

ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

ERIC A. CARLEN

Duality and Stability for Functional Inequalities

Tome XXVI, n° 2 (2017), p. 319-350.

http://afst.cedram.org/item?id=AFST_2017_6_26_2_319_0

© Université Paul Sabatier, Toulouse, 2017, tous droits réservés.

L'accès aux articles de la revue « Annales de la faculté des sciences de Toulouse Mathématiques » (<http://afst.cedram.org/>), implique l'accord avec les conditions générales d'utilisation (<http://afst.cedram.org/legal/>). Toute reproduction en tout ou partie de cet article sous quelque forme que ce soit pour tout usage autre que l'utilisation à fin strictement personnelle du copiste est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

cedram

Article mis en ligne dans le cadre du
Centre de diffusion des revues académiques de mathématiques
<http://www.cedram.org/>

Duality and Stability for Functional Inequalities

Dedicated to Dominique Bakry on the occasion of his 60th birthday

ERIC A. CARLEN ⁽¹⁾

ABSTRACT. — We develop a general framework for using duality to “transfer” stability results for a functional inequality to its dual inequality. As an application, we prove a stability bound for the Hardy–Littlewood–Sobolev inequality, which is related by duality, and the results proved here, to a stability inequality for the Sobolev inequality proved by Bianchi and Egnell, and extended by Chen, Frank and Weth. We also discuss how the results proved here can be combined with the proof of functional inequalities by means of flows to prove stability bounds with computable constants.

RÉSUMÉ. — Nous développons un cadre général pour l’utilisation d’une dualité permettant de “transférer” des résultats de stabilité pour une inégalité fonctionnelle à son inégalité duale. Comme application, nous donnons un résultat de stabilité pour l’inégalité de Hardy–Littlewood–Sobolev qui est lié, par la dualité et les résultats prouvés ici, à une inégalité de stabilité pour l’inégalité de Sobolev prouvée par Bianchi et Egnell, et prolongée par Chen, Frank et Weth. Nous discutons également comment les résultats donnés ici peuvent être combinés à la preuve d’inégalités fonctionnelles utilisant des flots, afin de démontrer les limites de stabilité avec des constantes calculables.

1. Introduction

Consider two functionals \mathcal{E} and \mathcal{F} on some normed real linear space X with values in $(-\infty, \infty]$ that are related by a functional inequality of the form

$$\mathcal{E}(x) \leq \mathcal{F}(x) \quad \text{for all } x \in X. \tag{1.1}$$

Keywords: uniform convexity, entropy.

Math. classification: 81V99, 82B10, 94A17.

⁽¹⁾ Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road Piscataway NJ 08854-8019, USA — carlen@math.rutgers.edu

Work partially supported by U.S. National Science Foundation grant DMS 1501007.

We may regard any complex linear space as a real linear space by restriction of the linear structure. The functional inequality $\mathcal{E} \leq \mathcal{F}$ on X is *sharp* if $\mathcal{E}(x) \leq \mathcal{F}(x)$ for all $x \in X$ and if for all $\lambda < 1$, there exist $x \in X$ such that $\mathcal{E}(x) > \lambda \mathcal{F}(x)$. The subset X_0 defined by

$$X_0 = \{x \in X : \mathcal{E}(x) = \mathcal{F}(x) \text{ and } \mathcal{F}(x) < \infty\} \quad (1.2)$$

is called the *set of optimizers* of the inequality. When $X_0 \neq \emptyset$, the inequality is *optimal*. An optimal functional inequality is necessarily sharp, but not *vice-versa*.

Suppose that the inequality (1.1) is optimal, and suppose that $\{x_n\}$ is a sequence in X such that

$$\lim_{n \rightarrow \infty} (\mathcal{F}(x_n) - \mathcal{E}(x_n)) = 0. \quad (1.3)$$

We are interested in conditions under which (1.3) implies that

$$\lim_{n \rightarrow \infty} d(x_n, X_0) = 0, \quad (1.4)$$

where d is some metric on X which may be different from the norm metric. We are also interested in how the rates of convergence in (1.3) and (1.4) may be related. The following class of convex functions on \mathbb{R}_+ is pertinent to discussion of such rates of convergence.

DEFINITION 1.1 (Rate function). — *A rate function is a strictly convex $\Phi : [0, \infty) \rightarrow [0, \infty)$ with $\Phi(0) = 0$.*

Remark 1.2. — The term *rate function* has a well established meaning in the context of large deviations problems. Our usage here is not meant to suggest any connection. It is simply motivated by the fact that theorems proved below will relate the rates of convergence in (1.3) and (1.4) in terms of such functions.

Remark 1.3. — Let $\Phi(t)$ be any rate function. Then, like all convex functions from \mathbb{R} to \mathbb{R} , Φ is differentiable almost everywhere, and for all $t \in [0, \infty)$,

$$\Phi(t) = \int_0^t \Phi'(s) \, ds.$$

Since Φ is strictly monotone increasing, $\Phi' > 0$ almost everywhere. Hence, if Φ_1 and Φ_2 are rate functions, $\min\{\Phi'_1, \Phi'_2\} > 0$ almost everywhere, and then

$$\Phi_0 := \int_0^t \min\{\Phi'_1(s), \Phi'_2(s)\} \, ds$$

is a rate function such that $\Phi_0(t) \leq \min\{\Phi_1(t), \Phi_2(t)\}$ for all t . Note that $\min\{\Phi_1(t), \Phi_2(t)\}$ need not be convex.

DEFINITION 1.4 (Stability bounds). — Let $\mathcal{E} \leq \mathcal{F}$ be an optimal functional inequality on a normed linear space X , and let X_0 be the set of optimizers. Let d be a metric on X , not necessarily the metric induced by the norm, and let Φ be a rate function. Then the functional inequality $\mathcal{E} \leq \mathcal{F}$ is (d, Φ) -stable in case

$$\mathcal{F}(x) - \mathcal{E}(x) \geq \Phi(d(x, X_0)) \quad \text{for all } x \in X. \quad (1.5)$$

In the special case that $d(x, y) = \|x - y\|_X$ and $\Phi(t) = \kappa t^2$ for some $\kappa > 0$ and all $t \in [0, \infty)$, so that

$$\mathcal{F}(x) - \mathcal{E}(x) \geq \kappa \inf_{x \in X_0} \|x - z\|_X^2 \quad \text{for all } x \in X, \quad (1.6)$$

we say the the functional inequality $\mathcal{E} \leq \mathcal{F}$ is κ -stable.

Evidently, when the optimal functional inequality is (Φ, d) -stable, the rate of convergence in (1.3) governs the rate of convergence in (1.4).

L^p norms on a measure space $(\Omega, \mathcal{B}, \mu)$ occur frequently in the examples that illustrate our results as well as in their applications. Throughout the exposition, for all $1 \leq p < \infty$, $\|f\|_p$ denotes $(\int_{\Omega} |f|^p d\mu)^{1/p}$, and $\|f\|_{\infty}$ denotes the essential supremum of $|f|$.

Before further developing the abstract theory, we turn to a concrete example, namely the Sobolev inequality. Let $n \geq 3$ and let $0 < \alpha < n/2$. There is a unique domain in $L^2(\mathbb{R}^n)$ containing the smooth, compactly supported functions on which $(-\Delta)^{\alpha/2}$ is self-adjoint, where Δ is the Laplacian on \mathbb{R}^n . See Lieb and Loss [13] for a concrete description of this domain. (For instance, $(-\Delta)^{-\alpha/2}$ may be written in term of the Fourier transform, and also as the composition of a Riesz potential with an integer power of the Laplacian.)

The Sobolev inequality in its optimal form says that for all f in the domain of $(-\Delta)^{\alpha/2}$,

$$\left(\int_{\mathbb{R}^n} |f(\eta)|^{2n/(n-2\alpha)} d\eta \right)^{(n-2\alpha)/n} \leq \mathcal{S}_{n,\alpha} \int_{\mathbb{R}^n} |(-\Delta)^{\alpha/2} f(\eta)|^2 d^n \eta. \quad (1.7)$$

where

$$\mathcal{S}_{n,\alpha} = \frac{\|h\|_{2n/(n-2\alpha)}^2}{\|(-\Delta)^{\alpha/2} h\|_2^2} \quad \text{and} \quad h(\eta) = (1 + |\eta|^2)^{-(n-2\alpha)/2}. \quad (1.8)$$

Evidently there is equality in (1.7) when $f = h$. More generally, taking into account the homogeneity, and the translation and dilation invariance of the inequality, there is equality when f is a multiple of a function of the form $h(a(\eta - \eta_0))$ for some $a > 0$ and some $\eta_0 \in \mathbb{R}^n$. In fact, these are the *only* cases of equality.

In the special case $\alpha = 1$, for which we have $\|(-\Delta)^{1/2}f\|_2^2 = \|\nabla f\|_2^2$, this result was proved by Aubin [1] and Talenti [16]. The general case is a direct consequence of Lieb’s theorem [12] on the optimal Hardy–Littlewood–Sobolev (HLS) inequality, to which (1.7) is equivalent by duality as we recall below. The integrals in $\|h\|_{2n/(n-2\alpha)}^2$ and $\|(-\Delta)^{\alpha/2}h\|_2^2$ are easily evaluated in terms of Γ functions, but the explicit value plays no role in what follows.

To relate this example to the general framework introduced above, let X denote the completion of smooth, compactly supported functions f on \mathbb{R}^n in the norm $\|\cdot\|_X$ defined by

$$\|f\|_X := \|(-\Delta)^{\alpha/2}f\|_2. \quad (1.9)$$

Define functionals \mathcal{E} and \mathcal{F} on X by

$$\mathcal{E}(f) := \|f\|_{2n/(n-2\alpha)} \quad \text{and} \quad \mathcal{F}(f) := \mathcal{S}_{n,\alpha}\|f\|_X^2. \quad (1.10)$$

Then we can rewrite (1.7) in the form $\mathcal{E}(f) \leq \mathcal{F}(f)$. By the characterization of the optimizers quoted above, this inequality is optimal, and the set X_0 of optimizers is given by

$$X_0 = \{zh(a(\eta - \eta_0)) : z \in \mathbb{C}, a > 0, \eta_0 \in \mathbb{R}^n\}. \quad (1.11)$$

A line of work starting with Bianchi and Egnell [3] for the case $\alpha = 1$, has addressed the stability of the Sobolev inequality. The recent paper of Chen, Frank and Weth [9] extends the result of Bianchi and Egnell to the full range of α . Theorem 1 of [9] states that for all $\alpha \in (0, n/2)$, there exists some $\kappa_{BE} > 0$ depending only on n and α such that for all $f \in X$,

$$\mathcal{F}(f) - \mathcal{E}(f) \geq \kappa_{BE} \inf_{z \in X_0} \|f - z\|_X^2. \quad (1.12)$$

In our terminology, this says that the Sobolev inequality is κ_{BE} -stable.

The papers [3] and [9] are separated by 22 years, and in-between a number of authors treated various integer values of α ; see [9] for the history. The breakthrough to general α in [9] turns on an insightful use of the stereographic projection to lift the problem to the sphere, which was also crucial to Lieb’s determination of the sharp form of the HLS inequality, though the use made of this in [9] is somewhat different; it facilitates an eigenvalue computation that has no analog in Lieb’s work.

All of the work discussed in the previous paragraph follows the general strategy initiated by Bianchi and Egnell of proving a “local” stability bound for functions that are sufficiently close to the set of optimizers, and then using a compactness argument to show by contradiction that there are no “almost optimizers” that are far from the set of optimizers, and thus to obtain the globally valid result (1.12). It is on account of this compactness argument

that these proofs do not give any means of explicitly estimating κ_{BE} from below.

Our main concern in this paper is with stability theorems for functional inequalities, such as the ones just discussed, and how they may be proved using *duality arguments* and *flow arguments*.

We begin with duality. As noted above, it has been known for some time that the optimal form of the Sobolev inequality is *dual* to Lieb's optimal Hardy–Littlewood–Sobolev (HLS) inequality; see e.g. [8], as we now recall.

Define $Y = L^{2n/(n+2\alpha)}(\mathbb{R}^n)$. For $g \in Y$, define $\mathcal{E}^*(g) = \|g\|_{2n/(n+2\alpha)}^2$. It is well known, and easy to check, that if we define, for all $f \in L^{2n/(n-2\alpha)}(\mathbb{R}^n)$ and all $g \in L^{2n/(n+2\alpha)}(\mathbb{R}^n)$, a real bilinear form by

$$\langle f, g \rangle = 2\Re \left(\int_{\mathbb{R}^n} f^* g \, dx \right), \tag{1.13}$$

then, using the notation of the previous example,

$$\mathcal{E}^*(g) = \sup_{f \in L^{2n/(n-2\alpha)}(\mathbb{R}^n)} \left\{ \langle f, g \rangle - \|f\|_{2n/(n-2\alpha)}^2 \right\} = \sup_{f \in X} \{ \langle f, g \rangle - \mathcal{E}(f) \}. \tag{1.14}$$

The reason for the factor of 2 in the bilinear form on (1.13) is that it avoids factors of 1/2 in the definitions of \mathcal{E} and \mathcal{E}^* that would otherwise be required and are inconvenient in our context. The calculations justifying this are explained below when we recall facts about the Legendre transform; (1.14) says that \mathcal{E}^* is the Legendre transform of \mathcal{E} . We point out that the second equality in (1.14) is valid since the Sobolev space X , defined in the previous example, is dense in $L^{2n/(n-2\alpha)}(\mathbb{R}^n)$.

