STRONG CONVERGENCE OF SEQUENCES WITH VANISHING RELATIVE ENTROPY

NUNO J. ALVES, JAKUB SKRZECZKOWSKI, AND ATHANASIOS E. TZAVARAS

ABSTRACT. We show that under natural growth conditions on the entropy function, convergence in relative entropy is equivalent to L^p-convergence. The main tool is the theory of Young measures, in a form that accounts for the formation of concentrations in weak limits.

1. INTRODUCTION

For a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, define $f(\cdot|\cdot)$ by the quadratic part of the Taylor series expansion

$$f(v|u) = f(v) - f(u) - Df(u) \cdot (v - u) \tag{1.1}$$

for $v, u \in \mathbb{R}^n$, and note that if f is strictly convex, then $f(v|u) > 0$ for $v \neq u$. When f is the entropy (energy), the function $f(\cdot|\cdot)$ is called *relative entropy* (*energy*). In several problems in mechanics, kinetic theory, or information theory, the function f is convex and one is able to derive information on the convergence of a sequence $\{u_n\}$ to a limit function u via an integral relation of the form

$$F[u_n|u] = \int_\Omega f(u_n|u) \, \mathrm{d}x \to 0 \qquad \text{as } n \to \infty \tag{1.2}$$

where $\Omega \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, is a bounded open set (see *e.g.* [7, 10, 13]). The question arises as to what type of convergence this information translates into. An attempt in this direction, employing real analysis techniques, may be found in [2].

The reader should note that $f(\cdot|\cdot)$ or its integral does not define a distance and is not even symmetric. The lack of symmetry may be alleviated by introducing the symmetrized form of the relative entropy:

$$f_{\mathrm{sym}}(v|u) = f(v|u) + f(u|v) = (Df(u) - Df(v)) \cdot (u - v)$$

which is symmetric and (strictly) positive for f strictly convex and $v \neq u$. However, $(v, u) \mapsto \int f_{\mathrm{sym}}(v|u) \, \mathrm{d}x$ still does not induce a distance, as it does not satisfy the triangle inequality. The intent of this note is to prove that under natural growth conditions of the entropy f, convergence in (1.2) implies strong convergence in L^p.

This result can be easily checked in some special cases. It is trivial in the quadratic case $f(u) = |u|^2$, since then $f(v|u) = |v - u|^2$. Another interesting case concerns $h(u) = u \ln u$ (for $n = 1$). The relative entropy then takes the form

$$h(v|u) = v \ln \frac{v}{u} - (v - u).$$

When v, u are probability densities, $\int v \, dx = \int u \, dx = 1$, the relevant quantity becomes the functional

$$H(v|u) = \int v \ln \frac{v}{u} \, dx$$

which is often used in information theory and statistics under the name Kullback-Leiber divergence. The Csiszár-Kullback-Pinsker inequality states that for $v, u \in L^1(\mathbb{R}^d)$ nonnegative functions with $\|v\|_1 = \|u\|_1 = 1$ it holds [10] that

$$\|v - u\|_1^2 \le 2H(v|u) \tag{1.3}$$

indicating that convergence in relative entropy implies L^1-convergence.

Inequality (1.3) can be generalized as

$$\frac{A}{2^{2/p}} \min(\|v\|_p^{p-2}, \|u\|_p^{p-2}) \|u - v\|_p^2 \le \int f(v|u) \, dx$$

where $u, v \ge 0$, $u, v \in L^1 \cap L^p$, $p \in [1, 2]$ and $f \in C^2(0, \infty)$ is a strictly convex nonnegative function with $f(1) = f'(1) = 0$ and $A := \inf_{s \in (0, \infty)} s^{2-p} f''(s) > 0$; see [5, Proposition 3.1].

Our approach employs the theory of Young measures and a representation theorem from [8] as methodological tools to address this question. The analysis shows that convergence in relative entropy precludes concentrations in weak limits.

2. Preliminaries and statement of results

We start by presenting two situations where one may conclude immediately. Suppose the function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the hypothesis that its Hessian is bounded from below

$$D^2 f(\lambda) \ge cI \tag{2.1}$$

for some $c > 0$ and all $\lambda \in \mathbb{R}^n$, and let $\{u_n\}$ be a sequence satisfying (1.2). We then conclude

$$c \int_\Omega |u_n - u|^2 \, dx \le \int_\Omega f(u_n|u) \, dx \to 0$$

as $n \to \infty$.

