FOKKER–PLANCK EQUATIONS IN THE MODELLING OF SOCIO-ECONOMIC PHENOMENA

GIULIA FURIOLI
DIGIP, University of Bergamo, viale Marconi 5, 24044 Dalmine, Italy
giulia.furioli@unibg.it

ADA PULVIRENTI
Department of Mathematics, University of Pavia, via Ferrata 1, Pavia, 27100 Italy
ada.pulvirenti@unipv.it

ELIDE TERRANEO
Department of Mathematics, University of Milan, via Saldini 50, 20133 Milano, Italy
elide.terraneo@unimi.it

GIUSEPPE TOSCANI
Department of Mathematics,, University of Pavia, via Ferrata 1, Pavia, 27100 Italy,
giuseppe.toscani@unipv.it

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

We present and discuss various one-dimensional linear Fokker–Planck type equations that have been recently considered in connection with the study of interacting multi-agent systems. In general, these Fokker–Planck equations describe the evolution in time of some probability density of the population of agents, typically the distribution of the personal wealth or of the personal opinion, and are mostly obtained by linear or bilinear kinetic models of Boltzmann type via some limit procedure. The main feature of these equations is the presence of variable diffusion, drift coefficients and boundaries, which introduce new challenging mathematical problems in the study of their long-time behavior.

Keywords: Kinetic models; Fokker–Planck equations; Relative entropies; Large-time behavior

AMS Subject Classification: 35Q84; 82B21; 91D10, 94A17

1. Introduction

Starting from the well-consolidated concept that many laws of nature have a statistical origin, the mathematical modelling of systems composed by a huge number of interacting agents has been often and fruitfully based on arguments typical of statistical physics. In the last two decades, there has been in fact a trend toward applications of statistical physics to interdisciplinary fields ranging from the classical biological context to the new aspects of socio-economic dynamics.[28,82,84]

In the biological context, a consistent part of the recent research in population and behavioral biology and ecology has been looking for the emergent behavior of bird flocks, fish schools or bacteria aggregations.[26,27,38,39,40,60,61,62] Other important examples of emergent behaviors describe building of tumors by cancer cells and their migration through the tissues.[8,9,10,11,70] Further examples refer to the classical Luria–Delbrück mutation problem,[70,74,76,77,99] and to various probabilistic models of genome evolution describing the size distribution of gene families. [7,51,58,65,69,87]

In socio-economic phenomena, systems are composed not by particles but by humans, and every individual usually interacts with a very limited number of peers, which appears negligible compared to the total number of people in the system. Nevertheless, the phenomena are characterized by unexpected global behaviors, like the formation of very stable curves for the wealth distribution or the emergence of consensus about a specific issue.

Economics is, by far, one of the human behaviors to which methods borrowed from statistical mechanics for particle systems have been systematically applied[16,29,30,31,32,36,42,63,67,90] (cf. also Ref. 23, 24 for relationships with free boundary problems, and Ref. 35, 78, 98 for connections with finance and micro-economy). Starting from the original idea developed by Angle,[3,4] a variety of both discrete and continuous models for wealth distribution has been proposed and studied in view of the relation between parameters in the microscopic rules and the resulting macroscopic statistics.[30,36,42,56,89]

Likewise, the dynamics of opinion formation in a multi-agent society has received growing attention.[12,13,14,15,19,20,21,34,41,46,52,53,54,55,91,96] In view of the relation between parameters in the microscopic interaction rules, the society develops a certain steady macroscopic opinion distribution,[82,84] which characterizes the formation of a relative consensus around certain opinions.

Other aspects of human behaviors treated by the same methodology include the formation of knowledge,[83] and conviction around a certain opinion [22] (cf. also Ref. 1 for a recent exhaustive survey).

The description of these apparently different phenomena has its common basis in statistical physics. In particular, methods borrowed from kinetic theory of rarefied gases have been successfully used to construct master equations of Boltzmann type, usually referred to as kinetic equations, describing the time-evolution of the number density of the population and, eventually, the emergence of universal behaviors through their equilibria.[25,82,84]

The building block of kinetic theory is represented by binary interactions, which, similarly to binary interactions between particles velocities in the classical kinetic theory of rarefied gases, describe the variation law of some selected agent trait, like its wealth or opinion. Then, the microscopic law of variation of the number density consequent to the (fixed-in-time) way of interaction, is able to capture both the time evolution and the steady profile, in presence of some conservation law. [82,84]

In the following, $v \in I$ will indicate the value of the selected trait of agents. Here trait will denote a distinguishing quality or characteristic of the agents (wealth, opinion, knowledge, conviction and others). Depending on the problem one is dealing with, the domain $I \subseteq \mathbb{R}$ usually denotes a fixed interval, or the positive half-line. The pair v, w will indicate the values of the selected trait of the pair of agents before their interaction. The most general binary interaction considered so far can be uniquely described as follows. When two agents interact, the values v, w of their traits change into the values v^*, w^* according to the rule

$$\begin{aligned} v^* &= v + P(v)(w-v) + Q(v)\eta, \\ w^* &= w + P(w)(v-w) + Q(w)\tilde{\eta}. \end{aligned} \tag{1.1}$$

The nonnegative functions P and Q contain the details of the interaction, and measure the intensities of the variation of the trait due to the presence of the other agent (the function P), and the variation of the trait due to random effects (the function Q). In particular, to outline the independence among agents, the random parameters η and $\tilde{\eta}$, usually of zero mean and bounded variance, are assumed to be independent and identically distributed. Since the trait v varies on the domain $I \subseteq \mathbb{R}$, which can be a fixed interval, or the positive half-line, the functions P and Q and the random parameters are subject to conditions which assure that the binary interaction (1.1) maintains the values v^*, w^* on the same domain I. In particular, this will imply that $Q(0) = 0$ if $I = \mathbb{R}^+$ and $Q(v) = 0$ on the boundary of I whenever I is a bounded interval.

The study of the time-evolution of the distribution of the selected trait v consequent to interactions of type (1.1) among individuals can be obtained by resorting to kinetic collision-like models.[84] Let $f = f(v,t)$ the density of agents which at time $t > 0$ are represented by their trait $v \in I$. Then, the time evolution of $f(v,t)$ obeys to a nonlinear Boltzmann-like equation. This equation can be fruitfully written in weak form. It corresponds to say that the solution $f(v,t)$ satisfies, for all smooth functions $\varphi(v)$ (the observable quantities)

$$\begin{aligned} &\frac{d}{dt}\int_I f(v,t)\varphi(v)\,dv = \\ &\frac{1}{2}\Big\langle \int_{I\times I} \left(\varphi(v^*) + \varphi(w^*) - \varphi(v) - \varphi(w)\right) f(v,t) f(w,t)\,dv\,dw \Big\rangle. \end{aligned} \tag{1.2}$$

In (1.2) $\langle \cdot \rangle$ represents mathematical expectation. Here expectation takes into account the presence of the random parameters $\eta, \tilde{\eta}$ in (1.1).

The meaning of the kinetic equation (1.2) is clear. At any positive time $t > 0$, the variation in time of the distribution of the trait v (the left-hand side) is due to the change in v resulting from interactions of type (1.1) that the system of agents has at time t. This change is measured by the interaction operator at the right-hand side.

Related linear kinetic equations of Boltzmann type are obtained by assuming that the change in the trait $v \in I$ is consequent to interactions with a fixed environment. In this case, the variation of individual trait in a single (microscopic) interaction is the result of three different contributes

$$v^* = v + P_E(v)z - P(v)v + Q(v)\eta. \tag{1.3}$$

Unlike the previous case, in (1.3) the intensity of the (nonnegative) variation P_E of the trait v due to the presence of the environment is in general assumed to be different from P. Here $z \in I$ indicates the amount of trait absorbed by the agent from the environment. In this case the density $f(v,t)$ satisfies, for all smooth functions $\varphi(v)$ (the observable quantities)

$$\frac{d}{dt}\int_I f(v,t)\varphi(v)\,dv = \Big\langle \int_{I\times I} (\varphi(v^*) - \varphi(v))\, f(v,t)\mathcal{E}(z)\,dv\,dz \Big\rangle. \tag{1.4}$$

The function $\mathcal{E}(z)$, $z \in I$ denotes the distribution of the environment. As precised before, to be physically consistent, the functions P, P_E and Q and the random parameter η are subject to conditions which ensure that the binary interaction (1.1) maintains the values v^*, w^* on the domain I. Likewise, the distribution $\mathcal{E}$ of the environment will take values on the same domain I.

The kinetic equations (1.2) and (1.4) then describe the evolution of the density, and allow us to study, at least numerically, the long-time behavior of the system of agents, by recovering its macroscopic universal behavior.

However, except in some simple case,[5,6,90] a precise analytic description of the emerging equilibria is very difficult to obtain. A further insight into the large-time behavior of the kinetic equation, and a more accessible description of the possible stationary states can be achieved by resorting to particular asymptotics which lead to Fokker-Planck type equations.[36,96]

This asymptotic procedure is a well-consolidated technique which is reminiscent of the so-called grazing collision limit, fruitfully applied to the classical Boltzmann equation,[100,101,102] and to the dissipative versions of Kac caricature of a Maxwell gas [50] introduced in Ref. 86.

In the rest of the paper we will introduce various Fokker–Planck type equations which arise in the study of socio-economic systems. Section 2 will deal with the description of the asymptotic procedure which allows us to pass from the kinetic model of Boltzmann type to the Fokker–Planck description. Various examples which refer to the evolution of different traits of multi-agent system will be presented and discussed. Next, Section 3 will be concerned with the main mathematical methods that have been introduced to study the asymptotic behavior of the solution, and the rate of convergence to equilibrium. Finally Sections 4 and 5 will be devoted to some open problems and concluding remarks.

2. Quasi-invariant limits and Fokker-Planck equations

To start with, let us consider the case of the linear kinetic equation (1.4), and let us fix the domain $I = \mathbb{R}_+$. This corresponds to assume that the trait v can only assume nonnegative values. Moreover, here and in the rest of the paper the random parameter η will be limited to assume values on a bounded set, and in addition to satisfy $\langle \eta \rangle = 0$ and $\langle \eta^2 \rangle = \lambda$. To avoid inessential difficulties, the distribution of the environment $\mathcal{E}(z)$, $z \geq 0$ will possess moments finite at least up to the fourth order. In particular, the finite average value of the environment will be denoted by M_E.

To justify computations, the functions $P_E(v)$, $vP(v)$ and $Q(v)$ which characterize interaction (1.3) will be limited to have at most a linear growth. The following definition will be used

Definition 2.1. We will say that a function $f = f(v)$, $v \in \mathbb{R}_+$ has linear growth if there is a positive constant C such that $|f(v)| \leq C(1 + v)$, $v \in \mathbb{R}_+$.

Among observable quantities, besides the mass which is conserved as we can easily check by letting $\varphi = 1$ in (1.4), the first representative ones to be studied are the average value of the density f, as well as its variance. To this aim, choosing $\varphi(v) = v$ in (1.4) and remarking that (1.3) implies

$$\langle v^* - v \rangle = P_E(v)z - P(v)v = A(v, z),$$

we obtain

$$\frac{d}{dt}\int_{\mathbb{R}_+} v\, f(v,t)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} A(v,z) f(v,t)\mathcal{E}(z)\, dv\, dz. \tag{2.1}$$

Likewise, since

$$\langle {v^*}^2 - v^2 \rangle = \lambda\, Q^2(v) + A^2(v,z) + 2vA(v,z),$$

$$\frac{d}{dt}\int_{\mathbb{R}_+} v^2\, f(v,t)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} \big(\lambda\, Q^2(v) + A^2(v,z) + 2vA(v,z)\big)\, f(v,t)\mathcal{E}(z)\, dv\, dz. \tag{2.2}$$

Note that, in view of the condition of linear growth, the boundedness of the moments of the distribution $\mathcal{E}$ is enough to guarantee that the moments at the first two orders of the solution to equation (1.4) remain bounded at any time $t > 0$, provided that they are bounded initially.

Let us suppose now that the interaction (1.3) produces a very small mean change of the trait. This can be easily achieved by multiplying the functions $P_E(\cdot)$ and $P(\cdot)$ by some value ϵ, with $\epsilon \ll 1$, and the function $Q(\cdot)$ by ϵ^α, where the exponent α is a positive constant. In other words, given a small parameter ϵ, we will consider the scaling

$$P_E(\cdot) \to \epsilon P_E(\cdot), \quad P(\cdot) \to \epsilon P(\cdot), \quad Q(\cdot) \to \epsilon^\alpha Q(\cdot). \tag{2.3}$$

Concerning the evolution of the average value (2.1) this scaling will produce a small variation of the average, independent of the value of the exponent α

$$\frac{d}{dt}\int_{\mathbb{R}_+} v\, f(v,t)\, dv = \epsilon \int_{\mathbb{R}_+\times\mathbb{R}_+} A(v,z) f(v,t)\mathcal{E}(z)\, dv\, dz.$$

It is clear that, if we want to observe an evolution of the average value independent of ϵ,we can resort to a scaling of time. If we set $\tau = \epsilon t$, $f_\epsilon(v,\tau) = f(v,t)$, then the evolution of the average value for $f_\epsilon(v,\tau)$ satisfies

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} v\, f_\epsilon(v,\tau)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} A(v,z) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz,$$

namely the same evolution law for the average value of f given by (2.1). The reason is clear. If we assume that the interactions are scaled to produce a very small change in the trait, to observe an evolution of the average value independent of the smallness, we need to wait enough time to restore the original evolution.

By using the same scaling (2.3) into the equation (2.2), the evolution equation for the second moment of $f_\epsilon(v,\tau)$ takes the form

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} v^2\, f_\epsilon(v,\tau)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} \left(\epsilon^{2\alpha-1}\lambda\, Q^2(v) + \epsilon A^2(v,z) + 2vA(v,z)\right) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz.$$

Hence, by choosing $\alpha = 1/2$ one shows that the evolution of the second moment of $f_\epsilon(v,\tau)$, for any given $\epsilon \ll 1$ depends on all the quantities appearing in the interaction (1.3), and

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} v^2\, f_\epsilon(v,\tau)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} \left(\lambda Q^2(v) + 2vA(v,z)\right) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz + R_\epsilon(\tau),$$

where the (small) remainder is given by

$$R_\epsilon(\tau) = \epsilon \int_{\mathbb{R}_+\times\mathbb{R}_+} A^2(v,z) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz.$$

If the remainder vanishes as $\epsilon \to 0$, one obtains a closed form for the evolution of the second moment, valid for any value of the small parameter ϵ. In this case, the limit evolution equation for a density $f(v,\tau)$ reads

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} v^2\, f(v,\tau)\, dv = \int_{\mathbb{R}_+\times\mathbb{R}_+} \left(\lambda\, Q^2(v) + 2vA(v,z)\right) f(v,\tau)\mathcal{E}(z)\, dv\, dz. \tag{2.4}$$

However, one has to note that, while the scaling (2.3), with $\alpha = 1/2$ is such that the evolution law of the average value is independent of ϵ, the limit evolution law of the second moment, as given by (2.4), is different from the evolution law (2.2). In particular, for a fixed density f the right-hand side of (2.2) is strictly bigger than the right-hand side of (2.4). This shows that the variance of the solution to the kinetic model is strictly bigger than the variance of the (possibile) limit density.

