ENTROPY DISSIPATION METHODS FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

ANSGAR JÜNGEL
INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING
VIENNA UNIVERSITY OF TECHNOLOGY, AUSTRIA

Lecture Notes, Spring School, Bielefeld, March, 2012

Contents

1. Motivation 2
2. Entropies 4
  2.1. Definitions 5
  2.2. The heat equation revisited 6
  2.3. The homogeneous Boltzmann equation 8
3. Fokker-Planck equations 10
  3.1. Relaxation to self-similarity 11
  3.2. The Fokker-Planck equation and logarithmic Sobolev inequality 13
  3.3. Nonlinear Fokker-Planck equations 17
4. Further applications 22
  4.1. Systematic entropy construction method 22
  4.2. Entropy variables and cross-diffusion systems 29
5. Summary and open problems 33
References 35

Date: March 23, 2012.

1. Motivation

Entropy dissipation methods have been developed recently to investigate the qualitative behavior of solutions to nonlinear partial differential equations (PDEs) and to derive explicit or even optimal constants in functional inequalities. The entropy was introduced by Rudolf Clausius in 1865 as a state function in thermodynamics. Later, Ludwig Boltzmann, Josiah W. Gibbs, and James C. Maxwell gave a statistical physics interpretation of entropy. In particular, Ludwig Boltzmann defined in 1877 the entropy of a system, e.g. consisting of ideal gas particles, to be proportional to the logarithm of the number of micro-states of the system. Claude Shannon developed in 1948 a concept of information entropy measuring information, choice, and uncertainty in order to quantify the statistical nature of phone-line signals.

The notion of entropy plays a fundamental role also in PDE theory. Loosely speaking, an entropy (in the mathematical sense) is a quantity (Lyapunov functional) which is non-increasing along the trajectories of an evolution equation. The entropy dissipation is the negative time derivative of the entropy. The concept of entropy was extended by Lax in 1973 to hyperbolic conservation laws [31] and by DiPerna in 1985 [17] to the framework of compensated compactness. In kinetic theory, the entropy provides a priori estimates which were used for an existence analysis (DiPerna-Lions 1989 [18]) and for compactness results in hydrodynamic limits (Bardos-Golse-Levermore 1993 [3], Golse-Levermore 2005 [22], Golse-Saint-Raymond 2004 [23]). The Boltzmann entropy is employed to derive some information about the long-time behavior of the solutions to the Boltzmann equation and their decay rates (Desvillettes-Villani 2001 [16]). In particular, connections to logarithmic Sobolev inequalities (Gross 1975 [24], Del Pino-Dolbeault [14]) and to stochastic diffusion processes (Bakry-Emery 1983) were discovered. The stochastic ansatz was re-interpreted by Toscani in 1997 [39] for kinetic Fokker-Planck diffusion using the notions of entropy and entropy dissipation.

The goal of these lecture notes is to introduce some aspects of entropy dissipation methods which give insight in the structure of nonlinear PDEs and the qualitative behavior of their solutions. In order to understand the idea of the methods, we consider first a simple example, the heat equation

$$u_t = \Delta u, \quad u(\cdot,0) = u_0 \ge 0 \quad \text{in } \mathbb{T}^d,\ t > 0,$$

where $\mathbb{T}^d$ is the d-dimensional torus. It is well known that for integrable nonnegative initial data u_0, there exists a smooth nonnegative solution satisfying $\int_{\mathbb{T}^d} u(x,t)dx = \int_{\mathbb{T}^d} u_0(x)dx =: \overline{u}$ for all $t > 0$. We normalize the initial mass by setting $\overline{w} = \overline{u}/\text{meas}(\mathbb{T}^d)$. For simplicity, we write $u(t) = u(\cdot,t)$. Then $u(t)$ is a function depending on the spatial variable, $u(t) : \mathbb{T}^d \to \mathbb{R}$. We introduce the following functionals:

$$H_1[u] = \int_{T^d} u \log\Big(\frac{u}{\overline{w}}\Big)dx, \quad H_2[u] = \frac{1}{2}\int_{T^d} (u-\overline{u})^2 dx.$$

Observe that both functions are nonnegative. Indeed, the elementary inequality $\log z + 1/z - 1 \ge 0$ for all $z > 0$ implies that, taking $z = u/\overline{w}$,

$$0 \le \int_{\mathbb{T}^d} \Big(u \log\Big(\frac{u}{\overline{w}}\Big) + \overline{w} - u\Big)dx = \int_{\mathbb{T}^d} u\log\Big(\frac{u}{\overline{w}}\Big)dx + \int_{\mathbb{T}^d} \overline{w}dx - \int_{\mathbb{T}^d} udx = H_1.$$

We claim that H_1 and H_2 are both Lyapunov functionals along the solutions of the heat equation. First, we consider H_2. By integration by parts, we find that

$$\frac{dH_2}{dt}[u(t)] = \int_{\mathbb{T}^d}(u-\overline{u})u_t dx = \int_{\mathbb{T}^d}(u-\overline{u})\Delta u dx = -\int_{\mathbb{T}^d}|\nabla u|^2 dx \le 0, \tag{1.1}$$

and thus, H_2 is a Lyapunov functional along solutions to the heat equation. The expression on the right-hand side is, up to the sign, the dissipation of the entropy H_2. This term allows us to deduce more than just the monotonicity of H_2. For this, we need the Poincaré inequality

$$\|u-\overline{u}\|^2_{L^2(\Omega)} \le C_P\|\nabla u\|^2_{L^2(\Omega)} \quad \text{for all } u \in H^1(\Omega), \tag{1.2}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain. The Poincaré constant C_P is the inverse of the first non-vanishing eigenvalue of the Laplace operator with homogeneous Neumann boundary conditions (Dautray-Lions 1988 [12], Corollary 3, p. 131). For some domains, the constant C_P can be determined explicitly or can at least be estimated. For instance, for bounded convex domains, $C_P \le C(d)\mathrm{diam}(\Omega)/\mathrm{meas}(\Omega)$ with $C(d) > 0$ only depending on the space dimension $d \ge 3$ (Dautray-Lions [12], Proposition 3, p. 132); for bounded convex domains with Lipschitz boundary, $C_P \le \mathrm{diam}(\Omega)/\pi$ (Payne-Weinberger 1960 [34]; Bebendorf 2003 [4]); for $\Omega = \mathbb{T}^d$ (with unit measure), $C_P = 1/(2\pi)$. The Poincaré inequality helps to relate the entropy H_2 to the entropy dissipation. Indeed, combining (1.1) and (1.2), we infer that

$$\frac{dH_2}{dt}[u(t)] = -\|\nabla u\|^2_{L^2(\mathbb{T}^d)} \le -C_P^{-1}\|u-\overline{u}\|^2_{L^2(\Omega)} = -2C_P^{-1}H_2.$$

By the Gronwall inequality (or just integrating this differential inequality),

$$\|u(t)-\overline{u}\|^2_{L^2(\mathbb{T}^d)} = H_2[u(t)] \le H_2[u_0]e^{-2t/C_P}, \quad t > 0. \tag{1.3}$$

Hence, the solution of the heat equation converges in the L^2 norm exponentially fast to the steady state $\overline{u}$ with explicit rate $1/C_P$.

Remark 1.1. This result is not surprising. Indeed, by semigroup theory, we can write u as the series

$$u(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t}(u_0, v_k)_{L^2} v_k,$$

where v_k is the k-th (normalized) eigenfunction to $-\Delta$ with periodic boundary conditions, λ_k is the corresponding eigenvalue with increasing $\lambda_k \to \infty$ as $k \to \infty$, and $(\cdot,\cdot)_{L^2}$ is the L^2 scalar product. The first eigenfunction v_1 is constant and the corresponding eigenvalue λ_1 vanishes. Therefore, since $\overline{u} = (u_0, v_1)_{L^2} v_1$,

$$\|u(t)-\overline{u}\|^2_{L^2(\Omega)} = \sum_{k=2}^{\infty} e^{-2\lambda_k t}(u_0, v_k)^2_{L^2} \le e^{-2\lambda_2 t}\|u_0\|^2_{L^2(\Omega)}.$$

The convergence rate $\lambda_2 = 1/C_P$ is the same as in (1.3). □

The strength of entropy dissipation methods is that such decay properties can be derived in other "norms" too which might be less acessible to semigroup theory and that nonlinear equations can be treated as well. We illustrate the first statement by computing the derivative of H_1:

$$\frac{dH_1}{dt}[u(t)] = \int_{\mathbb{T}^d} \Big(\log\Big(\frac{u}{w}\Big)+1\Big)u_t dx = -\int_{\mathbb{T}^d} \nabla\Big(\log\Big(\frac{u}{w}\Big)+1\Big)\cdot\nabla u dx = -4\int_{\mathbb{T}^d}|\nabla\sqrt{u}|^2 dx. \tag{1.4}$$

Again, we need an expression relating the entropy H_1 and the entropy dissipation. This is phrased by the logarithmic Sobolev inequality (which will be proven below, see Corollary 3.6 and the following comments)

$$\int_\Omega u\log\frac{u}{w}dx \le C_L\int_\Omega|\nabla\sqrt{u}|^2dx \quad \text{for all } \sqrt{u}\in H^1(\Omega),\ u\ge 0,$$

where $\Omega\subset\mathbb{R}^d$ is a bounded domain. If $\Omega=\mathbb{T}$ (with unit measure), the constant C_L equals $1/(2\pi^2)$ (Rothaus 1980 [35], Weissler 1980 [41], Dolbeault-Gentil-Jüngel 2006 [19]). This shows that

$$\frac{dH_1}{dt}[u(t)] \le 4C_L^{-1}H_1 \quad \text{and} \quad H_1[u(t)]\le H_1[u_0]e^{-4t/C_L}, \quad t>0.$$

The solution converges in the "norm" of H_1 exponentially fast to its constant steady state with rate $4/C_L$.

The above example shows that the entropy dissipation method presented above consists of the following ingredients:

- an entropy functional,
- an entropy dissipation inequality, and
- a relation between the entropy and the entropy dissipation.

Entropy methods are important tools not only to prove the long-time behavior of solutions to evolution equations. In fact, inequalities (1.1) and (1.4) provide a priori estimates, which can be used in proving the *global-in-time existence* of weak solutions. Employing other entropy functionals, the *regularity* of solutions may be proven. We will show below that an entropy method can substitute the *minimum/maximum principle* in systems of equations. This is of importance since the classical maximum principle for elliptic or parabolic second-order equations generally does not hold in such situations. Furthermore, new *functional inequalities* with explicit constants can be proven.

In the following section we will specify which entropy functionals are of interest. The above technique will be explained for Fokker-Planck equations in more detail in Section 3. Variants of entropy methods will be presented in Section 4.

2. Entropies

We define the notions of entropy and entropy dissipation and give some examples of entropies for the heat equation and the Boltzmann transport equation.

2.1. **Definitions.** We start with some definitions. Throughout this section, let X be a Banach space and let $A : D(A) \to X$ be some (possibly nonlinear) operator defined on its domain $D(A)$. We assume that there exists a smooth function $u(t) : D(A) \to \mathbb{R}$ satisfying

$$u_t + Au = 0 \quad \text{in } \Omega,\ t > 0, \quad u(0) = u_0, \tag{2.5}$$

where $u_0 \in D(A)$. The regularity conditions can be relaxed but they simplify the subsequent arguments. We assume that the stationary equation $Au = 0$ possesses a steady state $0 \le u_\infty \in D(A)$.

Definition 2.1 (Lyapunov functional)**.** *Let $H : D(A) \to \mathbb{R}$ be a functional satisfying*

$$\frac{dH}{dt}[u(t)] \le 0 \quad \textit{for all } t > 0.$$

Then we call H a Lyapunov functional *along the trajectory $u(t)$.*

As we stated already in the introduction, an entropy is a specific Lyapunov functional. In the literature, there does not exist a standardized definition of entropy. We give in the following a possible definition (taken from [32]) but we will use in these lecture notes the term "entropy" also without verification or as being a convex, nonnegative Lyapunov functional.

Definition 2.2 (Entropy)**.** *We call the functional $H : D(A) \to \mathbb{R}$ an* entropy *of* (2.5) *if the following conditions are satisfied:*

- *H is a Lyapunov functional;*
- *H is convex;*
- *There exists a continuous function $\Phi : \mathbb{R} \to \mathbb{R}$ such that $\Phi(0) = 0$ and*

$$d(u, u_\infty) \le \Phi(H[u] - H[u_\infty]) \quad \textit{for all } u \in D(A). \tag{2.6}$$

Definition 2.3 (Entropy dissipation)**.** *Let H be an entropy of* (2.5) *and let u be a (smooth) solution to this equation. Then the* entropy dissipation *D is defined as*

$$D[u(t)] = -\frac{dH}{dt}[u(t)], \quad t > 0.$$

We call an entropy to be of *k-th order* if it contains partial derivatives of k-th order. For instance, the following functionals may be zeroth-order entropies:

$$H_\alpha[u] = \frac{1}{\alpha(1-\alpha)} \int_\Omega u^\alpha dx, \quad \alpha > 0,\ \alpha \ne 1,$$
$$H_1[u] = \int_\Omega u(\log u - 1) dx,$$
$$H_0[u] = \int_\Omega (u - \log u) dx.$$

When we wish to prove the decay rate of a solution $u(t)$ to the stationary state u_∞, it is more appropriate to define *relative entropies*, e.g.

$$H_\alpha[u] = \frac{1}{\alpha(\alpha-1)} \int_\Omega (u^\alpha - u_\infty^\alpha) dx, \quad \alpha > 0,\ \alpha \ne 1,$$
$$H_1[u] = \int_\Omega u \log \frac{u}{u_\infty} dx,$$

but also other definitions are possible. Candidates of first-order entropies are:

$$E_\alpha[u] = \int_\Omega |\nabla u^{\alpha/2}|^2 dx, \quad \alpha > 0,$$
$$E_0[u] = \int_\Omega |\nabla \log u|^2 dx.$$

Sometimes, E_2 is called the energy of the equation. The functional E_1 is referred to as the *Fisher information* since it plays an important role in information theory. One may also consider second-order entropies, e.g.

$$F_\alpha[u] = \int_\Omega |\Delta u^{\alpha/2}|^2 dx, \quad \alpha > 0,$$

but the computations with these functionals become often very involved, and we will not consider such functionals here.

2.2. **The heat equation revisited.** We apply the definitions of the previous subsection to the solution to the heat equation with periodic boundary conditions,

$$u_t = \Delta u \quad \text{in } \mathbb{T}^d,\ t > 0, \quad u(0) = u_0 > 0. \tag{2.7}$$

To simplify the presentation, we assume that the initial datum is smooth and satisfies $\int_{\mathbb{T}^d} u_0(x)dx = 1$. The (constant) steady state of (2.7) is given by

$$u_\infty = \frac{1}{\text{meas}(\mathbb{T}^d)} \int_{\mathbb{T}^d} u_0 dx.$$

By the maximum principle, the solution $u(t)$ is positive. We wish to prove the following result.