Another direct case occurs when $f(\lambda) = |\lambda|^p$, $p > 1$, and we assume for the sequence $\{u_n\}$ that

$$u_n \rightharpoonup u \quad \text{weakly in } L^p(\Omega)$$
$$f(u_n|u) \to 0 \quad \text{in } L^1(\Omega) \tag{2.2}$$

for some $u \in L^p(\Omega)$. Then, from (1.1) and the hypothesis (2.2) we deduce

$$\int_\Omega |u_n|^p \, dx - \int_\Omega |u|^p \, dx = \int_\Omega p|u|^{p-2} u \cdot (u_n - u) \, dx + \int_\Omega f(u_n|u) \, dx \to 0$$

that is, $\|u_n\|_{L^p(\Omega)} \to \|u\|_{L^p(\Omega)}$. Hence, $u_n \to u$ in $L^p(\Omega)$.

The assumption of global uniform convexity (2.1) might be too restrictive (and often unattainable when L^∞ bounds are not available), while the case of the L^p norm is specialized. To address the problem for more general functions, we use the theory of Young measures introduced by L. Tartar, who formalized an idea for generalized functions initially proposed by L.C. Young. For thorough presentations of the relevant results, the reader is referred to Ball [3] and Pedregal [11, Chapter 6]. The second key component is the representation of potential concentrations in the entropy norm via a concentration measure, as developed in [8, Appendix A]. This representation is generalized here to hold for a convex entropy h.

Proposition 2.1. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a convex function that satisfies for $p \geq 1$ the bound*

$$h(\lambda) \geq \frac{1}{c}|\lambda|^p - c$$

for some constant $c > 0$ and every $\lambda \in \mathbb{R}^n$. Let $\{u_n\}$ be a sequence of functions $u_n : \Omega \to \mathbb{R}^n$, where Ω is an open bounded subset of $\mathbb{R}^d$, that satisfy the uniform bounds

$$\sup_n \int_\Omega h(u_n) \, dx \leq C < \infty. \tag{2.3}$$

There exist a subsequence $\{u_{n_k}\}$, a parametrized family of probability measures (Young measures) $\{\nu_x\}_{x \in \Omega}$, and a concentration measure $\mu(dx)$ such that

$$f(u_{n_k}) \rightharpoonup \int f(\lambda) \, d\nu_x(\lambda) \quad \text{weakly in } L^1(\Omega) \tag{2.4}$$

for every continuous function $f : \mathbb{R}^n \to \mathbb{R}$ satisfying $\lim_{|\lambda| \to \infty} \frac{|f(\lambda)|}{1+|\lambda|^p} = 0$, and

$$\int_\Omega h(u_{n_k}) \varphi \, dx \to \int_\Omega \left(\int h(\lambda) \, d\nu_x(\lambda) \right) \varphi \, dx + \int_\Omega \varphi \, \mu(dx) \quad \text{as } k \to \infty \tag{2.5}$$

for all $\varphi \in C(\bar{\Omega})$.

We remark that the representation (2.4) of the weak limit via Young measures in the most general setting holds for sequences $\{u_n\}$ and $\{\psi(x, u_n)\}$ such that $\{u_n\}$ is bounded in $L^q(\Omega)$ for some $q > 0$, and ψ is a Carathéodory function so that $\{\psi(x, u_n(x))\}$ is relatively weakly compact in $L^1(\Omega)$; see [11, Theorem 6.2]. In that case, we have

$$\psi(x, u_{n_k}(x)) \rightharpoonup \int \psi(x, \lambda) \, d\nu_x(\lambda) \quad \text{weakly in } L^1(\Omega).$$

Proposition 2.1, combined with the characterizations of uniform p-integrability for a sequence of functions $\{u_n\}$ (see Appendix), leads to an equivalence between L^p-convergence for $p > 1$ and convergence in relative entropy.

Theorem 2.2. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable, strictly convex function satisfying the bounds*

$$\frac{1}{c}|\lambda|^p - c \leq h(\lambda) \leq c|\lambda|^p + c \tag{2.6}$$

for some $p > 1$, $c > 0$ and every $\lambda \in \mathbb{R}^n$.

Let $u_n, u : \Omega \to \mathbb{R}^n$ belong to $L^p(\Omega)$, where Ω is an open bounded subset of $\mathbb{R}^d$. Then

$$\int_\Omega h(u_n | u) \, dx \to 0 \qquad \text{if and only if} \qquad u_n \to u \text{ in } L^p(\Omega)$$

as $n \to \infty$.