Now define a functional \mathcal{F}^* on Y by

$$\mathcal{F}^*(g) = \sup_{f \in X} \{ \langle f, g \rangle - \mathcal{F}(f) \} = \sup_{f \in X} \left\{ \langle f, g \rangle - \mathcal{S}_{n,\alpha} \|(-\Delta)^{\alpha/2} f\|_2^2 \right\}. \tag{1.15}$$

A simple completion of the square now shows that

$$\mathcal{F}^*(g) = \mathcal{S}_{n,\alpha}^{-1} \|(-\Delta)^{-\alpha/2} f\|_2^2.$$

The right hand side may be written in terms of the Green's function for $(-\Delta)^\alpha$; that is

$$\|(-\Delta)^{-\alpha/2} f\|_2^2 = \mathcal{C}_{n,\alpha} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} g(x) \frac{1}{|x-y|^{n-2\alpha}} g(y) \, dx dy \tag{1.16}$$

where an explicit expression for the constant $\mathcal{C}_{n,\alpha}$ may be found in [13]. This formula has the advantage that the right hand side is well-defined for any non-negative measurable function g .

By the Sobolev inequality, for all $f \in X$ and all $g \in Y$,

$$\langle f, g \rangle - \mathcal{F}(f) \leq \langle f, g \rangle - \mathcal{E}(f).$$

Taking the supremum in f over X on both sides, we obtain

$$\mathcal{F}^*(g) \leq \mathcal{E}^*(g) \quad \text{for all } g \in Y. \quad (1.17)$$

This is the well known *order reversal* property of the Legendre transform. Applying the Legendre transform to a functional inequality $\mathcal{E} \leq \mathcal{F}$ yields a new inequality for the Legendre transforms, namely $\mathcal{F}^* \leq \mathcal{E}^*$.

The inequality (1.17) is Lieb's optimal LHS inequality. More explicitly, it states that

$$\|(-\Delta)^{-\alpha/2}g\|_2^2 \leq \mathcal{S}_{n,\alpha}\|g\|_{2n/(n+2\alpha)}^2. \quad (1.18)$$

The fact that the same constant $\mathcal{S}_{n,\alpha}$ shows up in (1.18) and in (1.7) is no accident; it is essentially a consequence of the involutive nature of the Legendre transform: Taking the dual of $\mathcal{E} \leq \mathcal{F}$, we obtain $\mathcal{F}^* \leq \mathcal{E}^*$. If for some $\lambda < 1$, we could improve this to $\mathcal{F}^* \leq \lambda\mathcal{E}^*$, then taking the dual again we would obtain $\mathcal{E}^{**} \leq \lambda\mathcal{F}^{**}$. Under mild regularity conditions specified in the Fenchel–Moreau Theorem discussed below, the Legendre transform is involutive so that $\mathcal{E}^{**} = \mathcal{E}$ and $\mathcal{F}^{**} = \mathcal{F}$. Thus we would have $\mathcal{E} \leq \lambda\mathcal{F}$, but this is impossible when $\mathcal{E} \leq \mathcal{F}$ is optimal. In other words, when the Legendre transform acts as an involution, it takes sharp inequalities to sharp inequalities. In particular, (1.7) and (1.18) are equivalent inequalities, which for general α were first found by Lieb [12] in the form (1.18).

Moreover, the duality relating (1.7) and (1.18), as well as other dual pairs of optimal inequalities, goes much further. As we show here, when $\mathcal{E} \leq \mathcal{F}$ is not only optimal, but stable, then under suitable quantitative convexity assumptions on \mathcal{E} , one obtains both optimality and a stability result for the dual inequality $\mathcal{F}^* \leq \mathcal{E}^*$. As a particular consequence, we shall obtain the following stability theorem for the HLS inequality:

THEOREM 1.5. — *For all $\alpha \in (0, n/2)$, there is a constant $\kappa_{BE}^* > 0$ depending only on n and α such that for all $g \in L_{2n/(n-2\alpha)}(\mathbb{R}^n)$,*

$$\mathcal{S}_{n,\alpha}\|g\|_{2n/(n+2\alpha)}^2 - \|(-\Delta)^{-\alpha/2}g\|_2^2 \geq \kappa_{BE}^* \inf_{y \in Y_0} \|g - y\|_{2n/(n+2\alpha)}^2, \quad (1.19)$$

and the constant $\kappa_{BE}^ > 0$ is explicitly computable in terms of the constant κ_{BE} in (1.12).*

This result is new. If one attempts to prove it directly using the method of Bianchi and Egnel, the non-locality of $(-\Delta)^{-\alpha/2}$ complicates matters. Perhaps these complications could be managed, but as shown here, there is no need for this.

Notice that just as in (1.12), the norm in the remainder term is the *stronger* of the two norms involved in the functional inequality.

Duality methods for functional inequalities are useful because one or the other of the dual forms might be easier to prove directly. Indeed, except for the $\alpha = 1$ cases, Lieb’s optimal form of the HLS inequality (1.18) predates the optimal Sobolev inequality (1.7). Lieb’s proof of the HLS inequality makes important use of rearrangement inequalities. In particular, by the Riesz Rearrangement Inequality and (1.16), $\|(-\Delta)^{-\alpha/2} f\|_2^2$ does not decrease when f is replaced by its spherical decreasing rearrangement. This is also true of $\|(-\Delta)^{\alpha/2} f\|_2^2$ for $\alpha \leq 1$, but only for $\alpha \leq 1$. A pair of functional inequalities may be equivalent via the Legendre transform, but a wider range of tools and techniques may be applicable to one or the other of them.

As we shall see, this is also true in proving stability inequalities. One tool that is useful for this purpose is *monotone flows*. Let $\mathcal{E} \leq \mathcal{F}$ be an optimal functional inequality on X . Let $\{\Phi_t\}$ be a semigroup of transformations on X . Suppose that for all x in X , $\lim_{t \rightarrow \infty} \Phi_t(x)$ exists and belongs to X_0 . Suppose also that for all $t > s \geq 0$,

$$\mathcal{F}(\Phi_t(x)) - \mathcal{E}(\Phi_t(x)) \leq \mathcal{F}(\Phi_s(x)) - \mathcal{E}(\Phi_s(x)).$$

Then the flow associated to the semigroup $\{\Phi_t\}$ is *monotone for the functional inequality* $\mathcal{E} \leq \mathcal{F}$. Under certain additional conditions, monotone flows may be used to prove stability inequalities. Examples of this may be found in [6]. Here we shall combine such flow methods with duality to prove variants of the stability bounds for the Sobolev and HLS inequalities with computable constants.

The transfer of a stability bound for $\mathcal{E} \leq \mathcal{F}$ to its dual $\mathcal{F}^* \leq \mathcal{E}^*$ is not so simple as the almost automatic order reversal property of the Legendre transform that has been described above. Indeed, one way to write a stability bound for $\mathcal{E} \leq \mathcal{F}$ is $\mathcal{E} + \mathcal{R} \leq F$ where \mathcal{R} is a “remainder term”. For example, in (1.12) we would have $\mathcal{R}(f) = \kappa_{BE} \inf_{z \in X_0} \|f - z\|_X^2$. The Legendre transform of a sum is the *infimal convolution* of the Legendre transforms of the summands; see, e.g., [14]. That is, going back to the general setting of functional on a dual pair $(X, Y, \langle \cdot, \cdot \rangle)$,

$$(\mathcal{E} + \mathcal{R})^*(y) = \inf_{z \in Y} \{\mathcal{E}^*(y - z) + \mathcal{R}^*(z)\}.$$

Doing this for (1.12) does not yield anything at all like (1.19), or even anything that is likely to be at all useful. The passage from (1.12) to (1.19) is more complicated, and involves an interplay between the stability bound (1.12) an *quantitative convexity estimates* for the functional \mathcal{E} , as will be explained below. This interplay between stability bounds and quantitative convexity bounds was crucial in the paper [7] on stability of optimal

potentials for Schrödinger operators. In [7] an “adding and subtracting” argument is used to combine a Bianchi–Egnell like stability bound for a family of Gagliardo–Nirenberg inequalities with a stability bound for Hölder’s inequality to prove a stability bound for an inequality relating the fundamental eigenvalue $\lambda(V)$ of a Schrödinger operator $-\Delta + V(x)$ with the L^q norm of V for appropriate q . The eigenvalue inequality is the Legendre transform dual of the Gagliardo–Nirenberg inequality. The present work was partly motivated by the desire to understand this transfer of stability within a general framework of duality.

Acknowledgements

I would like to thank Dominique Bakry for the invitation to present a three lecture mini-course on functional inequalities and evolution equations in Toulouse in April 2014. The main ideas (and some other examples) were presented during this course. Dominique Bakry, Szymon Pezdat and Bogusław Zegarliński then arranged for me to give a second course on the topic at the CNRS-PAN Mathematics Summer Institute, June 28–July 4, 2015 in Krakow. The applications to quantitative stability for the Sobolev inequality were carried out with these lectures in mind. This work was partially done while I was visiting the I.M.A. in Minneapolis in the spring of 2015. I thank the I.M.A. for providing a stimulating and congenial environment for this work. Finally, I would like to thank Rupert Frank and Elliott Lieb for many discussions about stability, and Rupert Frank and two anonymous referees for a careful reading of the first version of this work and for making valuable comments.

2. Duality

DEFINITION 2.1. — *Let X and Y be two normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively. Let $\langle \cdot, \cdot \rangle$ be a separately continuous bilinear form on $X \times Y$ such that for all non-zero $x \in X$, there exists $y \in Y$ with $\langle x, y \rangle > 0$ and likewise, for all non-zero $y \in Y$, there exists $x \in X$ with $\langle x, y \rangle > 0$. Then the triple $(X, Y, \langle \cdot, \cdot \rangle)$ is a dual pair of normed linear spaces.*

When $(X, Y, \langle \cdot, \cdot \rangle)$ is a dual pair, the map $x \mapsto \langle x, \cdot \rangle$ is injective into Y^* , the topological dual to Y . Likewise, $y \mapsto \langle \cdot, y \rangle$ is injective into X^* , the topological dual to X . Throughout this section X and Y together with the bilinear form $\langle \cdot, \cdot \rangle$ is a fixed dual pair of normed linear spaces. Most of the material well-known, and can be found in Rockafellar [14], for example.

However, we include some proofs since they are very short and relevant to what follows.

A convex function \mathcal{E} from X to $(-\infty, \infty]$ is *proper* in case it is not identically $+\infty$. The *Legendre transform* of a convex functional $\mathcal{E} : X \rightarrow (-\infty, \infty]$ is the function \mathcal{E}^* on Y defined by

$$\mathcal{E}^*(y) = \sup_{x \in X} \{\langle x, y \rangle - \mathcal{E}(x)\}. \quad (2.1)$$

The function \mathcal{E}^* , being a supremum of a family of continuous linear functions is lower semicontinuous and convex.

DEFINITION 2.2 (Closed convex function). — *Let X be a normed linear space. A function \mathcal{F} on X with values in $(-\infty, \infty]$ is a closed convex function in case it is lower semicontinuous.*

The terminology is motivated by the fact that a convex function \mathcal{F} is lower semicontinuous if and only if its *epigraph* is closed; see [14].

The *Fenchel–Moreau Theorem* states that if \mathcal{E} is a proper closed convex function on X , and $(X, Y, \langle \cdot, \cdot \rangle)$ is the canonical dual pair associated to X , then $\mathcal{E}^{**} = \mathcal{E}$ where \mathcal{E}^{**} denotes the Legendre transform of \mathcal{E}^* .

It is a direct consequence of the definition that $\langle x, y \rangle \leq \mathcal{E}(x) + \mathcal{E}^*(y)$ for all $x \in X$ and all $y \in Y$. This is *Young’s inequality*. The cases of equality in this inequality are crucially important for what follows. This brings us to the notion of the subgradient of a convex function:

DEFINITION 2.3 (subgradient). — *Let $(X, Y, \langle \cdot, \cdot \rangle)$ be a dual pair, an let \mathcal{E} be a convex function on X . The subgradient of \mathcal{E} at $x_0 \in X$ is the subset $\partial\mathcal{E}(x_0)$ of Y consisting of all y such that*

$$\mathcal{E}(x_1) \geq \mathcal{E}(x_0) + \langle x_1 - x_0, y \rangle \quad \text{for all } x_1 \in X. \quad (2.2)$$

For $W \subset X$, er define $\partial\mathcal{E}(W) = \cup_{x \in W} \partial\mathcal{E}(x)$.

That is, $y \in \partial\mathcal{E}(x_0)$ if and only if the hyperplane $(x, \mathcal{E}(x_0) + \langle x - x_0, y \rangle)$ lies below the graph of $x \mapsto \mathcal{E}(x)$, and intersects it at $(x_0, \mathcal{E}(x_0))$, so that the hyperplane $(x, \mathcal{E}(x_0) + \langle x - x_0, y \rangle)$ is a supporting plane to the graph of $x \mapsto \mathcal{E}(x)$ at x_0 . At this level of generality, $\partial\mathcal{E}(x)$ may be empty. However, it is easy to see that if \mathcal{E} is differentiable at x , with derivative $\nabla\mathcal{E}(x) \in Y$, then $\partial\mathcal{E}(x) = \{\nabla\mathcal{E}(x)\}$.

The domain of a proper, closed convex functional \mathcal{F} on X is the set $\{x \in X \mid \mathcal{F}(x) < \infty\}$. As is well known, for a proper closed convex function \mathcal{F} , $\partial\mathcal{F}(x)$ is never empty at any x in the domain of \mathcal{F} ; see [14]. In particular, when we have an optimal functional inequality $\mathcal{E} \leq \mathcal{F}$ relating two proper,

closed convex functionals, the set of optimizers is, by the definition (1.1), included in the domains of both \mathcal{E} and \mathcal{F} .

