General Relative Entropy inequality for Cauchy problems preserving positivity in function spaces

Étienne Bernard

General Relative Entropy inequality for Cauchy problems preserving positivity in function spaces

Étienne Bernard

CERMICS, École des Ponts, Marne-la-Vallée, France

Abstract

The Generalized Relative Entropy inequality is a ubiquitous property in linear Cauchy problems conserving positivity of the solution over time. Yet, it is currently proved on a case-by-case basis in the literature. Here, we first prove that by considering the Cauchy problems in the framework of Riesz spaces, GRE is actually a generic consequence of a Jensen-type inequality applied to a vector-valued convex function associated to the relative entropy. Next, we extend the method to the simplest case of nonlinearity, i.e. the affine case, and we show that it also implies either GRE for a subclass of convex functions either a relaxed GRE for a larger subclass, which suggests a new avenue of research for the challenge of GRE in nonlinear problems arising in population dynamics.

Keywords: Generalized Relative Entropy, Riesz space, function space, Jensen inequality
2020 MSC: 00-01, 99-00

Declarations of interest: none

1. Introduction

Generalized Relative Entropy inequality (GRE in short) is a property that often occurs in linear Partial Differential Equations (PDE) preserving the positivity of the initial condition over time. We recall the principle of GRE: consider $\mathcal{L}$ be a (possibly unbounded) linear operator in an ambient space E, typically a (weighted) Lebesgue space, and let the associated Cauchy problem:

$$\begin{cases} \frac{\partial n}{\partial t} &= \mathcal{L}n, \\ n_{|t=0} &= n^{\text{in}}, \end{cases} \tag{1}$$

and assume that it preserves positivity over time, meaning that for any time t, the solution n of (1) is nonnegative whenever its initial condition n^{in} is nonnegative. Assume also that there exists (λ_0, ϕ, ψ) with $\lambda_0 \in \mathbb{R}$ and nonnegative functions $\phi \in E, \psi \in E^{'}$ solving the eigenproblem and its dual:

$$\begin{cases} \mathcal{L}\phi + \lambda_0\phi &= 0 \text{ in } E, \\ \mathcal{L}^*\psi + \lambda_0\psi &= 0 \text{ in } E'. \end{cases} \tag{2}$$

The Cauchy problem above is then said to satisfy GRE, if for any convex function H defined on $\mathbb{R}_+$ and provided that the integral $\int \phi(x)\psi(x)H\Big(e^{-\lambda_0 t}\frac{n(t,x)}{\phi(x)}\Big)\mu(dx)$ makes sense for any $t > 0$, the following functional, the so-called relative entropy, is nonincreasing over time:

$$t \mapsto \mathcal{H}_\psi(n\,|\,\phi) := \int \phi(x)\psi(x)H\left(e^{-\lambda_0 t}\frac{n(t,x)}{\phi(x)}\right)\mu(dx). \tag{3}$$

Email address: etienne.bernard@enpc.fr (Étienne Bernard)

Notice that by writing $u(t,x) = e^{-\lambda_0 t} n(t,x)$ and noticing that $\frac{\mathrm{d}}{\mathrm{d}t} u(t,x) = \frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-\lambda_0 t} n\big) = -\lambda_0 u(t,x) - \mathcal{L} u$ and $\frac{\mathrm{d}}{\mathrm{d}t} u(t,x) + (\lambda_0 + \mathcal{L}) u = 0$. So that, by replacing $\mathcal{L}$ by $(\lambda_0 + \mathcal{L})$, we can assume henceforth without loss of generality that $\lambda_0 = 0$. The GRE plays a fundamental role in mathematical models in biology or physics of polymerization (see [1, 2, 3, 4] and references therein) where it is used to get *a priori* estimates or long time convergence to a steady state or periodic solutions (e.g. [5, 6, 1, 7]). It has been established for various linear PDE but only on a case-by-case basis. Moreover, in a few papers such as in [7], it remains at a formal level. The present paper addresses this question by introducing a framework in which it rigorously defines the relative entropy and establishes the GRE as a generic property of positivity-preserving solutions of initial problems.

The heuristic observation is that the general relative entropy functional defined in (3) can be seen as

$$\mathcal{H}_\psi(u\,|\phi) = \langle \psi | H_\phi(u) \rangle_{E',E}$$

with the functional $H_\phi(u) := u \in E \mapsto \phi H\left(\frac{u}{\phi}\right)$ that is convex with value in E by convexity of the real valued function H. This, with the remark that for a large class of Cauchy problems, building a positive solution amounts to defining, for any $t > 0$, a positive operator T_t such that $u = T_t u^{\text{in}}$ is solution of the Cauchy problem with u^{in} as initial condition. All this together leads us to expecting a Jensen-type inequality applied to the E-valued convex function H_ϕ and the positive operator T_t:

$$H_\phi(T_t) \leqslant T_t H_\phi(u).$$

If such an inequality is valid in E, as ψ is nonnegative and as $\mathcal{L}^*\psi = 0$, which implies $T_t^*\psi = \psi$ for any $t \geqslant 0$;

$$\begin{aligned}
\mathcal{H}_\psi\big(T_t u^{\text{in}}\,|\phi\big) &= \big\langle \psi \big| H_\phi\big(T_t u^{\text{in}}\big)\big\rangle_{E',E} \\
&\leqslant \big\langle \psi \big| T_t H_\phi\big(u^{\text{in}}\big)\big\rangle_{E',E} \\
&\leqslant \big\langle T_t^*\psi \big| H_\phi\big(u^{\text{in}}\big)\big\rangle_{E',E} \\
&\leqslant \big\langle \psi \big| H_\phi\big(u^{\text{in}}\big)\big\rangle_{E',E} \\
&\leqslant \mathcal{H}_\psi\big(u^{\text{in}}\,|\phi\big).
\end{aligned}$$