An obvious example satisfying the conditions of Theorem 2.2 is the function $h(\lambda) = |\lambda|^p$. On the other hand, in applications of relative entropy, such as those arising in kinetic theory or information theory, one often encounters the function $\lambda \ln \lambda$. The following theorem covers this case by selecting $h(\lambda) = 1/e + \lambda \ln \lambda$, $\lambda \geq 0$.

Theorem 2.3. *Consider a function $h : [0, \infty) \to [0, \infty)$ and assume the following:*

    (i) $h \in C([0, \infty)) \cap C^2(0, \infty)$,

    (ii) $h''(\lambda) > 0$ for every $\lambda > 0$,

    (iii) $\dfrac{h(\lambda)}{\lambda} \to \infty$ as $\lambda \to \infty$,

    (iv) $c\lambda - c \leq h(\lambda)$ for some $c > 0$ and every $\lambda \geq 0$.

Let $\{u_n\} \subseteq L^1(\Omega)$ and $u \in L^\infty(\Omega)$ be nonnegative. If $Dh(u) \in L^\infty(\Omega)$,

$$\sup_n \int_\Omega h(u_n)\,\mathrm{d}x = C < \infty \tag{2.7}$$

and

$$\int_\Omega h(u_n|u)\,\mathrm{d}x \to 0 \qquad \text{as } n \to \infty$$

then $\{u_n\}$ converges to u in $L^1(\Omega)$.

3. Representation via Young measures and concentration measures

We provide the proof of Proposition 2.1. Let $\{u_n\}$ satisfy the uniform bounds (2.3). The standard theory of Young measures [3], [11] implies there exists a subsequence $\{u_{n_k}\}$ and a parametrized family of probability measures $\{\nu_x\}_{x\in\Omega}$ such that (2.4) holds for any continuous $f : \mathbb{R}^n \to \mathbb{R}$ that satisfies $\lim_{|\lambda|\to\infty} \frac{|f(\lambda)|}{1+|\lambda|^p} = 0$.

The representation formula does not cover the entropy function h. There are two alternatives that one may pursue. One can use the theory of concentration measures of DiPerna-Majda [9] or Alibert-Bouchitté [1]. This approach has the advantage that it can represent entire families of functions with critical growth but under the assumption that their recession functions satisfy continuity properties on the unit sphere. An alternative approach is pursued in [8] which can only represent the entropy function, but with no continuity assumptions on the recession function, only assuming convexity and positivity. We assume first that $h : \mathbb{R}^n \to \mathbb{R}$ is convex and nonnegative (the nonnegativity assumption will be removed later). The proof proceeds in two steps:

First, one considers the map $\langle \nu, h \rangle : x \mapsto \int h(\lambda)\,\mathrm{d}\nu_x(\lambda)$ for $x \in \Omega$ and shows that

$$\int h(\lambda)\,\mathrm{d}\nu_x(\lambda) \in L^1(\Omega). \tag{3.1}$$

To this end, for $R > 0$, let h_R be the truncation of h:

$$h_R(\lambda) = \begin{cases} h(\lambda) & \text{when } h(\lambda) \leq R, \\ R & \text{when } h(\lambda) > R. \end{cases}$$

Since $h_R \nearrow h$ as $R \to \infty$, the monotone convergence theorem implies

$$\int h(\lambda)\,\mathrm{d}\nu_x(\lambda) := \lim_{R\to\infty} \int h_R(\lambda)\,\mathrm{d}\nu_x(\lambda) \quad \text{for a.e. } x \in \Omega$$

providing a definition for the bracket $\langle \nu, h \rangle$. Observe that

$$\int_\Omega \left(\int h_R(\lambda)\,\mathrm{d}\nu_x(\lambda) \right) \mathrm{d}x = \lim_{k\to\infty} \int_\Omega h_R(u_{n_k})\,\mathrm{d}x$$

$$\leq \limsup_{k\to\infty} \int_\Omega h(u_{n_k})\,\mathrm{d}x \leq C.$$

Letting $R \to \infty$ and using $h_R \nearrow h$ yields (3.1).

