Publishers, New York-London, revised edition, 1961.
- [51] Y. Shenfeld. Matrix displacement convexity along density flows. *Arch. Ration. Mech. Anal.*, 248(5):Paper No. 74, 41, 2024.
- [52] S. Singh, J. Ginster, and A. Acharya. A hidden convexity of nonlinear elasticity. *J. Elasticity*, 156(3):975–1014, 2024.
- [53] W. A. Strauss. Nonlinear scattering theory at low energy. *J. Functional Analysis*, 41(1):110–133, 1981.
- [54] K.-T. Sturm. Generalized orlicz spaces and wasserstein distances for convex–concave scale functions. *Bulletin des sciences mathématiques*, 135(6-7):795–802, 2011.
- [55] C. Sulem and P.-L. Sulem. *The nonlinear Schrödinger equation: self-focusing and wave collapse*. Springer Science & Business Media, 2007.
- [56] C. Villani. *Topics in optimal transportation*. American Mathematical Soc., 2003.
- [57] D. Vorotnikov. Partial differential equations with quadratic nonlinearities viewed as matrix-valued optimal ballistic transport problems. *Arch. Ration. Mech. Anal.*, 243(3):1653–1698, 2022.
- [58] H. Ziegler. Some extremum principles in irreversible thermodynamics with application to continuum mechanics. In *Progress in solid mechanics, Vol. IV*, pages 91–193. North-Holland, Amsterdam, 1963.

(D. Vorotnikov) UNIVERSITY OF COIMBRA, CMUC, DEPARTMENT OF MATHEMATICS, 3001-501 COIMBRA, PORTUGAL

Email address: mitvorot@mat.uc.pt