Proposition 2.4. *The functionals H_α, defined in Section 2.1, are Lyapunov functionals to* (2.7). *Furthermore, H_1 is an entropy for* (2.7) *in the sense of Definition 2.2.*

The first claim is easily proved by integration by parts:

$$\frac{dH_\alpha}{dt} = \frac{1}{\alpha-1}\int_{\mathbb{T}^d} u^{\alpha-1}u_t dx = \frac{1}{\alpha-1}\int_{\mathbb{T}^d} u^{\alpha-1}\Delta u dx = -\int_{\mathbb{T}^d} u^{\alpha-2}|\nabla u|^2 dx \le 0,$$

where $\alpha \neq 0, 1$. An analogous computation shows that such an inequality also holds when $\alpha = 0$ or $\alpha = 1$. In order to show that H_1 is even an entropy we need an auxiliary result, the Csiszár-Kullback inequality.

Lemma 2.5 (Csiszár-Kullback)**.** *Let $\Omega \subset \mathbb{R}^d$ be a domain and let $f, g \in L^1(\Omega)$ satisfy $f \ge 0$, $g > 0$, and $\int_\Omega f dx = \int_\Omega g dx = 1$. Furthermore, let $\phi \in C^1(\mathbb{R})$ satisfy*

$$\phi(s) \ge \phi(1) + \phi'(1)(s-1) + \gamma^2(s-1)^2 \mathbf{1}_{\{s<1\}}$$

for all $s \in \mathbb{R}$ and some $\gamma > 0$, where $\mathbf{1}_A$ is the characteristic function on $A \subset \mathbb{R}$. Finally, let

$$H_\phi[f] = \int_\Omega \phi\Big(\frac{f}{g}\Big) g dx.$$

Then

$$\|f-g\|^2_{L^1(\Omega)} \le \frac{4}{\gamma^2}(H_\phi[f]-H_\phi[g]).$$

The assumptions on ϕ are satisfied if $\phi \in C^2(\mathbb{R})$ and $\phi''(s) \ge 2\gamma^2 > 0$ for $0 < s < 1$ and $\phi''(s) \ge 0$ else. The classical Csiszár-Kullback inequality [11, 30] is obtained for $\phi(s) = s(\log s - 1) + 1$ with

$$\|f-g\|_{L^1(\Omega)} \le \sqrt{8(H_\phi[f]-H_\phi[g])}.$$

Notice that the optimal constant is $\sqrt{2}$ instead of $\sqrt{8}$. For generalizations of Lemma 2.5 we refer to Carrillo-Jüngel-Markowich-Toscani-Untereiter 2001 [7], Section 4.2.

Proof. The proof is taken from [32]. Since f and g have both mass one, we find that

$$\begin{aligned}
\|f-g\|_{L^1(\Omega)} &= \int_{\{f<g\}} |f-g|dx + \int_{\{f\ge g\}} |f-g|dx \\
&= \int_{\{f<g\}} (g-f)dx + \int_{\{f\ge g\}} fdx - \int_{\{f\ge g\}} gdx \\
&= \int_{\{f<g\}} (g-f)dx + \Big(1-\int_{\{f<g\}} fdx\Big) - \Big(1-\int_{\{f<g\}} gdx\Big) \\
&= 2\int_{\{f<g\}} (g-f)dx.
\end{aligned}$$

Hence, by the Cauchy-Schwarz inequality and the condition $\int_\Omega gdx = 1$,

$$\begin{aligned}
\|f-g\|_{L^1(\Omega)} &= 2\int_{\{f<g\}} \Big|\frac{f}{g}-1\Big| gdx \le 2\Big(\int_{\{f<g\}} \Big|\frac{f}{g}-1\Big|^2 gdx\Big)^{1/2}\Big(\int_\Omega gdx\Big)^{1/2} \\
&= 2\Big(\int_{\{f<g\}} \Big|\frac{f}{g}-1\Big|^2 gdx\Big)^{1/2}.
\end{aligned}$$

Now we employ the assumption on ϕ to conclude that

$$\begin{aligned}
H_\phi[f]-H_\phi[g] &= \int_\Omega \Big(\phi\Big(\frac{f}{g}\Big)-\phi(1)\Big)gdx \ge \int_\Omega \Big(\phi'(1)\Big(\frac{f}{g}-1\Big)g + \gamma^2\Big(\frac{f}{g}-1\Big)^2 g\mathbf{1}_{\{f<g\}}\Big)dx \\
&= \phi'(1)\int_\Omega (f-g)dx + \gamma^2\int_{\{f<g\}} \Big(\frac{f}{g}-1\Big)^2 gdx \ge \frac{\gamma^2}{4}\|f-g\|^2_{L^1(\Omega)}.
\end{aligned}$$

In the last step, we used $\int_\Omega fdx = \int_\Omega gdx$. ◻

Proof of Proposition 2.4. Since $s \mapsto s(\log s - 1)$ is convex, the functional $H_1[u]$ is convex too. The solution to the heat equation satisfies $\int_\Omega u(t)dx = 1$ for all $t > 0$, and $u_\infty = 1/\mathrm{meas}(\mathbb{T}^d)$. Hence,

$$\begin{aligned}
H_1[u]-H_1[u_\infty] &= \int_{\mathbb{T}^d} u\log u dx - \int_{\mathbb{T}^d} u_\infty \log u_\infty dx = \int_{\mathbb{T}^d} u\log u dx - \log u_\infty \\
&= \int_{\mathbb{T}^d} u\log u dx - \log u_\infty \int_{\mathbb{T}^d} udx = \int_{\mathbb{T}^d} u\log\frac{u}{u_\infty}dx.
\end{aligned}$$

By Lemma 2.5 with $\phi(s) = s\log s$ and $\gamma = 1/\sqrt{2}$, we infer that

$$H_1[u] - H_1[u_\infty] \geq \frac{1}{8}\|u - u_\infty\|^2_{L^1(\mathbb{T}^d)},$$

i.e., H_1 satisfies (2.6) with $\Phi(s) = \sqrt{8s}$. □

2.3. **The homogeneous Boltzmann equation.** The entropy H_1 plays a key role in the homogeneous Boltzmann equation modeling a rarefied mono-atomic gas. It describes the temporal change of the probability to find molecules of a given velocity v. Let $f(v,t)$ be the probability density at time $t > 0$ to find molecules with velocity $v \in \mathbb{R}^d$. The homogeneous Boltzmann equation is derived under the assumptions that the molecules move freely and that they exchange momentum and energy in binary elastic collisions. Let v and w be the velocities of two molecules before a collision, and v^* and w^* the post-collisional velocities. Elastic collisions conserve momentum and energy, i.e.

$$v + w = v^* + w^*, \quad |v|^2 + |w|^2 = |v^*|^2 + |w^*|^2. \tag{2.8}$$

These are $d+1$ equations for the $2d$ unknowns v^* and w^*. Therefore, the solutions are given in terms of $d-1$ parameters. For instance, the solutions can be expressed as

$$v^* = \frac{1}{2}(v + w + |v - w|n), \quad w^* = \frac{1}{2}(v + w - |v - w|n),$$

where $n \in \mathbb{S}^{d-1}$ is a parameter on the unit sphere. Under these assumptions, Boltzmann derived in 1872 [5] the equation

$$\frac{\partial f}{\partial t} = Q(f) = \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|, n)(f(v^*)f(w^*) - f(v)f(w))dwdn. \tag{2.9}$$

The nonnegative function $B(z,\nu)$ is the Boltzmann collision kernel which depends on the collision angle via $\nu = (v-w)\cdot n/|v-w|$. The right-hand side can be split into a gain and a loss term. The loss term involving $f(v)f(w)$ counts all collisions in which a particle with velocity v encounters another particle with velocity w. After the collision, the particle will generally change its velocity, resulting in less particles with velocity v. When particles with velocities v^* and w^* collide, one particle may acquire the velocity v, resulting in a gain of particles with that velocity. This gives the gain term involving $f(v^*)f(w^*)$.

The Boltzmann equation can be written in a weak form. Indeed, multiplying the equation by a smooth test function $\Phi(v)$ and employing the changes of variables $(v,w) \mapsto (v^*,w^*)$ and $(v,w) \mapsto (w,v)$ (here we omit some details on how to modify the parameter n; see Villani 2003

[40]), we obtain

$$
\begin{aligned}
\int_{\mathbb{R}^d} Q(f)\Phi(v)dv &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|,n)(f(v^*)f(w^*)-f(v)f(w))\Phi(v)dvdwdn \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|,n)f(v)f(w)(\Phi(v^*)-\Phi(v))dvdwdn \\
&= \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|,n)f(v)f(w) \\
&\quad \times(\Phi(v^*)+\Phi(w^*)-\Phi(v)-\Phi(w))dvdwdn. \qquad (2.10)
\end{aligned}
$$

Symmetrizing this expression once more, it follows that

$$
\begin{aligned}
\int_{\mathbb{R}^d} Q(f)\Phi(v)dv = -\frac{1}{4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|,n)(f(v^*)f(w^*)-f(v)f(w)) \\
\times(\Phi(v^*)+\Phi(w^*)-\Phi(v)-\Phi(w))dvdwdn. \qquad (2.11)
\end{aligned}
$$

As a consequence of (2.10), whenever Φ satisfies

$$\Phi(v)+\Phi(w)=\Phi(v^*)+\Phi(w^*) \quad \text{for all } v,w,n, \qquad (2.12)$$

it holds formally

$$\frac{d}{dt}\int_{\mathbb{R}^d} f(v,t)\Phi(v)dv = \int_{\mathbb{R}^d} Q(f)\Phi(v)dv = 0.$$

By momentum and energy conservation (2.8), this holds true for the functions $\Phi(v)=1$, v_j, $|v|^2/2$ ($j=1,\ldots,d$). It can be shown that all solutions to (2.12) are linear combinations of these functions (see the book of Cercignani-Illner-Pulvirenti 1994 [8], pp. 36-42). This yields the conservation laws of the Boltzmann equation,

$$\frac{d}{dt}\int_{\mathbb{R}^d} f(v,t)(1,v,\tfrac{1}{2}|v|^2)dv = 0,$$

expressing conservation of mass, momentum, and energy of the gas.

The weak form (2.11) is used to prove that

$$H_1[f] = \int_{\mathbb{R}^d} f\log f dv$$

is an entropy for the Boltzmann equation on the space U of probability densities $f(v)$ satisfying

$$\int_{\mathbb{R}^d} f(v)dv = 1, \quad \int_{\mathbb{R}^d} f(v)v dv = 0, \quad \int_{\mathbb{R}^d} f(v)\frac{|v|^2}{2}dv = \frac{1}{2}.$$

Theorem 2.6 (Boltzmann's H theorem)**.** *The functional H_1 is an entropy (in the sense of Definition 2.2) for the homogeneous Boltzmann equation* (2.9) *on the domain U.*

Proof. The proof is taken from [32]. First, we observe that H_1 is a Lyapunov functional. Indeed, taking $\Phi(v)=\log f(v)$ in (2.11), we find that

$$
\begin{aligned}
\frac{dH_1}{dt}[f] = -\frac{1}{4}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B(|v-w|,n) \\
\times(\log(f(v^*)f(w^*))-\log(f(v)f(w)))(f(v^*)f(w^*)-f(v)f(w))dvdwdn \le 0,
\end{aligned}
$$

since $x \mapsto \log x$ is strictly increasing and the expression under the integral is nonnegative.

Next, we need to determine the steady state f_∞. We claim that the Maxwellian

$$f_\infty(v) = (2\pi)^{-d/2}\exp(-|v|^2/2)$$

is the unique stationary solution to the Boltzmann equation. First, we observe that for $f \in U$,

$$\begin{aligned}\int_{\mathbb{R}^d} f\log f_\infty dv &= -\int_{\mathbb{R}^d} f\Big(\frac{d}{2}\log(2\pi)+\frac{1}{2}|v|^2\Big)dv = -\frac{d}{2}\log(2\pi)-\frac{1}{2}\\ &= -\frac{d}{2}\log(2\pi)\underbrace{(2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-|v|^2/2}dv}_{=1}-\frac{1}{2}\underbrace{(2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-|v|^2/2}|v|^2dv}_{=1}\\ &= -\int_{\mathbb{R}^d}(2\pi)^{-d/2}e^{-|v|^2/2}\Big(\frac{d}{2}\log(2\pi)-\frac{1}{2}|v|^2\Big)dv\\ &= \int_{\mathbb{R}^d} f_\infty\log f_\infty dv = H_1[f_\infty].\end{aligned}$$

Then, with $\rho = f/f_\infty$, we have

$$H_1[f]-H_1[f_\infty] = \int_{\mathbb{R}^d} f(\log f-\log f_\infty)dv = \int_{\mathbb{R}^d}\rho\log(\rho)f_\infty dv = \int_{\mathbb{R}^d}\phi(\rho)f_\infty dv,$$

where $\phi(s) = s\log s$. We apply Jensen's inequality to the integral with measure $f_\infty dv$ to obtain

$$H_1[f]-H_1[f_\infty] \ge \phi\Big(\int_{\mathbb{R}^d}\rho f_\infty dv\Big) = \Big(\int_{\mathbb{R}^d}\rho f_\infty dv\Big)\log\Big(\int_{\mathbb{R}^d}\rho f_\infty dv\Big) = 0,$$

since $\int_{\mathbb{R}^d}\rho f_\infty dv = \int_{\mathbb{R}^d} f dv = 1$. Equality holds if and only if $\phi(\rho = 0$ or $\rho(v) = 1$ or $f(v) = f_\infty(v)$ for all $v \in \mathbb{R}^d$. Thus, f_∞ minimizes H_1.

Finally, the Csiszár-Kullback inequality (Lemma 2.5)

$$\|f-f_\infty\|_{L^1(\mathbb{R}^d)} \le \frac{2}{\gamma}(H_1[f]-H_1[f_\infty])^{1/2}$$

shows property (2.6) with the distance induced by the L^1 norm. □

3. Fokker-Planck equations

Fokker-Planck equations are drift-diffusion equations of the form

$$u_t = \operatorname{div}(\nabla f(u)+u\nabla V),$$

where $f(u)$ is some nonlinearity and V a potential. First, we analyze the long-time asymptotics of the linear Fokker-Planck equation (i.e. $f(u) = u$) and show relations to a specific functional inequality, the logarithmic Sobolev inequality. Second, the entropy technique is extended to nonlinear Fokker-Planck equations.

3.1. **Relaxation to self-similarity.** We consider the heat equation but now in the whole space,

$$u_t = \Delta u \quad \text{in } \mathbb{R}^d,\ t > 0, \quad u(0) = u_0 \geq 0 \quad \text{in } \mathbb{R}^d, \quad \int_{\mathbb{R}^d} u_0 dx = 1. \tag{3.13}$$

The solution $u(t) \geq 0$ can be written explicity:

$$u(x,t) = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-|x-y|^2/(4t)} u_0(y) dy.$$

In particular, it is strictly positive and conserves mass, $\int_{\mathbb{R}^d} u(t)dx = 1$ for all $t > 0$. From this formula follows that $u(t) \to 0$ in $L^\infty(\Omega)$ as $t \to \infty$. Furthermore, the functional

$$H_1[u] = \int_{\mathbb{R}^d} u(\log u - 1) dx$$

is a Lyapunov functional along solution u to (3.13). However,

$$H[u(t)] \leq \int_{\mathbb{R}^d} u(t) \log \|u(t)\|_{L^\infty(\Omega)} dx = \log \|u(t)\|_{L^\infty(\Omega)} \to -\infty,$$

and entropy estimates seem to be not applicable. In fact, this is not surprising, since the only (integrable) steady state to (3.13) is $u_\infty = 0$, and this function has not unit mass. The entropy is useful to study the relaxation of the solution to the self-similar solution

$$U(x,t) = \frac{1}{(2\pi(2t+1))^{d/2}} \exp\Big(-\frac{|x|^2}{2(2t+1)}\Big), \quad x \in \mathbb{R}^d,\ t > 0, \tag{3.14}$$

i.e., we wish to analyze how fast $u(t) - U(t)$ decays to zero. Clearly, this gives much more information than just the fact that $u(t) \to 0$ as $t \to \infty$.