Second, for the sequence $\{h(u_{n_k}) - \int h(\lambda)\,\mathrm{d}\nu_x(\lambda)\}$ we prove consecutively:

(i) $\sup_k \int_\Omega |h(u_{n_k}) - \int h(\lambda)\,\mathrm{d}\nu_x(\lambda)|\,\mathrm{d}x \leq C < \infty$,

(ii) There exists a concentration measure $\mu(\mathrm{d}x) \in \mathcal{M}(\Omega)$ such that along a subsequence

$$h(u_{n_k}) - \int h(\lambda)\,\mathrm{d}\nu_x(\lambda) \rightharpoonup \mu(\mathrm{d}x) \qquad \text{weak-}* \text{ in measures},$$

(iii) $\mu(\mathrm{d}x) \geq 0$.

Part (i) follows from (3.1) and (2.3). Then, due to (i), there exists a (signed) measure $\mu \in \mathcal{M}(\Omega)$ such that along a subsequence which we still denote by $\{u_{n_k}\}$ we have

$$\int_\Omega \left(h(u_{n_k}) - \int h(\lambda) \, \mathrm{d}\nu_x(\lambda) \right) \varphi(x) \, \mathrm{d}x \to \int_\Omega \varphi(x) \, \mu(\mathrm{d}x)$$

for any $\varphi \in C(\bar{\Omega})$. Finally, for $\varphi \in C(\bar{\Omega})$, $\varphi \geq 0$, we have

$$\int_\Omega \left(\int h(\lambda) \, \mathrm{d}\nu_x(\lambda) \right) \varphi(x) \, \mathrm{d}x = \sup_{R>0} \int_\Omega \left(\int h_R(\lambda) \, \mathrm{d}\nu_x(\lambda) \right) \varphi(x) \, \mathrm{d}x$$

$$= \sup_{R>0} \lim_{k\to\infty} \int_\Omega h_R(u_{n_k}) \, \varphi(x) \, \mathrm{d}x$$

$$\leq \lim_{k\to\infty} \int_\Omega h(u_{n_k}) \, \varphi(x) \, \mathrm{d}x$$

which implies $\mu := \text{weak} - * \lim \left(h(u_{n_k}) - \int h(\lambda) d\nu_x(\lambda) \right) \geq 0$.

This completes the proof when h is a convex and nonnegative function. If we only know that h is convex, then there exists $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ such that

$$h(\lambda) \geq -a \cdot \lambda - b.$$

The function $\hat{h}(\lambda) := h(\lambda) + a \cdot \lambda + b$ is convex and nonnegative. We apply the previous procedure to $\hat{h}(\lambda)$. Since $u_{n_k} \rightharpoonup u$ weakly in $L^1(\Omega)$ we can transfer the relation (2.5) from the function $\hat{h}$ to the function h and complete the proof.

4. Proof of Theorem 2.2

First, we prove that strong convergence in $L^p(\Omega)$ implies convergence in relative entropy. Suppose that $\{u_n\}$ converges to u in $L^p(\Omega)$. From [6, Proposition 2.32] and (2.6), it follows that $Dh(u) \in L^{p'}(\Omega)$, where $p' = \frac{p}{p-1}$ the conjugate exponent of p. Hence

$$\int_\Omega |Dh(u)||u_n - u| \, \mathrm{d}x \to 0 \quad \text{as } n \to \infty. \tag{4.1}$$

Using [6, Proposition 2.32] once more,

$$\int_\Omega |h(u_n) - h(u)| \, \mathrm{d}x \leq C \int_\Omega (1 + |u_n|^{p-1} + |u|^{p-1})|u_n - u| \, \mathrm{d}x$$

$$\leq C \left(\int_\Omega (1 + |u_n|^{p-1} + |u|^{p-1})^{p'} \, \mathrm{d}x \right)^{\frac{1}{p'}} \|u_n - u\|_{L^p(\Omega)} \tag{4.2}$$

$$\to 0 \quad \text{as } n \to \infty.$$

Combining (4.1) and (4.2) yields the desired convergence in relative entropy.

The converse direction is more intricate. We prove that every subsequence of $\{u_n\}$ has a further subsequence that converges in $L^p(\Omega)$ to u. Thus, it suffices to show the existence of a subsequence of $\{u_n\}$ that converges to u in $L^p(\Omega)$. Let $\{u_n\}$, u satisfy

$$\int_\Omega h(u_n|u) \, \mathrm{d}x \to 0. \tag{4.3}$$

Assume without loss of generality that $h(\lambda) \geq 0$. Since h is convex, it satisfies for some $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ the lower bound

$$h(\lambda) \geq -a \cdot \lambda - b.$$

If h assumes negative values, then it is replaced by the convex, nonnegative function $\hat{h}(\lambda) = h(\lambda) + a \cdot \lambda + b$ which still satisfies (4.3).