Example 2.4 (Squared L^p norms). — Let $(\Omega, \mathcal{B}, \mu)$ be a measure space. Let $1 < p < \infty$, and let $p' = p/(p-1)$. Let $X = L^p(\Omega, \mathcal{B}, \mu)$ regarded as a real Banach space (so that for nonzero $f \in L^p(\Omega, \mathcal{B}, \mu)$, f and if are linearly independent). Let $Y = L^{p'}(\Omega, \mathcal{B}, \mu)$ regarded as a real Banach space, and define $\langle \cdot, \cdot \rangle$ on $X \times Y$ by

$$\langle f, g \rangle = 2\Re \left(\int_{\Omega} f^* g \, d\mu \right).$$

Evidently with this bilinear form, X and Y are a dual pair. Consider the functional \mathcal{E} on X defined by $\mathcal{E}(f) = \|f\|_p^2$. Since $t \mapsto t^2$ is convex and increasing on $[0, \infty)$, it is a consequence of the Minkowski inequality that \mathcal{E} is convex.

By Hölder's inequality, for all non-zero $f \in X$ and $g \in Y$,

$$\begin{aligned} \langle f, g \rangle - \mathcal{E}(f) &= 2\Re \left(\int_{\Omega} f^* g \, d\mu \right) - \|f\|_p^2 \\ &\leq 2\|g\|_{p'}\|f\|_p - \|f\|_p^2 = \|g\|_{p'}^2 - (\|g\|_{p'} - \|f\|_p)^2. \end{aligned}$$

By the well-known cases of equality in Hölder's inequality, there is equality if and only if g is a multiple of $|f|^{p-1} \operatorname{sgn} f$. We conclude that $\langle f, g \rangle - \mathcal{E}(f) \geq \|g\|_{p'}^2$ with equality if and only if

$$g = \|f\|_p^{2-p} |f|^{p-1} \operatorname{sgn}(f). \quad (2.3)$$

This proves that $\mathcal{E}^*(g) = \|g\|_{p'}^2$, and that $\partial\mathcal{E}(f) = \{\|f\|_p^{2-p} |f|^{p-1} \operatorname{sgn}(f)\}$ for non-zero f . (It is evident that $\partial\mathcal{E}(0) = \{0\}$). In this example \mathcal{E} (together with \mathcal{E}^* by symmetry) is differentiable, and $\partial\mathcal{E}(f) = \{\nabla\mathcal{E}(f)\}$ where $\nabla\mathcal{E}(f) = \|f\|_p^{2-p} |f|^{p-1} \operatorname{sgn}(f)$.

Finally, note that $\nabla\mathcal{E}$ is homogeneous of degree 1, and is a norm preserving map from X to Y . That is,

$$\|\nabla\mathcal{E}(f)\|_{p'} = \|f\|_p. \quad (2.4)$$

LEMMA 2.5. — *Let \mathcal{E} and \mathcal{F} be proper lower semicontinuous convex functionals on X and suppose that the inequality $\mathcal{E} \leq \mathcal{F}$ is valid and optimal. Let X_0 be the set of optimizers for it. Then for all $x_0 \in X_0$,*

$$\partial\mathcal{E}(x_0) \subset \partial\mathcal{F}(x_0). \quad (2.5)$$

Proof. — If $y_0 \in \partial\mathcal{E}(x_0)$, then for all $x_1 \in X$, $\mathcal{E}(x_1) \geq \mathcal{E}(x_0) + \langle x_1 - x_0, y_0 \rangle$. Now using $\mathcal{F}(x_1) \geq \mathcal{E}(x_1)$ and $\mathcal{F}(x_0) = \mathcal{E}(x_0)$, we have

$$\mathcal{F}(x_1) \geq \mathcal{E}(x_1) \geq \mathcal{E}(x_0) + \langle x_1 - x_0, y_0 \rangle = \mathcal{F}(x_0) + \langle x_1 - x_0, y_0 \rangle,$$

and this proves $y_0 \in \partial\mathcal{F}(x_0)$. \square

Example 2.6. — Let $X = Y = \mathbb{R}$, and let $\langle a, x \rangle = xy$. Let $\mathcal{E}(x) = |x|$ and $\mathcal{F}(x) = 2|x|$. Then $\mathcal{E} \leq \mathcal{F}$ and $X_0 = \{0\}$, and we have $\partial\mathcal{E}(0) = [-1, 1]$ and $\partial\mathcal{F}(0) = [-2, 2]$, showing that the containment in Lemma 2.5 can be strict.

LEMMA 2.7 (Young's inequality). — *Let $(X, Y, \langle \cdot, \cdot \rangle)$ be a dual pair. Let \mathcal{E} be a closed convex function on X . Then for all $x \in X$ and all $y \in Y$,*

$$\langle x, y \rangle \leq \mathcal{E}(x) + \mathcal{E}^*(y). \quad (2.6)$$

Moreover:

- (1) *Whenever there is equality in (2.6), $y \in \partial\mathcal{E}(x)$ and $x \in \partial\mathcal{E}^*(y)$.*
- (2) *Whenever $y \in \partial\mathcal{E}(x)$, there is equality in (2.6), and consequently*

$$y \in \partial\mathcal{E}(x) \Rightarrow x \in \partial\mathcal{E}^*(y), \quad (2.7)$$

*and in case $\mathcal{E}^{**} = \mathcal{E}$, the reverse implication is also valid.*

Proof. — We need only explain the cases of equality. If there is equality in (2.6), then for all $y_1 \in Y$,

$$\langle x, y_1 \rangle - \langle x, y \rangle \leq (\mathcal{E}(x) + \mathcal{E}^*(y_1)) - (\mathcal{E}(x) + \mathcal{E}^*(y)) = \mathcal{E}^*(y_1) - \mathcal{E}^*(y),$$

which is the same as $\mathcal{E}^*(y_1) \geq \mathcal{E}^*(y) + \langle x, y_1 - y \rangle$, which means that $y \in \partial\mathcal{E}(x)$. An analogous argument shows that $x \in \partial\mathcal{E}^*(y)$ when there is equality in (2.6).

Conversely, suppose that $y \in \partial\mathcal{E}(x)$. Then for all $x_1 \in X$, $\mathcal{E}(x_1) \geq \mathcal{E}(x) + \langle x_1 - x, y \rangle$, and rearranging terms we obtain

$$\langle x, y \rangle \geq \mathcal{E}(x) + \langle x_1, y \rangle - \mathcal{E}(x_1).$$

Taking the supremum over x_1 , we obtain $\langle x, y \rangle \geq \mathcal{E}(x) + \mathcal{E}^*(y)$, and equality must hold in (2.6). An analogous argument shows that argument when $x \in \partial\mathcal{E}^*(y)$, then $\langle x, y \rangle \leq \mathcal{E}^{**}(x) + \mathcal{E}^*(y)$, which is the same as (2.6) when $\mathcal{E}^{**} = \mathcal{E}$. □

2.1. Subgradients and optimizers

It is well-known that the Legendre transform reverses order, so that if $\mathcal{E}(x) \leq \mathcal{F}(x)$ for all $x \in X$, then $\mathcal{F}^*(y) \leq \mathcal{E}^*(y)$ for all $y \in Y$. As mentioned in the introduction, one way to see this is to note that $\mathcal{E}(x) \leq \mathcal{F}(x)$ implies that $\langle x, y \rangle - \mathcal{F}(x) \leq \langle x, y \rangle - \mathcal{E}(x)$. Taking the supremum over x , we obtain $\mathcal{F}^*(y) \leq \mathcal{E}^*(y)$.

By the Fenchel–Moreau Theorem, if \mathcal{E} and \mathcal{F} be proper, closed convex functions on X , the Legendre transform is involutive, and in this case Legendre transforming twice yields:

$$\mathcal{E}(x) \leq \mathcal{F}(x) \text{ for all } x \in X \iff \mathcal{F}^*(y) \leq \mathcal{E}^*(y) \text{ for all } y \in Y, \quad (2.8)$$

which expresses the equivalence of the primal inequality $\mathcal{E} \leq \mathcal{F}$ and its dual inequality $\mathcal{F}^* \leq \mathcal{E}^*$.

However, something more is true: Under the condition that \mathcal{E} and \mathcal{F} are proper, closed convex functions on X , there is a general relation between the optimizers (if any) of $\mathcal{E} \leq \mathcal{F}$ and the optimizers (if any) of $\mathcal{F}^* \leq \mathcal{E}^*$.

THEOREM 2.8 (Duality for functional inequalities). — *Let $(X, Y, \langle \cdot, \cdot \rangle)$ be a dual pair. Let \mathcal{E} and \mathcal{F} be proper, closed convex functions on X so that $\mathcal{E}^{**} = \mathcal{E}$ and $\mathcal{F}^{**} = \mathcal{F}$. Then for all $y \in \partial\mathcal{F}(x)$,*

$$\mathcal{F}(x) - \mathcal{E}(x) \leq \mathcal{E}^*(y) - \mathcal{F}^*(y), \quad (2.9)$$

and for all $x \in \partial\mathcal{E}^*(y)$,

$$\mathcal{F}(x) - \mathcal{E}(x) \geq \mathcal{E}^*(y) - \mathcal{F}^*(y). \quad (2.10)$$

Suppose further that $\mathcal{E} \leq \mathcal{F}$, or equivalently, $\mathcal{F}^* \leq \mathcal{E}^*$. If X_0 is the set of optimizers of the inequality $\mathcal{E} \leq \mathcal{F}$ and Y_0 is the set of optimizers of the inequality $\mathcal{F}^* \leq \mathcal{E}^*$, then

$$Y_0 = \partial\mathcal{E}(X_0) \quad \text{and} \quad X_0 = \partial\mathcal{F}^*(Y_0). \quad (2.11)$$

Proof. — Let $y \in \partial\mathcal{F}(x)$. By Young’s inequality, and then the cases of equality in it,

$$\mathcal{F}(x) + \mathcal{F}^*(y) = \langle x, y \rangle \leq \mathcal{E}(x) + \mathcal{E}^*(y).$$

Rearranging terms we obtain (2.9). Likewise, let $x \in \partial\mathcal{E}^*(y)$. Then

$$\mathcal{F}(x) + \mathcal{F}^*(y) \geq \langle x, y \rangle = \mathcal{E}(x) + \mathcal{E}^*(y).$$

and rearranging terms, yields (2.10).

Now suppose that $X_0 \neq \emptyset$, and $x_0 \in X_0$. By the remark preceding Example 2.4, $\partial\mathcal{E}(x_0) \neq \emptyset$, and by Lemma 2.7, for any $y_0 \in \partial\mathcal{E}(x_0)$, $x_0 \in \partial\mathcal{E}^*(y_0)$, so that (2.10) implies that $0 \geq \mathcal{E}^*(y_0) - \mathcal{F}^*(y_0)$, so that $y_0 \in Y_0$. This shows that $\partial\mathcal{E}(X_0) \subset Y_0$.

Likewise, suppose that $Y_0 \neq \emptyset$, and $y_0 \in Y_0$. Then, as above, $\partial\mathcal{F}^*(y_0) \neq \emptyset$, and for any $x_0 \in \partial\mathcal{F}^*(y_0)$, $y_0 \in \partial\mathcal{F}(x_0)$, so that (2.9) implies that $0 \geq \mathcal{F}(x_0) - \mathcal{E}(x_0)$, so that $x_0 \in X_0$. This shows that $\partial\mathcal{F}^*(Y_0) \subset X_0$.

We next show that $Y_0 \subset \partial\mathcal{E}(X_0)$. Suppose that $y_0 \in Y_0$. By what we have proved just above, if $x_0 \in \partial\mathcal{F}^*(y_0)$, then $x_0 \in X_0$. Furthermore, by the cases of equality in Young’s inequality,

$$\langle x_0, y_0 \rangle = \mathcal{F}(x_0) + \mathcal{F}^*(y_0). \quad (2.12)$$

Then since $\mathcal{F}(x_0) = \mathcal{E}(x_0)$ and $\mathcal{F}^*(y_0) = \mathcal{E}^*(y_0)$, we also have $\langle x_0, y_0 \rangle = \mathcal{E}(x_0) + \mathcal{E}^*(y_0)$. Once more, by the cases of equality in Young's inequality,

$$y_0 \in \partial\mathcal{E}(x_0) \in \partial\mathcal{E}(X_0).$$

an entirely analogous argument shows that $X_0 \subset \partial\mathcal{F}^*(Y_0)$. □

2.2. Quantitative convexity

DEFINITION 2.9 ((d, Φ) -convexity). — *Let X be a Banach space, and let \mathcal{E} be a proper, closed convex function on X . Let Φ be a rate function, and let d be a metric on X . Let $U \subset X$ be a convex subset of the domain of \mathcal{E} . Then \mathcal{E} is (d, Φ) -convex on U in case for all $x_1, x_2 \in U$ and all $y \in \partial\mathcal{E}(x_1)$,*

$$\mathcal{E}(x_2) \geq \mathcal{E}(x_1) + \langle x_2 - x_1, y \rangle + \Phi(d(x_1, x_2)). \quad (2.13)$$

In case d is the metric induced by the norm, and $\Phi(t) = \lambda t^2$ for all t , so that for all $x_1, x_2 \in U$

$$\mathcal{E}(x_2) \geq \mathcal{E}(x_1) + \langle x_2 - x_1, y \rangle + \lambda \|x_1 - x_2\|_X^2, \quad (2.14)$$

we say that \mathcal{E} is λ -convex on U .