Therefore, GRE would be a consequence of the expected Jensen-type inequality. A first difficulty in this approach is the definition of H_ϕ. Indeed, the division of two functions is in general not defined in any function space E except in a few rare cases such as the space of measurable functions relative to a measurable space. A way to get around the obstacle is to work first in the space of measurable functions, and then to extend the result in a function space fairly close to the first one such that the inequality can remain valid in both. As we will see, the theory of Riesz spaces provides the framework and tools for such a task, but we have to lay the foundations before stating precisely the main results.

Besides, given the importance of GRE in the study of linear models in population dynamics, some authors have established to a certain extent GRE in some nonlinear cases (e.g [8, 9, 10]). But unlike the linear case, the GRE does not always hold in general for any convex function in the nonlinear case, e.g. Remark 8 in [4]. Finding in which case and for which convex function GRE holds is a challenging task. The proof of Lemma 3.2 gives a possible solution to the problem of finding the convex functions with which the GRE can hold, or even finding some inequalities in the same spirit as GRE without being so in the strict sense; in other words, a kind of relaxed GRE. What we will see in the affine case that is the simplest form of nonlinearity.

Consequently, the paper is organized as fellows: in Section 2, we set the framework in which the problem will be rigorously set, and we then state the main result. Afterward, we establish the various consequences of the Jensen-type inequality. In Section 3, we prove the Jensen-type inequality. Eventually, in Section 4, we show how the method of proof used in Section 3 can be adapted in the affine case to establish a Jensen-type inequality for a subclass of convex functions compatible with the nonlinear structure. We will also see how to get a variant of Jensen-type inequality who will leads to a kind of relaxed GRE if we have a law of conservation.

2. Reformulation of the problem and main results

Before giving the main results, we give its framework by recalling first the notions of Riesz spaces and of function spaces, then we specify the meaning of preserving positivity property for a Cauchy problem. Afterward, we define the relative entropy functional both in L^0 and in a function space.

2.1. Function spaces

Speaking of Cauchy problems preserving positivity necessitates an ambient space endowed with a (partial) order relation compatible with its topology. The theory of Riesz spaces gives a convenient framework. The literature upon the subject is huge, a good starting point is [11, 12], but for the sake of completeness, we recall briefly its main aspects.
A vector space E over $\mathbb{R}$ endowed with an order relation $\leqslant$ is a *ordered vector space* if these axioms are satisfied:

1. $x \leqslant y \Rightarrow x+z \leqslant y+z$ for all $x, y, z \in E$,
2. $x \leqslant y \Rightarrow \lambda x \leqslant \lambda y$ for all $x, y \in E$ and $\lambda \in \mathbb{R}_+$.

An element x of E is said to be *positive* if $0 \leqslant x$ and the subset $E_+ := \{x \in E : x \geqslant 0\}$ is called the *positive cone* of E. A linear map A of E is said to be *positive* if $A(E_+ \cap D(A)) \subset E_+$. It is moreover a *vector lattice* (also called *Riesz space*) if $x \vee y := \sup\{x, y\}$ and $x \wedge y := \inf\{x, y\}$ are well-defined in E. A subset U of E is called *solid* if $0 \leqslant x \leqslant y$ for some $y \in U$ and $x \in E$ implies that $x \in U$. Every solid subspace F of E is called an (order) *ideal* in E.

Let us give an important exemple of Riesz space. Let (X, Σ, μ) be a complete $\sigma-$finite measure space, then the set $L^0(X, \Sigma, \mu)$ of all $\Sigma-$measurable $\mu-$almost everywhere finite real valued functions modulo μ-null functions endowed with the pointwise order ($f \leqslant g$ if and only if $f(x) \leqslant g(x)$ $\mu-$a.e.) is a Riesz space (see [12] p. 12). To reduce the amount of notation, we drop henceforth (X, Σ, μ) and we write only L^0 for $L^0(X, \Sigma, \mu)$. In the same spirit, in the following, the abbreviation a.e. is implicitly relative to (X, Σ, μ). We also introduce the following notation:

$$L^0_{++} := \left\{ f \in L^0 \,|\, f(x) > 0 \ \mu - \text{a.e.} \right\}$$

The Lebesgue spaces $L^p(X, \Sigma, \mu)$ $(1 \leqslant p < \infty)$ are lattice subspaces of L^0 and are also ideals of L^0. Notice that the set L^0_{++} coincides with the one of *weak order units*, i.e. the set of positive vectors e such that $|x| \wedge e = 0$ implies $x = 0$ (Definition 23.5 in [12] p.163 and remark below it, we recall that L^0 and its ideals are Archimedean, see Example 9.2 (iv) in [12] p.40)

We now underline a difficulty in the definition of relative entropy (3). Generally, we consider Cauchy problems in spaces such as Lebesgue spaces or Sobolev spaces that are "smaller" than L^0. Yet, in L^0, the division of any function by a strictly positive function give a measurable function and is well-defined in L^0. Therefore, the definition of the relative entropy functional does not pose any difficulty in L^0. The situation can be different in the case of classic spaces used in the study of Cauchy problems. Consequently, we have to work both in L^0 and in one of its subspaces, which motivates the introduction of *function space* whose the definition is recalled below:

Definition 2.1. *A function space E is simply an ideal of L^0.*

It is a classic concept, although little known: see e.g. [15] p. 194. Function spaces include L^0, Lebesgue, Orlicz, Orlicz-Lorentz and Marcinkiewicz spaces. Notice that the spaces of (bounded, vanishing at infinity, and so on) continuous functions are not ideals of L^0 and therefore are not function spaces. Henceforth, we denote for any function space E :

$$E_{++} := E \cap L^0_{++}.$$

Notice that $E_{++} \neq \emptyset$ is not a trivial statement since it implies that the support of E is X and thus that E is order dense in L^0 (see Definition 1.93 and Lemma 1.94 p.60 in [15]). That leads to the first assumption:

Assumption 1. *The ambient space E of the Cauchy problem is a function space with $E_{++} \neq \emptyset$.*

2.2. The notion of preserving positivity over time

We now specify the notion of preserving the positivity of a solution for a linear Cauchy problem. It is a well-known fact that there are many ways to define a solution to a Cauchy problem such as (1) depending on the properties of the operator $\mathcal{L}$ and the ambient space E; for instance the mild solution if it generates a C_0-semigroup ([16]), the weak solution ([17]), the very weak solution (see for instance [18]), and so on. In order to take them all into account, we consider an abstract version of theory of solutions by considering it as a family of (possibly unbounded) operators:

Definition 2.2. *A theory of solution for a given Cauchy problem in a function space E is a family of (possibly unbounded) linear operators $(F_t)_{t \in I}$ in E with $0 \in I \subseteq \mathbb{R}_+$ such that:*

1. *for any $u \in D(F_0) \subseteq E$ $F_0 u = u$,*
2. *and $F_t u \in E$ is a solution of the Cauchy problem with u as initial condition.*

We underline that we do not make any assumptions on the regularity of $t \mapsto F_t$ nor the semigroup property that is $F_{t+s} = F_t \circ F_s = F_s \circ F_t$. That explains why we have avoided the term of (semi-)flow that necessitates the notion of semigroup. That enables us to take into account for instance solutions of Cauchy problems with memory such as integro-differential equations (see for instance [19]) which have not the semigroup property. Notice also that the assumptions on I includes the discrete-time case. All we need is a positive operator in a function space.

Within this framework, speaking of a Cauchy problem preserving positivity over time in a function space E amounts to considering a theory of solutions as a family of positive operators of E. That leads to our second assumption:

Assumption 2. *Let E be a function space, let $\mathcal{L}$ be a (possibly unbounded) operator. When we tell that the associated Cauchy problem is positivity-preserving, we mean that there exists a family of positive operators $F_t : D(F_t) \mapsto E$ that is a theory of solutions for the Cauchy problem associated to $\mathcal{L}$.*

2.3. The relative entropy

We can now make precise the notion of relative entropy. We first introduce the following class of convex functions:

Definition 2.3. *A real-valued function f belongs to the set $\mathcal{C}$ if and only if*

1. *the function f is lower semi-continuous.*
2. *The nonnegative reals are included in its effective domain, i.e. $\mathbb{R}_+ \subseteq dom f := \{x \,|\, f(x) < +\infty\}$.*
3. *the function f is bounded from below by a constant on $\mathbb{R}_+$: $\exists C \in \mathbb{R} : f(x) \geqslant C \ \forall x \in \mathbb{R}_+$.*

Notice that $\mathcal{C}$ is obviously a real vector space. Moreoever, the constant functions and the classical functions in relative entropy theory such as $x \mapsto |x|,\ x \mapsto \frac{1}{2}|x|^2, x \mapsto x\ln x + x - 1$ belong to $\mathcal{C}$. We denote $\mathcal{C}_+$ the subset of nonnegative functions belonging to $\mathcal{C}$.

Let $H \in \mathcal{C}$ and let $\phi \in L^0_{++}$, then we define

$$H_\phi : u \in L^0_+ \mapsto \phi H\left(\frac{u}{\phi}\right). \tag{4}$$

As $\phi > 0$ a.e., $\frac{1}{\phi}$ is finite a.e. and measurable, which implies that $\frac{u}{\phi}$ is measurable for any measurable u as pointwise product of measurable functions. As $H \in \mathcal{C}$, it is continuous on $\mathbb{R}_+$ (see Corollary 2.5 p. 13 in [21]), thus $H\left(\frac{u}{\phi}\right)$ is measurable and so is $\phi H\left(\frac{u}{\phi}\right)$ as pointwise product of measurable functions. Therefore H_ϕ is a well-defined functional on L^0_+ with value in L^0.

Now let E be a function space, and we introduce:

$$D\big(H^E_\phi\big) := \{u \in E_+ \,|H_\phi(u) \in E\,\},$$

and the map in E:

$$H^E_\phi : u \in D\big(H^E_\phi\big) \mapsto H_\phi(u).$$

Obviously, $\mathbb{R}_+\phi \subset D\left(H^E_\phi\right)$, thus $D\left(H^E_\phi\right) \neq \emptyset$. Moreover, the convexity of H implies the one of $D\left(H^E_\phi\right)$. We can now define for every $\psi \in E'$ and $\phi \in E_{++}$ the relative entropy in E:

$$\mathcal{H}_\psi(\cdot\,|\phi) : u \in D\big(H^E_\phi\big) \mapsto \big\langle \psi \big| H^E_\phi(u)\big\rangle_{E',E}$$