We first prove that $\int_\Omega h(u_n)\,\mathrm{d}x \leq C$ for some $C > 0$ and every $n \in \mathbb{N}$. From the convergence in relative entropy, the sequence $\left\{\int_\Omega h(u_n|u)\,\mathrm{d}x\right\}$ is bounded, so there exists $K > 0$ such that for every $n \in \mathbb{N}$,

$$
\begin{aligned}
\int_\Omega h(u_n)\,\mathrm{d}x &= \int_\Omega h(u_n|u) + h(u) + Dh(u)\cdot(u_n - u)\,\mathrm{d}x \\
&\leq K + c\int_\Omega |u|^p\,\mathrm{d}x + c|\Omega| + \int_\Omega |Dh(u)\cdot u|\,\mathrm{d}x + \int_\Omega |Dh(u)\cdot u_n|\,\mathrm{d}x.
\end{aligned}
\tag{4.4}
$$

We estimate the last term on the right-hand side of the expression above using Young's inequality with a parameter ϵ to obtain

$$
\int_\Omega |Dh(u)\cdot u_n|\,\mathrm{d}x \leq C(\epsilon)\int_\Omega |Dh(u)|^{p'}\,\mathrm{d}x + \epsilon c^2|\Omega| + \epsilon c\int_\Omega h(u_n)\,\mathrm{d}x
\tag{4.5}
$$

where $C(\epsilon) = (\epsilon p)^{-\frac{p'}{p}}(p')^{-1}$. Combining (4.4) with (4.5), taking $\epsilon = 1/(2c)$ and noting that $Dh(u)\cdot u \in L^1(\Omega)$ yields the desired uniform bound.

Let $\{u_{n_k}\}$ be a subsequence, $\{\nu_x\}$ parametrized probability measures and $\mu(\mathrm{d}x)$ a concentration measure as in Proposition 2.1. Observe that $u_{n_k} \rightharpoonup \int \lambda\,\mathrm{d}\nu_x(\lambda)$ weakly in $L^p(\Omega)$, and hence

$$
\lim_{k\to\infty} \int_\Omega Dh(u)\cdot(u_{n_k} - u)\,\mathrm{d}x = \int_\Omega \int Dh(u)\cdot(\lambda - u)\,\mathrm{d}\nu_x(\lambda)\,\mathrm{d}x.
$$

Therefore, using (2.5) with $\varphi = 1$, we have

$$
\int_\Omega \int h(\lambda|u)\,\mathrm{d}\nu_x(\lambda)\,\mathrm{d}x + \mu(\Omega) = 0
$$

which implies $\mu = 0$ and $\int h(\lambda|u)\,\mathrm{d}\nu_x(\lambda) = 0$ for a.e. $x \in \Omega$ since $h(\lambda|u) \geq 0$. The strict convexity of h then yields $\nu_x = \delta_{u(x)}$. Applying (2.4) to the function $\psi(\lambda, x) = |\lambda - u(x)|^q$, with $1 \leq q < p$, gives

$$
\int_\Omega |u_{n_k} - u|^q\,\mathrm{d}x \to 0
$$

and along subsequences $u_{n_k} \to u$ almost everywhere and in measure. Since all subsequences of $\{u_n\}$ have the same limit we conclude $u_n \to u$ in $L^q(\Omega)$, $q \in [1, p)$.

The final step is to establish the convergence in $L^p(\Omega)$ of a subsequence of $\{u_{n_k}\}$. For that subsequence (not relabeled), it suffices to prove that $\{u_{n_k}\}$ is uniformly p-integrable; see the Appendix for characterizations of this property. The result then follows by the Vitali convergence theorem.

For simplicity of notation, we relabel the sequence $\{u_{n_k}\}$ as $\{u_k\}$ for the remainder of the proof. Since $u_k \rightharpoonup u$ weakly in $L^p(\Omega)$ it follows that

$$
\int_\Omega Dh(u)\cdot(u_k - u)\,\mathrm{d}x \to 0 \quad \text{as } n \to \infty.
\tag{4.6}
$$

In turn, (4.3) implies

$$
\int_\Omega h(u_k)\,\mathrm{d}x \to \int_\Omega h(u)\,\mathrm{d}x.
\tag{4.7}
$$