Analogous computations enable us to study the evolution of moments in the nonlinear kinetic equation (1.2). In particular, it can be shown that very close properties for the average value and the second moment continue to hold in presence of the scaling of the parameters in (1.1). We refer the interested reader to Ref. 36 for further details.

2.1. *From Boltzmann to Fokker-Planck*

The short discussion about moments of the previous section contains the main motivations and the mathematical ingredients that justify the passage from the kinetic model (1.4) to its continuos counterpart given by the Fokker–Planck description. As before, we restrict here our analysis to the linear model.

Given a smooth function $\varphi(v)$, let us expand in Taylor series $\varphi(v^*)$ around $\varphi(v)$. Using the scaling (2.3) with $\alpha = 1/2$ it holds

$$\langle v^* - v\rangle = \epsilon\, A(v,z); \quad \langle (v^* - v)^2\rangle = \epsilon^2\, A^2(v,z) + \epsilon\lambda Q^2(v).$$

Therefore, in terms of powers of ϵ, we easily obtain the expression

$$\langle \varphi(v^*) - \varphi(v)\rangle = \epsilon \left(\varphi'(v)A(v,z) + \frac{1}{2}\,\varphi''(v)\lambda\, Q^2(v)\right) + R_\epsilon(v,z),$$

where the remainder term R_ϵ, for a suitable $0 \le \theta \le 1$ is given by

$$R_\epsilon(v,z) = \frac{1}{2}\epsilon^2\, \varphi''(v)A^2(v,z) + \frac{1}{6}\langle \varphi'''(v+\theta(v^*-v))(v^*-v)^3\rangle, \tag{2.5}$$

and it is vanishing at the order $\epsilon^{3/2}$ as $\epsilon \to 0$. Therefore, if as in Section 2 we set $\tau = \epsilon t$, $f_\epsilon(v,\tau) = f(v,t)$, we obtain that the evolution of the (smooth) observable quantity $\varphi(v)$ is given by

$$\begin{aligned}
&\frac{d}{d\tau}\int_{\mathbb{R}_+} \varphi(v)\, f_\epsilon(v,\tau)\, dv = \\
&\int_{\mathbb{R}_+\times\mathbb{R}_+} \left(\varphi'(v)A(v,z) + \frac{1}{2}\varphi''(v)\lambda\, Q^2(v)\right) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz + \frac{1}{\epsilon}\mathcal{R}_\epsilon(\tau) = \\
&\int_{\mathbb{R}_+} \left(\varphi'(v)(P_E(v)M_E - P(v)v) + \frac{1}{2}\varphi''(v)\lambda\, Q^2(v)\right) f_\epsilon(v,\tau)\, dv + \frac{1}{\epsilon}\mathcal{R}_\epsilon(\tau),
\end{aligned}$$

where

$$\mathcal{R}_\epsilon(\tau) = \int_{\mathbb{R}_+\times\mathbb{R}_+} R_\epsilon(v,z) f_\epsilon(v,\tau)\mathcal{E}(z)\, dv\, dz,$$

and R_ϵ is given by (2.5). Letting $\epsilon \to 0$ shows that in consequence of the scaling (2.3) the weak form of the kinetic model (1.4) is well approximated by the weak form of a linear Fokker–Planck equation (with variable coefficients)

$$\begin{aligned}
&\frac{d}{d\tau}\int_{\mathbb{R}_+} \varphi(v)\, g(v,\tau)\, dv = \\
&\int_{\mathbb{R}_+} \left(\varphi'(v)(P_E(v)M_E - P(v)v) + \frac{1}{2}\varphi''(v)\lambda\, Q^2(v)\right) g(v,\tau)\, dv.
\end{aligned} \tag{2.6}$$

In fact, provided the boundary terms produced by the integration by parts vanish, equation (2.6) coincides with the weak form of the Fokker–Planck equation

$$\frac{\partial g}{\partial \tau} = \frac{\lambda}{2}\frac{\partial^2}{\partial v^2}\left(Q^2(v)g\right) + \frac{\partial}{\partial v}\left((P(v)v - P_E(v)M_E)g\right). \tag{2.7}$$

A brief discussion which clarifies the vanishing of the boundary terms will be carried out in the next subsection. One of the main advantages in resorting to this asymptotic procedure is that in many cases it is possible to obtain from the Fokker–Planck equation (2.7) its explicit stationary solution, which is obtained by solving an ordinary differential equation of first order. We will provide in the remaining of this section various examples connected with socio-economic applications. As we will see, most of these applications refer to a standard drift term, which corresponds to fix both the functions P and P_E to be constant. In this simple but relevant case for applications, the Fokker–Planck equation (2.7) simplifies to

$$\frac{\partial g}{\partial \tau} = \frac{\lambda}{2}\frac{\partial^2}{\partial v^2}\left(Q^2(v)g\right) + \frac{\partial}{\partial v}\left((v - M_E)g\right), \quad v \in \mathbb{R}_+. \tag{2.8}$$

Then, the form of the diffusion coefficient Q determines the characteristics of the equilibrium distribution.

It is remarkable that the results of this section remain valid also in the case in which the variable v is allowed to vary on some interval $I \subseteq \mathbb{R}$,[96] and in the case of the bilinear kinetic model (1.2).[36]

2.2. *Boundary conditions*

As main example, let us take into account the bilinear model (1.2) with $I = \mathbb{R}_+$, and let us set $P(v) = 1$. In addition, let us fix the (constant in time) average value of the solution $f(v,t)$ equal to one. In this case, under the scaling (2.3), the weak form of the kinetic equation is well-approximated by the Fokker–Planck equation (in weak form)

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} \varphi(v)h(v,\tau)\,dv = \int_{\mathbb{R}_+} \varphi'(v)(1-v)h(v,\tau)\,dv + \frac{\lambda}{2}\int_{\mathbb{R}_+} h(v,\tau)Q^2(v)\varphi''(v)dv. \tag{2.9}$$

Integration by parts then shows that equation (2.9) coincides with the weak form of the Fokker–Planck equation

$$\frac{\partial h}{\partial \tau} = \frac{\lambda}{2}\frac{\partial^2}{\partial v^2}\left(Q^2(v)h\right) + \frac{\partial}{\partial v}\left((v-1)h\right), \tag{2.10}$$

provided the boundary terms produced by integration vanish. While the vanishing of the boundary term at infinity follows by choosing initial data with a smooth and rapid decay, a more detailed analysis is required at the boundary $v = 0$. In particular, on this boundary one obtains the conditions

$$Q^2(v)h(v,\tau)\,|_{v=0} = 0, \qquad \tau > 0 \tag{2.11}$$

and

$$(v-1)h(v,\tau) + \frac{\lambda}{2}\frac{\partial}{\partial v}\left(Q^2(v)h(v,\tau)\right)|_{v=0} = 0, \qquad \tau > 0. \tag{2.12}$$

While condition (2.11) is automatically satisfied for a sufficiently regular density h, condition (2.12) requires an exact balance between the so-called advective and diffusive fluxes on the boundary $v = 0$. This condition is usually referred to as the *no-flux* boundary condition.

This condition appears in a natural way by imposing that the solution to the Fokker–Planck equation (2.10) is mass preserving, which follows from (2.9) by taking $\varphi(v) = 1$. The same result holds for the standard form (2.10) if

$$\frac{d}{d\tau}\int_{\mathbb{R}_+} h(v,\tau)\,dv = \int_{\mathbb{R}_+} \frac{\partial}{\partial v}\left(\frac{\lambda}{2}\frac{\partial}{\partial v}\left(Q^2(v)h(v,\tau)\right) + (v-1)h(v,\tau)\right) dv = 0,$$

which is certainly true if (2.12) holds.

Likewise, condition (2.11) appears by imposing that the solution to the Fokker–Planck equation (2.10) preserves the mean value, provided that the mean value of the initial density is equal to one. This fact follows from (2.9) by taking $\varphi(v) = v$.

Remark 2.1. As this example clearly shows, boundary conditions are present when the space variable in the one-dimensional differential problem is ranging on a bounded interval, or, more generally, on a half-line. This is a problem that appears each time one resorts to an asymptotic procedure to pass from a kinetic description in terms of a Boltzmann-type collision operator to a kinetic description in terms of a Fokker–Planck-type operator, which involves derivatives with respect to the spatial variable. Then, as the previous discussion clarifies, the natural boundary conditions are found by imposing that the solution to the limit equation would maintain the same macroscopic properties of the solution to the kinetic equation. In reason of the fact that the solution to the kinetic equation is mass preserving, we can always assume that the solution to the various Fokker–Planck equations satisfy no-flux boundary conditions of type (2.12). Then, further conservation properties at the kinetic level will induce other *ad-hoc* boundary conditions.

Remark 2.2. Fokker–Planck equations with variable coefficients in presence of boundary conditions were first studied in a paper by Feller, who treated the case $v \in \mathbb{R}_+$ and $Q(v) = \sqrt{v}$, with a general drift term in Ref. 48 (cf. also the book Ref. 49 for a general view about boundary conditions for diffusion equation). In particular, the importance of the boundary conditions has been shown in Ref. 48 to be related to the action of the drift term.

2.3. *Examples*

2.3.1. *Wealth distribution*

The basic model discussed in this section has been introduced in 2005 in Ref. 36 within the framework of classical models of wealth distribution in economy, to understand the possible formation of heavy tails, as predicted by the economic analysis of the italian economist Vilfredo Pareto.[85] This model belongs to a class of models in which the interacting agents are indistinguishable. In most of these models an agent's *state* at any instant of time $t \geq 0$ is completely characterized by his current wealth $v \geq 0$.[43,44] When two agents encounter in a trade, their *pre-trade wealths* v, w change into the *post-trade wealths* v^*, w^* according to the rule [29,30,32]

$$v^* = p_1 v + q_1 w, \quad w^* = q_2 v + p_2 w.$$

The *interaction coefficients* p_i and q_i are non-negative random variables. While q_1 denotes the fraction of the second agent's wealth transferred to the first agent, the difference $p_1 - q_2$ is the relative gain (or loss) of wealth of the first agent due to market risks. It is usually assumed that p_i and q_i have fixed laws, which are independent of v and w, and of time. This means that the amount of wealth an agent contributes to a trade is (on the average) proportional to the respective agent's wealth.

In Ref. 36 the trade has been modelled to include the idea that wealth changes hands for a specific reason: one agent intends to *invest* his wealth in some asset, property etc. in possession of his trade partner. Typically, such investments bear some risk, and either provide the buyer with some additional wealth, or lead to the loss of wealth in a non-deterministic way. An easy realization of this idea consists in coupling the saving propensity parameter [29,30] with some *risky investment* that yields an immediate gain or loss proportional to the current wealth of the investing agent

$$v^* = \Big(1 - \gamma + \eta_1\Big) v + \gamma w, \quad w^* = \Big(1 - \gamma + \eta_2\Big) w + \gamma v, \tag{2.13}$$

where $0 < \gamma < 1$ is the parameter which identifies the saving propensity, namely the intuitive behavior which prevents the agent to put in a single trade the whole amount of his money. In this case

$$p_i = 1 - \gamma + \eta_i, \quad q_i = \gamma \quad (i = 1, 2).$$

The coefficients η_1, η_2 are random parameters, which are independent of v and w, and distributed so that always v^*, $w^* \geq 0$, i.e. η_1, $\eta_2 \geq \gamma - 1$. Unless these random variables are centered, i.e. $\langle \eta_1 \rangle = \langle \eta_2 \rangle = 0$, it is immediately seen that the mean wealth is not preserved, but it increases or decreases exponentially (see the computations in Ref. 36). For centered η_i,

$$\langle v^* + w^* \rangle = (1 + \langle \eta_1 \rangle) v + (1 + \langle \eta_2 \rangle) w = v + w,$$

implying conservation of the average wealth. Various specific choices for the η_i have been discussed in Ref. 81. The easiest one leading to interesting results is $\eta_i = \pm r$, where each sign comes with probability 1/2. The factor $r \in (0, \gamma)$ should be understood as the *intrinsic risk* of the market: it quantifies the fraction of wealth agents are willing to gamble on. Within this choice, one can display the various regimes for the steady state of wealth in dependence of γ and r, which follow from numerical evaluation. In the zone corresponding to low market risk, the wealth distribution shows again *socialistic* behavior with slim tails. Increasing the risk, one falls into *capitalistic*, where the wealth distribution displays the desired Pareto tail. A minimum of saving ($\gamma > 1/2$) is necessary for this passage; this is expected since if wealth is spent too quickly after earning, agents cannot accumulate enough to become rich. Inside the capitalistic zone , the Pareto index decreases from $+\infty$ at the border with *socialist* zone to unity. Finally, one can obtain a steady wealth distribution which is a Dirac delta located at zero. Both risk and saving propensity are so high that a marginal number of individuals manages to monopolize all of the society's wealth. In the long-time limit, these few agents become infinitely rich, leaving all other agents truly pauper.

The analysis of Ref. 81 essentially shows that the microscopic interaction (2.13) considered in Ref. 36 is such that the kinetic equation (1.2) is able to describe all interesting behaviors of wealth distribution in a multiagent society.

The choice in (2.13) corresponds to set in (1.1) $P(v) = \gamma$ and $Q(v) = v$. By assuming $\langle \eta_i^2 \rangle = \lambda$, for $i = 1, 2$, and a unitary average value of the initial density, we obtain that the scaled density satisfies the limit Fokker–Planck equation

$$\frac{\partial h}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial v^2} \left(v^2 h \right) + \gamma \frac{\partial}{\partial v} \left((v-1) h \right). \tag{2.14}$$

It is immediately recognizable that equation (2.14) has a unique stationary solution of unit mass, given by the Γ-like distribution [18,36]

$$h_\infty(v) = \frac{(\mu - 1)^\mu}{\Gamma(\mu)} \frac{\exp\left(-\frac{\mu-1}{v}\right)}{v^{1+\mu}}, \tag{2.15}$$

where

$$\mu = 1 + 2\frac{\gamma}{\lambda} > 1.$$

This stationary distribution exhibits a power-law tail for large values of the wealth variable.

Among the possible generalizations of the models presented in this section, one of them is related to the statistical description of agent-based models constituted by agents from n different countries or social groups of individuals which can trade with each other. These groups shall be identified with countries or social classes inside a country. The main outcome one expects from this type of models is to reach stationary profiles for wealth distribution able to capture phenomena which are present in the recent history of economies, but impossible to obtain on the basis of binary exchanges in a homogeneous group.