To see that this is a strengthened form of convexity, let $z_1, z_2 \in U$ and let $\alpha \in (0, 1)$. Define $x_1 = (1 - \alpha)z_1 + \alpha z_2$, and let $y \in \partial(\mathcal{E})(x_1)$. Applying (2.13) with this choice of x_1 , and then $x_1 = z_2$ and z_1 in turn, we obtain

$$\begin{aligned} \alpha\mathcal{E}(z_2) &\geq \mathcal{E}(x_1) + \alpha\langle z_2 - x_1, y \rangle + \alpha\Phi(d(x_1, z_2)) \\ (1 - \alpha)\mathcal{E}(z_1) &\geq \mathcal{E}(x_1) + (1 - \alpha)\langle z_1 - x_1, y \rangle + (1 - \alpha)\Phi(d(x_1, z_1)). \end{aligned}$$

Adding these inequalities we obtain

$$\begin{aligned} (1 - \alpha)\mathcal{E}(z_1) + \alpha\mathcal{E}(z_2) &\geq \mathcal{E}((1 - \alpha)z_1 + \alpha z_2) + \Phi((1 - \alpha)d(z_1, x_1) + \alpha d(x_1, z_2)) \\ &\geq \mathcal{E}((1 - \alpha)z_1 + \alpha z_2) + \Phi(\min\{\alpha, 1 - \alpha\}d(z_1, z_2)). \end{aligned} \quad (2.15)$$

In the case d is the metric induced by the norm on X and $\Phi(t) = \lambda t^2$ we can avoid the final inequality in (2.15) and compute

$$\begin{aligned} \Phi((1 - \alpha)d(z_1, x_1) + \alpha d(x_1, z_2)) &= \lambda((1 - \alpha)\|z_1 - x_1\|_X + \alpha\|z_2 - x_1\|_X)^2 \\ &= \lambda\alpha(1 - \alpha)\|z_1 - z_2\|_X^2 \end{aligned}$$

Hence, when \mathcal{E} is λ -convex on U , then for all $x_1, x_2 \in U$ and all $\alpha \in (0, 1)$,

$$\alpha\mathcal{E}(x_2) + (1 - \alpha)\mathcal{E}(x_1) \geq \mathcal{E}(\alpha x_2 + (1 - \alpha)x_1) + \lambda\alpha(1 - \alpha)\|x_2 - x_1\|_X^2. \quad (2.16)$$

This is the standard definition of λ -convexity. To see that the two formulations are equivalent, rewrite (2.16) as

$$\alpha(\mathcal{E}(x_2) - \mathcal{E}(x_1)) \geq \mathcal{E}(x_1 + \alpha(x_2 - x_1)) - \mathcal{E}(x_1) + \lambda\alpha(1 - \alpha)\|x_2 - x_1\|_X^2.$$

For $y \in \partial\mathcal{E}(x_1)$, $\mathcal{E}(x_1 + \alpha(x_2 - x_1)) - \mathcal{E}(x_1) \geq \alpha\langle x_2 - x_1, y \rangle$, and we obtain

$$\alpha(\mathcal{E}(x_2) - \mathcal{E}(x_1)) \geq \alpha\langle x_2 - x_1, y \rangle + \lambda\alpha(1 - \alpha)\|x_2 - x_1\|_X^2.$$

Dividing by α , and then taking the limit $\alpha \rightarrow 0$, we obtain

$$y \in \partial\mathcal{E}(x_1) \quad \Rightarrow \quad \mathcal{E}(x_2) \geq \mathcal{E}(x_1) + \langle x_2 - x_1, y \rangle + \lambda\|x_2 - x_1\|_X^2. \quad (2.17)$$

THEOREM 2.10 ((d, Φ) -convexity and the strong Young's inequality). — *Let $(X, Y, \langle \cdot, \cdot \rangle)$ be a dual pair. Let \mathcal{E} be a closed convex function on X . Let d be a metric on X , and let Φ be a rate function. Suppose that \mathcal{E} is (d, Φ) -convex on a convex set $U \subset X$. Then for all $y \in \partial\mathcal{E}(U)$, $\partial\mathcal{E}^*(y) \cap U$ is a singleton, and defining $\nabla\mathcal{E}^*(y)$ to be the single element in $\partial\mathcal{E}^*(y) \cap U$,*

$$\mathcal{E}(x) + \mathcal{E}^*(y) \geq \langle x, y \rangle + \Phi(d(x, \nabla\mathcal{E}^*(y))) \quad (2.18)$$

is satisfied for all $x \in U$ and all $y \in \partial\mathcal{E}(U)$.

Conversely, if $\partial\mathcal{E}^(y) \cap U$ is a singleton $\{\nabla\mathcal{E}^*(y)\}$ for all $y \in \partial\mathcal{E}(U)$ and (2.18) is valid for all $x \in U$ and all $y \in \partial\mathcal{E}(U)$, then \mathcal{E} is (d, Φ) -convex on U*

Remark 2.11. — We refer to inequality (2.18) as the *strong Young's inequality* for the (d, Φ) -convex function \mathcal{E} on U . Theorem 2.10 says that the validity of this inequality, together with the property that for all $y \in \partial\mathcal{E}(U)$, $\partial\mathcal{E}^*(y) \cap U$ is a singleton, characterizes (d, Φ) -convexity.

Proof of Theorem 2.10. — Let $x_1, x_2 \in U$, and let $y \in \partial\mathcal{E}(x_1)$. By the cases of equality in Young's inequality

$$\mathcal{E}(x_1) - \langle x_1, y \rangle = -\mathcal{E}^*(y)$$

Combining this with (2.13) yields $\mathcal{E}(x_2) + \mathcal{E}^*(y) \geq \langle x_2, y \rangle + \Phi(d(x_2, x_1))$. Since $y \in \partial\mathcal{E}(x_1) \iff x_1 \in \partial\mathcal{E}^*(y)$, we conclude that for all $x \in U$ and $y \in \partial\mathcal{E}(U)$,

$$\mathcal{E}(x) + \mathcal{E}^*(y) \geq \langle x, y \rangle + \sup_{z \in \partial\mathcal{E}^*(y) \cap U} \Phi(d(x, z)). \quad (2.19)$$

By the ordinary Young's inequality, when $y \in \partial\mathcal{E}(x)$, this reduces to

$$0 \geq \sup_{z \in \partial\mathcal{E}^*(y) \cap U} \Phi(d(x, z)).$$

Evidently this implies that $\partial\mathcal{E}^*(y) \cap U$ is a singleton for all $y \in \partial\mathcal{E}(U)$. Thus we may define a function $y \mapsto \nabla\mathcal{E}^*(y)$ on $\partial\mathcal{E}(U)$ by defining $\nabla\mathcal{E}^*(y)$ to be the single element in $\partial\mathcal{E}^*(y) \cap U$. Then (2.19) simplifies to become (2.18)

Conversely, let $x, x_1 \in U$ and $y \in \partial\mathcal{E}(x_1)$. then $x_1 \in \partial\mathcal{E}^*(y) \cap U$, and hence if this set is the singleton $\{\nabla\mathcal{E}^*(y)\}$, $x_1 = \nabla\nabla\mathcal{E}^*(y)$. Thus, when $\partial\mathcal{E}^*(y) \cap U$ is a singleton $\{\nabla\mathcal{E}^*(y)\}$ for all $y \in \partial\mathcal{E}(U)$, and (2.18) is valid for all $x \in U$ and $y \in \mathcal{E}(U)$, then

$$\mathcal{E}(x) + \mathcal{E}^*(y) \geq \langle x, y \rangle + \Phi(d(x, x_1)).$$

Since $\mathcal{E}^*(y) = \langle x_1, y \rangle - \mathcal{E}(x_1)$, this implies

$$\mathcal{E}(x) \geq \mathcal{E}(x_1)\langle x - x_1, y \rangle + \Phi(d(x, x_1)).$$

so that \mathcal{E} is (d, Φ) convex on U . □

The next theorem, which is proved in an appendix, provides an important class of examples:

THEOREM 2.12 (Strong convexity of squared L^p norms). — *Let $1 < p < \infty$ and let $p' = p/(p - 1)$. Let $X = L^p(\Omega, \mathcal{B}, \mu)$ and $Y = L^{p'}(\Omega, \mathcal{B}, \mu)$, For $f \in X$ and $g \in Y$ let $\langle f, g \rangle = 2\Re\left(\int_{\Omega} f^*g \, d\mu\right)$, so that X, Y equipped with $\langle \cdot, \cdot \rangle$ is a dual pair. Define \mathcal{E} on X by $\mathcal{E}(f) = \|f\|_p^2$. Then \mathcal{E} is differentiable, and we have:*

- (1) For $p \in (1, 2]$ and all $f_1, f_2 \in X$,

$$\mathcal{E}(f_2) \geq \mathcal{E}(f_1) + \langle f_2 - f_1, \nabla\mathcal{E}(f_1) \rangle + (p - 1)\|f_2 - f_1\|_p^2, \quad (2.20)$$

and hence for such p , \mathcal{E} is $(p - 1)$ -convex.

- (2) For $p \in (2, \infty)$, \mathcal{E} is not λ -convex for any $\lambda > 0$, however, for all $f_1, f_2 \in X$,

$$\begin{aligned} \mathcal{E}(f_2) \geq \mathcal{E}(f_1) + \langle f_2 - f_1, \nabla\mathcal{E}(f_1) \rangle \\ + \frac{1}{4p} \left(\frac{2}{3}\right)^{p-1} (\|f_1\|_p + \|f_2\|_p)^{2-p} \|f_2 - f_1\|_p^p. \end{aligned} \quad (2.21)$$

It is well-known that some form of strong convexity of a convex function \mathcal{E} implies certain regularity properties of the Legendre transform \mathcal{E}^* . The next lemma is a precise version of this in a context that is relevant here. This is followed by a general result.

THEOREM 2.13 (Continuity of the gradient maps in L^p). — *Let $1 < p < \infty$, and let $p' = p/(p - 1)$. Let $(\Omega, \mathcal{B}, \mu)$ be a measure space, and let $X = L^p(\Omega, \mathcal{B}, \mu)$ and $Y = L^{p'}(\Omega, \mathcal{B}, \mu)$. Let $\langle \cdot, \cdot \rangle$ be the bilinear form on $X \times Y$ given by $\langle f, g \rangle = 2\Re\left(\int_{\Omega} f^*g \, d\mu\right)$. Define the functional \mathcal{E} on X by $\mathcal{E}(f) = \|f\|_X^2$. Then \mathcal{E} and \mathcal{E}^* are differentiable, so that $\nabla\mathcal{E}$ and $\nabla\mathcal{E}^*$ are well defined.*

(1) For $1 < p \leq 2$, and all $g_1, g_2 \in Y$,

$$\|\nabla\mathcal{E}^*(g_1) - \nabla\mathcal{E}^*(g_2)\|_X \leq \frac{1}{p-1} \|g_1 - g_2\|_Y. \quad (2.22)$$

(2) For $2 < p < \infty$, and all $g_1, g_2 \in Y$ with $\max\{\|g_1\|_Y, \|g_2\|_Y\} \leq R/2$,

$$\|\nabla\mathcal{E}^*(g_1) - \nabla\mathcal{E}^*(g_2)\|_X \leq \left[\frac{2}{3} \left(\frac{1}{4p} \right)^{1/(p-1)} R^{(2-p)/(p-1)} \right] \|g_1 - g_2\|_Y^{1/(p-1)}. \quad (2.23)$$

Proof. — The differentiability of \mathcal{E} and \mathcal{E}^* is well-known. Suppose first that $1 < p \leq 2$. Let $g_1, g_2 \in Y$ and define $f_j = \nabla\mathcal{E}^*(g_j)$, $j = 1, 2$. Then $g_j = \nabla\mathcal{E}(f_j)$, $j = 1, 2$. By Theorem 2.12,

$$\mathcal{E}(f_2) - \mathcal{E}(f_1) \geq \langle f_2 - f_1, g_1 \rangle + (p-1) d_X(f_1, f_2)$$

and

$$\mathcal{E}(f_1) - \mathcal{E}(f_2) \geq \langle f_1 - f_2, g_2 \rangle + (p-1) d_X(f_1, f_2).$$

Summing these inequalities,

$$0 \geq \langle f_2 - f_1, g_1 - g_2 \rangle + 2(p-1) d_X(f_1, f_2).$$

Then since $2\|f_2 - f_1\|_X \|g_2 - g_1\|_Y \geq \langle f_2 - f_1, g_1 - g_2 \rangle$ we conclude

$$\|f_2 - f_1\|_X \|g_2 - g_1\|_Y \geq (p-1) \|f_1 - f_2\|_X^2.$$

This proves (2.22).