2.4. The main results

Before giving the main results, we underline that the GRE stated in (3) can have two different interpretations in the new framework. Either it means a control by the initial condition in the following sense:

$$\text{for any } u \in D\big(H^E_\phi\big) \text{ s.t. } F_t u \in D\big(H^E_\phi\big),\ \mathcal{H}_\psi(F_t u\,|\phi) \leqslant \mathcal{H}_\psi(u\,|\phi), \tag{5}$$

for a family of positive operators that is a theory of solution for a Cauchy problem. Or the functional is nonincreasing meaning that for any $t, s \geqslant 0$ s.t. $0 \leqslant s \leqslant t$ and for any $u \in D\left(H^E_\phi\right)$ such that $F_t u, F_s u \in D\left(H^E_\phi\right)$:

$$\mathcal{H}_\psi(F_t u\,|\phi) \leqslant \mathcal{H}_\psi(F_s u\,|\phi). \tag{6}$$

for a family of positive operators that is a theory of solution for a Cauchy problem. That is the most used in literature as in [6]. Obviously, if the theory of solutions $(F_t)_{t\in I}$ is in fact a semi-flow, hence having the semigroup property, then (6) is a consequence of (5). Here, we aim to establish GRE in the sense (5) and therefore, it is enough to find which conditions on an unbounded positive operator A in a function space E and on $(\phi,\psi) \in E \times E'$ entail for any $u \in D\left(H^E_\phi\right) \cap D(A)$ such that $Au \in D\left(H^E_\phi\right)$:

$$\mathcal{H}_\psi(Au\,|\phi) \leqslant \mathcal{H}_\psi(u\,|\phi). \tag{7}$$

Our first result is an abstract Jensen-type inequality applied to relative entropy:

Theorem 2.4. *Let E be a function space, and let A be a positive (possibly unbounded) linear operator. Let $H \in \mathcal{C}$, and $\phi \in E_{++}$ such that there exists $\tilde{\phi} \in D(A) \cap E_{++}$ with $\phi = A\tilde{\phi}$. Then we have the following stability property:*

$$\textit{for any } u \in D(A) \cap D\left(H^E_{\tilde{\phi}}\right) \textit{ such that } H_{\tilde{\phi}}(u) \in D(A),\ \textit{ we have } Au \in D\big(H^E_\phi\big),$$

and the Jensen-type inequality:

$$\textit{for any } u \in D(A) \cap D\left(H^E_{\tilde{\phi}}\right) \textit{ such that } H_{\tilde{\phi}}(u) \in D(A),\ H^E_\phi(Au) \leqslant AH^E_{\tilde{\phi}}(u).$$

Let's make several remarks.

1. As we work in both L^0 and E where the order relation is the same, we have to be careful about the context of any inequality between two terms: does it happen in L^0 or in E ? If both functions are nonnegative, then by Definition 2.1, if the upper function is in E then the lower function is also in E. But if the sign of the lower function is unknown, then we can't deduce in general that the inequality holds in E even if the upper function is in E. That explains why in (7) we have written as condition $Au \in D\left(H_\phi^E\right)$. But that condition is dropped in Theorem 2.4 because, as we will show in the proof of Lemma 3.2, the existence of a lower bound to H and the presence ϕ and $\tilde{\phi}$ as pivot functions imply that $H_\phi(Au) \in E$ provided that $AH_{\tilde{\phi}}(u) \in E$, no matter the sign of $H_\phi(Au)$.

2. It is not the first Jensen-type result in Banach lattice theory, see [22] and especially [13]. The main difference with these papers is that we have established it for a vector-valued convex function H_ϕ whose "pivot" functions enables us to consider a larger class of real-valued functions and positive operators than in [22, 13](see the proof of Lemma 3.2).

3. As $f \mapsto \langle \psi | f \rangle_{E',E}$ is a positive linear form for any $\psi \in E'_+$, Theorem 2.4 implies that with the same assumptions that for any $\psi \in E'_+ \cap D(A^*)$:

$$\forall u \in D(A) \text{ such that } H_{\tilde{\phi}}(u) \in D(A),\ \langle \psi | H_\phi(Au) \rangle \leqslant \left\langle A^*\psi \middle| H_{\tilde{\phi}}(u) \right\rangle_{E',E}$$

And therefore, we get the second main result:

Corollary 2.5 (Abstract GRE)**.** *Let E be a function space, and let A be a positive (possibly unbounded) linear operator. Let $H \in \mathcal{C}$, and $\phi \in E_{++}$ such that there exists $\tilde{\phi} \in D(A) \cap E_{++}$ with $\phi = A\tilde{\phi}$. Then for any $\psi \in E'_+ \cap D(A^*)$:*

$$\forall u \in D(A) \text{ such that } H_{\tilde{\phi}}(u) \in D(A),\ \mathcal{H}_\psi(Au \,|\phi) \leqslant \mathcal{H}_{A^*\psi}\left(u \,\middle|\tilde{\phi}\right)$$

Which implies immediately GRE in the form (7) when we have moreover $A^*\psi = \psi$ and $A\phi = \phi$.