For this, we transform the variables (x,t) to make U stationary in these coordinates. We set $y = x/\sqrt{2t+1}$, $s = \log\sqrt{2t+1}$, and

$$v(y,s) = e^{ds} u(e^s y, \tfrac{1}{2}(e^{2s} - 1)), \quad y \in \mathbb{R}^d,\ s > 0.$$

Then

$$\frac{\partial v}{\partial s} = d e^{ds} u + e^{ds} e^s \nabla_x u + e^{ds} e^{2s} u_t = dv + \nabla_y v + \Delta_y v = \operatorname{div}_y(\nabla_y v + yv),$$

and the function v satisfies the Cauchy problem

$$v_s = \operatorname{div}(\nabla v + yv) \quad \text{in } \mathbb{R}^d,\ s > 0, \quad v(0) = u_0. \tag{3.15}$$

This equation is of Fokker-Planck type with a quadratic potential $V(y) = \frac{1}{2}|y|^2$. The self-similar solution in the new coordinates becomes

$$M(y) = (2t+1)^{d/2} U(x,t) = (2\pi)^{-d/2} e^{-|y|^2/2},$$

which we call the Maxwellian (see Section 2.3). It is the unique steady state to (3.15). We choose functions v from the domain

$$X = \Big\{ v \in L^1(\mathbb{R}^d) : v \geq 0,\ |y|^2 v,\ v\log v \in L^1(\mathbb{R}^d),\ \int_{\mathbb{R}^d} v dy = 1 \Big\}.$$

In this setting, we work with the relative entropy

$$H_1[v] = \int_{\mathbb{R}^d} v \log \frac{v}{M} dy = \int_{\mathbb{R}^d} v \log v dy + \frac{1}{2} \int_{\mathbb{R}^d} (d \log(2\pi) + |y|^2) v dy.$$

Theorem 3.1 (Exponential decay for the Fokker-Planck equation)**.** *Let $u_0 \in L^1(\mathbb{R}^d)$ be nonnegative and satisfy $\int_{\mathbb{R}^d} u_0 dx = 1$. Let v be the solution to* (3.15)*. Then, with H_1 as defined above,*

$$0 \le H_1[v(s)] \le e^{-2s} H_1[u_0] \quad \text{for all } s > 0. \tag{3.16}$$

Moreover, $v(s)$ converges exponentially fast to the Maxwellian M,

$$\|v(s) - M\|_{L^1(\mathbb{R}^d)} \le e^{-s} \sqrt{8H_1[u_0]} \quad \text{for all } s > 0. \tag{3.17}$$

Proof. We differentiate, employ (3.13), and integrate by parts:

$$\begin{aligned} \frac{dH_1}{ds}[v(s)] &= \int_{\mathbb{R}^d} v_s \log v dy + \frac{1}{2} \int_{\mathbb{R}^d} |y|^2 v_s dy \\ &= -\int_{\mathbb{R}^d} \nabla \log v \cdot (\nabla v + yv) dy - \frac{1}{2} \int_{\mathbb{R}^d} \nabla |y|^2 \cdot (\nabla v + yv) dy \\ &= -\int_{\mathbb{R}^d} \Big(\frac{|\nabla v|^2}{v} + 2y \cdot \nabla v + |y|^2 v \Big) dy = -\int_{\mathbb{R}^d} v |\nabla \log v + y|^2 dy \le 0. \end{aligned}$$

Hence, H_1 is a Lyapunov functional. Another formulation of the right-hand side is, after integrating by parts in the mixed term and using $\int_{\mathbb{R}^d} v dx = 1$,

$$\frac{dH_1}{ds}[v(s)] = -\int_{\mathbb{R}^d} (4|\nabla \sqrt{v}|^2 - 2dv + |y|^2 v) dy = -\int_{\mathbb{R}^d} (4|\nabla \sqrt{v}|^2 + |y|^2 v) dx - 2d.$$

Now, assume that the following inequality holds:

$$2 \int_{\mathbb{R}^d} |\nabla \sqrt{v}|^2 dy \ge \int_{\mathbb{R}^d} v \log v dy + d(1 + \log \sqrt{2\pi}). \tag{3.18}$$

Then we find that

$$\frac{dH_1}{ds}[v(s)] \ge -2 \int_{\mathbb{R}^d} v \log v dy - \int_{\mathbb{R}^d} (|y|^2 + d \log(2\pi)) v dy = 2H_1[v(s)].$$

By Gronwall's inequality, we infer (3.16). Estimate (3.17) is a consequence of (3.16) and the Csiszár-Kullback inequality. □

Before we discuss (3.18), we go back to the original variables. Notice that the self-similar solution can be written in terms of the Maxwellian as follows:

$$U(x,t) = (2t+1)^{-d/2} M((2t+1)^{-1/2} x).$$

The left-hand side of (3.17) writes after the substitution $y = (2t+1)^{-1/2} x$ as

$$\|v(s) - M\|_{L^1(\mathbb{R}^d)} = \|u(t) - U(t)\|_{L^1(\mathbb{R}^d)},$$

whereas the right-hand side is formulated as

$$e^{-s} \sqrt{8H_1[u_0]} = (2t+1)^{-1/2} \sqrt{8H_1[u_0]}.$$

Thus, we have shown the following result.

Corollary 3.2 (Relaxation to self-similarity)**.** *Let* $u_0 \in L^1(\mathbb{R}^d)$ *be nonnegative and has unit mass,* $\int_{\mathbb{R}^d} u_0 dx = 1$*. Let* U *be defined in* (3.14) *and let* $u(t)$ *be the solution to* (3.13)*. Then*

$$\|u(t) - U(t)\|_{L^1(\mathbb{R}^d)} \le \frac{\sqrt{8H_1[u_0]}}{\sqrt{2t+1}} \quad \textit{for all } t > 0.$$

It remains to prove (3.18) written for $v = f^2$:

$$\int_{\mathbb{R}^d} f^2 \log f^2 dx + d(1 + \log \sqrt{2\pi}) \le 2 \int_{\mathbb{R}^d} |\nabla f|^2 dx \tag{3.19}$$

for $f \in H^1(\mathbb{R}^d)$, $f \ge 0$, $\|f\|_{L^2(\mathbb{R}^d)} = 1$. This inequality is called the *logarithmic Sobolev inequality*. It was first formulated by Federbush in 1969 [21] and exploited by Gross in 1975 [24]. We will give a proof in the next subsection.

3.2. **The Fokker-Planck equation and logarithmic Sobolev inequality.** The approach of the previous subsection has the drawback that the equilibration property relies on the logarithmic Sobolev inequality which needs to be proven separately. The entropy method can be modified in such a way that *both* statements – equilibration property and logarithmic Sobolev inequality – can be proved simultaneously. To this end, we consider a slightly more general Fokker-Planck equation than in the previous subsection:

$$u_t = \operatorname{div}(\nabla u + u\nabla V) \quad \text{in } \mathbb{R}^d,\ t > 0, \quad u(0) = u_0. \tag{3.20}$$

As in the previous subsection, we assume that the initial datum $u_0 \in L^1(\mathbb{R}^d)$ is nonnegative and has unit mass. The potential $V(x)$ is assumed to be smooth and satisfies $\lim_{|x|\to\infty} V(x) = \infty$. The Fokker-Planck equation possesses the steady state

$$0 = \nabla u_\infty + u_\infty \nabla V = u_\infty(\nabla \log u_\infty + V).$$

Hence, if $u_\infty > 0$, $\log u_\infty + V$ is constant. Thus, u_∞ is given by

$$u_\infty(x) = Ze^{-V(x)}, \quad Z = \Big(\int_{\mathbb{R}^d} e^{-V(y)} dy\Big)^{-1}. \tag{3.21}$$

In order to introduce the entropy, let $\phi : [0,\infty) \to [0,\infty)$ be a smooth and convex function satisfying $\phi(1) = \phi'(1) = 0$. An example for such a function is given by $\phi(s) = s(\log s - 1) + 1$, $s > 0$. Then we introduce the entropy functional:

$$H_\phi[u] = \int_{\mathbb{R}^d} \phi\Big(\frac{u}{u_\infty}\Big) u_\infty dx. \tag{3.22}$$

The entropy is a Lyapunov functional for the Fokker-Planck equation:

Proposition 3.3. *Let* ϕ *be defined as above. Then* H_ϕ *is a Lyapunov functional for the Fokker-Planck equation* (3.20).

Proof. Using the expression $\nabla u + u\nabla V = \nabla u - \nabla \log u_\infty = u_\infty \nabla(u/u_\infty)$, we compute

$$\frac{dH_\phi}{dt}[u(t)] = \int_{\mathbb{R}^d} \phi'\Big(\frac{u}{u_\infty}\Big)u_t dx = -\int_{\mathbb{R}^d} \nabla\phi'\Big(\frac{u}{u_\infty}\Big)\cdot(\nabla u + u\nabla V)dx$$
$$= -\int_{\mathbb{R}^d} \phi''\Big(\frac{u}{u_\infty}\Big)\nabla\Big(\frac{u}{u_\infty}\Big)\cdot\nabla\Big(\frac{u}{u_\infty}\Big)u_\infty dx = -\int_{\mathbb{R}^d} \phi''\Big(\frac{u}{u_\infty}\Big)\Big|\nabla\Big(\frac{u}{u_\infty}\Big)\Big|^2 u_\infty dx.$$

Since ϕ is convex, the right integral is nonnegative and hence, $dH_\phi[u(t)]/dt \le 0$. □

The idea of the entropy method here is to compute the *second* time derivative of H_ϕ. This is due to Bakry and Emery 1983 [2]. Let $D_\phi[u(t)] = -dH_\phi[u(t)]/dt \ge 0$ be the entropy dissipation. Notice that $D_\phi[u(t)] = 0$ if and only if $u(t) = u_\infty$.

Lemma 3.4. *Let $\nabla^2 V(x) - \lambda\mathbb{I}$ be positive semi-definite uniformly in $x \in \mathbb{R}^d$ for some $\lambda > 0$. Let $\phi \in C^4([0,\infty))$ be convex such that $1/\phi''$ is concave. Then, along solutions $u(t)$ to* (3.20),

$$\frac{dD_\phi}{dt}[u(t)] \le -2\lambda D_\phi[u(t)] \quad \textit{for } t > 0.$$

As a consequence of this lemma, if $D_\phi[u_0] < \infty$, we have exponential decay with rate 2λ:

$$D_\phi[u(t)] \le e^{-2\lambda t} D_\phi[u_0], \quad t > 0.$$

Proof. The proof is due to Arnold-Markowich-Toscani-Unterreiter 2001 [1] but the idea goes back to Bakry-Emery 1983 [2]. Let $\rho = u/u_\infty$. Then the Fokker-Planck equation can be written equivalently as $\rho_t = u_\infty^{-1} u_t = u_\infty^{-1}\operatorname{div}(u_\infty \nabla\rho)$. The proof of Proposition 3.3 shows that

$$D_\phi[u] = \int_{\mathbb{R}^d} \phi''(\rho)|\nabla\rho|^2 u_\infty dx.$$

We calculate, as in [32],

$$\frac{dD_\phi}{dt}[u(t)] = \int_{\mathbb{R}^d} \partial_t\phi''(\rho)|\nabla\rho|^2 u_\infty dx + 2\int_{\mathbb{R}^d} \phi''(\rho)\nabla\rho\cdot\partial_t\nabla\rho u_\infty dx. \tag{3.23}$$

The first integral equals

$$\int_{\mathbb{R}^d} \partial_t\phi''(\rho)|\nabla\rho|^2 u_\infty dx = \int_{\mathbb{R}^d} \phi'''(\rho)|\nabla\rho|^2 \operatorname{div}(u_\infty\nabla\rho)dx$$
$$= -\int_{\mathbb{R}^d} \nabla(\phi'''(\rho)|\nabla\rho|^2)\cdot\nabla\rho u_\infty dx$$
$$= -\int_{\mathbb{R}^d} (\phi''''(\rho)|\nabla\rho|^4 + 2\phi'''(\rho)\nabla\rho\nabla^2\rho\nabla\rho)u_\infty dx.$$

For the second integral, we observe that

$$\nabla\rho\cdot\nabla\rho_t = \nabla\rho\cdot\nabla(\Delta\rho - \nabla\rho\cdot\nabla V) = \operatorname{div}(\nabla^2\rho\cdot\nabla\rho) + |\nabla^2\rho|^2 - \nabla\rho\nabla^2 V\nabla\rho - \nabla\rho\nabla^2\rho\nabla V.$$

Using $\nabla\rho\nabla^2V\nabla\rho \geq \lambda|\nabla\rho|^2$ und integrating by parts in the term involving the divergence, the second integral becomes

$$\begin{aligned}2\int_{\mathbb{R}^d}\phi''(\rho)\nabla\rho\cdot\partial_t\nabla\rho u_\infty dx &\leq -2\lambda\int_{\mathbb{R}^d}\phi''(\rho)|\nabla\rho|^2u_\infty dx \\ &\quad + 2\int_{\mathbb{R}^d}\phi''(\rho)(\mathrm{div}(\nabla^2\rho\nabla\rho)+|\nabla^2\rho|^2-\nabla\rho\nabla^2\rho\nabla V)u_\infty dx \\ &= -2\lambda D_\phi[u] + 2\int_{\mathbb{R}^d}\phi''(\rho)(|\nabla^2\rho|^2-\nabla\rho\nabla^2\rho\nabla V)u_\infty dx \\ &\quad - 2\int_{\mathbb{R}^d}\nabla^2\rho\nabla\rho\nabla u_\infty dx - 2\int_{\mathbb{R}^d}\phi'''(\rho)\nabla\rho\nabla^2\rho\nabla\rho dx \\ &= -2\lambda D_\phi[u] + 2\int_{\mathbb{R}^d}\phi''(\rho)|\nabla^2\rho|^2u_\infty dx - 2\int_{\mathbb{R}^d}\phi'''(\rho)\nabla\rho\nabla^2\rho\nabla\rho dx,\end{aligned}$$

where we used $\nabla u_\infty + u_\infty\nabla V = 0$. Inserting these expressions into (3.23), we infer that

$$\frac{dD_\phi}{dt}[u(t)] \leq -2\lambda D_\phi[u] - \int_{\mathbb{R}^d}(\phi''''(\rho)|\nabla\rho|^4 + 4\phi'''(\rho)\nabla\rho\nabla^2\rho\nabla\rho + 2\phi''(\rho)|\nabla^2\rho|^2)u_\infty dx.$$

We claim that our assumptions on ϕ imply that the last integrand is pointwise nonnegative. Indeed, the convexity of ϕ gives $\phi'' \geq 0$, and the convavity of $1/\phi''$ is equivalent to $\phi''''\phi'' - 2(\phi''')^2 \geq 0$. These conditions ensure that the quadratic form

$$Q(x,y) = \phi''''x^2 - 4\phi'''xy + 2\phi''y^2, \quad x,y\in\mathbb{R},$$

is nonnegative. Hence, by the Cauchy-Schwarz inequality,

$$\begin{aligned}&\phi''''(\rho)|\nabla\rho|^4 + 4\phi'''(\rho)\nabla\rho\nabla^2\rho\nabla\rho + 2\phi''(\rho)|\nabla^2\rho|^2 \\ &\qquad \geq \phi''''(\rho)|\nabla\rho|^4 - 4\phi'''(\rho)\|\nabla^2\rho\||\nabla\rho|^2 + 2\phi''(\rho)\|\nabla^2\rho\|^2 \\ &\qquad = Q(|\nabla^2\rho|, \|\nabla\rho\|^2) \geq 0.\end{aligned}$$

This proves the claim. □

Now we can prove the exponential decay of $H_\phi[u(t)]$.