In general, to simplify models without affecting the possibility of a general outcome, it is adopted the hypothesis that all agents belonging to one group share a common saving rate parameter. This hypothesis can be further relaxed by assuming that the saving rate is a random quantity, with a statistical mean which is different for different social groups. A general model was proposed in Ref. 45. In case of an international trade, i.e. when two agents of *different countries* interact, each agent uses the transaction parameter which is characteristic for his country. Hence, when two agents, one from country i $(i = 1, 2, \dots, n)$ with pre-trade wealth v and the other from country j $(j = 1, 2, \dots, n)$ with pre-trade wealth w interact, their post-trade wealths v^* and w^* are given by

$$v^* = (1 - \gamma_i \gamma)v + \gamma_j \gamma w + \eta_{ij} v, \tag{2.16a}$$

$$w^* = (1 - \gamma_j \gamma)w + \gamma_i \gamma v + \eta_{ji} w. \tag{2.16b}$$

In (2.16), the trade depends on the transaction parameters γ and γ_i $(i = 1, \dots, n)$, while the risks of the market are described by η_{ij} $(i, j = 1, \dots, n)$, which are equally distributed random variables with zero mean and variance λ_{ij}. The different variances for domestic trades in each country and for international trades reflect different risk structures in these trades. For example, investments and trades inside different countries or markets may be subject to different types and quantities of risk, and international trading may face additional risks compared to domestic trades.

In the usual scaling (2.3) this general trade leads to a system of Fokker-Planck equations for the wealth densities $g_i(v, \tau)$, $(i = 1, \dots, n)$

$$\frac{\partial g_i}{\partial \tau} = \sum_{j=1}^{n} \Big[\frac{\lambda_{ij}}{2\tau_{ij}} \frac{\partial^2}{\partial v^2} \big(v^2 \rho_j g_i\big) + \frac{1}{\tau_{ij}} \frac{\partial}{\partial v} \big((\gamma_i v \rho_j - \gamma_j m_j) g_i \big) \Big],$$

where τ_{ij} represent suitable relaxation times, and the m_i's are the averages of the densities.

As observed in Ref. 45 the distributed saving gives rise to an additional interesting feature when a special case is considered where the saving parameter is assumed to take only two fixed values, preferably widely separated. In this case, the steady distribution of wealth can result in a bimodal distribution.[59] The numerical output evolves towards a robust and distinct two-peak distribution as the difference in the two saving parameters is increased systematically.

2.3.2. *Knowledge in a society*

A similar Fokker–Planck equation has been considered, in connection with the formation of knowledge in a multi-agent society, in Ref. 83. The main reason there was to understand the joint effect of knowledge and trade in the distribution of wealth.

Let us briefly explain the main motivations about microscopic interactions which determine the individual knowledge, which can be described as a familiarity with someone or something unknown, which can include information, facts, descriptions, or skills acquired through experience or education. Knowledge is in part inherited from the parents, but the main factor that can enrich it is the environment in which the individual grows and lives.[57,92] Indeed, the experiences that produce knowledge can not be fully inherited from the parents, such as the genome, but rather are acquired over a lifetime of several elements of the environment. The learning process is very complicated and produces different results for each individual in a population.

Although all individuals are given the same opportunities, at the end of the cognitive process every individual appears to have a different level of knowledge. Also, the personal knowledge is the result of a selection, which leads to retain mostly the notions that the individuals consider important, and to discard the rest. As noticed in Ref. 83, this aspect of the process of learning has been recently discussed in a convincing way by Umberto Eco,[47] one of the greatest philosophers and contemporary Italian writers. In his fascinating lecture, Eco outlines the importance of a drastic selection of the surrounding quantity of information, to maintain a certain degree of ingenuity.

If one agrees with these facts, each microscopic variation of knowledge is interpreted as an interaction where a fraction of the knowledge of the individual is lost by virtue of the selection, while at the same time the external background (the surrounding environment) can move a certain amount of its knowledge to the individual. If we quantify the nonnegative amount of knowledge of the individual with $v \in \mathbb{R}_+$, and with $z \in \mathbb{R}_+$ the knowledge achieved from the environment in a single interaction, the new amount of knowledge can be computed using the interaction (1.3), where the functions P and P_E quantify, respectively, the amounts of selection and external learning, while η is a random parameter which takes into account the possible unpredictable modifications of the knowledge process. In this model, it is assumed that the possible random variation of knowledge are proportional to the knowledge itself. Therefore, $Q(v) = v$. If one assumes that $\langle \eta^2 \rangle = \lambda$, and that the average value of the distribution of knowledge in the environment is equal to M_E, it is immediate to recognize that the Fokker–Planck equation for the density $k = k(v, \tau)$ of the agents which possess knowledge v at time $\tau > 0$ is given by equation (2.7), in which $Q(v) = v$

$$\frac{\partial k}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial v^2} \left(v^2 k\right) + \frac{\partial}{\partial v} \left((P(v)v - P_E(v) M_E) k\right). \tag{2.17}$$

2.3.3. *The formation of conviction*

Our third example concerns the formation of conviction considered in Ref. 22. Conviction is typically described as a certain resistance to modify a personal behavior. How the personal amount of conviction is formed is a very difficult question. In Ref. 22 it was argued that, among other reasons, responsible of conviction forming include familiar environment, personal contacts, readings or skills acquired through experience or education. Moreover, while conviction (at least concerning some aspects of life like religious or political beliefs) is in part inherited in the interior of family from the parents, it is also evident that the main factor that can influence it is the social background in which the individual grows and lives.[79,92] Like in knowledge formation, although all individuals are given the same opportunities, at the end of the process every individual appears to have a different level of conviction about something. Also, it is almost evident that the personal conviction is heavily dependent on the individual nature. A consistent part of us is accustomed to rethink, and to have continuous afterthoughts on many aspects of our daily decisions. This is particularly true nowadays, where the global access to information via web gives to each individual a huge number of different information, very often producing insecurity.

Analogously to the process of knowledge formation, each variation of conviction is interpreted as an interaction where a fraction of the conviction of the individual could be lost by virtue of afterthoughts and insecurities, while at the same time the individual can absorb a certain amount of conviction through the information and social pressure achieved from the external environment. The individual conviction is the trait, quantified in terms of a scalar continuous parameter v, ranging from zero to infinity. Small values of this parameter will characterize floating agents, while high values will characterize inflexible agents.

Owing to the previous considerations, the variation of individual conviction follows the law (1.3), where P quantifies the loss of individual conviction due to the action of afterthoughts and insecurities and P_E the amount of conviction absorbed from the social environment. Finally $Q(v)\eta$ quantifies the possible unpredictable modifications of the conviction process. The randomness present in the interaction is given by the random parameter η, while Q denotes an increasing function of conviction. This choice is driven by the assumption that random modifications of conviction are directly proportional to the conviction itself. The typical choice is to take $Q(v) = v^\nu$, with $0 < \nu \le 1$.

In Ref. 22 two different situations were considered, both corresponding to the choice of constant functions P and P_E. If $P = \gamma$ and $P_E = \gamma_E$, $\langle \eta^2 \rangle = \lambda$, and the average value of the distribution of conviction in the environment is equal to M_E, in the scaling (2.3) the density of conviction $c(v,\tau)$ satisfies the Fokker–Planck equation

$$\frac{\partial c}{\partial \tau} = \frac{\lambda}{2}\frac{\partial^2}{\partial v^2}\left(Q^2(v)c\right) + \frac{\partial}{\partial v}\left((\gamma v - \gamma_E M_E)c\right). \tag{2.18}$$

The explicit form of the steady distribution of conviction then depends on the choice of a particular function Q. The case in which $Q(v) = v$ leads to a Fokker–Planck equation similar to (2.14). One obtains

$$c_\infty(v) = \frac{C_0}{v^{2+2\gamma/\lambda}} \exp\left\{-\frac{2\gamma_E M_E}{\gamma v}\right\}. \tag{2.19}$$

In (2.19) the constant C_0 is chosen to fix the total mass of $c_\infty(v)$ equal to one. Note that the steady profile is heavy tailed, and the size of the polynomial tails is related to both λ and γ. Hence, the percentage of individuals with high conviction is increasing as soon as the parameter γ of insecurity is decreasing, and/or the parameter λ of self-thinking is increasing. It is moreover interesting to note that the size of the parameter γ_E is important only in the first part of the v-axis, and contributes to determine the size of the number of undecided. Like in the case of wealth distribution, this solution has a large *middle class*, namely a large part of the population with a certain degree of conviction, and a small *poor class*, namely a small part of undecided people.

The second case refers to the choice $Q(v) = \sqrt{v}$. Now, people with high conviction is more resistant to change (randomly) with respect to the previous case. On the other hand, if the conviction is small, $v < 1$, the individual is less resistant to change. Direct computations now show that the steady profile is given by

$$c_\infty(v) = C_0\, v^{-1+(2\gamma_E M_E)/\lambda} \exp\left\{-\frac{2\gamma}{\lambda} v\right\},$$

where again the constant C_0 is chosen to fix the total mass of $c_\infty(v)$ equal to one. In contrast with the previous case, the distribution decays exponentially to infinity, thus describing a population in which there are very few agents with a large conviction. Moreover, this distribution describes a population with a huge number of undecided agents. Note that, since the exponent of v in c_∞ is strictly bigger than -1, c_∞ is integrable for any choice of the relevant parameters.

Other choices of the exponent ν in the range $0 < \nu \leq 1$ do not lead to essential differences. The previous examples show that, despite the simplicity of the kinetic interaction, by acting on the coefficient of the random part η one can obtain very different types of steady conviction distributions.

2.3.4. *Opinion formation*

Our last model is concerned with the problem of opinion formation. A kinetic basis to this problem has been done in Ref. 96, and we refer to this paper for further details and references (cf. also Ref. 84). In the pertinent literature, the opinion trait is usually represented by a number v which takes values in the interval $I = \{|v| \leq 1\}$, where ± 1 represent the extremal opinions. In order to build a possibly realistic model, this severe limitation has to be coupled with a reasonable physical interpretation of the process of opinion forming. In other words, the impossibility of crossing the boundaries has to be a by-product of good modelling of binary interactions.

The two terms in the binary interaction (1.1) assume now the meaning of compromise and self-thinking. In other words, the functions P and Q take into account the local relevance of compromise and diffusion for a given opinion. Since extremal opinions can not be crossed, and opinions close to the extremal are considered more difficult to change, it is consistently assumed that both $P(v)$ and $Q(v)$ are nonincreasing functions of v^2, and in addition $Q(v)$ is equal to zero as $v^2 = 1$.

The most interesting example in Ref. 96 refers to the choice $P(v) = 1$, and $Q(v) = \sqrt{1 - v^2}$. In this case the weak form of the Fokker–Planck equation (2.8) for the opinion density $g(v, \tau)$ is given by

$$\frac{\partial g}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial v^2}\left((1 - v^2)g\right) + \frac{\partial}{\partial v}\left((v - m)g\right), \tag{2.20}$$

where $|v| \leq 1$, with suitable boundary conditions which guarantee mass and momentum conservation (cf. the discussion of Section 2.2). The interesting feature of the Fokker–Planck equation (2.20) is that it leads to close evolution of moments. The steady-state solution of equation (2.20) solves

$$\frac{\lambda}{2}\frac{\partial}{\partial v}\left((1-v^2)g\right) + (v-m)g = 0,$$

and equals

$$g_\infty(v) = c_{m,\lambda}\left(\frac{1}{1+v}\right)^{1-(1+m)/\lambda}\left(\frac{1}{1-v}\right)^{1-(1-m)/\lambda}.$$

The constant $c_{m,\lambda}$ is such that the mass of g_∞ is equal to one. Since $-1 < m < 1$, g_∞ is integrable on $\{|v| \leq 1\}$. Note that g_∞ has no peaks inside the interval, and as soon as $\lambda > 1 + |m|$ tends to infinity as $v \to \pm 1$.

2.4. *Further examples*

The basic models introduced in the previous sections are the building blocks to study more realistic situations. Indeed, one realizes that the distribution of wealth in a society can depend on further aspects of human behavior, like personal knowledge (cf. Ref. 64 for a detailed study of the relationships between early-life cognition and late-life financial knowledge). In this case one can fruitfully consider binary trades in which the outcome of the trade depends on both the wealth and the knowledge. This problem has been dealt with in Ref. 83, along the following lines. Knowledge is supposed to evolve along microscopic interactions like the ones described in Section 2.3.2. Then, given two agents A and B characterized by the pair (x, v) (respectively (y, w)) of knowledge and wealth, the new binary trade between A and B now reads

$$\begin{aligned} v^* &= \Big(1 - \Psi(x)\gamma + \Phi(x)\eta_1\Big)v + \Psi(y)\gamma w, \\ w^* &= \Big(1 - \Psi(y)\gamma + \Phi(y)\eta_2\Big)w + \Psi(x)\gamma v. \end{aligned} \tag{2.21}$$

In (2.21) the personal saving propensity and risk perception of the agents depends on their personal knowledge and are contained into the functions $\Psi = \Psi(x)$ and $\Phi = \Phi(x)$. In this way, the outcome of binary trade considered in Section 2.3.1 results from a combined effect of (personal) saving propensity, knowledge and wealth. Among other possibilities, one reasonable choice is to fix the functions Ψ and Φ as non-increasing functions. This reflects the idea that the knowledge could be fruitfully employed both to improve the result of the outcome and to reduce the risks. In Ref. 83, the numerical experiments have been done by choosing $\Psi(x) = (1+x)^{-\alpha}$ and $\Phi(x) = (1+x)^{-\beta}$, with various values of $\alpha, \beta > 0$.

The choice of a trade in the form (2.21) induces a limit procedure that generates in the scaling (2.3) a Fokker-Planck equation for the joint density $h = h(x, v, \tau)$ of knowledge and wealth

$$\frac{\partial h}{\partial \tau} = \frac{\delta}{2}\frac{\partial^2 h}{\partial x^2} + \Phi^2(x)\frac{\sigma}{2}\frac{\partial^2 h}{\partial v^2} + \frac{\partial}{\partial x}\left[(x\lambda(x) - \lambda_B M)h\right] + \gamma\frac{\partial}{\partial v}\left[(\Psi(x)v - M_W(\tau))h\right], \tag{2.22}$$

where we denoted

$$M_W(\tau) = \left\langle \int_{\mathbb{R}_+^2} w\Psi(y)h(y,w,\tau)\,dy\,dw \right\rangle. \tag{2.23}$$

In the simpler case in which the saving propensity remains a universal constant, so that $\Psi(y) = 1$, the drift term in the Fokker-Planck equation (2.22) simplifies, and, by resorting to the conservation of the mean wealth, one can show that the density $h = h(x, v, \tau)$ solves the equation

$$\frac{\partial h}{\partial \tau} = \frac{\delta}{2}\frac{\partial^2 h}{\partial x^2} + \Phi^2(x)\frac{\sigma}{2}\frac{\partial^2 h}{\partial v^2} + \frac{\partial}{\partial x}\left[(x\lambda(x) - \lambda_B M)h\right] + \gamma\frac{\partial}{\partial v}\left[(v - M_W)h\right],$$

where now M_W represents the (constant) value of the quantity in (2.23). Unlike the previous cases in which the steady state solution was explicitly found, in presence of two parameters it seems difficult to extract not only the explicit expression, but also its essential properties.