4. When $\mathcal{L}$ is generator of a C_0-semigroup $(T_t)_{t\in\mathbb{R}_+}$, then the solution $(\phi, \psi) \in E \times E'$ to the eigenproblem (2) with $\lambda_0 = 0$ satisfies $T_t\phi = \phi$ and $T_t^*\psi = \psi$ and thus Corollary 2.5 implies immediately GRE for the semigroup in the classical sense (5), thus the third main result:

Corollary 2.6 (GRE for positive C_0-semigroups)**.** *Let E be a function space, and an unbounded operator $\mathcal{L}$ being generator of a positive C_0-semigroup $(T_t)_{t\in\mathbb{R}_+}$ such that there exists nonnegative functions $(\phi, \psi) \in E \times E'$ being solutions of*

$$\begin{cases} \mathcal{L}\phi &= 0 \text{ in } E, \\ \mathcal{L}^*\psi &= 0 \text{ in } E'. \end{cases}$$

Then for any $u \in D\left(H_\phi^E\right)$ and any $t \geqslant 0$, we have $T_t u \in D\left(H_\phi^E\right)$ and

$$t \mapsto \mathcal{H}_\psi(T_t u \,|\phi) \text{ is nonincreasing.}$$

5. The main results presented above can be generalized in several directions.

   - We can retrieve the main results with $L^0(X, \Sigma, \mu)$ replaced by $C(X), C_b(X)$ or $C_0(X)$ that are Riesz spaces ([12] p.14) provided that we assume moreover that $D(H_\phi)$ is stable for A, since the spaces evoked above are not ideals of L^0.

- In Definition 2.3, we have considered convex functions whose effective domains include $\mathbb{R}_+$, but we can replace $\mathbb{R}_+$ by any interval $I \subseteq \mathbb{R}_+$ keeping the main results provided that we assume moreover that the linear operators satisfy the following property

$$\forall u \text{ such that } \forall x, u(x) \in I \Rightarrow \forall x, Au(x) \in I.$$

  That enables us to extend for instance the main results to the Kermack-McKendrick function $x \mapsto x - \ln x$ with $I = \mathbb{R}_+^*$ for linear Cauchy problem preserving the **strict** positivity.
- It can also be extended to stochastic partial equations preserving positivity such as the Black-Scholes equation.

3. A Jensen-type theorem for linear positive operators

The present section is devoted to the proof of Theorem 2.4. We first recall that for any $H \in \mathcal{C}$ and any $\phi \in L^0_{++}$, the L^0-valued functional H_ϕ set as in (4) is well-defined. The heuristic remark that motivates Theorem 2.4 is that by $\phi \in E_+$ and convexity of H, we have for any $u, v \in E_+$ and any $\lambda \in [0, 1]$:

$$H_\phi(\lambda u + (1 - \lambda)v) \leqslant \lambda H_\phi(u) + (1 - \lambda)H_\phi(v).$$

In other words, the L^0-valued function H_ϕ is convex on L^0_+. We therefore expect that it shares some properties with real-valued convex functions. The lemma below extends to it a well-known property of real-valued l.s.c. convex functions, i.e. being the pointwise supremum of affine functions:

Lemma 3.1. *Let $H \in \mathcal{C}$, and let $\phi \in L^0_{++}$ and let H_ϕ be defined on L^0 as in* (4)*. Then, there exists $\mathcal{F} \subset \mathbb{R}^2$ such that for any $u \in L^0$*

$$H_\phi(u(x)) = \sup_{(a,b)\in\mathcal{F}} a_n u(x) + b_n \phi(x),$$

Proof. As $H \in \mathcal{C}$, it is equal to the pointwise supremum of all its lower affine functions by Proposition 3.1 p.14 in [21]. Therefore, there exists $\equiv \mathcal{F} \subset \mathbb{R}^2$ such that

$$H(x) = \sup_{(a,b)\in\mathcal{F}} ax + b.$$

That implies that:

$$H\left(\frac{u(x)}{\phi(x)}\right) = \sup_{(a,b)\in\mathcal{F}} a\frac{u(x)}{\phi(x)} + b$$

That means that whatever $(a, b) \in \mathcal{F}$, we have. $a\frac{u(x)}{\phi(x)} + b \leqslant H\left(\frac{u(x)}{\phi(x)}\right)$ and for μ-a.e x there exists a sequence $(a_n, b_n) \in \mathcal{F}$ such that $a_n\frac{u(x)}{\phi(x)} + b_n \uparrow H\left(\frac{u(x)}{\phi(x)}\right)$. As ϕ is nonnegative, we have for any x $a_n u(x) + b_n\phi(x) \leqslant \phi(x)H\left(\frac{u(x)}{\phi(x)}\right)$ and for any x, for the same sequence above $a_n u(x) + b_n\phi(x) \uparrow \phi(x)H\left(\frac{u(x)}{\phi(x)}\right)$. Therefore there exists $\mathcal{F} \subset \mathbb{R}^2$ such that

$$H_\phi(u(x)) = \sup_{(a,b)\in\mathcal{F}} au(x) + b\phi(x),$$

which is the desired conclusion. □

We can now prove a Jensen-type inequality:

Lemma 3.2. *Let $H \in \mathcal{C}$ and $\phi \in E_{++}$. Let H_ϕ defined as in* (4)*. Let E be a function space. Let A be a positive linear map of E such that there exists $\tilde{\phi} \in E_{++}$ with $A\tilde{\phi} = \phi$ Then we have the Jensen-type following inequality: for any $u \in D(A) \cap L^0_+$ such that $H_{\tilde{\phi}}(u) \in D(A)$,*

$$H_\phi(Au) \leqslant A\Big(H_{\tilde{\phi}}(u)\Big) \text{ both in } L^0 \text{ and in } E.$$