Theorem 3.5 (Exponential decay in L^1)**.** *Let u be the solution to the Fokker-Planck equation* (3.20)*, let ϕ satisfy the conditions of Lemma 3.4, and let H_ϕ be defined by* (3.22)*. Then*

$$H_\phi[u(t)] \leq e^{-2\lambda t}H_\phi[u_0] \quad \text{for all } t>0.$$

Moreover, if ϕ satisfies the assumptions of Lemma 2.5 (Csiszár-Kullback inequality) then

$$\|u(t)-u_\infty\|_{L^1(\mathbb{R}^d)} \leq C_\phi\sqrt{H_\phi[u_0]}e^{-\lambda t} \quad \text{for all } t>0,$$

where the constant $C_\phi > 0$ only depends on ϕ.

Proof. In the following, we present only a formal proof. The calculations can be made rigorous by suitable but tedious density arguments, see Arnold-Markowich-Toscani-Unterreiter 2001 [1]. We rewrite the inequality of Lemma 3.4:

$$-\frac{dD_\phi}{dt}[u(t)] \geq 2\lambda D_\phi[u(t)] = -2\lambda\frac{dH_\phi}{dt}[u(t)], \quad t > 0.$$

Integrating both sides in $t \in (\tau, \infty)$, we obtain

$$D_\phi[u(\tau)] - \lim_{t\to\infty} D_\phi[u(t)] \geq 2\lambda\big(H_\phi[u(\tau)] - \lim_{t\to\infty} H_\phi[u(t)]\big). \tag{3.24}$$

Next, we show that the limits vanish implying that

$$\frac{dH_\phi}{dt}[u(\tau)] = -D_\phi[u(\tau)] \leq -2\lambda H_\phi[u(\tau)].$$

This gives the first claim.

The entropy dissipation D_ϕ is nonnegative, nonincreasing as a function of t, and it holds

$$\int_0^\infty D_\phi[u(t)]dt \leq D_\phi[u_0]\int_0^\infty e^{-2\lambda t}dt < \infty.$$

Hence, $D_\phi[u(t)]$ converges to zero as $t \to \infty$. Since $D_\phi[u(t)] \geq 0$, we find that

$$0 = \lim_{t\to\infty} D_\phi[u(t)] = D_\phi\Big[\lim_{t\to\infty} u(t)\Big].$$

The functional D_ϕ vanishes exactly at u_∞, which shows that $\lim_{t\to\infty} u(t) = u_\infty$. Therefore,

$$\lim_{t\to\infty} H_\phi[u(t)] = H_\phi\Big[\lim_{t\to\infty} u(t)\Big] = H_\phi[u_\infty] = 0,$$

and it remains to apply the Gronwall inequality.

By the Csiszár-Kullback inequality (Lemma 2.5),

$$\|u(t) - u_\infty\|_{L^1(\mathbb{R}^d)} \leq \frac{2}{\gamma}\sqrt{H_\phi[u(t)] - H_\phi[u_\infty]} = \frac{2}{\gamma}\sqrt{H_\phi[u(t)]} \leq \frac{2}{\gamma}\sqrt{H_\phi[u_0]}e^{-\lambda t},$$

which shows the second claim. □

It seems that in the above proof, we did not use the logarithmic Sobolev inequality. In fact, we did. Reformulating (3.24), we see that this inequality is *equivalent* to a convex Sobolev inequality.

Corollary 3.6 (Convex Sobolev inequality)**.** *Let $u \in L^1(\mathbb{R}^d)$ be nonnegative and has unit mass, let V and ϕ satisfy the conditions of Lemma 3.4. Furthermore, let u_∞ be given by* (3.21)*. Then*

$$H_\phi[u] = \int_{\mathbb{R}^d} \phi\Big(\frac{u_0}{u_\infty}\Big)u_\infty dx \leq \frac{1}{2\lambda}\int_{\mathbb{R}^d} \phi''\Big(\frac{u_0}{u_\infty}\Big)\Big|\nabla\Big(\frac{u}{u_\infty}\Big)\Big|^2 u_\infty dx = \frac{1}{2\lambda}D_\phi[u]. \tag{3.25}$$

Let $V(x) = \frac{1}{2}|x|^2$ and $\phi(s) = s(\log s - 1) + 1$. Then $\lambda = 1$, $u_\infty(x) = (2\pi)^{-d/2}\exp(-|x|^2/2)$ is the Maxwellian introduced in Section 3.1, and a calculation shows that, using $\int_{\mathbb{R}^d} u dx = \int_{\mathbb{R}^d} u_\infty dx = 1$,

$$\begin{aligned} H_\phi[u] &= \int_{\mathbb{R}^d} u \log u dx - \int_{\mathbb{R}^d} u \log u_\infty dx = \int_{\mathbb{R}^d e} u \log u dx + \frac{d}{2}\log(2\pi) + \frac{1}{2}\int_{\mathbb{R}^d} |x|^2 u dx, \\ D_\phi[u] &= \int_{\mathbb{R}^d} \frac{u_\infty^2}{u}\Big|\nabla \frac{u}{u_\infty}\Big|^2 dx = \int_{\mathbb{R}^d} (4|\nabla\sqrt{u}|^2 + 2x\cdot\nabla u + |x|^2 u) dx \\ &= 4\int_{\mathbb{R}^d} |\nabla\sqrt{u}|^2 dx - 2d + \int_{\mathbb{R}^d} |x|^2 u dx. \end{aligned}$$

Inserting these expressions in the convex Sobolev inequality (3.25), we find that

$$\int_{\mathbb{R}^d} u \log u dx + \frac{d}{2}\log(2\pi) + d \le 2\int_{\mathbb{R}^d} |\nabla\sqrt{u}|^2 dx,$$

which is exactly the logarithmic Sobolev inequality (3.19). Thus, the above proof simultaneously shows the exponential decay of the Fokker-Planck solutions and the convex Sobolev inequality.

Remark 3.7. In bounded domains without confining potential, the logarithmic Sobolev inequality is a consequence of the Sobolev and Poincaré inequalities. This argument is due to Stroock [37], and a short proof is given by Desvillettes and Fellner 2007 [15]. More precisely, let $\Omega \subset \mathbb{R}^d$ be a bounded domain and $u \in H^1(\Omega)$ such that

$$\begin{aligned} \Big\|u - \int_\Omega u dx\Big\|_{L^2(\Omega)} &\le C_P \|\nabla u\|_{L^2(\Omega)}, \\ \|u\|_{L^q(\Omega)} &\le C_S \|u\|_{H^1(\Omega)}, \end{aligned}$$

where $1/q = 1/2 - 1/d$. Then the logarithmic Sobolev inequality

$$\int_\Omega u^2 \log\Big(\frac{u^2}{\|u\|_{L^2(\Omega)}^2}\Big) dx \le C_L \|\nabla u\|_{L^2(\Omega)}^2$$

holds for some constant $C_L > 0$ which depends on Ω and d. □

3.3. **Nonlinear Fokker-Planck equations.** The arguments of the previous subsection can be generalized to nonlinear diffusions. We consider the nonlinear Fokker-Planck equation

$$u_t = \operatorname{div}(\nabla f(u) + u\nabla V) \quad \text{in } \Omega,\ t > 0, \quad u(0) = u_0 \ge 0. \tag{3.26}$$

Here, $\Omega \subset \mathbb{R}^d$ is either a bounded domain with smooth boundary or $\Omega = \mathbb{R}^d$. In the former case, we impose no-flux boundary conditions,

$$(\nabla f(u) + u\nabla V)\cdot\nu = 0 \quad \text{on } \partial\Omega.$$

The initial datum satisfies $u_0 \in L^1(\Omega)$, and we set $\int_\Omega u_0 dx =: M > 0$. To fix the ideas, we assume that the potential is quadratic, $V(x) = \frac{\lambda}{2}|x|^2$ ($\lambda > 0$), but more general choices are possible (as long as the potential is convex; see Carrillo-Jüngel-Markowich-Toscani-Unterreiter 2001 [7]). The nonlinar function is assumed to be smooth, strictly increasing, and $f(0) = 0$. Again, to avoid technicalities, we choose $f(s) = s^m$ with $m > 1$. We notice that the stationary solutions to (3.26) are the compactly supported Barenblatt profiles,

$$u_\infty(x) = \Big(N - \frac{m-1}{2m}|x|^2\Big)_+^{1/(m-1)}, \tag{3.27}$$

where $z_+ = \max\{0, z\}$ denotes the positive part of $z \in \mathbb{R}$. The constant N can be determined from the mass condition, for given $M > 0$,

$$M = \int_\Omega u_0 dx = \int_\Omega \Big(N - \frac{m-1}{2m}|x|^2\Big)_+^{1/(m-1)} dx.$$

In the following, we explain the main ideas of the entropy method for the whole-space situation $\Omega = \mathbb{R}^d$ only and proceed as in [32]. In principle, the strategy of the previous subsection can be extended to the nonlinear equation but there are some additional technical difficulties. One difficulty is that u_∞ may vanish, which makes it impossible to introduce the relative entropy as in (3.22). Therefore, one has to resort to the less convenient absolute entropy

$$H[u] = \int_{\mathbb{R}^d} u\Big(\frac{u^{m-1}}{m-1} + \frac{\lambda}{2}|x|^2\Big)dx,$$

and the difference

$$H^*[u] = H[u] - H[u_\infty].$$

Furthermore, we introduce the function $h : [0,\infty) \to \mathbb{R}$ by $h'(u) = f'(z)/z$ and $h(0) = 0$. In the present case, $h(u) = (m/(m-1))u^{m-1}$. This definition is motivated by the fact that (3.26) can be formulated as $u_t = \operatorname{div}(u\nabla(h(u) + V))$.

The main result is as follows.

Theorem 3.8 (Exponential decay)**.** *Let $u \geq 0$ satisfying $H[u] < \infty$. Then*

$$H^*[u] \leq \frac{1}{2\lambda}\int_{\mathbb{R}^d} u|\nabla(h(u) + V)|^2 dx. \tag{3.28}$$

Let $u(t)$ be a smooth solution to (3.26)*, where $H[u_0] < \infty$. Then there exists a constant $C > 0$ such that*

$$H[u(t)] \leq Ce^{-2\lambda t}, \quad D[u(t)] \leq Ce^{-2\lambda t}, \quad \|u(t) - u_\infty\|_{L^1(\mathbb{R}^d)} \leq Ce^{-\lambda t}$$

for $t > 0$, where $D[u(t)] = -dH[u(t)]/dt$ is the entropy dissipation of H.

We give only a sketch of the proof. There are a number of difficulties to overcome in the rigorous proof (which can be found in [7]). First, since the solution to the porous-medium equation is generally only Hölder continuous and not as smooth as the solution to the heat equation, the solutions to (3.26) have to be approximated by smooth and positive functions. Second, one has to justify that the boundary terms vanish in the integrations by parts, which is not trivial due to the potential which does not vanish as $|x| \to \infty$.

Proof. The proof is divided in several steps. We assume that (3.26) possesses a smooth positive solution $u(t)$.

Step 1: Entropy dissipation. We compute the entropy dissipation:

$$D[u(t)] = -\int_{\mathbb{R}^d} \Big(\underbrace{\frac{m}{m-1}u^{m-1}}_{=h(u)} + \underbrace{\frac{\lambda}{2}|x|^2}_{=V(x)} \Big) u_t dx = -\int_{\mathbb{R}^d} u|\nabla(h(u)+V)|^2 dx \le 0.$$

Step 2: Second entropy dissipation. The computation of $dD[u(t)]/dt$ is involved. We just remark that a straight-forward computation gives

$$\frac{dD}{dt}[u(t)] = -2\lambda \int_{\mathbb{R}^d} u|\nabla(h(u)+V)|^2 dx - \frac{2m}{m-1}R(t) = -2\lambda D[u(t)] - \frac{2m}{m-1}R(t),$$

where

$$R(t) = \int_{\mathbb{R}^d} u\nabla(h(u)+V)\nabla^2 u^{m-1}\nabla(h(u)+V)dx + (m-1)\int_{\mathbb{R}^d} u^{m-2}|\mathrm{div}(u\nabla(h(u)+V))|^2 dx.$$

By several integrations by parts, it follows that

$$R(t) = \frac{(m-1)^2}{m}\int_{\mathbb{R}^d} u^m(\Delta(h(u)+V))^2 dx + \frac{m-1}{m}\int_{\mathbb{R}^d} u^m \|\nabla^2(h(u)+V)\|^2 dx \ge 0.$$

Therefore,

$$\frac{dD}{dt}[u(t)] \le -2\lambda D[u(t)], \quad t>0. \tag{3.29}$$

Step 3: Functional inequality. Integration of (3.29) yields

$$D[u(t)] \le D[u_0]e^{-2\lambda t}, \quad t > 0,$$

from which (3.28) follows as in the linear case; see the proof of Theorem 3.5. Here, one has to prove that $D[u(t)] \to 0$ as $t \to \infty$ which is not obvious.

Step 4: Convergence in the L^1 norm. This part of the proof is surprisingly difficult due to the lack of positivity of the steady state u_∞. The idea is to estimate the L^1 norm of $u - u_\infty$ first for steady states u_∞ whose support is contained in some ball and then to control the behavior of u outside the support of u_∞. Here, the definition of the entropy H^* is needed. For details, we refer to [32, Section 1.7]. □

In Section 3.2, we have shown that the linear Fokker-Planck equation is related to a convex Sobolev inequality, including the logarithmic Sobolev inequality; see Corollary 3.6. One may ask if the nonlinear Fokker-Planck equation is related to a functional inequality too. The answer is yes and the corresponding inequality is the Gagliardo-Nirenberg inequality.