3. Large-time behavior of Fokker–Planck equations

The presentation of the previous section clarifies that socio-economic modelling leads to consider a variety of one-dimensional linear Fokker–Planck equations which are characterized by the presence of non constant diffusion coefficients, general drift terms and boundaries. The further interesting features of these Fokker–Planck equations is that they are closely related to kinetic models based on binary interactions with natural conservation properties. This characteristic introduces a correct physical structure on these equations, which in many cases have an explicit stationary solution, which possesses the essential features which are expected from the modelling assumptions.

The main example is represented by the Fokker–Planck equation for wealth distribution (2.14), which exhibits a steady state with Pareto tails. Even reducing the mathematical analysis to Fokker–Planck equations with a standard drift term, the presence of a diffusion term with variable diffusion coefficient is such that, while the explicit form of the equilibrium density is available in many cases, the standard methods which are usually introduced to control the large-time behavior of the solution, and the possible rate of convergence of the solution towards the equilibrium fail. A detailed analysis of the differential inequalities which are classically used to control convergence has been recently done in Ref. 80. This analysis shows that in general one can not expect, in presence of a diffusion term with non constant coefficient, exponential convergence in relative entropy of the solution to the Fokker–Planck equation towards the equilibrium solution.

In what follows, we enlighten in some details the available techniques, and the possible results that can be expected for this new group of Fokker–Planck equations.

For the sake of clarity, we will consider only the case of Fokker–Planck equations which are characterized by a linear drift term and a variable diffusion coefficient. Given a density $f = f(v,\tau)$, where the trait variable v belongs to the domain I, the prototype of these equations can be fruitfully written in divergence form as

$$\frac{\partial f}{\partial \tau} = \frac{\partial}{\partial v}\left[\frac{\partial}{\partial v}\left(\kappa(v) f\right) + (v-m)\, f\right]. \tag{3.1}$$

The diffusion coefficient $\kappa(v)$ is a nonnegative function. In general κ vanishes only for $v = 0$ if it is defined on the half line and in correspondence to the boundaries of the domain whenever this is bounded. Equation (3.1) covers most of the models introduced in Section 2.3. The Fokker–Planck equation (2.14) describing the evolution of wealth, corresponds to the choice $\kappa(v) = (\lambda v^2)/2$, $m = 1$ and $I = (0, +\infty)$. Similar choices lead to the Fokker–Planck equations for knowledge formation (2.17), and to equation (2.18) for conviction. Here, also the case $\kappa(v) = (\lambda v)/2$ is important to study. Finally, the choice $\kappa(v) = (\lambda(1-v^2))/2$, and $I = (-1,1)$ leads to the Fokker-Planck equation (2.20) treated in Section 2.3.4.

The stationary solution f_∞ of equation (3.1) is found by solving on $I \subseteq \mathbb{R}$ the differential equation

$$\frac{\partial}{\partial v}\left(\kappa(v) f_\infty\right) + (v-m)\, f_\infty = 0. \tag{3.2}$$

Equation (3.1) admits many equivalent formulations, each of them useful for various purposes. For example, since

$$\frac{\partial}{\partial v}\left(\kappa(v) f\right) + (v-m)\, f = \kappa(v) f\left(\frac{\partial}{\partial v}\log(\kappa(v) f) + \frac{v-m}{\kappa(v)}\right) =$$

$$\kappa(v) f\left(\frac{\partial}{\partial v}\log(\kappa(v) f) - \frac{\partial}{\partial v}\log(\kappa(v) f_\infty)\right) = \kappa(v) f \frac{\partial}{\partial v}\log\frac{f}{f_\infty} = \kappa(v) f_\infty \frac{\partial}{\partial v}\frac{f}{f_\infty},$$

we can write the Fokker–Planck equation (3.1) in the equivalent form

$$\frac{\partial f}{\partial \tau} = \frac{\partial}{\partial v}\left[\kappa(v) f \frac{\partial}{\partial v}\log\frac{f}{f_\infty}\right], \tag{3.3}$$

which enlightens the role of the logarithm of the quotient f/f_∞, and

$$\frac{\partial f}{\partial \tau} = \frac{\partial}{\partial v}\left[\kappa(v) f_\infty \frac{\partial}{\partial v}\frac{f}{f_\infty}\right]. \tag{3.4}$$

In particular, owing to (3.2), the form (3.4) allows us to obtain the evolution equation for the quotient $F = f/f_\infty$. Indeed

$$\frac{\partial f}{\partial \tau} = f_\infty \frac{\partial F}{\partial \tau} = \kappa(v) f_\infty \frac{\partial^2}{\partial v^2}\frac{f}{f_\infty} + \frac{\partial}{\partial v}(\kappa(v) f_\infty)\frac{\partial}{\partial v}\frac{f}{f_\infty} =$$

$$\kappa(v) f_\infty \frac{\partial^2 F}{\partial v^2} - (v-m) f_\infty \frac{\partial F}{\partial v},$$

which shows that F satisfies the equation

$$\frac{\partial F}{\partial \tau} = \kappa(v)\frac{\partial^2 F}{\partial v^2} - (v-m)\frac{\partial F}{\partial v}. \tag{3.5}$$

Remark 3.1. As discussed in Section 2.2, if mass conservation is imposed on equation (3.1), and $I = (i_-, i_+) \subseteq \mathbb{R}$ is the allowed domain, we obtain the boundary conditions

$$\frac{\partial}{\partial v}(\kappa(v)f(v,\tau)) + (v-m)f(v,\tau)\Big|_{v=i_\pm} = 0, \quad \tau > 0.$$

In analogous way, the boundary conditions of the two equivalent forms of the Fokker-Planck equation (3.1), given by (3.3) and (3.4), follow by imposing mass conservation. In this case for $\tau > 0$

$$k(v)f(v,\tau)\frac{\partial}{\partial v}\log\frac{f(v,\tau)}{f_\infty(v)}\Big|_{v=i_\pm} = 0. \tag{3.6}$$

Likewise, if mass conservation is imposed on equation (3.4), the natural boundary condition reads for $\tau > 0$

$$k(v)f_\infty(v)\frac{\partial}{\partial v}\frac{f(v,\tau)}{f_\infty(v)}\Big|_{v=i_\pm} = k(v)f_\infty(v)\frac{\partial F(v,\tau)}{\partial v}\Big|_{v=i_\pm} = 0. \tag{3.7}$$

Remark 3.2. For a suitable choice of the coefficients, equation (3.5) allows us to obtain in various cases uniform bounds on the ratio $F(\tau) = f(\tau)/f_\infty$, provided that the same bounds hold for the initial value f_0/f_∞. In particular, when the maximum principle is shown to hold for the solution to equation (3.5), all the computations that will be done in the forthcoming sections are rigorously justified. This is the case, for example, of the Fokker–Planck equation for wealth distribution (2.14) introduced in Section 2.3.1, which corresponds to the choice $\kappa(v) = v^2$, studied in Ref. 94, and of the Fokker–Planck equation for opinion formation (2.20) considered in Section 2.3.4. This last equation indeed can be treated in the framework of the theory developed by Lions and Le Bris in Ref. 75. For this reason, even if a general result is still missing, we will always assume in the following that the solution to (3.5) satisfies a maximum principle, and that consequently all the forthcoming computations are rigorously justified.

3.1. *Lyapunov functionals*

As it happens for the standard Fokker-Planck equation (cf. for example Ref. 95), convergence to the stationary state can be achieved by looking at the monotonicity in time of various Lyapunov functionals of the solution. The typical one is the relative Shannon entropy. Let $f, g : I \subset \mathbb{R} \longrightarrow \mathbb{R}^+$ denote two probability densities. Then, the relative Shannon entropy of f and g is defined by the formula

$$H(f,g) = \int_I f(v)\log\frac{f(v)}{g(v)}\,dv. \tag{3.8}$$

As proven in Ref. 2, whenever $g = f_\infty$, and the classical Fokker–Planck equation is considered (i.e. $\kappa(v) = 1$ and $m = 0$), other Lyapunov functionals are shown to decrease monotonically in time.

The same property holds true for the Fokker–Planck equation (3.1). This follows easily from the following

Proposition 3.1. *Let $F(v,\tau)$ be the solution to equation (3.5) in $I = (i_-, i_+) \subseteq \mathbb{R}$. Then, if $\Psi(v)$ is a smooth function such that*

$$|\Psi(i_\pm)| \le c < \infty, \tag{3.9}$$

the following equality holds

$$\int_I f_\infty(v)\Psi(v)\frac{\partial F(v,\tau)}{\partial \tau}\,dv = -\int_I \kappa(v) f_\infty(v)\frac{\partial \Psi(v)}{\partial v}\frac{\partial F(v,\tau)}{\partial v}\,dv. \tag{3.10}$$

Proof. The proof of (3.10) is immediate. Indeed, since $F(v,\tau)$ satisfies (3.5), then

$$\begin{aligned}
&\int_I f_\infty(v)\Psi(v)\frac{\partial F(v,\tau)}{\partial \tau}\,dv = \\
&\int_I \left(\Psi(v)\kappa(v)f_\infty(v)\frac{\partial^2 F(v,\tau)}{\partial v^2} - (v-m)f_\infty(v)\Psi(v)\frac{\partial F(v,\tau)}{\partial v}\right)dv = \\
&\Psi(v)\kappa(v)f_\infty(v)\frac{\partial F(v,\tau)}{\partial v}\bigg|_{i_-}^{i_+} - \int_I \frac{\partial F(v,\tau)}{\partial v}\frac{\partial}{\partial v}\left(\kappa(v)f_\infty(v)\Psi(v)\right)dv + \\
&-\int_I \Psi(v)(v-m)f_\infty(v)\frac{\partial F(v,\tau)}{\partial v}\,dv = \\
&-\int_I \kappa(v)f_\infty(v)\frac{\partial F(v,\tau)}{\partial v}\frac{\partial \Psi(v)}{\partial v}\,dv + \\
&-\int_I \Psi(v)\left(\frac{\partial}{\partial v}\left(\kappa(v)f_\infty(v)\right) + (v-m)f_\infty(v)\right)\frac{\partial F(v,\tau)}{\partial v}\,dv = \\
&-\int_I \kappa(v)f_\infty(v)\frac{\partial F(v,\tau)}{\partial v}\frac{\partial \Psi(v)}{\partial v}\,dv.
\end{aligned}$$

Indeed, the border term vanishes in view of conditions (3.9) and (3.7), and we used (3.2) on the last line to conclude. □

Proposition 3.1 has important consequences, we prove in the following

Theorem 3.1. *Let the smooth function $\Phi(x)$, $x \in \mathbb{R}_+$ be convex. Then, if $F(v,\tau)$ is the solution to equation (3.5) in $I = (i_-, i_+) \subseteq \mathbb{R}$, and $c \le F(v,\tau) \le C$ for some positive constants $c < C$, the functional*

$$\Theta(F(\tau)) = \int_I f_\infty(v)\Phi(F(v,\tau))\,dv$$

is monotonically decreasing in time, and the following equality holds

$$\frac{d}{d\tau}\Theta(F(\tau)) = -I_\Theta(F(\tau)), \tag{3.11}$$

where I_Θ *denotes the nonnegative quantity*

$$I_\Theta(F(\tau)) = \int_I \kappa(v) f_\infty(v) \Phi''(F(v,\tau)) \left| \frac{\partial F(v,\tau)}{\partial v} \right|^2 dv. \tag{3.12}$$

Proof. Since the integral defining $\Theta(F(\tau))$ is uniformly bounded, we get

$$\frac{d}{d\tau}\Theta(F(\tau)) = \int_I f_\infty(v) \Phi'(F(v,\tau)) \frac{\partial F(v,\tau)}{\partial \tau}\, dv.$$

Then, we apply Proposition 3.1 with $\Psi(v) = \Phi'(F(v,\tau))$ with fixed $\tau > 0$. □

Remark 3.3. Theorem 3.1 shows that the decay of convex functionals along the solution to the Fokker–Planck equation (3.1) does not depend on the presence of the weight κ. However, the weight κ is present in the rate of decay. As we will see, this represents a major obstacle in determining precise rates of convergence.

We list below various leading examples of functionals that can be used to look for the large-time behavior of the solution to equation (3.1).