Now, since $\{u_k\}$ converges to u in measure, then along a subsequence if necessary, $\{u_k\}$ converges to u almost everywhere. For each $M > 0$, define $\varphi_M : \mathbb{R}^n \to \mathbb{R}$ by

$$
\varphi_M(\lambda) = \begin{cases} 1, & \text{if } 0 \leq |\lambda| \leq M, \\ M - |\lambda| + 1, & \text{if } M < |\lambda| \leq M + 1, \\ 0, & \text{if } |\lambda| \geq M + 1. \end{cases}
\tag{4.8}
$$

The sequence of functions $\{h(u_k)\,\varphi_M(u_k)\}$ is dominated by the constant function $\max_{0 \leq |\lambda| \leq M+1} h(\lambda)$. By the dominated convergence theorem,

$$\int_\Omega h(u_k)\varphi_M(u_k)\,\mathrm{d}x \to \int_\Omega h(u)\varphi_M(u)\,\mathrm{d}x \quad \text{as } k \to \infty,$$

which together with (4.7) implies

$$\int_\Omega h(u_k)(1 - \varphi_M(u_k))\,\mathrm{d}x \to \int_\Omega h(u)(1 - \varphi_M(u))\,\mathrm{d}x \quad \text{as } k \to \infty.$$

We conclude that for each $M > 0$ there exists $K_M \in \mathbb{N}$ such that for $k \geq K_M$ it holds

$$\int_\Omega h(u_k)(1 - \varphi_M(u_k))\,\mathrm{d}x \leq 2 \int_\Omega h(u)(1 - \varphi_M(u))\,\mathrm{d}x. \tag{4.9}$$

Next, using (2.6), we select $M_* > 1$ such that

$$h(\lambda) \geq \frac{1}{2c}|\lambda|^p \quad \text{for } |\lambda| \geq M_* .$$

For $M > M_*$ we may select K_M such that for $k > K_M$ using (4.9) we have

$$\int_{\{|u_k| \geq M+1\}} |u_k|^p\,\mathrm{d}x \leq 2c \int_{\{|u_k| \geq M+1\}} h(u_k)\,\mathrm{d}x$$

$$= 2c \int_{\{|u_k| \geq M+1\}} h(u_k)(1 - \varphi_M(u_k))\,\mathrm{d}x$$

$$\leq 4c \int_\Omega h(u)(1 - \varphi_M(u))\,\mathrm{d}x. \tag{4.10}$$

Now, let $\varepsilon > 0$ be fixed and select $\tilde{M} = \tilde{M}(\varepsilon) > M_*$ so that

$$4c \int_\Omega h(u)(1 - \varphi_{\tilde{M}}(u))\,\mathrm{d}x < \varepsilon^p.$$

For $k \geq K_{\tilde{M}}$, (4.10) gives

$$\int_{\{|u_k| \geq \tilde{M}+1\}} |u_k|^p\,\mathrm{d}x < \varepsilon^p. \tag{4.11}$$

Since $u_k \in L^p(\Omega)$, for each $k = 1, \ldots, K_{\tilde{M}} - 1$ there exists $M_k > 0$ so that

$$\int_{\{|u_k| \geq M_k\}} |u_k|^p\,\mathrm{d}x \leq \varepsilon^p, \quad k = 1, \ldots, K_{\tilde{M}} - 1. \tag{4.12}$$

Setting $M = \max\{\tilde{M} + 1, M_1, \ldots, M_{K_{\tilde{M}}} - 1\}$, from (4.11) and (4.12) it follows that

$$\sup_k \int_{\{|u_k| \geq M\}} |u_k|^p\,\mathrm{d}x < \varepsilon^p$$

which proves that $\{u_k\}$ is uniformly p-integrable. This completes the proof of Theorem 2.2.

5. Proof of Theorem 2.3

Similarly to the proof of Theorem 2.2, we show that $\{u_n\}$ has a subsequence that converges to u in $L^1(\Omega)$. For simplicity, none of the subsequences will be relabeled. We remark that we are working here with a nonnegative sequence $u_n \geq 0$.