Theorem 3.9 (Gagliardo-Nirenberg inequality)**.** *Let* $\frac{1}{2} < p < 1$. *Then there exists* $C_0 > 0$ *only depending on p and d such that for all* $w \in H^1(\mathbb{R}^d) \cap L^{2p}(\mathbb{R}^d)$,

$$\|w\|_{L^{p+1}(\mathbb{R}^d)} \le C_0 \|\nabla w\|^{\theta}_{L^2(\mathbb{R}^d)} \|w\|^{1-\theta}_{L^{2p}(\mathbb{R}^d)}, \quad \textit{where } \theta = \frac{d(1-p)}{(1+p)(2p+d(1-p))} \in (0,1).$$

Proof. The proof is not difficult but technical. The idea is to insert the explicit expressions for $h(u)$, $V(x)$, and u_∞ into the relation (3.28) between the entropy and the entropy dissipation. A computation leads to

$$\int_{\mathbb{R}^d} u^m dx \le A \int_{\mathbb{R}^d} |\nabla u^{m-1/2}|^2 dx + B\Big(\int_{\mathbb{R}^d} u dx\Big)^\gamma,$$

for some positive constants A, B, γ and with $m = (p+1)/(2p) > 1$. Next, defining $u(x) = \lambda^{d/m} v(\lambda x)$ and passing to the new integration variable $y = \lambda x$, a suitable choice of λ gives

$$\int_{\mathbb{R}^d} v^m dx \le C\Big(\int_{\mathbb{R}^d} |\nabla v^{m-1/2}|^2 dy\Big)^\mu \Big(\int_{\mathbb{R}^d} v dy\Big)^{\gamma(1-\mu)}.$$

Finally, the choice $w = v^{m-1/2}$ yields the result. Thus, we divide the proof into three steps.

Step 1: Reformulation of (3.28). Choosing $V(x) = \frac{1}{2}|x|^2$, the right-hand side of (3.28) becomes

$$\begin{aligned}\int_{\mathbb{R}^d} u|\nabla(h(u)+V)|^2 dx &= \int_{\mathbb{R}^d} u\Big|\frac{m}{m-1}\nabla u^{m-1} + x\Big|^2 dx \\ &= \Big(\frac{m}{m-1}\Big)^2 \int_{\mathbb{R}^d} u|\nabla u^{m-1}|^2 dx + \int_{\mathbb{R}^d} u|x|^2 dx + \frac{2m}{m-1}\int_{\mathbb{R}^d} ux\cdot\nabla u^{m-1} dx \\ &= \Big(\frac{m}{m-1/2}\Big)^2 \int_{\mathbb{R}^d} |\nabla u^{m-1/2}|^2 dx + \int_{\mathbb{R}^d} u|x|^2 dx + 2\int_{\mathbb{R}^d} x\cdot\nabla u^m dx.\end{aligned}$$

By integration by parts, the last integral writes as

$$2\int_{\mathbb{R}^d} x\cdot\nabla u^m dx = -2\int_{\mathbb{R}^d} \operatorname{div}(x) u^m dx = -2d\int_{\mathbb{R}^d} u^m dx.$$

Therefore, (3.28) is equivalent to

$$\begin{aligned}\int_{\mathbb{R}^d}\Big(\frac{u^m}{m-1} + \frac{1}{2}u|x|^2\Big)dx - H[u_\infty] = H^*[u] &\le \frac{1}{2}\int_{\mathbb{R}^d} u|\nabla(h(u)+V)|^2 dx \\ &= \frac{1}{2}\Big(\frac{m}{m-1/2}\Big)^2 \int_{\mathbb{R}^d} |\nabla u^{m-1/2}|^2 dx + \frac{1}{2}\int_{\mathbb{R}^d} u|x|^2 dx - d\int_{\mathbb{R}^d} u^m dx.\end{aligned}$$

Rearranging terms on both sides leads to

$$\Big(d + \frac{1}{m-1}\Big)\int_{\mathbb{R}^d} u^m dx \le \frac{1}{2}\Big(\frac{m}{m-1/2}\Big)^2 \int_{\mathbb{R}^d} |\nabla u^{m-1/2}|^2 dx + H[u_\infty]. \tag{3.30}$$

It remains to compute $H[u_\infty]$. The steady state u_∞ is the Barenblatt profile (3.27). With the transformation $y = x/\sqrt{N}$, we obtain

$$u_\infty(x) = \Big(N - \frac{m-1}{2m}|x|^2\Big)_+^{1/(m-1)} = N^{1/(m-1)} \underbrace{\Big(1 - \frac{2m}{m-1}|y|^2\Big)_+^{1/(m-1)}}_{=:U_m(y)}.$$

Using $|x|^2 = N|y|^2$ and $dx = N^{d/2}dy$, we compute

$$\begin{aligned} H[u_\infty] &= \frac{1}{m-1}\int_{\mathbb{R}^d} u_\infty^m dx + \frac{1}{2}\int_{\mathbb{R}^d}|x|^2 u_\infty dx \\ &= \frac{1}{m-1}N^{m/(m-1)}\int_{\mathbb{R}^d} U_m(y)^m dx + \frac{1}{2}N^{1/(m-1)}\int_{\mathbb{R}^d}|x|^2 U_m(y)dx \\ &= \frac{1}{m-1}N^{m/(m-1)+d/2}\int_{\mathbb{R}^d} U_m(y)^m dy + \frac{1}{2}N^{1/(m-1)+1+d/2}\int_{\mathbb{R}^d}|y|^2 U_m(y)dy \\ &= N^{m/(m-1)+d/2}\underbrace{\Big(\frac{1}{m-1}\int_{\mathbb{R}^d} U_m(y)^m dy + \frac{1}{2}\int_{\mathbb{R}^d}|y|^2 U_m(y)dy\Big)}_{=:K_m}, \end{aligned}$$

where $K_m > 0$ is a constant which depends only on m. Furthermore, we have

$$\int_{\mathbb{R}^d} u_\infty(x)dx = N^{1/(m-1)+d/2}\int_{\mathbb{R}^d} U_m(y)dy.$$

Solving for N gives

$$N^{m/(m-1)+d/2} = \Big(\int_{\mathbb{R}^d} U_m(y)dy\Big)^{-\gamma}\Big(\int_{\mathbb{R}^d} u_\infty(x)dx\Big)^{\gamma}, \quad \text{where } \gamma := \frac{m/(m-1)+d/2}{1/(m-1)+d/2},$$

Therefore, setting

$$L_m := \Big(\int_{\mathbb{R}^d} U_m(y)dy\Big)^{-\gamma},$$

which is another constant only depending on m, it follows that

$$H[u_\infty] = L_m\Big(\int_{\mathbb{R}^d} u_\infty(x)dx\Big)^{\gamma} = L_m\Big(\int_{\mathbb{R}^d} u dx\Big)^{\gamma},$$

since $u(t)$ has the same mass as u_∞. Going back to (3.30), we have shown that

$$\int_{\mathbb{R}^d} u^m dx \le A\int_{\mathbb{R}^d}|\nabla u^{m-1/2}|^2 dx + B\Big(\int_{\mathbb{R}^d} u dx\Big)^{\gamma}, \tag{3.31}$$

where

$$A := \frac{1}{2}\Big(d + \frac{1}{m-1}\Big)^{-1}\Big(\frac{m}{m-1/2}\Big)^2, \quad B := \Big(d + \frac{1}{m-1}\Big)^{-1}L_m.$$

Step 2: Optimization. We optimize (3.31) by defining a function v by $u(x) = \lambda^{d/m}v(\lambda x)$, $\lambda > 0$. Then, substituting $y = \lambda x$ with $dy = \lambda^d dx$,

$$\begin{aligned} \int_{\mathbb{R}^d} u^m(x)dx &= \lambda^d\int_{\mathbb{R}^d} v^m(\lambda x)dx = \int_{\mathbb{R}^d} v^m(y)dy, \\ \int_{\mathbb{R}^d} u(x)dx &= \lambda^{d/m}\int_{\mathbb{R}^d} v(\lambda x)dx = \lambda^{d/m-d}\int_{\mathbb{R}^d} v(y)dy, \\ \int_{\mathbb{R}^d}|\nabla_x u^{m-1/2}|^2 dx &= \lambda^{d(2m-1)/m}\int_{\mathbb{R}^d}\lambda^2|\nabla_y v^{m-1/2}|^2\lambda^{-d}dy = \lambda^{d-d/m+2}\int_{\mathbb{R}^d}|\nabla_y v^{m-1/2}|^2 dy. \end{aligned}$$

Inequality (3.31) can be written in terms of v as

$$\int_{\mathbb{R}^d} v^m dy \le \lambda^{d-d/m+2} \underbrace{A\int_{\mathbb{R}^d} |\nabla_y v^{m-1/2}|^2 dy}_{=:A_0} + \lambda^{-\gamma(d-d/m)} \underbrace{B\Big(\int_{\mathbb{R}^d} v dy\Big)^{\gamma}}_{=:B_0}.$$

Ww choose an appropriate $\lambda > 0$ which minimizes the right-hand side. The minimum of the function $f(\lambda) = \lambda^a A_0 + \lambda^{-b} B_0$ with $a = d - d/m + 2$, $b = \gamma(d - d/m)$ is given by $\lambda^* = cA_0^{-1/(a+b)} \times B_0^{1/(a+b)}$, where $c = (b/a)^{1/(a+b)}$. Choosing $\lambda = \lambda_*$ in the above inequality, we arrive at

$$\begin{aligned}\int_{\mathbb{R}^d} v^m dy &\le \lambda_*^a A_0 + \lambda_*^b B_0 \\ &\le c^a A_0^{-a/(a+b)+1} B_0^{a/(a+b)} + c^{-b} A_0^{b/(a+b)} B_0^{-b/(a+b)+1} \\ &\le \max\{c^a, c^{-b}\} A_0^{b/(a+b)} B_0^{a/(a+b)},\end{aligned}$$

where

$$\frac{b}{a+b} = \frac{\gamma d(m-1)}{(\gamma+1)d(m-1)+2m}, \quad \frac{a}{a+b} = \frac{d(m-1)+2m}{(\gamma+1)d(m-1)+2m}.$$

Hence, there exists $C_1 > 0$ only depending on d and m such that

$$\int_{\mathbb{R}^d} v^m dy \le C_1 \Big(\int_{\mathbb{R}^d} |\nabla_y v^{m-1/2}|^2 dy\Big)^{b/(a+b)} \Big(\int_{\mathbb{R}^d} v dy\Big)^{\gamma a/(a+b)}.$$

Step 3: Transformation. Changing the function to $w := v^{m-1/2} = v^{1/(2p)}$, where $p = 1/(2m-1)$, we infer that $v^m = w^{2m/(2m-1)} = w^{p+1}$ and $v = w^{2/(2m-1)} = w^{2p}$, and hence,

$$\int_{\mathbb{R}^d} w^{p+1} dy \le C_1 \Big(\int_{\mathbb{R}^d} |\nabla w|^2 dy\Big)^{b/(a+b)} \Big(\int_{\mathbb{R}^d} w^{2p} dy\Big)^{\gamma a/(a+b)} = C_1 \|\nabla w\|_{L^2(\mathbb{R}^d)}^{2b/(a+b)} \|w\|_{L^{2p}(\mathbb{R}^d)}^{2p\gamma a/(a+b)}.$$

Taking the $(p+1)$-th root, we conclude that

$$\|w\|_{L^{p+1}(\mathbb{R}^d)} \le C_1^{1/(p+1)} \|\nabla w\|_{L^2(\mathbb{R}^d)}^{2b/((a+b)(p+1))} \|w\|_{L^{2p}(\mathbb{R}^d)}^{2p\gamma a/((a+b)(p+1))},$$

and the theorem follows for $C_0 = C_1^{1/(p+1)}$ and $\theta = 2b/((a+b)(p+1))$. □

4. Further applications

In this section we discuss some additional topics.

4.1. **Systematic entropy construction method.** In Lemma 3.4 we have proved an estimate for the time derivative of the entropy dissipation. The proof is based on suitable integrations by parts. One may ask if the integrations by parts can be made systematic. This is indeed possible; see Jüngel-Matthes 2006 [26].

In order to motivate the method, we consider the thin-film equation

$$u_t = -(u^\beta u_{xxx})_x, \quad u(\cdot,0) = u_0 \quad \text{in } \mathbb{T},\ t > 0, \tag{4.32}$$

where $\mathbb{T}^d$ is the one-dimensional torus. This equation models the flow of a thin liquid along a solid surface with film height $u(x,t)$ (for $\beta = 2$ or $\beta = 3$) or the thin neck of a Hele-Shaw flow in the lubrication approximation (for $\beta = 1$); see the review of Myers 1998 [33]. We wish to determine

Lyapunov functionals $H_\alpha[u] = 1/(\alpha(\alpha-1))\int_{\mathbb{T}} u^\alpha dx$, $\alpha \neq 0, 1$, and $H_1[u] = \int_{\mathbb{T}} u(\log u - 1)dx$. We assume that there exists a smooth positive solution u to (4.32) and compute formally:

$$\frac{dH_\alpha}{dt}[u(t)] = \frac{1}{\alpha-1}\int_{\mathbb{T}} u^{\alpha-1}u_t dx = \int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xxx}u_x dx.$$

It is reasonable to eliminate the third-order derivative by integrating by parts:

$$\frac{dH_\alpha}{dt}[u(t)] = -(\alpha+\beta-2)\int_{\mathbb{T}} u^{\alpha+\beta-3}u_x^2 u_{xx}dx - \int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xx}^2 dx. \tag{4.33}$$

The second integral on the right-hand side has a good sign. For the first integral, we observe that $u_x^2 u_{xx} = u_x^3/3$ and integrate by parts again:

$$\frac{dH_\alpha}{dt}[u(t)] = -\frac{1}{3}(\alpha+\beta-2)(\alpha+\beta-3)\int_{\mathbb{T}} u^{\alpha+\beta-4}u_x^4 dx - \int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xx}^2 dx.$$

Thus, $H_\alpha[u]$ is a Lyapunov functional for the thin-film equation if $(\alpha+\beta-2)(\alpha+\beta-3) \geq 0$ or, equivalently, $2 \leq \alpha+\beta \leq 3$. Is this condition optimal? No, it is not. The reason is that the second integral can be used to estimate the first integral even when it is positive. The optimal result is as follows.

Proposition 4.1 (Lyapunov functionals for the thin-film equation)**.** *Let $\alpha, \beta > 0$ such that $\frac{3}{2} \leq \alpha+\beta \leq 3$ and let $u(t)$ be a smooth positive solution to* (4.32)*. Then $H_\alpha[u]$ is a Lyapunov functional for* (4.32).