Next, suppose that $2 < p < \infty$. Let $g_1, g_2 \in Y$ and define $f_j = \nabla\mathcal{E}^*(g_j)$, $j = 1, 2$. Then $g_j = \nabla\mathcal{E}(f_j)$, $j = 1, 2$. By Theorem 2.12,

$$\mathcal{E}(f_2) - \mathcal{E}(f_1) \geq \langle f_2 - f_1, g_1 \rangle + \frac{1}{4p} \left(\frac{2}{3} \right)^{p-1} (\|f_1\|_p + \|f_2\|_p)^{2-p} \|f_2 - f_1\|_p^p$$

and

$$\mathcal{E}(f_1) - \mathcal{E}(f_2) \geq \langle f_1 - f_2, g_2 \rangle + \frac{1}{4p} \left(\frac{2}{3} \right)^{p-1} (\|f_1\|_p + \|f_2\|_p)^{2-p} \|f_2 - f_1\|_p^p.$$

Summing these inequalities,

$$0 \geq \langle f_2 - f_1, g_1 - g_2 \rangle + \frac{1}{2p} \left(\frac{2}{3} \right)^{p-1} (\|f_1\|_p + \|f_2\|_p)^{2-p} \|f_2 - f_1\|_p^p.$$

Proceeding as above, we conclude

$$\|g_1 - g_2\|_Y \geq \frac{1}{4p} \left(\frac{2}{3} \right)^{p-1} (\|f_1\|_p + \|f_2\|_p)^{2-p} \|f_2 - f_1\|_p^{p-1}.$$

Therefore, using the norm preserving property $\|f_j\|_X = \|g_j\|_Y, j = 1, 2$,

$$\begin{aligned} & \|\nabla\mathcal{E}^*(g_2) - \nabla\mathcal{E}^*(g_1)\|_p \\ & \leq \left[\frac{2}{3} \left(\frac{1}{4p} \right)^{1/(p-1)} (\|g_1\|_p + \|g_2\|_{p'})^{(2-p)/(p-1)} \right] \|g_2 - g_1\|_{p'}^{1/(p-1)}. \quad \square \end{aligned}$$

The next lemma is a general result asserting that quantitative convexity of a functional implies quantitative smoothness of its Legendre transform.

LEMMA 2.14 ((d, Φ) -convexity of \mathcal{E} and regularity of \mathcal{E}^*). — *Let X and Y be a dual pair of Banach spaces. Let \mathcal{E} be a proper lower semicontinuous convex function on X . Let U be an open convex subset of X . Suppose that \mathcal{E} is (d, Φ) -convex on U for some metric d on X and some rate function Φ . Then $\partial\mathcal{E}^*(y)$ is a singleton $\{\nabla\mathcal{E}^*(y)\}$, and for all $y_1, y_2 \in \partial\mathcal{E}(U)$,*

$$\frac{\Phi(d(\mathcal{E}^*(y_2), \mathcal{E}^*(y_1)))}{\|\nabla\mathcal{E}^*(y_2) - \nabla\mathcal{E}^*(y_1)\|_X} \leq \|y_2 - y_1\|_Y. \quad (2.24)$$

Proof. — We have already seen that the map $y \mapsto \nabla\mathcal{E}^*(y)$ is well-defined. We first show that (2.24) is satisfied for all $y_1, y_2 \in \partial\mathcal{E}(U)$.

Let $y_1, y_2 \in \partial\mathcal{E}(U)$. Let $x_j = \nabla\mathcal{E}^*(y_j), j = 1, 2$. Then $y_j \in \partial\mathcal{E}(x_j), j = 1, 2$, and hence

$$\mathcal{E}(x_2) - \mathcal{E}(x_1) \geq \langle x_2 - x_1, y_1 \rangle + \Phi(d(x_1, x_2))$$

and

$$\mathcal{E}(x_1) - \mathcal{E}(x_2) \geq \langle x_1 - x_2, y_2 \rangle + \Phi(d(x_1, x_2)).$$

Adding these inequalities and rearranging terms,

$$2\Phi(d(x_1, x_2)) \leq \langle x_2 - x_1, y_2 - y_1 \rangle \leq 2\|x_2 - x_1\|_X \|y_2 - y_1\|_Y, \quad (2.25)$$

which implies (2.24).

Now suppose that U is open and \mathcal{E} itself is differentiable everywhere on U , so that $x \mapsto \nabla\mathcal{E}(x)$ is a well defined map on U . We have seen in Lemma 2.7 that for any proper lower semicontinuous convex function $\mathcal{E}, y \in \partial\mathcal{E}(x)$ which is the case if and only if $x \in \partial\mathcal{E}^*(y)$. Since $y = \nabla\mathcal{E}(x)$ is in $\partial\mathcal{E}(x), x \in \partial\mathcal{E}^*(y) = \{\nabla\mathcal{E}^*(y)\}$. That is, $x = \nabla\mathcal{E}^*(y) = \nabla\mathcal{E}^*(\nabla\mathcal{E}(x))$. Combining $y = \nabla\mathcal{E}(x)$ with $x = \nabla\mathcal{E}^*(y)$, we also obtain $y = \nabla\mathcal{E}(\nabla\mathcal{E}^*(y))$. \square

As indicated by Theorem 2.13, the inequality (2.24) simplifies when d is the metric induced by the norm on X . Define the function Ψ in terms of the rate function Φ by

$$\Psi(t) = \frac{\Phi(t)}{t} \quad (2.26)$$

for all $t > 0$. Since for any strictly convex function φ on \mathbb{R} ,

$$t_2 > t_1 > t_0 \quad \rightarrow \quad \frac{\varphi(t_2) - \varphi(t_0)}{t_2 - t_0} \geq \frac{\varphi(t_1) - \varphi(t_0)}{t_1 - t_0},$$

Ψ is a strictly increasing function on $(0, \infty)$. Suppose that

$$\lim_{t \downarrow 0} \Psi(t) = 0. \tag{2.27}$$

Using this notation, (2.24) becomes

$$\Psi(\|\nabla \mathcal{E}^*(y_2) - \nabla \mathcal{E}^*(y_1)\|_X) \leq \|y_2 - y_1\|_Y. \tag{2.28}$$

Moreover, under the condition (2.27), (2.28) implies that that $y \mapsto \nabla \mathcal{E}^*(y)$ is continuous on $\partial \mathcal{E}(U)$. Hence, when \mathcal{E} is differentiable, so that $y \mapsto \nabla \mathcal{E}^*(y)$ is invertible, $x \mapsto \nabla \mathcal{E}$ is open on U , and so $\nabla \mathcal{E}(U)$ is open. In this case, $y \mapsto \nabla \mathcal{E}^*(y)$ is differentiable on $\nabla \mathcal{E}(U)$.

To see this, let $y_1, y_2 \in \nabla \mathcal{E}(U)$. Then there are uniquely determined $x_1, x_2 \in U$ such that $x_j = \nabla \mathcal{E}^*(y_j)$, $j = 1, 2$. Therefore,

$$\mathcal{E}^*(y_2) \geq \mathcal{E}^*(y_1) + \langle x_1, y_2 - y_1 \rangle \quad \text{and} \quad \mathcal{E}^*(y_1) \geq \mathcal{E}^*(y_2) + \langle x_2, y_1 - y_2 \rangle. \tag{2.29}$$

Since

$$\langle x_2, y_1 - y_2 \rangle = \langle x_1, y_1 - y_2 \rangle + \langle x_2 - x_1, y_1 - y_2 \rangle,$$

the inequality (2.28) says that

$$|\mathcal{E}^*(y_2) - \mathcal{E}^*(y_1) - \langle x_1, y_2 - y_1 \rangle| \leq \|y_1 - y_2\|_Y \Psi(\|y_1 - y_2\|_Y),$$

which shows that \mathcal{E}^* is differentiable at y_1 , and the derivative is $x_1 = \nabla \mathcal{E}^*(y_1)$, which finally justifies our notation.

3. Duality and stability

When *either* \mathcal{E}^* or \mathcal{E} is (d, Φ) convex on their respective domains, we may use the strengthened form of Young's inequality (2.18) to strengthen the conclusion of Lemma 2.8.

LEMMA 3.1. — *Let \mathcal{E} and \mathcal{F} be proper, lower semicontinuous convex functions on X .*

- (1) *If for some metric d_Y on Y , and some rate function Φ_Y , \mathcal{E}^* is (d_Y, Φ_Y) -convex, then for all y in the domain if \mathcal{F}^* and all $x \in \partial \mathcal{F}^*(y)$,*

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \mathcal{F}(x) - \mathcal{E}(x) + \Phi_Y(d_Y(y, \nabla \mathcal{E}(x))). \tag{3.1}$$

- (2) If for some metric d_X on X , and some rate function Φ_X , \mathcal{E} is (d_X, Φ_X) -convex, then for all x in the domain of \mathcal{F} and all $y \in \partial\mathcal{F}(x)$,

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \mathcal{F}(x) - \mathcal{E}(x) + \Phi_Y(d_Y(y, \nabla\mathcal{E}(x))). \quad (3.2)$$

Proof. — We first prove (1). Suppose that \mathcal{E}^* is (d_Y, Φ_Y) -convex. Since $x \in \partial\mathcal{F}^*(y)$, there is equality in Young’s inequality so that

$$\mathcal{F}(x) + \mathcal{F}^*(y) = \langle x, y \rangle. \quad (3.3)$$

Then by the (d_Y, Φ_Y) -convexity of \mathcal{E}^* , expressed in the form (2.18),

$$\langle x, y \rangle \leq \mathcal{E}(x) + \mathcal{E}^*(y) - \Phi_Y(d_Y(y, \nabla\mathcal{E}(x))).$$

Combining this with the identity (3.3), we obtain

$$\mathcal{F}(x) + \mathcal{F}^*(y) \leq \mathcal{E}(x) + \mathcal{E}^*(y) - \Phi_Y(d_Y(y, \nabla\mathcal{E}(x))).$$

By rearranging the terms, we obtain (3.1). The proof of (3.2) is entirely analogous. \square

Now suppose that \mathcal{E} and \mathcal{F} are proper, closed convex functions, and that the functional inequality $\mathcal{E} \leq \mathcal{F}$ is not only valid, but is (d_X, Φ_X) stable. Suppose further that \mathcal{E}^* is (d_Y, Φ_Y) -convex. Then by Lemma 3.1(1), for all y in the domain of \mathcal{F}^* ,

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \Phi_X(d_X(x, X_0)) + \Phi_Y(d_Y(y, \nabla\mathcal{E}(x))). \quad (3.4)$$

For the right hand side to be small, we must have $y \approx \nabla\mathcal{E}(x)$ and $x \approx x_0 \in X_0$. The strong convexity of \mathcal{E}^* implies that $\nabla\mathcal{E}$ is continuous, and so $x \approx x_0$ implies $\nabla\mathcal{E}(x) \approx \nabla\mathcal{E}(x_0)$. Then $y \approx \nabla\mathcal{E}(x_0)$, and since $\nabla\mathcal{E}(x_0) \in Y_0$, y is necessarily close to an optimizer for the dual inequality $\mathcal{F}^* \leq \mathcal{E}^*$.

As shown below, going through this line of argument with precise estimates “transfers” a stability result for $\mathcal{E} \leq \mathcal{F}$ to its dual $\mathcal{F}^* \leq \mathcal{E}^*$.

As we have seen in the case of squared L^p norms, it may be that \mathcal{E} has better quantitative convexity properties than \mathcal{E}^* . In this case, we may combine the (d_X, Φ_X) stability of $\mathcal{E} \leq \mathcal{F}$ with (d_X, Φ_X) -convexity of \mathcal{E} using Lemma 3.1(2) to obtain

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \Phi_X(d_X(x, X_0)) + \Phi_X(d_X(x, \nabla\mathcal{E}^*(y))). \quad (3.5)$$

The sort of reasoning described just above again leads to the conclusion that if $\mathcal{E}^*(y) \approx \mathcal{F}^*(y)$, then $\nabla\mathcal{E}^*(y) \approx x_0 \in X_0$. But since $X_0 = \nabla\mathcal{E}^*(Y_0)$, this means that a stability bound for $\mathcal{F}^* \leq \mathcal{E}^*$, and even more simply since now we are comparing distances in the same space. In the next subsection, we carry out a general development of this idea. First however, we use Lemma 3.1 to prove Theorem 1.19

Proof of Theorem 1.19. — We use the notation from the introduction:

$$\|f\|_X = \|f\|_{2n/(n-2\alpha)} \quad \text{and} \quad \|f\|_X = \|(-\Delta)^{\alpha/2} f\|_2, \quad (3.6)$$

and

$$\|g\|_Y = \|g\|_{2n/(n+2\alpha)} \quad \text{and} \quad \|g\|_Y = \|(-\Delta)^{-\alpha/2} g\|_2. \quad (3.7)$$

Let $\mathcal{F}(f) = \mathcal{S}_{n,\alpha} \|f\|_X^2$ and $\mathcal{E}(f) = \|f\|_X^2$. With respect to the bilinear form $\langle f, g \rangle = 2\Re \left(\int_{\mathbb{R}^n} f^* g \, dm \right)$, $\mathcal{F}^*(g) = \mathcal{S}_{n,\alpha}^{-1} \|g\|_Y^2$ and $\mathcal{E}^*(g) = \|g\|_Y^2$.

Recall that general version of the Bianchi–Egnell Theorem proved [9] states that there exists a constant $\kappa_{BE} > 0$ depending only on n and α such that

$$\mathcal{F}(f) - \mathcal{E}(f) \geq \kappa_{BE} \inf_{z \in X_0} \{ \|f - z\|_X^2 \}. \quad (3.8)$$

That is, the Sobolev inequality is κ_{BE} stable.