Proof. We know that by Lemma 3.1, there exists a subset $\mathcal{F}$ of $\mathbb{R}^2$ such that

$$H_\phi(u(x)) = \sup_{(a,b)\in\mathcal{F}} au(x) + b\phi(x). \tag{8}$$

We first consider A as a positive unbounded operator defined on L^0 with value in L^0. We have then for any $u \in D(A) \cap L^0_+$,

$$\begin{aligned} H_\phi(Au)(x) &= \sup_{(a,b)\in\mathcal{F}} a(Au(x)) + b\phi(x), \\ &= \sup_{(a,b)\in\mathcal{F}} a(Au(x)) + b\Big(A\tilde{\phi}\Big)(x), \\ &= \sup_{(a,b)\in\mathcal{F}} A\Big(au + b\tilde{\phi}\Big)(x) \text{ (by linearity of } A), \\ &\leqslant A\left(\sup_{(a,b)\in\mathcal{F}} \Big(au + b\tilde{\phi}\Big)\right) \text{ (by order-preserving property of } A), \\ &= A\Big(H_{\tilde{\phi}}(u)\Big) \text{ (by (8)).} \end{aligned} \tag{9}$$

That implies the first inequality in L^0. Now, we show that it holds also in E. Let $u \in D\Big(H^E_{\tilde{\phi}}\Big) \cap E_+$ such that $H^E_{\tilde{\phi}}(u) \in D(A)$. As A is also an unbounded operator of L^0 with values in L^0 we have by (9) :

$$H_\phi(Au) \leqslant A\Big(H_{\tilde{\phi}}(u)\Big) \text{ in } L^0.$$

First assume that $H \in \mathcal{C}_+$, then since $Au \in E_+$ by positivity of A and since $u \in D(A)$ with $H^E_{\tilde{\phi}}(u) \in D(A)$ and $\phi \in E$, we have:

$$0 \leqslant H_\phi(Au) \leqslant A\Big(H_{\tilde{\phi}}(u)\Big) \text{ in } L^0,$$

with $A\Big(H_{\tilde{\phi}}(u)\Big) \in E$. That implies, since E is by definition an ideal of L^0, that $H_\phi(Au) \in E$ and therefore

$$0 \leqslant H_\phi(Au) \leqslant A\Big(H_{\tilde{\phi}}(u)\Big) \text{ in } E.$$

Now consider the general case: $H \in \mathcal{C}$. There exists a constant C such that $H + C \in \mathcal{C}_+$ and we can take $C \geqslant 0$. Notice that for any $\forall u \in L^0$,

$$(H + C)_\phi(u) = H_\phi(u) + C\phi.$$

Therefore, as $\phi \in E$, we have $u \in D\Big(H^E_\phi\Big)$ if and only if $u \in D\Big((H+C)^E_\phi\Big)$. As $\tilde{\phi} \in D(A)$ and $H_{\tilde{\phi}}(u) \in D(A)$, we have $(H+C)_{\tilde{\phi}}(u) \in D(A)$ and thus by the first case

$$0 \leqslant (H+C)_\phi(Au) \leqslant A(H+C)_{\tilde{\phi}}(u) \text{ in } L^0 \text{ and in } E \tag{10}$$

That implies by the same argument as in the case $H \in \mathcal{C}_+$ that:

$$H_\phi(Au) + C\phi \in E,$$

thus, since $\phi \in E$:

$$H_\phi(Au) \in E.$$

Consequently, $Au \in D\Big(H^E_\phi\Big)$ and by (10):

$$H_\phi(Au) \leqslant A\Big(H_{\tilde{\phi}}(u)\Big) \text{ in } E,$$

which is the desired conclusion. □

Remark 3.3. *1. The proof is reminiscent of a classical way of proving Jensen inequality that is often taught at university [23] (Theorem 1.8.1 p. 45 or Theorem 3.6.4. p.125) And that trick of convex function as supremum of affine functions has been used in [24] (see Proposition 3.4 p. 82) and in [25] to extend the classical Jensen inequality to (possibly infinite) measure space provided that H is moreover positively homogeneous.*

2. Lemma 3.2 can be seen as a generalization of $(i) \Rightarrow (ii)$ in Proposition 1 in [22] or Proposition 3.4 in [24]. We underline that in [22, 24], the authors have assumed that the operator is (sub)markovian, meaning that $AC \leqslant C$ for any constant C in order to handle the term b in (9). In our work, the markovianity is dropped as assumption thanks to the presence of the terms ϕ and $\tilde{\phi}$ as pivot functions as showed in the proof above.

4. A Jensen-type theorem for relative entropy in the affine case

Given the importance of GRE in mathematical models in biology of population, some authors have tried to establish it in a few nonlinear Cauchy problems. However, it presents a number of challenges. One of which is the impossibility of defining an adjoint operator for a nonlinear one. In order to have GRE from a Jensen-inequality for a given nonlinear operator A in the same spirit as in Theorem 2.4, we have to assume moreover a law of conservation, i.e. a positive vector $\psi \in E'$ such that:

$$\text{For any } \phi \in E_+,\ \langle\psi|A\phi\rangle = \langle\psi|\phi\rangle_{E',E}\,.$$

That being said, we will focus on the Jensen-type inequality and its variants. More precisely, we will see how the proof of Lemma 3.2 gives a method to find the convex functions with which the GRE can hold in its strict form or relaxed in the affine case.