Before we prove the proposition, we explain the entropy construction method. We wish to prove that the entropy dissipation $D_\alpha = -dH_\alpha[u]/dt$ is nonnegative. First, we need to systematize integration by parts. The first integration by parts leading to (4.33) consists in the following identity:

$$D_\alpha = -\int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xxx}u_x dx = (\alpha+\beta-2)\int_{\mathbb{T}} u^{\alpha+\beta-3}u_x^2 u_{xx}dx + \int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xx}^2 dx. \tag{4.34}$$

This can be written equivalently as

$$I_2 = \int_{\mathbb{T}} u^{\alpha+\beta}\Big((\alpha+\beta-2)\Big(\frac{u_x}{u}\Big)^2\frac{u_{xx}}{u} + \Big(\frac{u_{xx}}{u}\Big)^2 + \frac{u_x}{u}\,\frac{u_{xxx}}{u}\Big)dx = \int_{\mathbb{T}}\Big(u^{\alpha+\beta}\frac{u_x}{u}\,\frac{u_{xx}}{u}\Big)_x dx = 0.$$

Thus, (4.34) is equivalent to

$$D_\alpha = D_\alpha + c\cdot I_2 \quad \text{with } c = 1.$$

This expression is trivial since $I_2 = 0$, but after expanding I_2, both sides of the equation have different forms. We call I_2 an integration-by-parts formula. How many formulas do exist? There are two other integrals:

$$I_1 = \int_{\mathbb{T}} u^{\alpha+\beta}\Big((\alpha+\beta-3)\Big(\frac{u_x}{u}\Big)^4 + 3\Big(\frac{u_x}{u}\Big)^2\frac{u_{xx}}{u}\Big)dx = \int_{\mathbb{T}}\Big(u^{\alpha+\beta}\Big(\frac{u_x}{u}\Big)^3\Big)_x dx = 0,$$
$$I_3 = \int_{\mathbb{T}} u^{\alpha+\beta}\Big((\alpha+\beta-1)\frac{u_x}{u}\,\frac{u_{xxx}}{u} + \frac{u_{xxxx}}{u}\Big) = \int_{\mathbb{T}}\Big(u^{\alpha+\beta}\frac{u_{xxx}}{u}\Big)_x = 0.$$

The number of integration-by-parts formulas is determined by all integers p_1, p_2, and p_3 such that $1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 = 3$, and thus

$$(p_1, p_2, p_3) \in \{(3,0,0), (1,1,0), (0,0,1)\},$$

and there are exactly three such formulas. We conclude that *all* possible integrations by parts are given by the linear combinations

$$D_\alpha = D_\alpha + c_1 I_1 + c_2 I_2 + c_3 I_3.$$

Again, this expression is trivial since $I_1 = I_2 = I_3 = 0$. The goal is to find constants c_1, c_2, $c_3 \in \mathbb{R}$ such that the right-hand side is nonnegative:

$$\exists c_1, c_2, c_3 \in \mathbb{R}: \ D_\alpha + c_1 I_1 + c_2 I_2 + c_3 I_3 \geq 0. \tag{4.35}$$

Our main idea to prove such an inequality is to identify the derivatives u_x/u, u_{xx}/u, etc. with the polynomial variables ξ_1, ξ_2, etc.:

$$\xi_k = \frac{1}{u} \frac{\partial^k u}{\partial x^k}, \quad k \in \mathbb{N}_0.$$

Setting $\xi = (\xi_1, \xi_2, \xi_3, \xi_4)$,

$$\begin{aligned}
D_\alpha &\quad \text{corresponds to} \quad S(\xi) = -\xi_1 \xi_3, \\
I_1 &\quad \text{corresponds to} \quad T_1(\xi) = (\alpha + \beta - 3)\xi_1^4 + 3\xi_1^2 \xi_2, \\
I_2 &\quad \text{corresponds to} \quad T_2(\xi) = (\alpha + \beta - 2)\xi_1^2 \xi_2 + \xi_2^2 + \xi_1 \xi_3, \\
I_3 &\quad \text{corresponds to} \quad T_3(\xi) = (\alpha + \beta - 1)\xi_1 \xi_3 + \xi_4.
\end{aligned}$$

We call T_j shift polynomials since they allow us to "shift" partial derivatives. Then $D_\alpha \geq 0$ is proved if we are able to show that

$$\exists c_1, c_2, c_3 \in \mathbb{R}: \ \forall \xi \in \mathbb{R}^4: \ (S + c_1 T_1 + c_2 T_2 + c_3 T_3)(\xi) \geq 0. \tag{4.36}$$

Such problems are called polynomial decision problems which are well known in real algebraic geometry. It was shown by Tarksi in 1951 [38] that they can be always treated in the following way:

> A quantified statement about polynomials can be reduced to a quantifier-free statement about polynomials in an algorithmic way.

Solution algorithms for the above quantifier elimination problem have been implemented, for instance, in `Mathematica`. There are also tools specialized on quantifier elimination, like `QEPCAD` (Quantifier Elimination using Partial Cylindrical Algebraic Decomposition), see Collins-Hong 1991 [10]. The advantage of these algorithms is that the solution is complete and exact. The disadvantage is that the complexity of the algorithms is doubly exponential in the number of variables ξ_i and c_i. An alternative approach is given by the sum-of-squares (SOS) method. This method tries to write the polynomial as a sum of squares. The answer may be not complete since there are polynomials which are nonnegative but which cannot be written as a sum of squares. Fortunately, for the above decision problem, we can solve it directly without going into real algebraic geometry.

Notice that problems (4.35) and (4.36) are not equivalent since we estimate the integrands in (4.36) which is stronger than estimating the integrals in (4.35). However, it is possible to prove, at least in specific situations (for instance, for the one-dimensional thin-film equation), that no estimates are lost [26].

Now, we are able to prove the above proposition.

Proof of Proposition 4.1. We need to prove that, for some choice of c_j, for all $\xi \in \mathbb{R}^4$,

$$\begin{aligned}(S + c_1T_1 + c_2T_2 + c_3T_3)(\xi) = c_1(\alpha+\beta-3)\xi_1^4 + (3c_1 + c_2(\alpha+\beta-2))\xi_1^2\xi_2 \\ + c_2\xi_2^2 + (c_2 + c_3(\alpha+\beta-1)-1)\xi_1\xi_3 + c_3\xi_4 \geq 0.\end{aligned}$$

The variable ξ_4 appears only in the term $c_3\xi_4$ which is indefinite. Therefore, we take $c_3 = 0$. Furthermore, ξ_3 only appears in the term $\xi_1\xi_3$ which is also indefinite. We choose $c_2 = 1$ to eliminate this term. It remains to show that there exists $c_1 \in \mathbb{R}$ such that

$$(S + c_1T_1 + 1\cdot T_2 + 0\cdot T_3)(\xi) = c_1(\alpha+\beta-3)\xi_1^4 + (3c_1 + (\alpha+\beta-2))\xi_1^2\xi_2 + \xi_2^2 \geq 0.$$

We employ the following elementary result (see Jüngel-Matthes 2006 [26] for a proof).

Lemma 4.2. *The inequality*

$$a_1\xi_1^4 + a_2\xi_1^2\xi_2 + a_3\xi_2^2 \geq 0$$

is satisfied for all $(\xi_1,\xi_2) \in \mathbb{R}^2$ *if and only if*

$$\textit{either} \quad a_3 > 0,\ 4a_1a_3 - a_2^2 \geq 0, \quad \textit{or} \quad a_3 = 0,\ a_2 = 0,\ a_1 \geq 0.$$

Since the coefficient for ξ_2^2 is positive, the nonnegativity is guaranteed if and only if

$$\begin{aligned}0 \leq 4a_1a_3 - a_2^2 &= 4c_1(\alpha+\beta-3) - (\alpha+\beta-2+3c_1)^2 \\ &= -9\Big(c_1 + \frac{1}{9}(\alpha+\beta)\Big)^2 - \frac{8}{9}(\alpha+\beta)^2 + 4(\alpha+\beta) - 4.\end{aligned}$$

Choosing the maximizing value $c_1 = -(\alpha+\beta)/9$, this inequality is satisfied if and only if

$$0 \leq -8(\alpha+\beta)^2 + 36(\alpha+\beta) - 36.$$

The roots of the polynomial $x \mapsto -8x^2 + 36x - 36$ are $x_1 = 3/2$ and $x_2 = 3$. Therefore, the inequality is valid if and only if $\frac{3}{2} \leq \alpha+\beta \leq 3$, which proves the claim. □

In the following, we discuss some extensions and generalizations of the entropy construction method.

Entropy dissipation estimates. Estimates of the entropy dissipation are very useful to derive a priori estimates which are needed, for instance, in the existence or long-time analysis.

Proposition 4.3 (Entropy dissipation estimates)**.** *Let* $\alpha, \beta > 0$ *be such that* $\frac{3}{2} < \alpha+\beta \leq 3$ *and let* $u(t)$ *be a smooth positive solution to* (4.32)*. Then there exists* $0 < c < 1$ *depending only on* α *and* β *such that*

$$\frac{dH_\alpha}{dt}[u(t)] + c\int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xx}^2 dx \leq 0.$$

Proof. The idea is to identify the integral $\int_{\mathbb{T}} u^{\alpha+\beta-2}u_{xx}^2 dx$ with a polynomial and to apply the entropy construction method. In the present case, $P(\xi) = \xi_2^2$, and thus, we have to show:

$$\exists c_1, c_2, c_3 \in \mathbb{R},\ c > 0:\ \forall \xi \in \mathbb{R}^4:\ (S + c_1T_1 + c_2T_2 + c_3T_3 - cP)(\xi) \geq 0.$$

As in the proof of the previous proposition, we choose $c_3 = 0$ and $c_2 = 1$. Then our problem reduces to

$$\begin{aligned}&\exists c_1 \in \mathbb{R},\ c > 0 : \forall (\xi_1, \xi_2) \in \mathbb{R}^2 :\\ &c_1(\alpha+\beta-3)\xi_1^4 + (3c_1 + (\alpha+\beta-2))\xi_1^2\xi_2 + (1-c)\xi_2^2 \geq 0.\end{aligned}$$

A necessary condition for nonnegativity is $c < 1$. Then Lemma 4.2 gives the additional condition

$$\begin{aligned}0 &\leq 4a_1a_3 - a_2^2\\ &= -9\Big(c_1 + \frac{1}{9}((1+2c)(\alpha+\beta) - 6c)\Big)^2 + \frac{1}{9}((1+2c)(\alpha+\beta)-6c)^2 - (\alpha+\beta-2)^2.\end{aligned}$$

We choose c_1 such that the first bracket vanishes:

$$\begin{aligned}0 &\leq \frac{1}{9}((1+2c)(\alpha+\beta)-6c)^2 - (\alpha+\beta-2)^2\\ &= \frac{4}{9}(c-1)((\alpha+\beta)-3)((\alpha+\beta)(2+c)-3(c+1)).\end{aligned}$$

This inequality is satisfied of and only if

$$\frac{3(1-c)}{2+c} \leq \alpha+\beta \leq 3.$$

If we choose $\frac{3}{2} < \alpha+\beta \leq 3$, there exists $c > 0$ such that this inequality holds. □

In a similar way, we can prove that, for $\frac{3}{2} < \alpha+\beta < 3$, there exists $c > 0$ such that

$$\begin{aligned}\frac{dH_\alpha}{dt}[u(t)] + c\int_{\mathbb{T}} (u^{(\alpha+\beta)/2})_{xx}^2 dx \leq 0\\ \frac{dH_\alpha}{dt}[u(t)] + c\int_{\mathbb{T}} (u^{(\alpha+\beta)/4})_x^4 dx \leq 0.\end{aligned}$$

Higher-order entropies. Another extension of the method concerns higher-order entropies, like the first-order entropies

$$E_\alpha[u(t)] = \int_{\mathbb{T}} (u^{\alpha/2})_x^2 dx, \quad \alpha > 0.$$

Taking the time derivative, we find that

$$\begin{aligned}\frac{dE_\alpha}{dt}[u(t)] &= 2\int_\Omega (u^{\alpha/2})_x(u^{\alpha/2})_{tx}dx = -2\int_\Omega (u^{\alpha/2})_{xx}\frac{\alpha}{2}u^{\alpha/2-1}u_t dx\\ &= -\alpha\int_\Omega (u^{\alpha/2})_{xx}u^{\alpha/2-1}(u^\beta u_{xxx})_x dx.\end{aligned}$$

Thus, we have to find all integration-by-parts rules involving a total of six derivatives. It can be seen that there are seven integration-by-parts formulas giving seven shift polynomials. Some of the shift polynomials do not need to be taken into account, like

$$T(\xi) = (\alpha+\beta-1)\xi_1\xi_5 + \xi_6,$$

since the term ξ_6 is indefinite. Eventually, the decision problem becomes

$$\begin{aligned}\exists c_1, c_2 \in \mathbb{R} : \forall \xi \in \mathbb{R}^3 : \quad & (\alpha+\beta-5)c_1\xi_1^6 + (5c_1 + (\alpha+\beta-4)c_2)\xi_1^4\xi_2 + 3c_2\xi_1^2\xi_2^2 \\ & + (\tfrac{1}{2}(\alpha^2-5\alpha+6)+c_2)\xi_1^3\xi_3 + (2\alpha-4)\xi_1\xi_2\xi_3 + \xi_3^2 \geq 0.\end{aligned}$$

It can be solved by employing the following lemma whose proof can be found in Jüngel-Matthes 2006 [26].

Lemma 4.4. *Let the real polynomial*

$$P(\xi_1,\xi_2,\xi_3) = a_1\xi_1^6 + a_2\xi_1^4\xi_2 + a_3\xi_1^3\xi_3 + a_4\xi_1^2\xi_2^2 + a_5\xi_1\xi_2\xi_3 + \xi_3^2$$

be given. Then the quantified formula

$$\forall \xi_1,\xi_2,\xi_3 \in \mathbb{R} : \; P(\xi_1,\xi_2,\xi_3) \geq 0$$

is equivalent to the quantifier free formula

$$\begin{aligned} &\textit{either } 4a_4 - a_5^2 > 0 \quad \textit{and} \quad 4a_1a_4 - a_1a_5^2 - a_2^2 - a_3^2a_4 + a_2a_3a_5 \geq 0 \\ &\textit{or } 4a_4 - a_5^2 = 2a_2 - a_3a_5 = 0 \quad \textit{and} \quad 4a_1 - a_3^2 \geq 0.\end{aligned}$$

The result is displayed in 4.1. (The same result has been found first by Laugesen 2005 using a different method.) Notice that there is always a trivial first-order entropy corresponding to $\alpha = 2$, $E_2[u] = \int_{\mathbb{T}} u_x^2 dx$. We summarize the result in the following proposition.

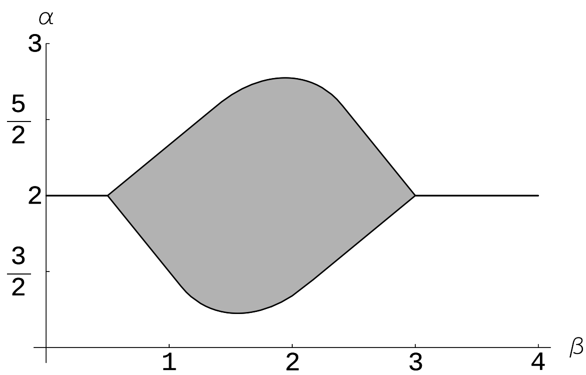

FIGURE 4.1. Values of α and β providing an entropy for the one-dimensional thin-film equation.

Proposition 4.5 (First-order entropies for the thin-film equation)**.** *Let* $(\alpha,\beta) \in \mathbb{R}^2$ *be an element of the gray region of Figure 4.1. Then* $H_\alpha[u]$ *is a Lyapunov functional for* (4.32).

Multi-dimensional equations. So far, we have discussed the one-dimensional case only. In principle, the above strategy can be generalized in a straightforward way to multi-dimensional equations. In this situation, we introduce polynomial variables for all the partial derivatives and shift polynomials for all integration-by-parts formulas by differentiating products in all variables. Practically, this strategy is useless since it leads to polynomial expressions in many variables ξ_k and a huge number of shift polynomials T_j. A better approach is not to incorporate all products of differential expressions but only those which have "symmetry" properties, like $|\nabla u|/u$, $\Delta u/u$, or $\nabla^2 u/u$.