Since $\frac{2n}{n+2\alpha} < 2$ and $\frac{2n}{n+2\alpha} - 1 = \frac{n-2\alpha}{n+2\alpha}$, Theorem 2.12 says that \mathcal{E}^* is $\frac{n-2\alpha}{n+2\alpha}$ -convex. Therefore, we have what we need to apply Lemma 3.1(1). Doing so we obtain that for $f = \nabla \mathcal{F}^*(g)$,

$$\mathcal{E}^*(g) - \mathcal{F}^*(g) \geq \mathcal{F}(f) - \mathcal{E}(f) + \frac{n-2\alpha}{n+2\alpha} \|g - \nabla \mathcal{E}(f)\|_Y^2. \quad (3.9)$$

By the Sobolev inequality itself, for all $f_0 \in X_0$,

$$\|f - f_0\|_X \geq \mathcal{S}_{n,\alpha}^{-1/2} \|f - f_0\|_X \geq \mathcal{S}_{n,\alpha}^{-1/2} \inf_{z \in X_0} \{ \|f - z\|_X \}.$$

Therefore,

$$\inf_{z \in X_0} \{ \|f - z\|_X^2 \} \geq \mathcal{S}_{n,\alpha}^{-1/2} \inf_{z \in X_0} \{ \|f - z\|_X \}.$$

Combining this with (3.8) and (3.9) yields

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \mathcal{S}_{n,\alpha}^{-1} \kappa_{BE} \inf_{z \in X_0} \{ \|f - z\|_X^2 \} + \frac{n-2\alpha}{n+2\alpha} \|g - \nabla \mathcal{E}(f)\|_Y^2.$$

Hence, for all $\epsilon > 0$, there is an $f_0 \in X_0$ such that

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \mathcal{S}_{n,\alpha}^{-1} \kappa_{BE} \|f - z\|_X^2 + \frac{n-2\alpha}{n+2\alpha} \|g - \nabla \mathcal{E}(f)\|_Y^2 - \epsilon. \quad (3.10)$$

By Theorem 2.13, since \mathcal{E}^* is $\frac{n-2\alpha}{n+2\alpha}$ -convex, $\nabla \mathcal{E}$ is Lipschitz from X to Y with Lipschitz constant $\frac{n+2\alpha}{n-2\alpha}$. Hence

$$\|f - f_0\|_X \geq \frac{n-2\alpha}{n+2\alpha} \|\nabla \mathcal{E}(f) - \nabla \mathcal{E}(f_0)\|_Y.$$

Using this in (3.10) yields:

$$\begin{aligned}
 & \mathcal{E}^*(y) - \mathcal{F}^*(y) \\
 & \geq \frac{n-2\alpha}{n+2\alpha} \min \left\{ \mathcal{S}_{n,\alpha}^{-1}(n)\kappa_{BE} \frac{n-2\alpha}{n+2\alpha}, 1 \right\} (\|f - f_0\|_X^2 + \|g - \nabla\mathcal{E}(f)\|_Y^2) - \epsilon \\
 & \geq \frac{n-2\alpha}{n+2\alpha} \min \left\{ \mathcal{S}_{n,\alpha}^{-1}(n)\kappa_{BE} \frac{n-2\alpha}{n+2\alpha}, 1 \right\} \frac{1}{2} (\|f - f_0\|_X + \|g - \nabla\mathcal{E}(f)\|_Y)^2 - \epsilon.
 \end{aligned}$$

Therefore, by the triangle inequality and the fact that $\nabla\mathcal{E}(f_0) \in Y_0$,

$$\begin{aligned}
 \|f - f_0\|_X + \|g - \nabla\mathcal{E}(f)\|_Y & \geq (\|\nabla\mathcal{E}(f) - \nabla\mathcal{E}(f_0)\|_Y + \|g - \nabla\mathcal{E}(f)\|_Y) \\
 & \geq (\|g - \nabla\mathcal{E}(f_0)\|_Y) \geq \inf_{w \in Y_0} \{\|g - w\|_Y\}.
 \end{aligned}$$

Noting that ϵ can be chosen arbitrarily small, we obtain the result with

$$\kappa_{BE}^* = \frac{1}{2} \frac{n-2\alpha}{n+2\alpha} \min \left\{ \mathcal{S}_{n,\alpha}^{-1}(n)\kappa_{BE} \frac{n-2\alpha}{n+2\alpha}, 1 \right\}. \quad (3.11)$$

□

3.1. Duality for stability inequalities in general

In this subsection we return to the general setting of an abstract dual pair X, Y . Suppose that for some metric d_X on X , and some rate function Φ_X , the inequality $\mathcal{E} \leq \mathcal{F}$ is (d_X, Φ_X) stable, and for some rate function Φ_Y on Y , \mathcal{E}^* is (Φ_Y, d_Y) stable.

Let Φ be any rate function such that $\Phi(t) \leq \min\{\Phi_X(t), \Phi_Y(t)\}$ for all t . We have seen that such a Φ always exists. Thus it is no loss of generality to suppose that $\Phi_X = \Phi_Y = \Phi$. Then, by the convexity of Φ , for any $a, b \geq 0$, $\Phi(a) + \Phi(b) \geq 2\Phi((a+b)/2)$. Thus, (3.4) implies that for all $y \in Y$ and $x = \nabla\mathcal{F}^*(y)$,

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq 2\Phi \left(\frac{1}{2} d_X(x, X_0) + \frac{1}{2} d_Y(y, \nabla\mathcal{E}(x)) \right). \quad (3.12)$$

Note that if $x \in X_0$, then $y \in Y_0$, and we may assume that this is not the case. For any $0 < \epsilon < d_X(x, X_0)$, we can find $x_0 \in X_0$ such that $d_X(x, X_0) \geq d_X(x, x_0) - \epsilon$, and hence

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq 2\Phi \left(\frac{1}{2} d_X(x, x_0) + \frac{1}{2} d_Y(y, \nabla\mathcal{E}(x)) - \epsilon \right)^2. \quad (3.13)$$

We have seen in Lemma 2.14 and the remarks following it that if d_X and d_Y are the metrics induced by the norms, and if \mathcal{E}^* is κ -convex, $\nabla\mathcal{E}$ is Lipschitz. In our present more general setting, a variety of assumptions

on the metric d_Y and the rate function Φ_Y such that \mathcal{E}^* is (d_Y, Φ_Y) -convex imply the existence of a rate function Ψ such that for all $x_1, x_2 \in X$,

$$d_X(x_1, x_2) \geq \Psi(d_Y(\nabla\mathcal{E}(x_1), \nabla\mathcal{E}(x_2))). \quad (3.14)$$

Using this in (3.12), we obtain

$$\begin{aligned} \mathcal{E}^*(y) - \mathcal{F}^*(y) & \\ & \geq 2\Phi\left(\frac{1}{2}\Psi(d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0))) + \frac{1}{2}d_Y(y, \nabla\mathcal{E}(x)) - \epsilon\right)^2. \end{aligned} \quad (3.15)$$

By the triangle inequality and monotonicity of Ψ ,

$$\Psi(d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0))) \geq \Psi(d_Y(y, \nabla\mathcal{E}(x_0)) - d_Y(\nabla\mathcal{E}(x), y)).$$

Therefore,

$$\Psi(d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0)) + d_Y(y, \nabla\mathcal{E}(x))) \geq \inf_{s>0} \{\Psi(d_Y(y, \nabla\mathcal{E}(x_0)) - s) + s\}.$$

Define

$$\chi(s) = \begin{cases} s & s \geq 0 \\ \infty & s < 0 \end{cases} \quad \text{and} \quad \widehat{\Psi}(t) = \inf_{s \in \mathbb{R}} \{\Psi(t - s) + \chi(s)\}. \quad (3.16)$$

The function $\widehat{\Psi}$ is the *infimal convolution* of Ψ and χ , Ψ . Since the infimal convolution of convex functions is convex, $\widehat{\Psi}$ is convex, and evidently it is strictly positive on $(0, \infty)$. Hence $\widehat{\Psi}$ is a rate function. We have proved:

$$\begin{aligned} d_X(x, x_0) &= \|x - x_0\|_X \\ &\geq \kappa \|\nabla\mathcal{E}(x) - \nabla\mathcal{E}(x_0)\|_Y = \kappa d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0)). \end{aligned} \quad (3.17)$$

In this case, by the triangle inequality,

$$\begin{aligned} d_X(x, x_0) + d_Y(y, \nabla\mathcal{E}(x)) & \\ & \geq \kappa d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0)) + d_Y(y, \nabla\mathcal{E}(x)) \\ & \geq \min\{\kappa, 1\} (d_Y(\nabla\mathcal{E}(x), \nabla\mathcal{E}(x_0)) + d_Y(y, \nabla\mathcal{E}(x))) \\ & \geq \min\{\kappa, 1\} d_Y(y, \nabla\mathcal{E}(x_0)) \end{aligned}$$

By Lemma 2.8, $\nabla\mathcal{E}(x_0) \in Y_0$, the set of optimizers for $\mathcal{F}^* \leq \mathcal{E}^*$. Thus,

$$d_X(x, x_0) + d_Y(y, \nabla\mathcal{E}(x)) \geq \min\{2\kappa, 1\} d_Y(y, Y_0).$$

Using this in (3.12) and using the fact that $\epsilon > 0$ is arbitrary, we finally obtain

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq 2\Phi(\min\{2\kappa, 1\} d_Y(y, Y_0)). \quad (3.18)$$

we have proved:

THEOREM 3.2. — *Let \mathcal{E} and \mathcal{F} be proper, lower semicontinuous convex functions on X . Suppose that the functional inequality $\mathcal{E} \leq \mathcal{F}$ is valid and optimal. Let Φ be a rate function. Suppose also that:*

- (1) *For some metric d_Y on Y , and some rate function Φ , \mathcal{E}^* is (d_Y, Φ) -convex.*
- (2) *For some metric d_X on X , $\mathcal{E} \leq \mathcal{F}$ is (d_X, Φ) -stable.*

Suppose that for some rate function Ψ ,

$$d_X(x_1, x_2) \geq \Psi(d_Y(\nabla\mathcal{E}(x_1), \nabla\mathcal{E}(x_2))) \quad (3.19)$$

for all $x_1, x_2 \in X$. Define the function $\widehat{\Psi}$ on $[0, \infty)$ by

$$\widehat{\Psi}(t) = \inf_{s>0} \{ \Psi(t-s) + s \}.$$

Then $\widehat{\Psi}$ is a rate function and for all $y \in Y$,

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq \Phi(\widehat{\Psi}(d_Y(y - Y_0))). \quad (3.20)$$

There is a variant in which the lower bound is in terms of $d_X(\nabla\mathcal{E}^*(y), X_0)$ instead of $d_Y(y, Y_0)$. This has the same qualitative meaning since by Theorem 2.8, $\nabla\mathcal{E}^*(y) \in X_0$ if and only if $y \in Y_0$.

To obtain this, we assume that $y \mapsto \nabla\mathcal{E}^*(y)$ satisfies an inequality of the form

$$\Psi(d_X(\mathcal{E}^*(y_1), \nabla\mathcal{E}^*(y_2))) \leq d_Y(y_1, y_2), \quad (3.21)$$

This will of course be the case if \mathcal{E} has good convexity properties. This variant will be useful when \mathcal{E} is λ convex for some $\lambda > 0$; e.g., when $\mathcal{E}(f) = \|f\|_p^2$ for some $1 < p \leq 2$.

THEOREM 3.3. — *Let \mathcal{E} and \mathcal{F} be proper, lower semicontinuous convex functions on X . Suppose that the functional inequality $\mathcal{E} \leq \mathcal{F}$ is valid and optimal. Let Φ be a rate function. Suppose also that:*

- (1) *For some metric d_Y on Y , and some rate function Φ , \mathcal{E}^* is (d_Y, Φ) -convex.*
- (2) *For some metric d_X on X , $\mathcal{E} \leq \mathcal{F}$ is (d_X, Φ) -stable.*
- (3) *For some rate function Ψ , (3.21) is satisfied for all $y_1, y_2 \in Y$.*

Define the function $\widehat{\Psi}$ on $[0, \infty)$ by $\widehat{\Psi}(t) = \inf_{s \in \mathbb{R}} \{ \Psi(|t-s|) + |s| \}$. Then $\widehat{\Psi}$ is a rate function and for all $y \in Y$

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq 2\Phi(\widehat{\Psi}(d_X(\nabla\mathcal{E}^*(y), X_0))). \quad (3.22)$$

Proof. — Arguing as before, we see that for every $\epsilon > 0$, there is an $x_0 \in X_0$ such that

$$\mathcal{E}^*(y) - \mathcal{F}^*(y) \geq 2\Phi \left(\frac{1}{2} d_X(x, x_0) + \frac{1}{2} d_Y(y, \nabla \mathcal{E}(x)) + \epsilon \right). \quad (3.23)$$

Under our hypotheses, $\nabla \mathcal{E}$ and $\nabla \mathcal{E}^*$ are inverse to one another. Define $x_1 = \nabla \mathcal{E}^*(y)$ so that $y = \nabla \mathcal{E}(x_1)$. Then by (3.21),

$$\begin{aligned} d_Y(y, \nabla \mathcal{E}(x)) &= d_Y(\nabla \mathcal{E}(x_1), \nabla \mathcal{E}(x)) \\ &\geq \Psi(d_X(\nabla \mathcal{E}^*(\nabla \mathcal{E}(x_1)), \nabla \mathcal{E}^*(\nabla \mathcal{E}(x))) = \Psi(d_X(x_1, x)). \end{aligned}$$

Using this in (3.23), we obtain

$$\begin{aligned} \mathcal{E}^*(y) - \mathcal{F}^*(y) &\geq 2\Phi \left(\frac{1}{2} \|x - x_0\|_X + \frac{1}{2} \Psi \|\nabla \mathcal{E}^*(y) - x\|_X + \epsilon \right) \\ &\geq \inf_{x_0 \in Y_0} \inf_{z \in X} \left\{ 2\Phi \left(\frac{1}{2} \Psi (\|\nabla \mathcal{E}^*(y) - z\|_X) + \frac{1}{2} \|x_0 - z\|_X + \epsilon \right) \right\} \\ &\geq \inf_{x_0 \in Y_0} \inf_{z \in X} \left\{ 2\Phi \left(\frac{1}{2} \Psi (\|\nabla \mathcal{E}^*(y) - x_0\|_X - \|x_0 - z\|_X) + \frac{1}{2} \|x_0 - z\|_X \right) \right\} \\ &\geq 2\Phi(\widehat{\Psi}(d_X(\nabla \mathcal{E}^*(y), X_0))). \quad \square \end{aligned}$$

3.2. Duality and local stability

The method used by Bianchi and Egnell to prove (3.8) can be elaborated to give a *quantitative* stability bound for the Sobolev inequality, but only locally, near to X_0 : As discussed above, their local estimate comes from an eigenvalue calculation and control of remainder terms in a Taylor expansion, and no essential use of qualitative compactness arguments is made up to this point. Though they did not obtain quantitative control over the remainder terms in the Taylor expansion, this can be done using uniform convexity inequalities; see [15, Section 2].