Let's first explain the approach. Let's consider in (1) the nonhomogeneous case, i.e. when the map $\mathcal{L}$ is an affine operator, i.e. $\mathcal{L} : u \mapsto Au + f$:

$$\begin{cases} \frac{\partial}{\partial t}u &= Au + f, \\ u_{|t=0} &= u^{\text{in}}, \end{cases} \tag{11}$$

with A is a (possibly unbounded) linear operator and a prescribed function f with values in E, called the *source term*. Assume moreover that A is a generator of a C_0-semigroup T_t and assume that $f \in L^1(\mathbb{R}_+, E)$ then a function $u \in C(\mathbb{R}_+; E)$ is a mild solution of 11 if

$$u(t) = T_t u^{in} + A\int_0^t u(s)ds + \int_0^t T_{t-s}f(s)ds,$$

(see [16] p. 451 or [26]). We can thus define for any $t \geqslant 0$:

$$F_t : u^{\text{in}} \mapsto T_t u^{in} + A\int_0^t u(s)ds + \int_0^t T_{t-s}f(s)ds.$$

It is obviously an affine map with $u^{\text{in}} \mapsto T_t u^{in} + A\int_0^t u(s)ds$ as linear part and $\int_0^t T_{t-s}f(s)ds$ as fixed part. Notice that the linear part is a positive operator and that as f is nonnegative, the fixed part is also a nonnegative function. Therefore, in the same spirit as in the linear case, we can consider a GRE as a consequence of a Jensen-type theorem for affine map both in L_0 and in a function space E. But, as we will see, the nonlinear feature is an additional constraint on the class of admissible convex function. We introduce the following

Definition 4.1. *Let $K > 0$. A real-valued function f belongs to the set $\mathcal{C}[K]$ if and only if*

1. *the function f is nonnegative.*
2. *the function f belongs to the set $\mathcal{C}$.*

3. There exists $\mathcal{F} \subset \left\{(a,b) \in \mathbb{R}^2 \,|a+b \leqslant K\right\}$ such that

$$\text{for any } x \in \mathbb{R}_+,\ H(x) = \sup_{(a,b)\in\mathcal{F}} ax + b.$$

Remark 4.2. *The first assumption that is the nonnegativity is to avoid some technicalities, as the aim of the present section is to show the possible extension to the nonlinear case of the method exposed in Section 3.*

We can now state the main result:

Proposition 4.3. *Consider a function space E and the affine operator $A = A_0 + f$ with A_0 being a positive unbounded linear operator and $f \in E_+$. Let $\phi \in E_{++}$ such there exists $\tilde{\phi} \in E_{++} \cap D(A)$ with $\phi = A\tilde{\phi}$. Then for any $H \in \mathcal{C}[K]$, any $u \in D(A)$ such that $H_{\tilde{\phi}}(u) \in E$ we have:*

$$H_\phi(Au) \in E,$$

and

$$H_\phi(Au) \leqslant (K \vee 1)A\Big(H_{\tilde{\phi}}(u)\Big) \text{ both in } L^0 \text{ and } E.$$

That implies

Corollary 4.4. *With the same assumptions as in Proposition 4.3 and assume moreover that the affine operator has $\psi \in E'$ as law of conservation, then:*

$$\mathcal{H}_\psi(Au\,|\phi) \leqslant (K \vee 1)\mathcal{H}_\psi\left(u \,\middle|\tilde{\phi}\right).$$

When $K \leqslant 1$, we retrieve GRE in its classical form. Otherwise, we get a kind of relaxed GRE, one of which we can still deduce some *a priori* estimates. And provided a careful study of the multiplicative constant over time, the relaxed GRE can be useful for the study of the long-time behavior. That suggests a program of research in the nonlinear case: not only look for some convex functions for which GRE holds, but also variants of GRE that can still be useful.

That being said, we now prove Proposition 4.3

Proof. By Lemma 3.1, we have

$$\begin{aligned}
H_\phi(Au)(x) &= \sup_{a,b\in\mathcal{F}} \phi(x)a\frac{Au(x)}{\phi(x)} + b,\\
&= \sup_{a,b\in\mathcal{F}} aAu(x) + b\phi(x),\\
&= \sup_{a,b\in\mathcal{F}} aA_0u(x) + bA_0\tilde{\phi}(x) + (a+b)f(x),\\
&\leqslant \sup_{a,b\in\mathcal{F}} A_0\left(\tilde{\phi}\left(a\frac{u}{\tilde{\phi}} + b\right)\right)(x) + Kf(x), \text{ since } H \in \mathcal{C}[K],\\
&\leqslant (K \vee 1)\Big(A_0\Big(H_{\tilde{\phi}}(u)\Big)(x) + f(x)\Big) \text{ by positivity of } A_0,\\
&\leqslant (K \vee 1)A\Big(H_{\tilde{\phi}}(u)\Big)(x) \text{in } L^0.
\end{aligned}$$

As $H \in \mathcal{C}[K]$, it is nonnegative, and thus since $u \in E_+$ and $Au \in E_+$, we have

$$0 \leqslant H_\phi(Au) \leqslant (K \vee 1)A\Big(H_{\tilde{\phi}}\Big) \text{ in } L^0$$

with $A\Big(H_{\tilde{\phi}}\Big) \in E_+$, which implies that, E being an ideal of L^0, that $H_\phi(Au) \in E$ and thus the inequality above is also valid in E. Whence the desired result. □

Acknowledgements

The author is grateful to Benoît Perthame for bringing to his attention the subject of Generalized Relative Entropy inequality and for useful discussions. The author also thank Pierre Gabriel for useful discussions and Tony Lelièvre for also useful discussions and for having pinpointed [4].