The relative Shannon entropy. The Shannon entropy of f relative to f_∞, defined by (3.8) with $g = f_\infty$, is obtained by choosing $\Phi(x) = x \log x$. In this case

$$I_\Theta(F(\tau)) = \int_I \kappa(v) f_\infty(v) \frac{1}{F(v,\tau)} \left| \frac{\partial F(v,\tau)}{\partial v} \right|^2 dv.$$

This quantity is usually referred to as entropy production and we will denote it by $I_\kappa(f(\tau), f_\infty)$. So formula (3.11) applied to the relative Shannon entropy reads

$$\frac{d}{d\tau} H(f(\tau), f_\infty) = -I_\kappa(f(\tau), f_\infty). \tag{3.13}$$

It is interesting to remark that $I_\kappa(f(\tau), f_\infty)$ can be written in two equivalent ways. A first expression is obtained by observing that $f_\infty = f/F$. In this case we obtain

$$\begin{aligned} I_\kappa(f(\tau), f_\infty) &= \int_I \kappa(v) f_\infty(v) \frac{1}{F(v,\tau)} \left| \frac{\partial F(v,\tau)}{\partial v} \right|^2 dv = \\ \int_I \kappa(v) f(v,\tau) \frac{1}{F^2(v,\tau)} &\left| \frac{\partial F(v,\tau)}{\partial v} \right|^2 dv = \\ \int_I \kappa(v) f(v,\tau) &\left| \frac{\partial \log F(v,\tau)}{\partial v} \right|^2 dv = \\ \int_I \kappa(v) f(v,\tau) &\left(\frac{\partial_v f(v,\tau)}{f(v,\tau)} - \frac{\partial_v f_\infty(v)}{f_\infty(v)} \right)^2 dv. \end{aligned} \tag{3.14}$$

If $\kappa = 1$, the entropy production induced by the relative Shannon entropy is known with the name of Fisher information of f relative to f_∞. In general, one can introduce the following

Definition 3.1. Let f and g be two (smooth) probability densities, and let κ be a nonnegative function. Then, the Fisher information of f relative to g weighted by κ is

$$I_\kappa(f,g) = \int_I \kappa(v) f(v) \left(\frac{\partial_v f(v)}{f(v)} - \frac{\partial_v g(v)}{g(v)} \right)^2 \, dv. \tag{3.15}$$

In addition to (3.14), a second fruitful expression of the Fisher information of f relative to f_∞ is obtained in the following way[68]

$$\int_I \kappa(v) f_\infty(v) \frac{1}{F(v,\tau)} \left| \frac{\partial F(v,\tau)}{\partial v} \right|^2 \, dv = \int_I \kappa(v) f_\infty(v) \left| \frac{1}{F^{1/2}(v,\tau)} \frac{\partial F(v,\tau)}{\partial v} \right|^2 \, dv =$$
$$4 \int_I \kappa(v) f_\infty(v) \left| \frac{\partial \sqrt{F(v,\tau)}}{\partial v} \right|^2 \, dv = 4 \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v} \sqrt{\frac{f(v,\tau)}{f_\infty}} \right|^2 \, dv. \tag{3.16}$$

Hence, for a given probability density f

$$I_\kappa(f(\tau), f_\infty) = 4 \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v} \sqrt{\frac{f(v,\tau)}{f_\infty(v)}} \right|^2 \, dv. \tag{3.17}$$

Weighted L^2–distance. A second example is furnished by the choice $\Phi(x) = (x-1)^2$. In this case, the quantity $\Theta(F(\tau))$ gives the (weighted by f_∞^{-1}) L^2–distance of the solution to the Fokker–Planck equation from f_∞ itself.[2]

$$L_2(f(\tau), f_\infty) = \int_I f_\infty(v) \left(F(v,\tau) - 1 \right)^2 \, dv = \int_I f_\infty^{-1}(v) \left(f(v,\tau) - f_\infty(v) \right)^2 \, dv. \tag{3.18}$$

Then, formula (3.11) shows that

$$\frac{d}{d\tau} L_2(f(\tau), f_\infty) = -J_\kappa(f(\tau), f_\infty), \tag{3.19}$$

where the entropy production relative to L_2 is given by the nonnegative expression

$$J_\kappa(f(\tau), f_\infty) = 2 \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v} \frac{f(v,\tau)}{f_\infty(v)} \right|^2 \, dv. \tag{3.20}$$

Hellinger distance. A third example of Lyapunov functional refers to the choice $\Phi(x) = (\sqrt{x} - 1)^2$. In this case the Lyapunov functional $\Theta(F(\tau))$ coincides with the square of the Hellinger distance of $f(\tau)$ and f_∞

$$d_H^2(f(\tau), f_\infty) = \int_I f_\infty(v) \left(\sqrt{F(v,\tau)} - 1 \right)^2 \, dv = \int_I \left(\sqrt{f(v,\tau)} - \sqrt{f_\infty(v)} \right)^2 \, dv.$$

Indeed, the Hellinger distance is defined as follows.

Definition 3.2. For any given pair of nonnegative functions f and g defined on a subset $I \subset \mathbb{R}$, the Hellinger distance $d_H(f,g)$ is[66,103]

$$d_H(f,g) = \left(\int_I \left(\sqrt{f(v)} - \sqrt{g(v)} \right)^2 \, dv \right)^{\frac{1}{2}}. \tag{3.21}$$

Applying formula (3.12) shows that

$$\frac{d}{d\tau} d_H^2(f(\tau), f_\infty) = -D_\kappa(f(\tau), f_\infty), \tag{3.22}$$

where the entropy production relative to d_H^2 is easily computed to give the nonnegative expression

$$D_\kappa(f(\tau), f_\infty) = 8 \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v} \sqrt[4]{\frac{f(v,\tau)}{f_\infty(v)}} \right|^2 \, dv. \tag{3.23}$$

Reverse relative Shannon entropy. Our last example refers to the choice $\Phi(x) = -\log x$. In contrast with the case $\Phi(x) = x \log x$ the Lyapunov functional $\Theta(F(\tau))$ is given by the Shannon entropy of f_∞ relative to $f(\tau)$

$$H(f_\infty, f(\tau)) = \int_I f_\infty(v) \log \frac{f_\infty(v)}{f(v,\tau)} \, dv. \tag{3.24}$$

Applying formula (3.11) to (3.24) shows that

$$\frac{d}{d\tau} H(f_\infty, f(\tau)) = -\tilde{I}_\kappa(f_\infty, f(\tau)),$$

where this time the entropy production $\tilde{I}_\kappa$ has the nonnegative expression

$$\tilde{I}_\kappa(f_\infty, f(\tau)) = \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v} \log \frac{f(v,\tau)}{f_\infty(v)} \right|^2 \, dv. \tag{3.25}$$

For a given constant α, with $0 < \alpha < 1$ let us consider the particular solution to the Fokker–Planck equation (3.1) given by $\alpha f(\tau) + (1-\alpha) f_\infty$ and its relative entropy

$$H_\alpha(f_\infty, f(\tau)) := H(f_\infty, \alpha f(\tau) + (1-\alpha) f_\infty). \tag{3.26}$$

If $\alpha = 1/2$, the entropy H_α is known with the name of Jensen-Shannon entropy.[73] This class of entropies has been introduced and studied in information theory mainly in view of their equivalence to the Hellinger distance.[73,88,93] We will be back on these entropies in Section 4.

Remark 3.4. The notion of Jensen-Shannon relative entropy suggests to consider closely related relative entropies, like the following

Definition 3.3. For any pair of probability densities f and g taking values on $I \subseteq \mathbb{R}$ and $\alpha \in [0,1)$ we define the weighted Jensen–Shannon entropy by

$$\tilde{H}_\alpha(f,g) = \frac{1}{(1-\alpha)^2} \int_I g(v) \log \frac{g(v)}{\alpha g(v) + (1-\alpha) f(v)} \, dv.$$

This class of entropies connects the relative Shannon entropy (3.8), achieved when $\alpha = 0$

$$\tilde{H}_0(f,g) = H(g,f) = \int g(v) \log \frac{g(v)}{f(v)} \, dv,$$

to a weighted L^2–distance between f and g (see Ref.2). Indeed we have

Proposition 3.2. *For any pair of probability densities f and g defined on a subset $I \subset \mathbb{R}$ it holds*

$$\tilde{H}_1(f,g) := \lim_{\alpha\to 1} \tilde{H}_\alpha(f,g) = \frac{1}{2} L_2(f,g) = \frac{1}{2}\int_I \frac{(f(v)-g(v))^2}{g(v)}\,dv.$$

Proof. We apply De l'Hôpital formula twice and we get

$$\lim_{\alpha\to 1} \tilde{H}_\alpha(f,g) = \lim_{\alpha\to 1} \frac{1}{2(1-\alpha)} \int_I g(v)\,\frac{g(v)-f(v)}{\alpha g(v)+(1-\alpha)f(v)}\,dv =$$
$$\lim_{\alpha\to 1} \frac{1}{2}\int_I g(v)\,\frac{(g(v)-f(v))^2}{(\alpha g(v)+(1-\alpha)f(v))^2}\,dv =$$
$$\frac{1}{2}\int_I \frac{(f(v)-g(v))^2}{g(v)}\,dv = \frac{1}{2}L_2(f,g),$$

where $L_2(f,g)$ is given by (3.18). □

Remark 3.5. All the expressions of the entropy production we found above share a common feature. Indeed, they all are expressed as

$$I_\Theta(F(\tau)) = \int_I \kappa(v) f_\infty(v) \left| \frac{\partial}{\partial v}\phi(F(v,\tau))\right|^2 dv, \tag{3.27}$$

where $\phi(x) = c_\delta x^\delta$ in the first three cases (respectively $\delta = 1/2, 1$ and $1/4$), and $\phi(x) = \log x$ in (3.25).

Remark 3.6. Concerning the standard relative Shannon entropy (3.8), it is remarkable that $H(f(\tau), f_\infty)$ and $H(f_\infty, f(\tau))$ give rise to different entropy productions, given by (3.17) and (3.25) respectively. While the former has been investigated in details,[68,95] the study of the latter, at least to our knowledge, is not present in the pertinent literature.

Remark 3.7. If we set $I = \mathbb{R}$, $\kappa(v) = 1$ and $m = 0$, we reduce to the classical one-dimensional Fokker–Planck equation

$$\frac{\partial g}{\partial \tau} = \frac{\partial^2 g}{\partial v^2} + \frac{\partial}{\partial v}(vg), \quad v \in \mathbb{R}. \tag{3.28}$$

In this case, the steady state is given by the Maxwellian (Gaussian) density

$$f_\infty(v) = M(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}, \tag{3.29}$$

and the relation

$$\frac{d}{d\tau} H(f(\tau), M) = -I(f(\tau), M) \tag{3.30}$$

coupled with the log–Sobolev inequality (cf. for example Ref. 95)

$$H(f(\tau), M) \le \frac{1}{2} I(f(\tau), M)$$

leads to exponential decay to zero of the relative entropy. [95,97] We denoted by $I(f,g)$ the standard relative Fisher information between the two densities f and g

$$I(f,g) = \int_I f(v) \left(\frac{\partial_v f(v)}{f(v)} - \frac{\partial_v g(v)}{g(v)} \right)^2 \, dv. \tag{3.31}$$

Then, Csiszár–Kullback–Pinsker inequality[37,72]

$$\|f-g\|_{L^1}^2 \le 2H(f,g)$$

permits to prove exponential convergence in L^1 to the Maxwellian density.

If the diffusion weight $\kappa(v)$ is not constant, the analogous of the log–Sobolev inequality is not available.[80] This leads to the challenging problem of investigating whether it is still possible to control the relative Shannon entropy with the weighted relative Fisher information, or, in alternative, if one can obtain a lower bound on the relative weighted Fisher information in order to prove convergence of a generic solution $f(\tau)$ to the stationary state in some sense and at a certain rate.

3.2. *Steady states and Chernoff-type inequalities*

In Ref. 68, Johnson and Barron showed that for the classical Fokker–Planck equation (3.28) it is possible to prove convergence in L^1 to the Maxwellian equilibrium without resorting to the log–Sobolev inequality, at the price of loosing the exponential rate of convergence. Indeed, since the relative entropy is decreasing and relation (3.30) holds true, for any $\tau > 0$ it holds

$$H(f(\tau), M) - H(f(0), M) = -\int_0^\tau I(f(s), M) \, ds.$$

In particular, for all $\tau > 0$

$$\int_0^\tau I(f(s), M) \, ds \le H(f(0), M),$$

which shows that $I(f(\tau), M) \in L^1(0,\infty)$. Consequently there is at least a diverging sequence of times $\{\tau_k\}$ such that

$$\lim_{k\to\infty} I(f(\tau_k), M) = 0. \tag{3.32}$$

In order to avoid the restriction to a sequence, and to pass from the vanishing of the relative Fisher information into a genuine distance, Johnson and Barron made use of a well-known inequality in probability theory, the Chernoff inequality[33], coupled with the Hellinger distance. Let us first recall the original result of Chernoff.

Theorem 3.2 (Chernoff). *Let X be a Gaussian random variable distributed with density M given as in (3.29). If the function ϕ is absolutely continuous and $\phi(X)$ has finite variance, then*

$$Var[\phi(X)] \le E[\phi'(X)]^2 \tag{3.33}$$

with equality if and only if $\phi(X)$ is linear in X.

In analytical terms, inequality (3.33) reads

$$\int_{\mathbb{R}} \left(\phi(v) - \left(\int_{\mathbb{R}} \phi(v)M(v)\ dv\right)\right)^2 M(v)\ dv \le \int_{\mathbb{R}} M(v)\left(\phi'(v)\right)^2\ dv. \tag{3.34}$$

Inequality (3.34) allows us to connect the relative Fisher information (3.31) with the Hellinger distance.[68] In view of expression (3.17), written for $\kappa = 1$ and $f_\infty = M$, let us apply inequality (3.34) with $\phi(v) = \sqrt{f(v,\tau)/M(v)}$ for fixed $\tau > 0$. Since the solution $f(\tau)$ of the classical Fokker–Planck equation has an explicit expression as a convolution between the initial density f_0 and the Gaussian kernel, $\sqrt{f(v,\tau)/M(v)}$ is smooth enough to satisfy the assumptions in Chernoff theorem. Recalling that $f(\tau)$ and M are probability density functions, we get

$$\begin{aligned}
&\int_{\mathbb{R}} M(v)\left(\partial v\sqrt{\frac{f(v,\tau)}{M(v)}}\right)^2\ dv \ge \\
&\int_{\mathbb{R}} \left(\sqrt{\frac{f(v,\tau)}{M(v)}} - \left(\int_{\mathbb{R}} \sqrt{\frac{f(v,\tau)}{M(v)}} M(v)\ dv\right)\right)^2 M(v)\ dv = \\
&\int_{\mathbb{R}} \left(\sqrt{\frac{f(v,\tau)}{M(v)}}\right)^2 M(v)\ dv - \left(\int_{\mathbb{R}} \sqrt{\frac{f(v,\tau)}{M(v)}} M(v)\ dv\right)^2 = \\
&\int_{\mathbb{R}} f(v,\tau)\ dv - \left(\int_{\mathbb{R}} \sqrt{f(v,\tau)M(v)}\ dv\right)^2 = \\
&1 - \left(\int_{\mathbb{R}} \sqrt{f(v,\tau)\,M(v)}\ dv\right)^2.
\end{aligned}$$

Hence we have

$$I(f(\tau), M) \ge 4\left(1 - \left(\int_{\mathbb{R}} \sqrt{f(v,\tau)\,M(v)}\ dv\right)^2\right). \tag{3.35}$$

It is immediate to relate the right-hand side of the previous inequality to the Hellinger distance. Indeed, whenever f and g are probability density functions

$$\begin{aligned}
&\int_{\mathbb{R}} \left(\sqrt{f(v)} - \sqrt{g(v)}\right)^2\ dv = \int_{\mathbb{R}} \left(f(v) + g(v) - 2\sqrt{f(v)\,g(v)}\right) dv = \\
&2\left(1 - \int_{\mathbb{R}} \sqrt{f(v)\,g(v)}\ dv\right) \le 2\left(1 - \left(\int_{\mathbb{R}} \sqrt{f(v)\,g(v)}\ dv\right)^2\right).
\end{aligned} \tag{3.36}$$

The last inequality in (3.36) follows by Cauchy–Schwartz inequality. Finally, (3.35) and (3.36) imply

$$I(f(\tau), M) \ge 2d_H(f(\tau), M)^2, \quad \tau > 0 \tag{3.37}$$

and, by (3.32)

$$\lim_{k\to\infty} d_H(f(\tau_k), M) = 0.$$

To use a similar approach for the Fokker–Planck equation (3.1), one needs a lower bound for the relative weighted Fisher information (3.15) in terms of the Hellinger distance. Indeed, if

$$I_\kappa(f(\tau), f_\infty) \geq 2d_H(f(\tau), f_\infty)^2, \quad \tau > 0 \tag{3.38}$$

by (3.13) we would conclude. Note that in reason of (3.16), inequality (3.38) is a weighted Chernoff-type inequality for the probability density f_∞.