The optimal logarithmic HLS inequality on \mathbb{R}^2 can be obtained from the HLS inequality by taking the limit $\alpha \rightarrow 1$; see [8]. It would be tempting to try and extract a stability result for the Logarithmic HLS inequality from Theorem 1.19 using the expression for κ_{BE}^* given in (3.11). However, κ_{BE}^* depends on κ_{BE} , and there is no information available on how κ_{BE} depends on α .

It is therefore of interest that the *local* Bianchi–Egnell result, which can be made quantitative, also transfers by duality. Application of this to stability

for the Logarithmic HLS inequality will be made elsewhere, but it is natural to explain here how the transfer of stability by duality localizes.

In our notation, the local stability result of Bianchi and Egnell says that there exist computable constants $r, \lambda > 0$ such that

$$\inf_{z \in X_0} \{\|f - z\|_X\} \leq r \|f\|_X \Rightarrow \mathcal{F}(f) - \mathcal{E}(f) \geq \lambda \inf_{z \in X_0} \{\|f - z\|_X^2\}, \quad (3.24)$$

where we are using the notation introduced in (1.9) and (1.10). In the rest of this section, we use the notation from the introduction in the discussion there of the Sobolev and HLS inequalities, as in the proof of Theorem 1.19. (In [3], (3.24) was proved for $\alpha = 1$ and any $r < 1$, but with a $o(\inf_{z \in X_0} \{\|f - z\|_X^2\})$ Taylor expansion remainder term on the right. Decreasing r further, one may absorb the Taylor remainder term into the main term since $\lambda > 0$. See also [9, Theorem 3] for general α .)

We can express (3.24) as saying that the inequality $\mathcal{E} \leq \mathcal{F}$ has is *locally λ -stable with respect to the $\|\cdot\|_X$ metric*. We now show that this local stability result “dualizes” to yield a local quantitative stability result for the HLS inequality. We do this for $\alpha = 1$ only for simplicity, and write \mathcal{S}_n for $\mathcal{S}_{n,1}$.

LEMMA 3.4 (Local quantitative local stability for the HLS inequality). *Let $\lambda, r > 0$ be such that the local Bianchi–Egnell inequality (3.24) is valid. Then for all $g \in Y = L^{2n/(n+2)}$ such that $2\mathcal{F}^*(g) \geq \mathcal{E}^*(g)$,*

$$\begin{aligned} d_Y(g, Y_0) &\leq \frac{r}{2} \|g\|_Y \\ \Rightarrow \mathcal{E}^*(g) - \mathcal{F}^*(g) &\geq \frac{1}{2} \frac{n-2}{n+2} \min \left\{ \lambda \frac{4}{\mathcal{S}_n} \frac{n-2}{n+2}, 1 \right\} d_Y^2(g, Y_0). \end{aligned} \quad (3.25)$$

Proof. — It is easy to see that $\nabla \mathcal{F}^*(g) = (-\Delta)^{-1}g$, and consequently, for all $g_1, g_2 \in Y$,

$$\|\nabla \mathcal{F}^*(g_1) - \nabla \mathcal{F}^*(g_2)\|_X \leq \|g_1 - g_2\|_Y \leq \sqrt{\mathcal{S}_n} \|g_1 - g_2\|_Y, \quad (3.26)$$

where the last inequality is the HLS inequality. That is,

$$\inf_{z \in X_0} \{\|\nabla \mathcal{F}^*(g) - z\|_X\} \leq \sqrt{\mathcal{S}_n} \inf_{w \in Y_0} \{\|g - w\|_Y\}.$$

Next, since $2\mathcal{F}^*(g) \geq \mathcal{E}^*(g)$,

$$\begin{aligned} \mathcal{S}_n \|g\|_Y^2 \\ \leq 2 \|(-\Delta)^{-1/2}g\|_2^2 = 2 \|(-\Delta)^{1/2} [(-\Delta)^{-1}g]\|_2^2 = 2 \|(-\Delta)^{1/2} [\nabla \mathcal{F}^*(g)]\|_2^2 \end{aligned}$$

Next, since $2\mathcal{F}^*(g) \geq \mathcal{E}^*(g)$, It follows from (3.24) and (3.26) that

$$\inf_{w \in Y_0} \{\|g - w\|_Y\} \leq \frac{r}{\sqrt{2}} \|g\|_Y \Rightarrow \inf_{z \in X_0} \{\|\nabla \mathcal{F}^*(g) - z\|_X\} \leq r \|\nabla \mathcal{F}^*(g)\|_X, \quad (3.27)$$

and then as a consequence of (3.24),

$$\begin{aligned} \inf_{w \in Y_0} \{ \|g - w\|_Y \} &\leq \frac{r}{\sqrt{2}} \\ \Rightarrow \mathcal{F}(\nabla \mathcal{F}^*(g)) - \mathcal{E}(\nabla \mathcal{F}^*(g)) &\geq \lambda \left(\inf_{z \in X_0} \{ \|\nabla \mathcal{F}^*(g) - z\|_X \} \right)^2. \end{aligned} \quad (3.28)$$

Since \mathcal{E}^* is $\frac{n-2}{n+2}$ -convex, Lemma 3.1(1) says that whenever

$$\inf_{w \in Y_0} \{ \|g - w\|_Y \} \leq \frac{r}{\sqrt{2}} \|g\|_Y,$$

for all $\epsilon > 0$, there is an $f_0 \in X_0$ such that with $f = \nabla \mathcal{F}^*(g)$,

$$\mathcal{E}^*(g) - \mathcal{F}^*(g) \geq \lambda \|f - f_0\|_X^2 - \epsilon + \frac{n-2}{n+2} \|g - \nabla \mathcal{E}(f)\|_Y^2. \quad (3.29)$$

Since \mathcal{E}^* is $\frac{n-2}{n+2}$ -convex, $\nabla \mathcal{E}$ is Lipschitz from $L^{2n/(n-2)}(\mathbb{R}^n)$ to $L^{2n/(n+2)}(\mathbb{R}^n)$ with Lipschitz constant $\frac{1}{2} \frac{n+2}{n-2}$, and by the Sobolev inequality, $\|f - f_0\|_X \geq \mathcal{S}_n^{-1/2} \|f - f_0\|_{2n/(n-2)}$. Therefore,

$$\|f - f_0\|_X \geq \frac{2}{\mathcal{S}_n^{1/2}} \frac{n-2}{n+2} \|\nabla \mathcal{E}(f) - \nabla \mathcal{E}(f_0)\|_{2n/(n+2)}. \quad (3.30)$$

Using this in (3.29), we obtain (recalling that $Y = L^{2n/(n+2)}$):

$$\begin{aligned} \mathcal{E}^*(g) - \mathcal{F}^*(g) &\geq \frac{n-2}{n+2} \left(\lambda \frac{4}{\mathcal{S}_n} \frac{n-2}{n+2} \|\nabla \mathcal{E}(f) - \nabla \mathcal{E}(f_0)\|_Y^2 + \|g - \nabla \mathcal{E}(f)\|_Y^2 \right) - \epsilon \\ &\geq \frac{1}{2} \frac{n-2}{n+2} \min \left\{ \lambda \frac{4}{\mathcal{S}_n} \frac{n-2}{n+2}, 1 \right\} (\|g - \nabla \mathcal{E}(f_0)\|_Y)^2 - \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the proof is complete. \square

4. Monotonicity and stability

In this section we discuss how the results of the previous section may be combined with functional flows to obtain quantitative stability results.

The HLS Sobolev inequalities for $\alpha = 1$ are intimately connected with the fast diffusion equation. The content of this section draws on results from [5] and especially [10].

For a smooth function $f(\eta)$ on \mathbb{R}^n , $\Delta f(\eta) = \sum_{j=1}^n \frac{\partial^2}{\partial \eta_j^2} f(\eta)$. A smooth function $u(\eta, t)$ on $\mathbb{R}^n \times (0, \infty)$ satisfies the *fast diffusion equation* in case

for some $m \in (0, 1)$,

$$\frac{\partial u}{\partial t}(\eta, t) = \Delta u^m(\eta, t) \quad (4.1)$$

(For $m = 1$, this is the usual diffusion equation).

Simple calculations show that $u(\eta, t)$ solves (4.1) if and only if $v(\eta, t) := e^{td}u(e^t\eta, e^{\beta t})$ with $\beta = 2 - d(1 - m)$ satisfies the equation

$$\frac{\partial v}{\partial t}(\eta, t) = \beta \Delta v^m(\eta, t) + \nabla \cdot [\eta v(\eta, t)]. \quad (4.2)$$

This equation has a steady state:

$$v_{\infty, M}(\eta) := \left(D(M) + \frac{1 - m}{2\beta m} |\eta|^2 \right)^{-1/(1-m)}, \quad (4.3)$$

where $D(M)$ is a computable constant. The rescaling induced by the additional drift term cancels the spreading effect of the diffusion. For $m = 1$, this is the Fokker–Planck equation, and (4.2) is a non-linear relative of the Fokker–Planck equation.

The Cauchy problem for the FDE (4.1) has been studied extensively; see Vázquez [18]. Herrero and Pierre [11] proved that the range of mass conservation for the fast diffusion equation is $1 - 2/d < m < 1$, which is exactly the range of $m < 1$ in which integrable self-similar solutions exist. Within this range, the flow associated to the fast diffusion equation is in many ways *even better* than the flow associated to the heat equation; see Vázquez [17]. The solutions of (4.1) with positive integrable initial data are C^∞ and strictly positive everywhere instantaneously, just as for the heat flow.

Moreover, for non-negative initial data f of mass M satisfying

$$\sup_{|x| > R} f(x) |x|^{2/(1-m)} < \infty \quad (4.4)$$

for some $R > 0$, which means that f is decaying at infinity at least as fast as the Barenblatt profile $v_{\infty, M}$, the solution $v(x, t)$ of (4.1) with initial data f satisfies the following remarkable bounds: For any $t_* > 0$, there exists a constant $C = C(t_*) > 0$ such that

$$\frac{1}{C} \leq \frac{v(x, t)}{v_{\infty, M}} \leq C, \quad (4.5)$$

for all $t \geq t_*$ and $x \in \mathbb{R}^d$. This shows that fast diffusion is “really fast” in that it spreads mass out to infinity to produce the “right tails” instantly.

Finally, it is well-known [4] that

$$\lim_{t \rightarrow \infty} \|v(t) - v_{\infty, M}\|_{L^1(\mathbb{R}^d)} = 0. \quad (4.6)$$

Now consider the case $m = 1 - 1/n$, $n \geq 3$. Then

$$v_{\infty, M}(\eta) = \left(D(M) + \frac{1-m}{2\beta m} |\eta|^2 \right)^{-n}.$$

Thus, $v_{\infty, M}(\eta)^{(n-2)/2n}$ is a Sobolev optimizer (for $\alpha = 1$). Demange [10] shows in that for non-negative $f \in L^{2n/(n-2)}(\mathbb{R}^n)$, $n \geq 3$, if one defines f_t for $t > 0$ by taking the $(n-2)/2n$ power of the solution at time t , of the Cauchy problem for (4.2 with initial data $v(\eta) = f^{2n/(n-2)}(\eta)$ then $\mathcal{S}_{n,1} \|\nabla f_t\|_2^2 - \|f_t\|_{2n/(n-2)}^2$ is monotone decreasing in t . By what we have explained above, as $t \rightarrow \infty$, f_t converges to $v_{\infty, M}(\eta)^{(n-2)/2n}$ which is a non-zero Sobolev optimizer. Hence if $\mathcal{S}_{n,1} \|\nabla f\|_2^2 - \|f\|_{2n/(n-2)}^2$ is sufficiently small, the local Bianchi–Egnell bound becomes applicable at some large t . But then $\mathcal{S}_{n,1} \|\nabla f_t\|_2^2 - \|f_t\|_{2n/(n-2)}^2$ cannot be too small at this point or the local bound would be violated. By Demange’s monotonicity result, then $\mathcal{S}_{n,1} \|\nabla f\|_2^2 - \|f\|_{2n/(n-2)}^2$ cannot have been too small either. In this way, one obtains a *global* quantitative version of the Bianchi–Egnell Theorem with a computable lower bound. By duality, this transfers to the HLS inequality as well.