References

[1] B. Perthame, Transport Equations in Biology: Benoit Perthame., Frontiers in Mathematics, Birkhäuser, Basel, 2007.

[2] B. Perthame, L. Ryzhik, Exponential decay for the fragmentation or cell-division equation, Journal of Differential Equations 210 (1) (2005) 155–177. doi:10.1016/j.jde.2004.10.018.

[3] J. Calvo, M. Doumic, B. Perthame, Long-Time Asymptotics for Polymerization Models, Communications in Mathematical Physics 363 (1) (2018) 111–137. doi:10.1007/s00220-018-3218-5.

[4] B. Jourdain, C. Le Bris, T. Lelièvre, F. Otto, Long-Time Asymptotics of a Multiscale Model for Polymeric Fluid Flows, Archive for Rational Mechanics and Analysis 181 (1) (2006) 97–148. doi:10.1007/s00205-005-0411-4.

[5] P. Michel, S. Mischler, B. Perthame, General entropy equations for structured population models and scattering, Comptes Rendus Mathematique 338 (9) (2004) 697–702. doi:10.1016/j.crma.2004.03.006.

[6] P. Michel, S. Mischler, B. Perthame, General relative entropy inequality: An illustration on growth models, Journal de Mathématiques Pures et Appliquées 84 (9) (2005) 1235–1260. doi:10.1016/j.matpur.2005.04.001.

[7] É. Bernard, M. Doumic, P. Gabriel, Cyclic asymptotic behaviour of a population reproducing by fission into two equal parts, Kinetic & Related Models 12 (3) (2019) 551. doi:10.3934/krm.2019022.

[8] P. Michel, General relative entropy in a nonlinear McKendrick model, in: G.-Q. Chen, E. Hsu, M. Pinsky (Eds.), Stochastic Analysis and Partial Differential Equations, Vol. 429 of Contemporary Mathematics, American Mathematical Soc., 2007, pp. 205–232. doi:10.1090/conm/429/08238.

[9] P. Michel, B. K. Kakumani, GRE methods for nonlinear model of evolution equation and limited ressource environment, Discrete & Continuous Dynamical Systems - B 24 (12) (2019) 6653. doi:10.3934/dcdsb.2019161.

[10] P. Gabriel, Long-time asymptotics for nonlinear growth-fragmentation equations, Communications in Mathematical Sciences 10 (3) (2012) 787–820. doi:10.4310/CMS.2012.v10.n3.a4.

[11] H. H. Schaefer, Banach Lattices and Positives Operators, 1st Edition, Die Grundlehren Der Mathematischen Wissenschaften, Springer, Berlin, Heidelberg, 1974. doi:10.1007/978-3-642-65970-6_1.

[12] A. C. Zaanen, Introduction to Operator Theory in Riesz Spaces, Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. doi:10.1007/978-3-642-60637-3.

[13] M. Haase, Convexity Inequalities for Positive Operators, Positivity 11 (1) (2007) 57–68. doi:10.1007/s11117-006-1975-4.

[14] A. C. Zaanen, Integration, completely revised edition Edition, North-Holland Wiley, Amsterdam New York, 1967.

[15] Y. Abramovich, C. Aliprantis, An Invitation to Operator Theory, Vol. 50 of Graduate Studies in Mathematics, American Mathematical Soc., 2002.

[16] K.-J. Engel, R. Nagel (Eds.), One-Parameter Semigroups for Linear Evolution Equations, 1st Edition, Graduate Texts in Mathematics, Springer, New York, NY, 2000. `doi:10.1007/0-387-22642-7_2`.

[17] L. C. Evans, Partial Differential Equations, American Mathematical Soc., 2010.

[18] C. Garetto, M. Ruzhansky, Hyperbolic Second Order Equations with Non-Regular Time Dependent Coefficients, Archive for Rational Mechanics and Analysis 217 (1) (2015) 113–154. `doi:10.1007/s00205-014-0830-1`.

[19] J. Prüss, Evolutionary Integral Equations and Applications, Vol. 87 of Monographs in Mathematics, Birkhäuser Basel, Basel, 1993. `doi:10.1007/978-3-0348-8570-6`.

[20] J. K. Hale, S. M. V. Lunel, Introduction to Functional Differential Equations, Vol. 99 of Applied Mathematical Sciences, Springer New York, New York, NY, 1993. `doi:10.1007/978-1-4612-4342-7`.

[21] I. Ekeland, R. Témam, Convex Analysis and Variational Problems, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1999. `doi:10.1137/1.9781611971088`.

[22] W. Arendt, P. Benilán, lnégalités de Kato et Semi-Groupes Sous- Markoviens, Revista Matemática Complutense (1992) 279–279`doi:10.5209/rev_REMA.1992.v5.n2.17914`.

[23] C. P. Niculescu, L.-E. Persson, Convex Functions and Their Applications - A Contemporary Approach, CMS Books in Mathematics, Springer International Publishing, Cham, 2018. `doi:10.1007/978-3-319-78337-6_1`.

[24] E.-M. Ouhabaz, Analysis of Heat Equations on Domains. (LMS-31), Princeton University Press, 2009. `doi:10.1515/9781400826483`.

[25] P. Roselli, M. Willem, A Convexity Inequality, The American Mathematical Monthly 109 (1) (2002) 64–70. `doi:10.2307/2695768`.

[26] J. M. Ball, Strongly Continuous Semigroups, Weak Solutions, and the Variation of Constants Formula, Proceedings of the American Mathematical Society 63 (2) (1977) 370–373. `doi:10.2307/2041821`.