As an example, we consider the Derrida-Lebowitz-Speer-Spohn (DLSS) equation

$$u_t + \nabla^2 : (u\nabla^2 \log u) = 0, \quad u(\cdot,0) = u_0 \geq 0 \quad \text{in } \mathbb{T}^d,\ t > 0. \tag{4.37}$$

Here $\nabla^2 u$ denotes the Hessian of u and $A : B = \sum_{i,j} A_{ij}B_{ij}$ for two matrices $A = (A_{ij})$ and $B = (B_{ij})$. The function $u(x,t)$ models the electron density in a quantum semiconductor in which electron-lattice interactions are strong; see Degond-Méhats-Ringhofer 2005 [13] for a derivation. It was shown in Jüngel-Matthes 2008 [27] that there exists a nonnegative weak solution to (4.37). We assume that u is positive and smooth to simplify the presentation. The arguments can be made rigorous for nonnegative weak solutions; see Jüngel-Matthes 2008 [27].

We differentiate formally the entropy functional

$$H_\alpha[u(t)] = \frac{1}{\alpha(\alpha-1)}\int_{\mathbb{T}^d} u^\alpha dx, \quad \alpha \neq 0, 1,$$

yielding, after integration by parts,

$$D_\alpha = -\frac{dH_\alpha}{dt}[u] = -\frac{1}{\alpha-1}\int_{\mathbb{T}^d} u^{\alpha-1}u_t dx = \frac{1}{\alpha-1}\int_{\mathbb{T}^d} u\nabla^2(u^{\alpha-1}) : \nabla^2 \log u dx.$$

We set $u = v^2$ which is possible since $u \geq 0$. Then a tedious computation shows that

$$D_\alpha = 4\int_{\mathbb{T}^d} v^{2\alpha}\Big(\frac{\|\nabla^2 v\|^2}{v^2} - 2(2-\alpha)\frac{\nabla v}{v}\,\frac{\nabla^2 v}{v}\,\frac{\nabla v}{v} + (3-2\alpha)\frac{|\nabla v|^4}{v^4}\Big)dx.$$

This motivates us to introduce the functions θ, λ, μ, respectively, by

$$\theta = \frac{|\nabla v|}{v}, \quad \lambda = \frac{1}{d}\,\frac{\Delta v}{v}, \quad (\lambda+\mu)\theta^2 = \frac{\nabla v}{v}\,\frac{\nabla^2}{v}\,\frac{\nabla v}{v},$$

and $\rho \geq 0$ by

$$\|\nabla^2 v\|^2 = \Big(d\lambda^2 + \frac{d}{d-1}\mu^2 + \rho^2\Big)v^2.$$

It can be shown that ρ is well defined. This follows from the inequality

$$\|\nabla^2 v\|^2 \geq \Big(d\lambda^2 + \frac{d}{d-1}\mu^2\Big)v^2,$$

which is proved in Jüngel-Matthes 2008 [27]. The integral D_α is expressed by these functions as

$$D_\alpha = \int_{\mathbb{T}^d} v^{2\alpha}\Big(d\lambda^2 + \frac{d}{d-1}\mu^2 + \rho^2 - 2(2-\alpha)(\lambda+\mu)\theta^2 + (3-2\alpha)\theta^4\Big)dx.$$

It turns out that just two integration-by-parts formulas are sufficient to prove the nonnegativity od D_α for certain α:

$$I_1 = \int_{\mathbb{T}^d} \operatorname{div}(v^{2\alpha-2}(\nabla^2 v - \Delta v \mathbb{I})\nabla v)dx,$$
$$I_2 = \int_{\mathbb{T}^d} \operatorname{div}(v^{2\alpha-3}|\nabla v|^2 \nabla v)dx,$$

where $\mathbb{I}$ denotes the unit matrix in $\mathbb{R}^{d\times d}$. In view of the periodic boundary conditions, $I_1 = I_2 = 0$. The goal is to find constants c_1, $c_2 \in \mathbb{R}$ such that

$$D_\alpha = D_\alpha + c_1 I_1 + c_2 I_2 \geq 0.$$

In terms of the above variables, this sum can be written as

$$D_\alpha = \int_{\mathbb{T}^d} v^{2\alpha}[d\lambda^2(1-(d-1)c_1) + \lambda\theta^2(2(\alpha-1)(1-(d-1)c_1) + (d-2)c_2 - 2)$$
$$+ Q(\theta,\mu,\rho)]dx,$$

where Q is a polynomial in θ, μ, and ρ with coefficients depending on c_1 and c_2 not not on λ. We choose to eliminate λ from the above integrand (although this may be not the optimal choice). Thus, we choose (c_1, c_2) as the solution to the linear system

$$1-(d-1)c_1 = 0, \quad 2(\alpha-1)(1-(d-1)c_1) + (d-2)c_2 - 2 = 0.$$

With this choice, the polynomial Q can be estimated by

$$Q(\theta,\mu,\rho) = b_1\mu^2 + 2b_2\mu\theta^2 + b_3\theta^4 + b_4\rho^2 \geq b_1\mu^2 + 2b_2\mu\theta^2 + b_3\theta^4,$$

since $b_4 = d(d+2)(d-1) \geq 0$. Here, b_1, b_2, and b_3 are coefficients which depend only on d and α. It remains to determine the conditions on these coefficients such that the quadratic polynomial in μ and θ^2 is nonnegative. A computation shows that this is the case if $0 < \alpha < 2(d+1)/(d+2)$. We have proved the following result.

Theorem 4.6 (Entropies for the DLSS equation)**.** *Let* $d \geq 1$, $0 < \alpha < 2(d+1)/(d+2)$. *Then* $H_\alpha[u]$ *is a Lyapunov functional for* (4.37).

4.2. **Entropy variables and cross-diffusion systems.** In the previous sections, we have considered scalar PDEs only. Stronly coupled systems of PDEs are much more difficult to treat since some standard tools available for scalar equations (maximum principle for second-order equations, regularity theory) often cannot be used. In this section, we show how the concept of entropy can help to analyze cross-diffusion systems. These are systems of parabolic or elliptic PDEs whose diffusion matrix is dense (i.e., it is neither diagonal nor tridiagonal). We consider only those systems which possess a logarithmic entropy (the reason will become clear later).

To fix ideas, let us investigate a model from population dynamics. Let u, v be the densities of two competing species. Their dynamics is governed by the continuity equations

$$u_t + \operatorname{div} J_u = 0, \quad v_t + \operatorname{div} J_v = 0 \quad \text{in } \Omega,\ t > 0,$$

where $\Omega \subset \mathbb{R}^d$ is a bounded domain. For simplicity, we neglect source terms (of Lotka-Volterra type). We assume that the flux J_u is completely defined by the diffusion of the two species: the self-diffusion $(a+v)\nabla u$ and the cross-diffusion $u\nabla v$. The basic idea is that the primary cause of dispersal is migration to avoid crowding instead of just random motion (modeled by the diffusion term $a\nabla u$). In particular, the spatial variation of the competiting species, ∇v, influences the flux of species u. Similarly, we define $J_v = (b+u)\nabla v + v\nabla u$. This leads to the following system:

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} - \operatorname{div}\left(\begin{pmatrix} a+v & u \\ v & b+u \end{pmatrix} \nabla \begin{pmatrix} u \\ v \end{pmatrix} \right) = 0 \quad \text{in } \Omega,\ t>0. \tag{4.38}$$

We supplement this system by initial and homogeneous Neumann boundary conditions:

$$u(\cdot,0) = u_0,\ u(\cdot,0) = v_0 \quad \text{in } \Omega, \quad \nabla u\cdot\nu = \nabla v\cdot\nu = 0 \quad \text{on } \partial\Omega,\ t>0, \tag{4.39}$$

where ν denotes the exterior unit normal vector to $\partial\Omega$. Equations (4.38)-(4.39) are a simplified version of a population model first suggested by Shigesada, Kawasaki, and Teramoto 1979 [36]. This model has attracted the attention of many mathematicians since it may have spatial pattern exhibiting segregation of species.

There are a number of mathematical problems. First, the diffusion matrix is generally neither symmetric nor positive definite, and hence, even the local-in-time existence of solutions is not obvious. Second, the strong coupling prohibits the application of the maximum or minimum principle such that the positivity of the population densities u and v cannot be proved. The solution to these problems is to employ the concept of entropy. The entropy is defined by

$$H[u,v] = \int_\Omega h[u,v]dx = \int_\Omega (u(\log u - 1) + v(\log v - 1))dx,$$

where $h[u,v]$ is the entropy density. This is indeed a Lyapunov functional since

$$\begin{aligned} \frac{dH}{dt}[u,v] &= \int_\Omega (u_t \log u + v_t \log v)dx \\ &= -\int_\Omega \Big(((a+v)\nabla u + u\nabla v)\cdot \frac{\nabla u}{u} + ((b+u)\nabla v + v\nabla u)\cdot\frac{\nabla v}{v}\Big)dx \\ &= -4\int_\Omega (a|\nabla\sqrt{u}|^2 + b|\nabla\sqrt{v}|^2 + |\nabla\sqrt{uv}|^2)dx \le 0. \end{aligned}$$

The estimate provides H^1 bounds for $\sqrt{u}$ and $\sqrt{v}$.

These bounds make only sense if u and v are nonnegative. This problem can be overcome by introducing the so-called entropy variables, which symmetrize the above system:

$$y = \frac{\partial h}{\partial u} = \log u, \quad z = \frac{\partial h}{\partial v} = \log v.$$

In the new variables, system (4.38) reads as ´

$$\partial_t \begin{pmatrix} e^y \\ e^z \end{pmatrix} - \operatorname{div}\left(\begin{pmatrix} ae^y + e^{y+z} & e^{y+z} \\ e^{y+z} & be^z + e^{y+z} \end{pmatrix} \nabla \begin{pmatrix} y \\ z \end{pmatrix} \right) = 0. \tag{4.40}$$

It turns out that the new diffusion matrix

$$B(y,z) = \begin{pmatrix} ae^y + e^{y+z} & e^{y+z} \\ e^{y+z} & be^z + e^{y+z} \end{pmatrix}$$

is symmetric and positive definite:

$$x^\top Bx = ae^y x_1^2 + be^z x_2^2 + e^{y+z}(x_1+x_2)^2 \ge \min\{ae^y, be^z\}\|x\|^2 \quad \text{for all } x=(x_1,x_2)^\top \in \mathbb{R}^2.$$

(Notice, however, that B is *not* uniformly positive definite in y and z.) Thus, if we are able to prove the existence of *bounded* solutions (y,z) to (4.40), the functions $u=e^y$ and $v=e^z$ are automatically positive and solutions to the original system (4.38).

Summarizing, the cross-diffusion system (4.38) can be "symmetrized", by a change of unknowns, and it possesses an entropy functional. Both properties are not a coincidence but they are related. In fact, it is well known from the theory of hyperbolic conservation laws that the existence of a symmetric formulation is equivalent to the existence of an entropy functional; see Kawashima and Shizuta 1988 [29]. Interestingly, the result of Kawashima and Shizuta also includes parabolic systems.

Using the above tools, the global-in-time existence of solutions can be proved.

Theorem 4.7 (Global existence of solutions to (4.38).)**.** *Let $\partial\Omega$ be smooth, $u_0, v_0 \in L^\infty(\Omega)$, $u_0 \ge 0$, $v_0 \ge 0$, and $a, b > 0$. Then there exists a weak solution (u,v) to (4.38)-(4.39) satisfying $u, v \ge 0$ in $\Omega\times(0,\infty)$ and*

$$u_t, v_t \in L^1_{\rm loc}(0,\infty;(H^s(\Omega))'), \quad uv \in L^1_{\rm loc}(0,\infty;W^{1,1}(\Omega)), \quad u,v \in L^{4/3}_{\rm loc}(0,\infty;W^{1,4/3}(\Omega)),$$

where $s = 1+d^2/(2d+2)$.

Proof. The proof is lengthy; therefore, we give only the main ideas. The complete proof is given in Chen-Jüngel 2006 [9]. We write (4.40) symbolically as $f(w)_t - \operatorname{div}(B(w)\nabla w) = 0$ for $w=(y,z)\in\mathbb{R}^2$ and $f(w) = (e^y, e^z)$.

- *Definition of the approximated system:* The system is approximated in time by a backward Euler scheme and in space by a Galerkin method (alternatively, one may add a regularizing term $\Delta^m(u,v)$ with sufficiently large $m\in\mathbb{N}$):

  $$\frac{1}{\triangle t}(f(w_N^k) - f(w_N^{k-1})) - \operatorname{div}(B(w_N^k)\nabla w_N^k) = 0 \quad \text{in } \Omega, \quad \nabla w_N^k\cdot\nu = 0 \quad \text{on } \partial\Omega,$$

  where w_N^k approximates $w(x,t_k)$, $t_k = k\triangle t$, and $N\in\mathbb{N}$ is the dimension of the Galerkin space. Employing the convexity of f, it is possible to show a discrete version of the entropy inequality:

  $$H[w_N^k] + P_N^k \le H[w_N^{k-1}],$$

  and P_N^k contains the L^2 norm of $\nabla\sqrt{u_N^k}$ and $\nabla\sqrt{v_N^k}$. Solving this recursive inequality, we find a priori bounds for $\sqrt{u_N^k}$ and $\sqrt{v_N^k}$ in $H^1(\Omega)$.
- *Existence for the approximated system:* The idea is to apply the Leray-Schauder fixed-point theorem. For this, the approximate system is linearized. For given $\widetilde{w}\in L^\infty(\Omega;\mathbb{R}^2)$, solve

  $$\frac{1}{\triangle t}(f(w_N^k) - f(w_N^{k-1})) - \operatorname{div}(B(\widetilde{w})\nabla w_N^k) = 0 \quad \text{in } \Omega, \quad \nabla w_N^k\cdot\nu = 0 \quad \text{on } \partial\Omega.$$

  Since B is symmetric and positive definite, the existence of weak solutions to this linear problem follows from the Lax-Milgram lemma. This defines the fixed-point operator

$\widetilde{w} \mapsto w_N^k$ on appropriate spaces. (They have been chosen such that $w_N^k \in L^\infty(\Omega;\mathbb{R}^2)$.) The discrete entropy inequality provides the uniform estimate needed to apply the fixed-point theorem.

- *Derivation of uniform estimates:* The entropy inequality also gives a priori estimates uniform in the approximation parameters N and $\triangle t$. Then, by the Gagliardo-Nirenberg inequality, further estimates for $w_N = (w_N^k)$ in some Sobolev spaces can be proved. Denoting by $\partial_t^{\triangle t}$ the discrete time derivative, it is possible to conclude from the approximated equations that also $\partial_t^{\triangle t} w_N$ is uniformly bounded in some suitable space.
- *Limit in the approximation parameter:* Finally, we pass to the limit in the approximation parameters N and $\triangle t$. In order to obtain strong convergence of w_N^k, we apply the Aubin lemma in the version of Dreher-Jüngel 2012 [20]. This lemma shows that, if some appropriate estimates on the (discrete) temporal and spatial derivatives on w_N are available, a subsequence of (w_N) converges strongly in some Lebesgue space, say L^2, to a function w, as $N \to \infty$ and $\triangle t \to 0$. The difficulty is to prove that $B(w_N)\nabla w_N$ converges to $B(w)\nabla w$. This is done by using the estimates derived from the Gagliardo-Nirenberg inequality and weak compactness results. The limit function $w = (u, v)$ is shown to be a solution to the original system (4.38)-(4.39).