The interest in this type of proof is related to the fact that, while log-Sobolev type inequalities are very difficult to achieve for a general steady state f_∞, Chernoff-type inequalities have been shown to be more flexible, and to hold for various probability density functions.[71] Surprisingly enough, the Fokker–Planck equation (3.1) permits to establish a key differential relationship between its steady state f_∞ and the weight κ which furnishes a direct way to reobtain the Chernoff-type inequalities of Klaassen in Ref. 71 for the probability density f_∞ with a sharp constant.

In other words, the differential equation which defines the steady state f_∞ of the Fokker–Planck equation (3.1) is the relationship which enables us to prove the Chernoff-type inequality for the probability density f_∞, and characterizes in a sharp way the weight associated to f_∞. We prove

Theorem 3.3 (Chernoff with weight). *Let X be a random variable distributed with density $f_\infty(v)$, $v \in I \subseteq \mathbb{R}$, where the probability density function f_∞ satisfies the differential equality*

$$\frac{\partial}{\partial v}\left(\kappa(v) f_\infty\right) + (v - m)\, f_\infty = 0, \quad v \in I. \tag{3.39}$$

If the function ϕ is absolutely continuous on I and $\phi(X)$ has finite variance, then

$$Var[\phi(X)] \leq E\left\{\kappa(X)[\phi'(X)]^2\right\} \tag{3.40}$$

with equality if and only if $\phi(X)$ is linear in X.

Remark 3.8. While generalized weighted Chernoff–type inequalities involving probability densities different from the normal density already exist in literature (cf. for example Ref. 71), the main novelty here is to point out the deep relationship between the weight κ and the density f_∞, which naturally appears by looking for the stationary state (of unit mass) of the Fokker–Planck equation (3.1).

Proof. For any given $v \in I$ and constant value $m \in I$, let $r(v,t)$ be defined as

$$r(v,t) = (v - m)t + m, \qquad 0 \leq t \leq 1$$

so that $r(v,0) = m$ and $r(v,1) = v$. Let $\dot r(v,t)$ denote the partial derivative of r with respect to t. Thanks to the gradient theorem, for any given (smooth) function $\phi(v)$ we have

$$\phi(v) - \phi(m) = \int_0^1 \phi'(r(v,t))\dot r(v,t)dt = \int_0^1 \phi'(r(v,t))(v-m)dt,$$

so that, by Jensen's inequality

$$(\phi(v)-\phi(m))^2 \le \int_0^1 \left[\phi'(r(v,t))(v-m)\right]^2 dt. \tag{3.41}$$

On the other hand, for any probability density function f_∞, we have the elementary inequality

$$\int_I \phi^2(v) f_\infty(v)\ dv - \left(\int_I \phi(v) f_\infty(v)\ dv\right)^2 \le \int_I (\phi(v)-\phi(m))^2 f_\infty(v)\ dv.$$

Hence, by (3.41)

$$Var[\phi(X)] \le \int_I (\phi(v)-\phi(m))^2 f_\infty(v)\ dv \le \int_I f_\infty(v)\,(v-m)^2 \int_0^1 \left(\phi'(r(v,t))\right)^2\ dt\,dv. \tag{3.42}$$

By virtue of (3.39)

$$\begin{aligned}
&\int_I (v-m)^2 f_\infty(v) \int_0^1 (\phi'(r(v,t)))^2\ dt\ dv = \\
&-\int_I (v-m)\frac{\partial}{\partial v}\left[\kappa(v) f_\infty(v)\right] \int_0^1 (\phi'(r(v,t)))^2\ dt\,dv \le \\
&\int_I \kappa(v) f_\infty(v) \frac{\partial}{\partial v}\left[(v-m)\int_0^1 (\phi'(r(v,t)))^2\,dt\right]\ dv.
\end{aligned} \tag{3.43}$$

In (3.43), we used integration by parts to get the last line. Indeed here the border term satisfies

$$-\ (v-m)\kappa(v) f_\infty(v) \int_0^1 (\phi'(r(v,t)))^2\ dt\bigg|_{i_-}^{i_+} \le 0. \tag{3.44}$$

When I is a bounded interval, inequality (3.44) is a consequence of the fact that $i_- < m < i_+$, while $\kappa \ge 0$. When $I = \mathbb{R}_+$ inequality (3.44) is proven by considering that, since m is a positive bounded constant,

$$\lim_{v\to\infty} (v-m)\kappa(v) f_\infty(v) \ge 0,$$

while

$$\lim_{v\to 0} (v-m)\kappa(v) f_\infty(v) \le 0.$$

The proof is completed by observing that, for any given function $\psi(r(v,t))$ one has the identity

$$\frac{d}{dt}\left(t\psi((v-m)t+m)\right) = \frac{\partial}{\partial v}\left[(v-m)\psi((v-m)t+m)\right]. \tag{3.45}$$

Clearly, (3.45) implies

$$\frac{\partial}{\partial v}\left[(v-m)\int_0^1 (\phi'(r(v,t)))^2\,dt\right] = \int_0^1 \frac{d}{dt}\left(t(\phi'(r(v,t)))^2\right)\ dt = (\phi'(v))^2.$$

Substituting into (3.43) gives the result. □

Remark 3.9. Inequality (3.40) improves the analogous one obtained by Klaassen in Ref. 71. Indeed, the relationship (3.39) enables us to identify the correct weight κ to set into this inequality. Surprisingly enough, the result of Klaassen has to our knowledge never been improved. Unfortunately, this proof is restricted to one-dimensional Fokker–Planck type equations. Related results in higher dimensions have been recently obtained in Ref. 17.

If we apply the new Chernoff–type inequality (3.40) with $\phi(v) = \sqrt{f(v,\tau)/f_\infty(v)}$ with fixed $\tau > 0$ (assuming $\sqrt{f(v,\tau)/f_\infty(v)}$ smooth enough, as we will point out in the next section) we get

$$I_\kappa(f(\tau), f_\infty) = 4\int_I \kappa(v) f_\infty(v) \left(\partial_v \sqrt{\frac{f(v,\tau)}{f_\infty(v)}}\right)^2 \, dv \geq$$
$$4\left(\int_I \frac{f(v,\tau)}{f_\infty(v)} f_\infty(v)\, dv - \left(\int_I \sqrt{\frac{f(v,\tau)}{f_\infty(v)}} f_\infty(v)\, dv\right)^2\right) = \tag{3.46}$$
$$4\left(1 - \left(\int_{\mathbb{R}} \sqrt{f(v,\tau)\, f_\infty(v)}\, dv\right)^2\right).$$

This is exactly the same relation as (3.35), which implies inequality (3.38), namely the same inequality (3.37) found in the classical Fokker–Planck case. In the case in which the decay of the relative Shannon entropy can be rigorously proven, this leads again to

$$\lim_{k\to\infty} d_H(f(\tau_k), f_\infty) = 0.$$

3.3. *Rates of convergence*

In this short section we will give few indications on the way in which the results we found in Section 3 can be used to determine rates of convergence to equilibrium for the solution to the Fokker–Planck equation (3.1). The general strategy is the following.[94] Let us consider the solution to the initial value problem for equation (3.1) corresponding to an initial probability density f_0 with some relative (to the equilibrium solution) Lyapunov functional bounded, typically the relative Shannon entropy (3.8). If the assumptions of Theorem 3.1 do not hold, we consider a suitable lifting of the initial value $f_{0,\epsilon}$, with the same mass of f_0, bounded Lyapunov functional, but such that the subsequent solution to equation (3.5), say F_ϵ (cf. Remark 3.2) satisfies $c \leq F_\epsilon \leq C$ for some finite positive constants $c < C$ and it is smooth enough to justify application of the generalized Chernoff inequality (3.40). Starting from the lifted initial datum, we can rigorously prove the time-decay of the Lyapunov functional, and subsequently apply the generalized Chernoff inequality (3.40) proven in Theorem 3.3. Once the time-rate of convergence (independent of ϵ) has been derived, one can eliminate the lifting. As a first example, let us start from an initial probability density with bounded Shannon entropy relative to the stationary solution. We have proved in the previous section that, regardless to the form of the dissipation coefficient κ and of the stationary state f_∞, the solution converges to the stationary state f_∞ at least on a sequence of diverging times τ_k and in Hellinger distance

$$\lim_{k\to\infty} d_H(f(\tau_k), f_\infty) = 0.$$

In order to obtain this result, it is enough to have a weighted Chernoff–type inequality (Theorem 3.3) which enables us to relate the relative Fisher information (3.15) with the Hellinger distance (3.21). Indeed, Chernoff inequality with weight is the key point to relate the relative entropy production (in this case the relative Fisher information) to the square of the Hellinger distance.

Moreover, resorting to the monotonicity of the Hellinger distance between $f(\tau)$ and f_∞, one can get rid of the restriction on a sequence of times in the convergence. Indeed we have

Theorem 3.4. *Let $f(\tau)$ be the solution of the Fokker–Planck equation* (3.1)*, corresponding to an initial value f_0 such that the relative Shannon entropy $H(f_0, f_\infty)$ is bounded, and the decay of relative entropy can be rigorously justified. Then,*

$$d_H(f(\tau), f_\infty)^2 = o\left(\frac{1}{\tau}\right), \quad \tau \to \infty.$$

Proof. The decay of the relative Shannon entropy coupled with the weighted Chernoff inequality in this case implies

$$\frac{d}{d\tau} H(f(\tau), f_\infty) = -I_\kappa(f(\tau), f_\infty) \le -2d_H(f(\tau), f_\infty)^2,$$

as we have showed in (3.46). Integrating with respect to time from 0 to ∞ we get the bound

$$\int_0^\infty 2d_H(f(\tau), f_\infty)^2 \, dt \le H(f_0, f_\infty). \tag{3.47}$$

Consequently, $d_H(f(\tau), f_\infty)^2$ belongs to $L^1(\mathbb{R}_+)$. On the other hand, by (3.22) the Hellinger distance is a monotone function of time. Hence

$$d_H(f(\tau), f_\infty)^2 = o\left(\frac{1}{\tau}\right), \quad \tau \to \infty. \qquad \square$$

In the case in which we need to apply a lifting to the initial value, inequality (3.47) still holds with the right-hand side given by the relative entropy $H(f_{0,\epsilon}, f_\infty)$. Letting ϵ to zero we then arrive to the same conclusion.

Remark 3.10. The previous result shows that, unlike the case of the classical Fokker–Planck equation, where exponential convergence in L^1 is known to hold for initial data with bounded relative entropy, in this case, by considering the same class of initial data, only a weaker time decay (essentially a rate of order $1/\tau^{1+\delta}$, with $\delta \ll 1$) is proven to hold. However, as we shall see, exponential convergence is found for initial data that are suitably close to the steady state.

This stronger result follows by considering as Lyapunov functional the weighted L^2-distance defined in (3.18). We have

Theorem 3.5. *Let f_0 a probability density satisfying*

$$L_2(f_0, f_\infty) < \infty$$

where f_∞ is the stationary solution of (3.1) *and L_2 is the weighted L^2–distance defined in* (3.18). *Then for any solution $f(\tau)$ of the Fokker–Planck equation* (3.1) *with f_0 as initial data, such that the decay of the L_2 functional can be rigorously justified, the following holds true*

$$L_2(f(\tau), f_\infty) \le e^{-2\tau} L_2(f_0, f_\infty), \quad \tau > 0.$$

Proof. The proof is based on Theorem 3.1 applied to the L_2 functional. Indeed by (3.19), (3.20) we get

$$\frac{d}{d\tau} \int_I \frac{(f(v,\tau) - f_\infty(v))^2}{f_\infty(v)} \, dv = -2 \int_I \kappa(v) f_\infty(v) \left(\frac{\partial}{\partial v} \frac{f(v,\tau)}{f_\infty(v)} \right)^2 \, dv.$$

Then, applying Theorem 3.3 with $\phi(v) = f(v,\tau)/f_\infty(v)$ with fixed $\tau > 0$, leads to

$$\begin{aligned} &\int_I \kappa(v) f_\infty(v) \left(\frac{\partial}{\partial v} \frac{f(v,\tau)}{f_\infty(v)} \right)^2 \, dv \ge \\ &\int_I \left(\frac{f(v,\tau)}{f_\infty(v)} - \left(\int_I \frac{f(v,\tau)}{f_\infty(v)} f_\infty(v) \, dv \right) \right)^2 f_\infty(v) \, dv = \\ &\int_I \left(\frac{f(v,\tau)}{f_\infty(v)} - 1 \right)^2 f_\infty(v) \, dv = \int_I \frac{(f(v,\tau) - f_\infty(v))^2}{f_\infty(v)} \, dv. \end{aligned}$$

Hence

$$\frac{d}{d\tau} \int_I \frac{(f(v,\tau) - f_\infty(v))^2}{f_\infty(v)} \, dv \le -2 \int_I \frac{(f(v,\tau) - f_\infty(v))^2}{f_\infty(v)} \, dv,$$

and this ends the proof. □

Remark 3.11. The boundedness of the quantity $L_2(f_0, f_\infty)$ is in general a quite restrictive condition, which limits the result to initial values very *close* to the stationary solution. For example, if one considers the Fokker–Planck equation for wealth distribution of Section 2.3.1, given by (2.14), the steady state (2.15) has a very rapid decay at $v = 0$. Hence, the boundedness of the weighted L^2- distance is equivalent to choosing well-behaved initial data, which have the same decay at $v = 0$, thus excluding more general (and natural) initial values for the problem. The situation in the case of the Fokker–Planck model for opinion formation presented in Section 2.3.4 is somewhat different. Here, the stationary solution of equation (2.20) is such that g_∞^{-1} is a bounded function, and the weighted L^2-norm is very close to the usual

L^2-norm. In this case, exponential convergence to equilibrium is found to hold for a larger class (with respect to equation (2.14)) of initial densities.

3.4. *Relative entropies between solutions are decaying*

We end this section by enlightening further the structure of the Fokker–Planck equation (3.1). Indeed, for any two solutions, say $f(\tau)$ and $g(\tau)$, departing from initial probability densities f_0 and g_0 such that some relative Lyapunov functional between them is initially bounded, the time-evolution of the same Lyapunov functional is monotonically decreasing in time. This property generalizes the analogous one proven for the pair f_0, $g_0 = f_\infty$. By virtue of Proposition 3.1 we proved that the relative Shannon entropy $H(f(\tau), f_\infty)$, the square of the Hellinger distance $d_H^2(f(\tau), f_\infty)$, the weighted L^2–distance $L_2(f(\tau), f_\infty)$ and the reverse relative Shannon entropy $H(f_\infty, f(\tau))$ are decreasing in time. Indeed, the same holds true for the same quantities evaluated on two generic solutions of the Fokker–Planck equation (3.1).