First, we notice that the sequence $\{u_n\}$ is bounded in $L^1(\Omega)$. This follows from condition (iv) in conjunction with (2.7). Next, we prove that $\{u_n\}$ is uniformly integrable. This follows from the De La Vallé-Poussin criterion; we provide the details for the reader's convenience. Let $\varepsilon > 0$ and

choose $M_\varepsilon > 0$ such that $C/M_\varepsilon < \varepsilon$. From (iii) we infer the existence of a constant $\Lambda_\varepsilon > 0$ such that whenever $\lambda > \Lambda_\varepsilon$ one has $h(\lambda) \geq M_\varepsilon \lambda$. Therefore

$$\int_{\{u_n \geq \Lambda_\varepsilon\}} u_n \, dx \leq \frac{1}{M_\varepsilon} \int_{\{u_n \geq \Lambda_\varepsilon\}} h(u_n) \, dx$$

$$\leq \frac{C}{M_\varepsilon} < \varepsilon$$

which establishes the uniform integrability of $\{u_n\}$. Thus, along a subsequence if necessary, $\{u_n\}$ converges weakly in $L^1(\Omega)$. There exists a Young measure $\{\nu_x\}_{x \in \Omega}$ generated by a subsequence of $\{u_n\}$, such that for any Carathéodory function $\psi : \Omega \times \mathbb{R}^+ \to \mathbb{R}$, if $\{\psi(x, u_n(x))\}$ converges weakly in $L^1(\Omega)$, its weak limit is represented by the map

$$x \mapsto \int_{\mathbb{R}^+} \psi(x, \lambda) \, d\nu_x(\lambda).$$

Noting that the arguments in the proof of Proposition 2.1 also work in this case with $p = 1$, there exists a (nonnegative) concentration measure $\mu(dx)$ such that

$$\int_\Omega h(u_n) \, \varphi \, dx \to \int_\Omega \left(\int_{\mathbb{R}^+} h(\lambda) \, d\nu_x(\lambda) \right) \varphi \, dx + \int_\Omega \varphi \, \mu(dx) \tag{5.1}$$

as $n \to \infty$, for all $\varphi \in C(\bar\Omega)$. As in the proof of Theorem 2.2 we deduce that $\nu_x = \delta_{u(x)}$ for almost every $x \in \Omega$.

Now, we apply [11, Proposition 6.5] to the Carathéodory function $\psi(x, \lambda) = |\lambda - u(x)|$. Since $u \in L^\infty(\Omega)$, we have $\psi(x, \lambda) \leq K(1 + \lambda)$ for some positive constant K. We claim that the sequence $\{|u_n - u|\}$ is relatively weakly compact in $L^1(\Omega)$. Clearly, it is bounded in $L^1(\Omega)$. Moreover, for each $k > K$, let $m_k = \frac{k}{K} - 1$ so that

$$\{\lambda \in \mathbb{R}^+ \; : \; K(1 + \lambda) \geq k\} \subseteq \{\lambda \in \mathbb{R}^+ \; : \; \lambda \geq m_k\}.$$

From condition (iii) it follows that

$$\lim_{\lambda \to \infty} \frac{h(\lambda)}{K(1 + \lambda)} = \infty$$

and so we can find $M_k > 0$, with $M_k \to \infty$ as $k \to \infty$, such that $h(\lambda) \geq M_k K(1 + \lambda)$ whenever $\lambda \geq m_k$. Consequently,

$$\int_{\{\psi(x, u_n(x)) \geq k\}} \psi(x, u_n(x)) \, dx \leq \int_{\{\psi(x, u_n(x)) \geq k\}} K(1 + u_n) \, dx$$

$$\leq \int_{\{u_n \geq m_k\}} K(1 + u_n) \, dx$$

$$\leq \frac{1}{M_k} \int_{\{u_n \geq m_k\}} h(u_n) \, dx$$

$$\leq \frac{C}{M_k} \to 0 \qquad \text{as } k \to \infty$$

which establishes the claim. Hence, along a subsequence if necessary, $\{|u_n - u|\}$ is weakly convergent in $L^1(\Omega)$, which in particular implies that

$$\int_\Omega |u_n - u| \, dx \to \int_\Omega \int_{\mathbb{R}^+} |\lambda - u| \, d\delta_{u(x)}(\lambda) \, dx = 0 \qquad \text{as } n \to \infty$$

and concludes the proof.

Appendix

Definition A.1. Let $p \geq 1$ and (X, Σ, μ) be a finite measure space. A sequence $\{u_n\}$ in $L^p(X)$ is said to be *uniformly p-integrable* if for every $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$\sup_n \int_E |u_n|^p \, d\mu < \varepsilon^p$$

for all $E \in \Sigma$ with $\mu(E) < \delta_\varepsilon$. A sequence that is uniformly 1-integrable is simply called *uniformly integrable*.