Monotonicity under fast diffusion with $m = n/(n+2)$ flow for the $\alpha = 1$ HLS functional has been proved in [5]. However, this was done for (4.1) and not (4.2), and there is no limiting optimizer in this case (except in for the Logarithmic HLS inequality, which has better scaling properties, so that for it there is monotonicity along (4.2) as well). This provides another approach to quantitative stability for the Logarithmic HLS inequality. Quantitative stability for the Logarithmic HLS inequality under additional assumptions on f has been proved in [6], the additional assumptions are natural for the application made there to the Keller–Segel equation, but the application would be strengthened by relaxing them. This will be taken up elsewhere.

Appendix

We now give a detailed and self-contained proof of Theorem 2.12. The cases $1 < p \leq 2$ and the case $p > 2$ are treated quite differently, and of course the results are quite different. The inequalities in Theorem 2.12 are closely related to the uniform convexity properties of the L^p spaces. A proof of the optimal 2-uniform convexity inequality for L^p , $1 < p \leq 2$, leads directly to the $(p-1)$ -convexity of the squared L^p norm for such p . For $p > 2$, the “easy” Clarkson inequality gives a sharp uniform convexity bound, but more work is required to obtain the local p th power convexity of the squared L^p in this range.

Proof of Theorem 2.12. — First let $1 < p \leq 2$. Let $X = L^p(\Omega, \mathcal{B}, \mu)$ as in Example 2.4. What follows is adapted from [2], with more detail provided. Let f and g be simple functions of the form

$$f(x) = \sum_{j=1}^n z_j 1_{A_j}(x) \quad \text{and} \quad g(x) = \sum_{j=1}^n w_j 1_{A_j}(x),$$

where $\mu(A_j) < \infty$ for each j . and for each j , $z_j w_j^*$ is not real. This latter condition guarantees that $z_j + t w_j \neq 0$ for any real t , and thus for all $x \in \cup_{j=1}^n A_j$, and all $t \in \mathbb{R}$, $f(x) + t g(x) \neq 0$. Define

$$Y(t) = \|f + t g\|_p^p \quad \text{and} \quad q = \frac{p}{2},$$

so that $\|f + t g\|_p^2 = Y^{1/q}(t)$. Differentiating twice,

$$\frac{d^2}{dt^2} \|f + t g\|_p^2 \frac{1}{q} \left(\frac{1}{q} - 1 \right) Y^{1/q-2} (Y')^2 + \frac{1}{q} Y^{1/q-1} Y'' \frac{1}{q} Y^{1/q-1} Y''$$

(Note that already this estimate would fail for $p > 2$.)

Next, a simple calculation, and the fact that

$$|f + t g|^{2q-2-4} (\Re((f + t g)^* g)) \geq |f + t g|^{2q-2-2} |g|^2$$

yields

$$Y''(t) \geq p(p-1) \int |f + t g|^{2q-2} |g|^2 \, d\mu.$$

The *reverse Hölder inequality* says that for $0 < r < 1$ and $s = r/(r-1)$, whenever $a_j \geq 0$ for $j = 1, \dots, n$, and $b_j > 0$ for $j = 1, \dots, n$,

$$\sum_{j=1}^n a_j b_j \geq \left(\sum_{j=1}^n a_j^r \right)^{1/r} \left(\sum_{j=1}^n b_j^s \right)^{1/s}.$$

Applying this to our integral of simple functions, the result is that for all t ,

$$\frac{d^2}{dt^2} \|f + t g\|_p^2 \geq 2(p-1) \|g\|_p^2. \tag{A.7}$$

Now the restriction the f and g be (special) simple functions is easily removed by density, and hence (A.7) is valid for all $f, g \in X$.

A much more direct calculation shows that $f, g \in X$, $\left. \frac{d}{dt} \|f + t g\|_p^2 \right|_{t=0} = \langle g, \nabla \mathcal{E}(f) \rangle$. Therefore, integration yields $\|f + g\|_p^2 \geq \|f\|_p^2 + \langle g, \nabla \mathcal{E}(f) \rangle + (p-1) \|g\|_p^2$. Defining $f_1 = f$ and $f_2 = f + g$, this is equivalent to

$$\mathcal{E}(f_2) \geq \mathcal{E}(f_1) + \langle f_2 - f_1, \nabla \mathcal{E}(f_1) \rangle + (p-1) \|f_2 - f_1\|_p^2. \tag{A.8}$$

This completes the proof that \mathcal{E} is $(p-1)$ -convex on $X = L^p(\Omega, \mathcal{B}, \mu)$ for $1 < p \leq 2$.

For $2 < p < \infty$, we cannot proceed via a lower bound on the second derivative of $\|f + tg\|_p^2$. We shall instead use a duality argument. For $p \geq 2$, let $f \in X$ and $g \in Y$ using the notation of Example 2.4. We showed there that

$$\begin{aligned} \mathcal{E}(f) + \mathcal{E}^*(g) &\geq \langle f, g \rangle - (\|f\|_p - \|g\|_{p'})^2 + \left(2\|f\|_p \|g\|_p - 2\Re \left(\int_{\Omega} f^* g \, d\mu \right) \right). \end{aligned} \quad (\text{A.9})$$

Now fix $f_1, f_2 \in X$, and let $g = \nabla \mathcal{E}(f_1) = \|f_1\|_p^{2-p} |f_1|^{p-1} \operatorname{sgn}(f_1)$. Then taking $f = f_2$ and this choice of g in (A.9), we obtain

$$\begin{aligned} \mathcal{E}(f_2) + \mathcal{E}^*(\nabla \mathcal{E}(f_1)) &\geq \langle f_2, \nabla \mathcal{E}(f_1) \rangle + (\|f_2\|_p - \|\nabla \mathcal{E}(f_1)\|_{p'})^2 \\ &\quad + \left(2\|f_2\|_p \|\nabla \mathcal{E}(f_1)\|_{p'} - 2\Re \left(\int_{\Omega} f_2^* \nabla \mathcal{E}(f_1) \, d\mu \right) \right). \end{aligned} \quad (\text{A.10})$$

By the cases of equality in Young's inequality,

$$\mathcal{E}(f_1) + \mathcal{E}^*(\nabla \mathcal{E}(f_1)) = \langle f_1, \nabla \mathcal{E}(f_1) \rangle.$$

Combining this with (A.9), and the norm preservation property (2.4) of $\nabla \mathcal{E}$, we obtain

$$\begin{aligned} \mathcal{E}(f_2) &\geq \mathcal{E}(f_1) + \langle f_2 - f_1, \nabla \mathcal{E}(f_1) \rangle + (\|f_2\|_p - \|f_1\|_p)^2 \\ &\quad + \left(2\|f_2\|_p \|f_1\|_p - 2\Re \left(\int_{\Omega} f_2^* \nabla \mathcal{E}(f_1) \, d\mu \right) \right). \end{aligned} \quad (\text{A.11})$$

Define the unit vectors u and v by $u = \|f_2\|_p^{-1} f_2$ and

$$v = \nabla \mathcal{E}(\|f_1\|_p^{-1} f_1) = \|f_1\|_p^{1-p} |f_1|^{p-1} \operatorname{sgn}(f_1). \quad (\text{A.12})$$

Then

$$\begin{aligned} &\left(2\|f_2\|_p \|f_1\|_p - 2\Re \left(\int_{\Omega} f_2^* \nabla \mathcal{E}(f_1) \, d\mu \right) \right) \\ &= 2\|f_2\|_p \|f_1\|_p \left(1 - \Re \left(\int_{\Omega} uv^* \, d\mu \right) \right). \end{aligned} \quad (\text{A.13})$$

It is proved in [7], as a simple consequence of the ‘‘easy Clarkson inequality’’ that for all unit vectors $u \in X$ and $v \in Y$,

$$1 - \Re \left(\int_{\Omega} uv^* \, d\mu \right) \geq \frac{1}{p2^{p-1}} \|u - \nabla \mathcal{E}^*(v)\|_p^p. \quad (\text{A.14})$$

Next, using (A.12),

$$\nabla \mathcal{E}^*(v) = |v|^{1/(p-1)} \operatorname{sgn}(v) = \|f_1\|_p^{-1} f_1.$$

Thus,

$$\begin{aligned} 2\|f_2\|_p\|f_1\|_p \left(1 - \Re \left(\int_{\Omega} uv^* \, d\mu \right)\right) \\ \geq \frac{\|f_2\|_p^{1-p}\|f_1\|_p^{1-p}}{p2^{p-1}} \|\|f_1\|_p f_2 - \|f_2\|_p f_1\|_p^p. \end{aligned} \quad (\text{A.15})$$

Thus, (A.11) becomes

$$\begin{aligned} \mathcal{E}(f_2) \geq \mathcal{E}(f_1) + \langle f_2 - f_1, \nabla \mathcal{E}(f_1) \rangle + (\|f_2\|_p - \|f_1\|_p)^2 \\ + \frac{\|f_2\|_p^{1-p}\|f_1\|_p^{1-p}}{p2^{p-1}} \|\|f_1\|_p f_2 - \|f_2\|_p f_1\|_p^p. \end{aligned} \quad (\text{A.16})$$

Define

$$m := \frac{\|f_1\|_p + \|f_2\|_p}{2}, \quad a := \frac{\|f_1\|_p - \|f_2\|_p}{\|f_1\|_p + \|f_2\|_p} \quad \text{and} \quad \epsilon = \frac{\|f_1 - f_2\|_p}{\|f_1\|_p + \|f_2\|_p}.$$

Then

$$\|f_1\|_p\|f_2\|_p = m^2(1 - a^2)$$

and

$$\|\|f_1\|_p f_2 - \|f_2\|_p f_1\|_p = m\|(f_2 - f_1) + a(f_1 + f_2)\|_p \geq 2m^2(\epsilon - a) \geq 0.$$

we can bound the right hand side of (A.16) from below by as

$$4m^2 a^2 + \frac{2m^2(1 - a^2)^{1-p}}{p}(\epsilon - a)^p.$$

If $a > \frac{1}{2}\epsilon^{p/2}$, this is bounded below by $m^2\epsilon^p$. If $a \leq \frac{1}{2}\epsilon^{p/2}$, this is bounded below by $\frac{1}{p} \left(\frac{2}{3}\right)^{p-1} m^2\epsilon^p$. This proves (2.23). \square

Bibliography

- [1] T. AUBIN, “Problèmes isoperimétriques et espaces de Sobolev”, *J. Differ. Geom.* **11** (1976), p. 573-598.
- [2] K. BALL, E. A. CARLEN & E. H. LIEB, “Sharp uniform convexity and smoothness inequalities for trace norms”, *Invent. Math.* **115** (1997), no. 3, p. 463-482.
- [3] G. BIANCHI & H. EGNELL, “A note on the Sobolev inequality”, *J. Funct. Anal.* **100** (1991), no. 1, p. 18-24.
- [4] A. BLANCHET, M. BONFORTE, J. DOLBEAULT, G. GRILLO & J. L. VÁZQUEZ, “Asymptotics of the fast diffusion equation via entropy estimates”, *Arch. Ration. Mech. Anal.* **191** (2009), no. 3, p. 347-385.
- [5] E. A. CARLEN, J. A. CARRILLO & M. LOSS, “Hardy–Littlewood–Sobolev inequalities via fast diffusion flows”, *Proc. Natl. Acad. Sci. USA* **107** (2010), no. 46, p. 19696-19701.
- [6] E. A. CARLEN & A. FIGALLI, “Stability for a GNS inequality and the log-HLS inequality, with application to the critical mass Keller–Segel equation”, *Duke Math. J.* **162** (2013), no. 3, p. 579-625.

- [7] E. A. CARLEN, R. L. FRANK & E. H. LIEB, “Stability estimates for the lowest eigenvalue of a Schrödinger operator”, *Geom. Funct. Anal.* **24** (2014), no. 1, p. 63-84.
- [8] E. A. CARLEN & M. LOSS, “Competing symmetries, the logarithmic Hardy–Littlewood–Sobolev inequality and Onofri’s inequality on S^n ”, *Geom. Funct. Anal.* **2** (1992), no. 1, p. 90-104.
- [9] S. CHEN, R. L. FRANK & T. WETH, “Remainder terms in the fractional Sobolev inequality”, *Indiana Univ. Math. J.* **62** (2013), no. 4, p. 1381-1397.
- [10] J. DEMANGE, “Porous media equation and Sobolev inequalities under negative curvature”, *Bull. Sci. Math.* **129** (2005), no. 10, p. 804-830.
- [11] M. A. HERRERO & M. PIERRE, “The Cauchy problem for $u_t = \Delta u^m$ when $0 < m < 1$ ”, *Trans. Am. Math. Soc.* **291** (1985), p. 145-158.
- [12] E. H. LIEB, “Sharp constants in the Hardy–Littlewood–Sobolev and related inequalities”, *Ann. Math.* **118** (1983), p. 349-374.
- [13] E. H. LIEB & M. LOSS, *Analysis*, Graduate Studies in Mathematics, vol. 14, American Mathematical Society, 1996, 278 pages.
- [14] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, 1970, xviii+451 pages.
- [15] F. SEUFFERT, “An Extension of the Bianchi–Egnell Stability Estimate to Bakry, Gentil, and Ledoux’s Generalization of the Sobolev Inequality to Continuous Dimensions”, <https://arxiv.org/abs/1512.06121>, 2016.
- [16] G. TALENTI, “Best constant in Sobolev inequality”, *Ann. Mat. Pura Appl.* **110** (1976), p. 353-372.
- [17] J. L. VÁZQUEZ, “Asymptotic behaviour for the porous medium equation posed in the whole space”, *J. Evol. Equ.* **3** (2003), no. 1, p. 67-118.
- [18] ———, *The Porous Medium Equation. Mathematical theory*, Oxford Mathematical Monographs, Oxford University Press, 2007, xxii+624 pages.