Proposition 3.3. *Let $f(\tau)$ and $g(\tau)$ denote two solutions of the Fokker–Planck equation (3.1), departing from initial data f_0 and g_0. Then*

$$\frac{d}{d\tau}\int_I f(v,\tau)\log\frac{f(v,\tau)}{g(v,\tau)}dv = -\int_I \kappa(v) f(v,\tau)\left(\frac{\partial}{\partial v}\log\frac{f(v,\tau)}{g(v,\tau)}\right)^2 dv, \tag{3.48}$$

$$\frac{d}{d\tau}\int_I \left(\sqrt{f(v,\tau)} - \sqrt{g(v,\tau)}\right)^2 dv = \\ -\frac{1}{2}\int_I \kappa(v)\sqrt{f(v,\tau)\,g(v,\tau)}\left(\frac{\partial}{\partial v}\log\frac{f(v,\tau)}{g(v,\tau)}\right)^2 dv, \tag{3.49}$$

$$\frac{d}{d\tau}\int_I \frac{(f(v,\tau)-g(v,\tau))^2}{g(v,\tau)}\,dv = -2\int_I \kappa(v) g(v,\tau)\left(\frac{\partial}{\partial v}\frac{f(v,\tau)}{g(v,\tau)}\right)^2 dv. \tag{3.50}$$

Proof. In order to make the proof easier to read, we will drop all the variables v or v,τ unless they are useful for comprehension. Let us begin by proving (3.48). To simplify computations, we resort to the proof of Proposition 3.1. Hence, denoting again $F = f/f_\infty$ and $G = g/f_\infty$, we prove (3.48) in an equivalent form in terms of F and G. Resorting to the conservation of the mass of f we get

$$\frac{d}{d\tau}\int_I f_\infty F\log\frac{F}{G}dv = \int_I f_\infty\frac{\partial F}{\partial\tau}\,\log\frac{F}{G}\,dv + \int_I f_\infty G\frac{\partial}{\partial\tau}\left(\frac{F}{G}\right)\,dv = \\ \int_I f_\infty\log\frac{F}{G}\frac{\partial F}{\partial\tau}\,dv - \int_I f_\infty\frac{F}{G}\frac{\partial G}{\partial\tau}\,dv.$$

Next, Proposition 3.1 gives

$$\int_I f_\infty \log\frac{F}{G}\frac{\partial F}{\partial \tau}\,dv - \int_I f_\infty \frac{F}{G}\frac{\partial G}{\partial \tau}\,dv =$$
$$-\int_I \kappa(v) f_\infty \left(\frac{\partial}{\partial v}\left(\log\frac{F}{G}\right)\frac{\partial F}{\partial v} - \frac{\partial}{\partial v}\left(\frac{F}{G}\right)\frac{\partial G}{\partial v}\right) dv =$$
$$-\int_I \kappa(v) f_\infty \left(F\frac{\partial}{\partial v}(\log F)\frac{\partial}{\partial v}\left(\log\frac{F}{G}\right) - F\frac{\partial}{\partial v}\left(\log\frac{F}{G}\right)\frac{\partial}{\partial v}(\log G)\right) dv =$$
$$\int_I \kappa(v) f_\infty F\left(\frac{\partial}{\partial v}\left(\log\frac{F}{G}\right)\right)^2 dv = \int_I \kappa(v) f\left(\frac{\partial}{\partial v}\left(\log\frac{f}{g}\right)\right)^2 dv.$$

Note that the last term coincides with the relative Fisher information of f relative to g weighted by κ, defined in (3.15).

Let us now prove (3.49). As before, we write the square of the Hellinger distance $d_H^2(f,g) = \int_I \left(\sqrt{f} - \sqrt{g}\right)^2 dv$, in terms of F and G. Owing to mass conservation we obtain

$$\frac{d}{d\tau}\int_I \left(\sqrt{F} - \sqrt{G}\right)^2 f_\infty dv = \int_I \left(\sqrt{F} - \sqrt{G}\right)\left(\frac{\partial_\tau F}{\sqrt{F}} - \frac{\partial_\tau G}{\sqrt{G}}\right) f_\infty\, dv =$$
$$-\int_I \left(\sqrt{\frac{G}{F}}\partial_\tau F + \sqrt{\frac{F}{G}}\partial_\tau G\right) f_\infty\, dv.$$

Next, Proposition 3.1 implies

$$-\int_I \left(\sqrt{\frac{G}{F}}\partial_\tau F + \sqrt{\frac{F}{G}}\partial_\tau G\right) f_\infty\, dv =$$
$$\int_I \kappa(v) f_\infty \left(\frac{\partial}{\partial v}\sqrt{\frac{G}{F}}\;\partial_v F + \frac{\partial}{\partial v}\sqrt{\frac{F}{G}}\;\partial_v G\right) dv =$$
$$\int_I \kappa(v) f_\infty \sqrt{FG}\left(\frac{\partial}{\partial v}\sqrt{\frac{G}{F}}\;\frac{\partial_v F}{\sqrt{FG}} + \frac{\partial}{\partial v}\sqrt{\frac{F}{G}}\;\frac{\partial_v G}{\sqrt{FG}}\right) dv =$$

$$-\frac{1}{2}\int_I \kappa(v) f_\infty \sqrt{FG}\left(\left(\frac{\partial_v F}{F}\right)^2 + \left(\frac{\partial_v G}{G}\right)^2 - 2\left(\frac{\partial_v F}{F}\right)\left(\frac{\partial_v G}{G}\right)\right) dv =$$
$$-\frac{1}{2}\int_I \kappa(v) f_\infty \sqrt{FG}\left(\frac{\partial_v F}{F} - \frac{\partial_v G}{G}\right)^2 dv =$$
$$-\frac{1}{2}\int_I \kappa(v) f_\infty \sqrt{FG}\left(\frac{\partial}{\partial v}\left(\log\frac{F}{G}\right)\right)^2 dv = -\frac{1}{2}\int_I \kappa(v)\sqrt{fg}\left(\frac{\partial}{\partial v}\left(\log\frac{f}{g}\right)\right)^2 dv.$$

Note that, in constrast with the previous computation, the relative entropy production is completely symmetric. Simple computations then show that this expression can be rewritten in a completely equivalent form. One has indeed

$$\frac{1}{2}\int_I \kappa(v)\sqrt{fg}\left(\frac{\partial}{\partial v}\left(\log\frac{f}{g}\right)\right)^2 dv = D_\kappa(f,g) = D_\kappa(g,f),$$

where D_κ has been defined in (3.23).

Last, we prove relation (3.50) on the weighted L^2–distance. To take advantage from Proposition 3.1, we write the distance in its equivalent form in terms of F and G. We obtain

$$\frac{d}{d\tau}\int_I \left(\frac{F}{G}-1\right)^2 G f_\infty\, dv = -2\int_I \kappa(v) f_\infty G \left(\frac{\partial}{\partial v}\left(\frac{F}{G}\right)\right)^2\, dv.$$

In details,

$$\begin{aligned}
&\frac{d}{d\tau}\int_I \left(\frac{F}{G}-1\right)^2 G f_\infty\, dv = \\
&2\int_I \left(\frac{F}{G}-1\right)\partial_\tau\left(\frac{F}{G}\right) G f_\infty\, dv + \int_I \left(\frac{F}{G}-1\right)^2 \partial_\tau G\, f_\infty\, dv = \\
&2\int_I \left(\frac{F}{G}-1\right)\left(\frac{1}{G}\,\partial_\tau F - \frac{F}{G^2}\partial_\tau G\right) G f_\infty\, dv + \int_I \left(\frac{F}{G}-1\right)^2 \partial_\tau G\, f_\infty\, dv = \\
&2\int_I \left(\frac{F}{G}-1\right)\partial_\tau F\, f_\infty\, dv - 2\int_I \left(\frac{F}{G}-1\right)\frac{F}{G}\partial_\tau G\, f_\infty\, dv + \int_I \left(\frac{F}{G}-1\right)^2 \partial_\tau G\, f_\infty\, dv = \\
&2\int_I \left(\frac{F}{G}-1\right)\partial_\tau F\, f_\infty\, dv + \\
&\int_I \left(\left(\frac{F}{G}-1\right)^2 - 2\left(\frac{F}{G}-1\right)\frac{F}{G} + \left(\frac{F}{G}\right)^2 - \left(\frac{F}{G}\right)^2\right)\partial_\tau G\, f_\infty\, dv = \\
&2\int_I \left(\frac{F}{G}-1\right)\partial_\tau F\, f_\infty\, dv + \int_I \left(1-\left(\frac{F}{G}\right)^2\right)\partial_\tau G\, f_\infty\, dv.
\end{aligned}$$

Finally, by Proposition 3.1 we get

$$\begin{aligned}
&2\int_I \left(\frac{F}{G}-1\right)\partial_\tau F\, f_\infty\, dv + \int_I \left(1-\left(\frac{F}{G}\right)^2\right)\partial_\tau G\, f_\infty\, dv = \\
&-2\int_I \kappa(v) f_\infty \frac{\partial}{\partial v}\left(\frac{F}{G}\right)\partial_v F\, dv + 2\int_I \kappa(v) f_\infty \partial_v G\left(\frac{F}{G}\right)\frac{\partial}{\partial v}\left(\frac{F}{G}\right)\, dv = \\
&-2\int_I \kappa(v) f_\infty \frac{\partial}{\partial v}\left(\frac{F}{G}\right)\left(\partial_v F - \partial_v G\left(\frac{F}{G}\right)\right)\, dv = \\
&-2\int_I \kappa(v) f_\infty G\left(\frac{\partial}{\partial v}\left(\frac{F}{G}\right)\right)^2\, dv = -2\int_I \kappa(v) f\left(\frac{\partial}{\partial v}\left(\frac{f}{g}\right)\right)^2\, dv.
\end{aligned}$$

□

4. Open problems

The mathematical analysis of Sections 3.1–3.3 outlined the importance of Lyapunov functionals in the study of the large-time behavior of the solution to the Fokker–Planck equation (3.1). Also, the presence of the weight κ, closely related to the shape of the equilibrium density f_∞, has been at the basis of the generalization of Chernoff inequality considered in Section 3.2. This establishes a deep connection among the Fokker–Planck equation (3.1), its steady state and various relative entropy functionals and relative entropy productions. Among others, this investigation enlightened the leading role of Shannon relative entropy between two solutions of the Fokker–Planck equation (3.1)

$$H(f,g) = \int_I f(v) \log \frac{f(v)}{g(v)} \, dv,$$

its relative weighted Fisher Information

$$I_\kappa(f,g) = \int_I \kappa(v) f(v) \left(\frac{\partial_v f(v)}{f(v)} - \frac{\partial_v g(v)}{g(v)} \right)^2 \, dv$$

and Hellinger distance

$$d_H(f,g) = \left(\int_I \left(\sqrt{f(v)} - \sqrt{g(v)} \right)^2 dv \right)^{\frac{1}{2}}.$$

Hence, the Fokker–Planck equation (3.1) appears to be a useful vehicle to establish connections between Shannon entropy of a probability density function f relative to a particular density f_∞ (the stationary solution) with the weighted Fisher information of f relative to f_∞. Then, in view of the Chernoff-type inequality (3.40) a link between the weighted Fisher information and the Hellinger distance follows.

However, while the Hellinger distance bounds from above the usual L^1 distance between probability densities, it does not bound from above the relative Shannon entropy. Consequently, in contrast with the well-known case of the standard Fokker–Planck equation, briefly discussed in Remark 3.7, it is not possible to conclude with the exponential convergence to equilibrium in relative Shannon entropy. This suggests to look for relative entropy functionals which are equivalent to Hellinger distance.

In the pertinent literature, the link between some relative entropies and the Hellinger distance has been known for many years now, and it has been object of several studies in the framework of information theory.[73] It is not surprising that the Fokker–Planck equation (3.1) represents a new interesting approach to study relationships among relative entropies introduced so far to generalize the concept of relative Shannon entropy, their relative entropy production and Hellinger distance.

In particular, a class of relative entropies, known as symmetric Jensen–Shannon relative entropies,[73,88,93] seems useful to be studied in this context.

Definition 4.1. Given a pair of probability densities f and g, taking values on $I \subseteq \mathbb{R}$, and $\alpha \in [0,1]$ the symmetric Jensen–Shannon relative entropy is defined by

$$K_\alpha(f,g) = H(f, \alpha f + (1-\alpha)g) + H(g, \alpha g + (1-\alpha) f) = \int_I f(v) \log \frac{f(v)}{\alpha f(v) + (1-\alpha) g(v)} \, dv + \int_I g(v) \log \frac{g(v)}{\alpha g(v) + (1-\alpha) f(v)} \, dv. \tag{4.1}$$

We remind that a non symmetrized version of (4.1) was considered in (3.26). Clearly

$$K_0(f,g) = H(f,g) + H(g,f)$$

the symmetric Shannon relative entropy, while

$$K_1(f,g) = 0.$$

Thanks to the linearity of the Fokker–Planck equation (3.1), any linear combination $\alpha f(\tau) + (1-\alpha)g(\tau)$ with $\alpha \in [0,1]$ of two different solutions $f(\tau)$ and $g(\tau)$ is a solution as well. Consequently

$$\frac{d}{d\tau}K_\alpha(f(\tau), f_\infty) = -I_\kappa(f(\tau), \alpha f(\tau)+(1-\alpha)f_\infty) - I_\kappa(f_\infty, \alpha f_\infty + (1-\alpha)f(\tau)). \quad (4.2)$$

The interesting feature of the symmetric Jensen–Shannon relative entropies K_α is that, for any given $\alpha \neq 0$, differently from the relative Shannon entropy, they are equivalent to the square of the Hellinger distance. This result is easy to obtain.