Proposition A.2. *Let $p \geq 1$ and (X, Σ, μ) be a finite measure space. If $\{u_n\}$ is a bounded sequence in $L^p(X)$ then $\{u_n\}$ is uniformly p-integrable if, and only if, for every $\varepsilon > 0$ there exists $M_\varepsilon > 0$ such that*

$$\sup_n \int_{\{|u_n| \geq M_\varepsilon\}} |u_n|^p \, dx < \varepsilon^p.$$

The proof of Proposition A.2 for $p = 1$ can be found in [12, Chapter 19.5]. For $p > 1$ it extends easily by noting that $\{u_n\}$ is uniformly p-integrable if and only if $\{|u_n|^p\}$ is uniformly integrable.

Theorem A.3 (Vitali Convergence Theorem [4]). *Let $p \geq 1$ and (X, Σ, μ) be a finite measure space. Consider a sequence $\{u_n\}$ in $L^p(X)$ and a measurable function u. Then $\{u_n\}$ converges to u in $L^p(X)$ if, and only if, $\{u_n\}$ is uniformly p-integrable and converges to u in measure.*

Theorem A.4 (Dunford-Pettis Theorem [12]). *Let (X, Σ, μ) be a finite measure space. If $\{u_n\}$ is a bounded sequence in $L^1(X)$, then $\{u_n\}$ is relatively weakly compact if, and only if, $\{u_n\}$ is uniformly integrable.*

Proposition A.5. *Let $p \geq 1$ and Ω be a bounded open subset of $\mathbb{R}^d$, $d \in \mathbb{N}$. Consider a sequence $\{u_n\}$ of functions $u_n : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$, and assume that $\{u_n\}$ is bounded in $L^p(\Omega)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function satisfying*

$$\frac{|f(u)|}{1 + |u|^p} \to 0 \qquad \text{as } |u| \to \infty.$$

Then, the sequence $\{f(u_n)\}$ is relatively weakly compact in $L^1(\Omega)$.

References

[1] J.J. Alibert and G. Bouchitté, Non-uniform integrability and generalized Young measures *J. Convex Anal.* **4** (1997), 129-147.

[2] N. J. Alves and J. Paulos. A mode of convergence arising in diffusive relaxation. *Q. J. Math.*, 75(1):143–159, 2024.

[3] J. M. Ball. A version of the fundamental theorem for Young measures. In *PDEs and continuum models of phase transitions (Nice, 1988)*, volume 344 of *Lecture Notes in Phys.*, pages 207–215. Springer, Berlin, 1989.

[4] R. G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1995.

[5] M. J. Cáceres, J. A. Carrillo, and J. Dolbeault. Nonlinear stability in L^p for a confined system of charged particles. *SIAM J. Math. Anal.*, 34(2):478–494, 2002.

[6] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2008.

[7] C. Dafermos, The second law of thermodynamics and stability. *Arch. Rational Mech. Anal.* **70** (1979), 167–179.

[8] S. Demoulini, D. M. A. Stuart, and A. E. Tzavaras. Weak-strong uniqueness of dissipative measure-valued solutions for polyconvex elastodynamics. *Arch. Ration. Mech. Anal.*, 205(3):927–961, 2012.

[9] R. DiPerna and A. J. Majda, Oscillations and concentrations in weak solutions of the incompressible fluid equations, *Comm. Math. Phys.* **108** (1987), 667-689.

[10] G. L. Gilardoni. On Pinsker's and Vajda's type inequalities for Csiszár's f-divergences. *IEEE Transactions on Information Theory*, 56(11):5377–5386, 2010.

[11] P. Pedregal. *Parametrized measures and variational principles*, volume 30 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1997.

[12] H. Royden and P. M. Fitzpatrick. *Real analysis*, fourth edition. China Machine Press, 2010.

[13] A. Unterreiter, A. Arnold, P. Markowich, and G. Toscani. On generalized Csiszár-Kullback inequalities. *Monatsh. Math.*, 131(3):235–253, 2000.

(Nuno J. Alves) University of Vienna, Faculty of Mathematics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.

Email address: nuno.januario.alves@univie.ac.at

(Jakub Skrzeczkowski) University of Oxford, Mathematical Institute, Woodstock Road, Oxford, OX2 6GG, United Kingdom.

Email address: jakub.skrzeczkowski@maths.ox.ac.uk

(Athanasios E. Tzavaras) King Abdullah University of Science and Technology, CEMSE Division, Thuwal 23955-6900, Saudi Arabia.

Email address: athanasios.tzavaras@kaust.edu.sa