This finishes the proof. □

In the above example, the entropy method allows us to prove the nonnegativity of the density without applying any maximum principle. One may ask if there are diffusion systems for which not only lower bounds but also upper bounds can be proved by that method. This is indeed the case. As an example, we consider a tumor-growth model. We assume that the tumor is described by the volume fractions of the tumor cells $c(x,t)$, the extra-cellular matrix (ECM) $m(x,t)$, and water $w(x,t)$. The ECM consists of a complex mixture of carbohydrates and proteins (e.g. collagen) providing structural support to the cells. Supposing that the mixture is saturated, we have $w = 1 - c - m$. Jackson and Byrne 2002 [25] derived from a fluiddynamical approach the following (simplified) diffusion system:

$$\partial_t \begin{pmatrix} c \\ m \end{pmatrix} - \operatorname{div}\left(A(c,m)\nabla \begin{pmatrix} c \\ m \end{pmatrix} \right) = 0 \quad \text{in } \Omega,\ t > 0, \tag{4.41}$$

where

$$A(c,m) = \begin{pmatrix} c(1-c) & -\beta cm \\ -cm & \beta m(1-m) \end{pmatrix}$$

is the diffusion matrix, $\beta > 0$ is the ratio of the ECM to the cell pressure constants and $\Omega \subset \mathbb{R}^d$ is a bounded domain. The equations are supplemented by initial and homogeneous boundary conditions for c and m,

$$c(\cdot,0) = c_0, \quad m(\cdot,0) = m_0, \quad \nabla c \cdot \nu = \nabla m \cdot \nu = 0 \quad \text{on } \partial\Omega,\ t > 0.$$

More precisely, the original model is posed on a one-dimensional intervall, assuming some symmetry of the tumor, but the above generalization does not complicate the problem.

We claim that the following entropy functional is a Lyapunov functional:

$$H = \int_\Omega h dx = \int_\Omega (c(\log c - 1) + m(\log m - 1) + (1-c-m)(\log(1-c-m)-1))dx.$$

The entropy density h is the sum of the logarithmic entropies of the three phases c, m, and $w = 1-c-m$. We differentiate formally:

$$\begin{aligned}
\frac{dH}{dt} &= \int_\Omega \Big(c_t \log\frac{c}{1-c-m} + m_t \log\frac{m}{1-c-m}\Big)dx \\
&= -\int_\Omega \Big((c(1-c)\nabla c - \beta c m \nabla m)\cdot\frac{(1-m)\nabla c + c\nabla m}{c(1-c-m)} \\
&\quad + (-cm\nabla c + \beta m(1-m)\nabla m)\cdot\frac{m\nabla c + (1-c)\nabla m}{m(1-c-m)}\Big)dx \\
&= -\int_\Omega (|\nabla c|^2 + \beta|\nabla m|^2)dx.
\end{aligned}$$

Inspired by the above considerations, it is reasonable to introduce the entropy variables

$$y = \frac{\partial h}{\partial c} = \log\frac{c}{1-c-m}, \quad z = \frac{\partial h}{\partial m} = \log\frac{m}{1-c-m}.$$

Conversely, c and m can be interpreted as functions of (y,z), given by

$$c(y,z) = \frac{e^y}{1+e^y+e^z}, \quad m(y,z) = \frac{e^z}{1+e^y+e^z}.$$

In the new variables, system (4.41) can be written as

$$\partial_t\begin{pmatrix} c \\ m\end{pmatrix} - \operatorname{div}\left(A(c,m)(\nabla^2 h)^{-1}\nabla\begin{pmatrix} y \\ z\end{pmatrix}\right) = 0,$$

since $\nabla(y,z)^\top = \nabla^2 h \nabla(c,m)^\top$, where $\nabla^2 h$ is the Hessian of the entropy density. The new matrix $B = A(c,m)\nabla^2 h$ has a rather complicated structure but it can be shown that it is symmetric and positive definite as long as $c > 0$, $m > 0$, and $c+m < 1$. The interesting feature of this change of unknowns is that the exponential transformation not only provides the positivity of the volume fractions but also an upper bound since

$$c(y,z) < 1, \quad m(y,z) < 1, \quad \text{and} \quad c(y,z) + m(y,z) < 1.$$

Therefore, applying similar proof techniques as above, one can prove the global-in-time existence of solutions (c,m) satisfying the above lower and upper bounds; see Jüngel-Stelzer 2012 [28].

5. Summary and open problems

In these lecture notes, we have investigated various aspects of entropy dissipation methods for evolution equations. Generally, if $u(t)$ is a (smooth) solution to the evolution equation $u_t + Au = 0$, where A is some differential operator, the aim is to derive identities of the type

$$\frac{dH}{dt}[u(t)] + D[u(t)] = 0, \quad t > 0,$$

and to show that the entropy dissipation $D[u]$ is nonnegative or that it can be estimated from below by some integral $P \geq 0$ containing squared derivatives of $u(t)$:

$$\frac{dH}{dt}[u(t)] + P \leq 0, \quad t > 0.$$

We have shown that this estimate may have a number of consequences:

- The entropy functional $H[u]$ is a Lyapunov functional.
- The estimate on P yields a priori estimates in certain Sobolev spaces.
- If the entropy dissipation can be estimated from below by some multiple of the (relative) entropy, $D[u] \geq \lambda H[u]$, the Gronwall inequality implies expontial decay of the solutions to the steady state.
- For convex entropies $H[u]$, the inequality $D[u] \geq \lambda H[u]$ often corresponds to a convex Sobolev inequality which can be proved simultaneously with the time decay property.
- The entropy construction method may help to prove inequalities of the type $dH/dt + P \leq 0$ or $D[u] \geq \lambda H[u]$.
- If a diffusion system possesses an entropy functional, the formulation in the entropy variables usually leads to a symmetric and positive definite diffusion matrix, which is useful for an existence analysis.
- If a diffusion system possesses a logarithmic entropy functional, the entropy variables are of exponential type such that the nonnegativity or even L^∞ bounds can be proved.

Entropy dissipation methods áre still under investigation, and there is a number of open problems. We mention some of them:

- The entropy construction method has been applied to some *multi-dimensional equations* (we have just mentioned the DLSS equation as an example) but there is still no systematic formulation of the method in the multi-dimensional case. The difficulty is to define a reduced number of polynomial variables corresponding to derivatives like $|\nabla u|/u$, $\Delta u/u$, $\nabla^2 u/u$, etc. and to select the useful integration-by-parts formulas.
- Entropy variables help to derive entropy dissipation inequalities for certain cross-diffusion systems, namely those which possess a logarithmic entropy. The reason is that these systems can be understood from thermodynamic principles for which the logarithmic entropy plays an important role. Are there other (important) *entropy functionals for cross-diffusion systems*? Can this be made more general?
- The energy-transport system

  $$u_t + \Delta(uT) = 0, \quad \tfrac{3}{2}(uT)_t + \tfrac{5}{2}\Delta(uT^2) = 0 \quad \text{in } \Omega,\ t > 0,$$

  describes the evolution of the particle density u and particle temperature T in a thermodynamic diffusion system. This system possesses the entropy functional

  $$H[u] = \int_\Omega u \log\Big(\frac{u}{T^{3/2}}\Big) dx.$$

  The *global-in-time existence* of solutions to this system (with initial and boundary conditions) is an open problem. The difficulty is to control the temperature in regions where the particle density vanishes.

- Consider the (simplified) Navier-Stokes equations with density-dependent viscosity,
$$\rho_t + \mathrm{div}(\rho u) = 0, \quad (\rho u)_t + \mathrm{div}(\rho u \otimes u) + \nabla\rho = 2\nu\,\mathrm{div}(\rho D(u)) \quad \text{in } \Omega,\ t > 0,$$
where ρ is the particle density, u the velocity, $\nu > 0$ the viscosity constant, and $D(u) = \frac{1}{2}(\nabla u^\top + \nabla u)$ the symmetric velocity gradient. Assuming appropriate boundary conditions, the energy identity for this system reads as
$$\frac{d}{dt}\int_\Omega \Big(\frac{\rho}{2}|u|^2 + \rho(\log\rho - 1)\Big)dx + \nu\int_\Omega \rho\|D(u)\|^2 dx = 0.$$
Surprisingly, the system possesses another energy identity, found by Bresch-Desjardins 2004 [6]:
$$\frac{d}{dt}\int_\Omega \Big(\frac{\rho}{2}|u + 2\nabla\log\rho|^2 + \rho(\log\rho - 1)\Big)dx + \nu\int_\Omega \Big(4|\nabla\sqrt{\rho}|^2 + \frac{\rho}{2}\|\nabla u^\top - \nabla u\|^2\Big)dx = 0.$$
The question is why are there two energy (entropy) identities? Is the reason related to a "Noether symmetry"? Are there other fluiddynamical models which possess *several energy identities*? This is important for the analysis of such equations.
- For numerical purposes, the evolution equations are discretized in time and space. In order to obtain stable and efficient numerical schemes, it is desirable to design numerical approximations which possess as many properties of the continuous problem as possible. In particular, entropy-stable, entropy-dissipating, and positivity-preserving schemes are needed. To what extend entropy tools can be generalized to *discrete entropy dissipation methods*?

References

[1] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Commun. Part. Diff. Eqs.* 26 (2001), 43-100.

[2] D. Bakry and M. Emery. Diffusions hypercontractives. *Séminaire de probabilités XIX*, 1983/84, pp. 177-206. Lect. Notes Math. 1123. Springer, Berlin, 1985.

[3] C. Bardos, F. Golse, and C. D. Levermore. Fluid dynamic limits of kinetic equations. II. Convergence proofs for the Boltzmann equation. *Commun. Pure Appl. Math.* 46 (1993), 667-753.

[4] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anw.* 22 (2003), 751-756.

[5] L. Boltzmann. Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wiener Berichte* 66 (1872), 275-370.

[6] D. Bresch and B. Desjardins.Some diffusive capillary models of Korteweg type. *C. R. Acad. Sci. Paris, Section Mécanique*, 332 (2004), 881-886.

[7] J. Carrillo, A. Jüngel, P. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.* 133 (2001), 1-82.

[8] C. Cercignani, R. Illner, and M. Pulvirenti. *The Mathematical Theory of Dilute Gases*. Springer-Verlag, New York, 1994.

[9] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Differ. Eqs.* 224 (2006), 39-59.

[10] G. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *J. Symbolic Comput.* 12 (1991), 299-328.

[11] I. Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Stud. Sci. Math. Hung.* 2 (1967), 299-318.
[12] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*, Volume 2. Springer-Verlag, Berlin, 1993.
[13] P. Degond, F. Méhats, and C. Ringhofer. Quantum energy-transport and drift-diffusion models. *J. Stat. Phys.* 118 (2005), 625-665.
[14] M. Del Pino and J. Dolbeault. Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. *J. Math. Pures Appl.* 81 (2002), 847-875.
[15] L. Desvillettes and K. Fellner. Entropy methods for reaction-diffusion equations in degenerate diffusion arising in reversible chemistry. Preprint, 2007. Available at http://www.uni-graz.at/~fellnerk.
[16] L. Desvillettes and C. Villani. On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems. Part I: The linear Fokker-Planck equation. *Commun. Pure Appl. Math.* 54 (2001), 1-42.
[17] R. DiPerna. Compensated compactness and general systems of conservation laws. *Trans. Amer. Math. Soc.* 292 (1985), 383-420.
[18] R. DiPerna and P.-L. Lions. On the Cauchy problem for the Boltzmann equation: Global existence and weak stability. *Ann. Math.* 130 (1989), 312-366.
[19] J. Dolbeault, I. Gentil, and A. Jüngel. A nonlinear fourth-order parabolic equation and related logarithmic Sobolev inequalities. *Commun. Math. Sci.* 4 (2006), 275-290.
[20] M. Dreher and A. Jüngel. Compact families of piecewise constant functions in $L^p(0,T;B)$. *Nonlin. Anal.* 75 (2012), 3072-3077.
[21] P. Federbush. A partially alternate derivation of a result by Nelson. *J. Math. Phys.* 10 (1969), 50-52.
[22] F. Golse and C. D. Levermore. Hydrodynamic limits of kinetic models. Topics in kinetic theory, pp. 1-75, *Fields Inst. Commun.* 46. Amer. Math. Soc., Providence, 2005.
[23] F. Golse and L. Saint-Raymond. The Navier-Stokes limit of the Boltzmann equation for bounded collision kernels. *Invent. Math.* 155 (2004), 81-161.
[24] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* 97 (1975), 1061-1083.
[25] T. Jackson and H. Byrne. A mechanical model of tumor encapsulation and transcapsular spread. *Math. Biosci.* 180 (2002), 307-328.
[26] A. Jüngel and D. Matthes. An algorithmic construction of entropies in higher-order nonlinear PDEs. *Nonlinearity* 19 (2006), 633-659.
[27] A. Jüngel and D. Matthes. The Derrida-Lebowitz-Speer-Spohn equation: existence, non-uniqueness, and decay rates of the solutions. *SIAM J. Math. Anal.* 39 (2008), 1996-2015.
[28] A. Jüngel and I. V. Stelzer. Entropy structure of a cross-diffusion tumor-growth model. To appear in *Math. Mod. Meth. Appl. Sci.*, 2012.
[29] S. Kawashima and Y. Shizuta. On the normal form of the symmetric hyperbolic-parabolic systems associated with the conservation laws. *Tohoku Math. J., II. Ser.* 40 (1988), 449-464.
[30] S. Kullback. A lower bound for discrimination information in terms of variation. *IEEE Trans. Inform. Theory* 4 (1967), 126-127.
[31] P. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. CBMS-NSF Regional Conference Series in Applied Mathematics 11. SIAM, Philadelphia, 1973.
[32] D. Matthes. Entropy Methods and Related Functional Inequalities. Lecture Notes, Pavia, Italy, 2007. http://www-m8.ma.tum.de/personen/matthes/papers/lecpavia.pdf.
[33] T. Myers. Thin films with high surface tension. *SIAM Rev.* 40 (1998), 441-462.
[34] L. Payne and H. Weinberger. An optimal Poincaré inequality for convex domains. *Arch. Rational Mech. Anal.* 5 (1960) 286-292.
[35] O. Rothaus. Logarithmic Sobolev inequalities and the spectrum of Sturm-Liouville operators. *J. Funct. Anal.* 39 (1980), 42-56.
[36] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.* 79 (1979), 83-99.
[37] D. Stroock. Logarithmic Sobolev inequalities for Gibbs states. *Lec. Notes Math.* 1563 (1993), 194-228.

[38] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. Univ. Calif. Press, Berkeley, 1951.
[39] G. Toscani. Sur l'inégalité logarithmique de Sobolev. *C. R. Acad. Sci. Paris Sér. I Math.* 324 (1997), 689-694.
[40] C. Villani. *Topics in Optimal Transportation*. Amer. Math. Soc., Providence, 2003.
[41] F. Weissler. Logarithmic Sobolev inequalities and hypercontractive estimates on the circle. *J. Funct. Anal.* 37 (1980), 218-234.