Lemma 4.1. *For any $\alpha \in (0,1)$, and a pair of probability densities f and g there are positive constants c_α and C_α such that*

$$c_\alpha \left(d_H(f,g)\right)^2 \le K_\alpha(f,g) \le C_\alpha \left(d_H(f,g)\right)^2.$$

Proof. Let us consider probability densities f and g which allow for rigorous computations. The general result then follows by a density argument. To make the proof easier to read, let us drop again the v variable inside the integrals. For $\beta \in (0,1)$ we obtain

$$\frac{d}{d\beta}K_\beta(f,g) = -\int_I \frac{f(f-g)}{\beta f + (1-\beta)g}\,dv - \int_I \frac{g(g-f)}{\beta g + (1-\beta)f}\,dv =$$

$$-\int_I \frac{(f-g)(1-\beta)(f^2-g^2)}{(\beta f + (1-\beta)g)(\beta g + (1-\beta)f)}\,dv =$$

$$-(1-\beta)\int_I \frac{(f-g)^2}{f+g}\frac{(f+g)^2}{\beta(1-\beta)(f^2+g^2) + (\beta^2+(1-\beta)^2)fg}\,dv.$$

Now, since

$$\beta^2 + (1-\beta)^2 \ge 2\beta(1-\beta),$$

we obtain the bounds

$$\begin{aligned} &\frac{2}{\beta^2+(1-\beta)^2}\frac{1}{(f+g)^2} \le \\ &\frac{1}{\beta(1-\beta)(f^2+g^2)+(\beta^2+(1-\beta)^2)fg} \le \frac{1}{\beta(1-\beta)}\frac{1}{(f+g)^2}, \end{aligned} \quad (4.3)$$

which implies

$$-\frac{1}{\beta}\int_I \frac{(f-g)^2}{f+g}\,dv \le \frac{d}{d\beta}K_\beta(f,g) \le -\frac{2(1-\beta)}{\beta^2+(1-\beta)^2}\int_I \frac{(f-g)^2}{f+g}\,dv.$$

Since $K_1(f,g) = 0$, integrating (4.3) between $\alpha \in (0,1)$ and 1 we get

$$c_\alpha \int_I \frac{(f-g)^2}{f+g}\, dv \le K_\alpha(f,g) \le C_\alpha \int_I \frac{(f-g)^2}{f+g}\, dv,$$

where

$$c_\alpha = \int_\alpha^1 \frac{1}{\beta}\, d\beta = \log\frac{1}{\alpha} > 0$$
$$C_\alpha = \int_\alpha^1 \frac{2(1-\beta)}{\beta^2 + (1-\beta)^2}\, d\beta > 0.$$

On the other hand, since

$$(d_H(f,g))^2 = \int_I \left(\sqrt{f} - \sqrt{g}\right)^2\, dv = \int_I \frac{(f-g)^2}{(\sqrt{f}+\sqrt{g})^2}\, dv,$$

one has the bounds

$$\frac{1}{2}\int_I \frac{(f-g)^2}{f+g}\, dv \le \int_I \frac{(f-g)^2}{(\sqrt{f}+\sqrt{g})^2}\, dv \le \int_I \frac{(f-g)^2}{f+g}\, dv,$$

and the proof is completed. □

In reason of Lemma 4.1, each symmetric Jensen–Shannon relative entropy (4.1) appears to be a good candidate to replace the classical relative Shannon entropy in order to eventually achieve exponential convergence in Jensen–Shannon relative entropy (or in the equivalent Hellinger distance) to the steady state of the Fokker–Planck equation (3.1). Indeed, due to (4.2), this convergence would follow any time one has the bound

$$I_\kappa(f(\tau), \alpha f(\tau) + (1-\alpha)f_\infty) + I_\kappa(f_\infty, \alpha f_\infty + (1-\alpha)f(\tau)) \\ \ge C(\alpha) d_H(f(\tau), \alpha f(\tau) + (1-\alpha)f_\infty)^2$$

which appears as a generalization of (3.38). Recalling the definition of the relative Fisher information (3.15) we have

$$\begin{aligned} &I_k(f(\tau), \alpha f(\tau) + (1-\alpha)f_\infty) + I_k(f_\infty, \alpha f_\infty + (1-\alpha)f(\tau)) = \\ &4\int_I \kappa(v)(\alpha f(v,\tau) + (1-\alpha)f_\infty)\left(\frac{\partial}{\partial v}\sqrt{\frac{f(v,\tau)}{\alpha f(v,\tau) + (1-\alpha)f_\infty(v)}}\right)^2\, dv + \\ &\int_I \kappa(v) f_\infty(v)\left(\frac{\partial}{\partial v}\log\frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha)f(v,\tau)}\right)^2\, dv. \end{aligned} \tag{4.4}$$

While both terms are of the type

$$\int_I \kappa(v) g(v,\tau)\left(\frac{\partial\phi(v,\tau)}{\partial v}\right)^2\, dv,$$

in the first one g is not the stationary state but a convex combination of f_∞ and $f(\tau)$, and the Chernoff-type inequality (3.40) is not directly applicable. In the second one it is possible to apply Chernoff inequality (3.40) to get

$$\int_I \kappa(v) f_\infty(v) \left(\frac{\partial}{\partial v} \log \frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha) f(v,\tau)} \right)^2 \, dv \geq$$
$$\int_I \left(\log \frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha) f(v,\tau)} + \right.$$
$$\left. - \left(\int_I \left(\log \frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha) f(v,\tau)} \right) f_\infty(v) \, dv \right) \right)^2 f_\infty(v) \, dv =$$
$$\int_I \left(\log \frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha) f(v,\tau)} \right)^2 f_\infty(v) \, dv +$$
$$- \left(\int_I \left(\log \frac{f_\infty(v)}{\alpha f_\infty(v) + (1-\alpha) f(v,\tau)} \right) f_\infty(v) \, dv \right)^2 .$$

Also in this case it is not clear how to relate this term with the Hellinger distance between f_∞ and $\alpha f_\infty + (1-\alpha) f(\tau)$. However, the question of finding lower bounds for the weighted entropy production given in (4.4) remains an open interesting question.

5. Conclusions

The mathematical modelling of various social and economic aspects of the modern society, in reason of the huge number of the population involved, is fruitfully described by resorting to statistical mechanics. Among others, kinetic theory of multi-agent systems became in the last decade a leading modelling approach to understand the formation of emerging phenomena, consequent to microscopic interactions between agents.

The possibility to obtain explicitly the expression of the underlying equilibrium densities is mainly based on the possibility to extract from kinetic models of Boltzmann type their mean field description, in the form of Fokker–Planck like equations.

In the present paper, we introduced and discussed various theoretical aspects of these Fokker–Planck equations, which, in contrast with the classical one coming from physics, are characterized by the presence both of a variable diffusion coefficient and of boundaries. As discussed in this paper, the study of the large-time behavior of the solution to these equations introduces a number of challenging problems, with intersections with classical Shannon's information theory and probability.

Acknowledgement

This work has been written within the activities of GNFM and GNAMPA groups of INdAM (National Institute of High Mathematics), and partially supported by MIUR project "Optimal mass transportation, geometrical and functional inequalities with applications".

References

1. G. Ajmone Marsan, N. Bellomo, L. Gibelli, Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics *Math. Mod. Meth. Appl. Scie.* **26**, (6) (2016) 1051–1093.
2. A. Arnold, P. Markowich, G. Toscani, A. Unterreiter, On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations, *Commun. Partial Diff. Equa.* **26** (2001), 43–100.
3. J. Angle, The surplus theory of social stratification and the size distribution of personal wealth. *Social Forces* **65** (1986) 293–326.
4. J. Angle, The inequality process as a wealth maximizing process. *Physica A* **367** (2006) 388–414.
5. F. Bassetti and G. Toscani. Explicit equilibria in a kinetic model of gambling. *Phys. Rev. E* **81** (2010) 066115.
6. F. Bassetti and G. Toscani. Explicit equilibria in bilinear kinetic models for socio-economic interactions. *ESAIM: Proc. and Surveys* **47** (2014) 1–16.
7. F. Bassetti and G. Toscani. Mean field dynamics of collisional processes with duplication, loss and copy, *Math. Mod. Meth. Appl. Scie.* **25** (10) (2015) 1887–1925.
8. N. Bellomo, D. Knopoff and J. Soler. On the difficult interplay between life, "complexity", and mathematical sciences. *Math. Models Methods Appl. Sci.* **23** (2013) 1861–1913.
9. N. Bellomo, N.K. Li and P.K. Maini. On the foundations of cancer modelling: selected topics, speculations, and perspectives. *Math. Models Methods Appl. Sci.* **18** (2008) 593–646.
10. N. Bellomo and J. Soler. On the mathematical theory of the dynamics of swarms viewed as complex systems. *Math. Models Methods Appl. Sci.* Suppl. **22** (2012) 1140006 (29 pages).
11. A. Bellouquid, E. De Angelis and D. Knopoff. From the modelling of the immune hallmarks of cancer to a black swan in biology. *Math. Models Methods Appl. Sci.* **23** (2013) 949–978.
12. E. Ben-Naim, P. L. Krapivski and S. Redner, Bifurcations and patterns in compromise processes. *Physica D* **183**, (2003) 190–204.
13. E. Ben-Naim, P. L. Krapivski, R. Vazquez and S. Redner, Unity and discord in opinion dynamics. *Physica A* **330**, (2003) 99–106.
14. E. Ben-Naim, Opinion dynamics: rise and fall of political parties. *Europhys. Lett.* **69**, (2005) 671–677.
15. M.L. Bertotti and M. Delitala, On a discrete generalized kinetic approach for modelling persuader's influence in opinion formation processes. *Math. Comp. Model.* **48**, (2008) 1107Â-1121.
16. M. Bisi, G. Spiga and G. Toscani, Kinetic models of conservative economies with wealth redistribution. *Commun. Math. Sci.* **7** (4) 901–916 (2009)
17. G. Bobkov and M. Ledoux, Weighted Poincaré-type Inequalities for Cauchy and other convex measures. *Ann. Probability* **37** (2) (2009) 403–427
18. J.F. Bouchaud and M. Mézard, Wealth condensation in a simple model of economy. *Physica A*, **282** (2000) 536–545.
19. L. Boudin and F. Salvarani, The quasi-invariant limit for a kinetic model of sociological collective behavior. *Kinetic Rel. Mod.* **2**, (2009) 433–449.
20. L. Boudin and F. Salvarani, A kinetic approach to the study of opinion formation. *ESAIM: Math. Mod. Num. Anal.* **43**, (2009) 507–522.
21. L. Boudin, A. Mercier and F. Salvarani, Conciliatory and contradictory dynamics in opinion formation. *Physica A* **391**, (2012) 5672Â–5684.

22. C. Brugna and G. Toscani, Kinetic models of opinion formation in the presence of personal conviction, *Phys. Rev. E* **92**, (2015) 052818.
23. M. Burger, L. Caffarelli, P.A. Markowich and M-T. Wolfram. On a Boltzmann-type price formation model. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **469** (2013).
24. M. Burger, L. Caffarelli, P.A. Markowich and M-T. Wolfram. On the asymptotic behavior of a Boltzmann-type price formation model. *Commun. Math. Sci.* **12** (2014) 1353–1361.
25. M.J. Cáceres and G. Toscani. Kinetic approach to long time behavior of linearized fast diffusion equations. *J. Stat. Phys.* **128** (2007) 883–925.
26. J.A. Carrillo, M. Fornasier, J. Rosado and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM J. Math. Anal.* **42** (2010) 218–236.
27. J.A. Carrillo, M. Fornasier, G. Toscani and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modelling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Inc., Boston, MA, 2010.
28. C. Castellano, S. Fortunato and V. Loreto, Statistical physics of social dynamics. *Rev. Mod. Phys.* **81** (2009) 591–646.
29. A. Chakraborti, Distributions of money in models of market economy. *Int. J. Modern Phys. C* **13**, (2002) 1315–1321.
30. A. Chakraborti and B.K. Chakrabarti, Statistical Mechanics of Money: Effects of Saving Propensity. *Eur. Phys. J. B* **17**, (2000) 167–170.
31. A. Chatterjee, B.K. Chakrabarti and S.S. Manna, Pareto law in a kinetic model of market with random saving propensity. *Physica A* **335** (2004), 155–163.
32. A. Chatterjee, B.K. Chakrabarti and R.B. Stinchcombe, Master equation for a kinetic model of trading market and its analytic solution. *Phys. Rev. E* **72**, (2005) 026126.
33. H. Chernoff, A note on an inequality involving the normal distribution. *Ann. Probab.* **9** (3) (1981) 533–535.
34. V. Comincioli, L. Della Croce and G. Toscani, A Boltzmann-like equation for choice formation. *Kinetic Rel. Mod.* **2**, (2009) 135–149.
35. S. Cordier, L. Pareschi and C. Piatecki, Mesoscopic modelling of financial markets. *J. Stat. Phys.* **134** (1), (2009) 161–184.
36. S. Cordier, L. Pareschi and G. Toscani, On a kinetic model for a simple market economy. *J. Stat. Phys.* **120**, (2005) 253-277.
37. I. Csiszár, Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis von Markoffschen Ketten. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8**, (1963) 85–108.
38. F. Cucker and E. Mordecki, Flocking in noisy environments. *J. Math. Pures Appl.* **89** (2008) 278–296.
39. F. Cucker and S. Smale, Emergent behavior in flocks. *IEEE Trans. Automat. Control* **52** (2007) 852–862.
40. F. Cucker and S. Smale. On the mathematics of emergence. *Jpn. J. Math.* **2** (2007) 197–227.
41. G. Deffuant, F. Amblard, G. Weisbuch and T. Faure, How can extremism prevail? A study based on the relative agreement interaction model. *J. Art. Soc. Soc. Sim.* **5**, (2002) 1.
42. A. Drăgulescu and V.M. Yakovenko, Statistical mechanics of money. *Eur. Phys. Jour. B* **17**, (2000) 723-729.
43. B. Düring, D. Matthes and G. Toscani, Kinetic Equations modelling Wealth Redistribution: A comparison of Approaches. *Phys. Rev. E*, **78**, (2008) 056103.
44. B. Düring, D. Matthes and G. Toscani, A Boltzmann-type approach to the formation

of wealth distribution curves. (Notes of the Porto Ercole School, June 2008) *Riv. Mat. Univ. Parma* (1) **8** (2009) 199–261.
45. B. Düring and G. Toscani, International and domestic trading and wealth distribution. *Commun. Math. Sci.*, **6**/4: (2008) 1043–1058.
46. B. Düring, P.A. Markowich, J-F. Pietschmann and M-T. Wolfram. Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **465** (2009) 3687–3708.
47. U. Eco, Memoria e dimenticanza. (2011), in http://www.3dnews.it/node/1824.
48. W. Feller, Two singular diffusion problems. *Ann. Math.* **54** (2) (1951) 173–182.
49. W. Feller, *An introduction to probability theory and its applications. Vol. I.* (John Wiley & Sons Inc. 1968).
50. G. Furioli, A. Pulvirenti, E. Terraneo and G. Toscani. The grazing collision limit of the inelastic Kac model around a Lévy-type equilibrium. *SIAM J. Math. Anal.* **44** (2012) 827–850.
51. E. Gabetta and E. Regazzini, About the gene families size distribution in a recent model of genome evolution. *Math. Models Methods Appl. Sci.* **20** (2010) 1005–1020.
52. S. Galam, Y. Gefen and Y. Shapir, Sociophysics: A new approach of sociological collective behavior. I. Mean-behaviour description of a strike *J. Math. Sociology* **9**, (1982) 1–13.
53. S. Galam and S. Moscovici, Towards a theory of collective phenomena: consensus and attitude changes in groups. *Euro. J. Social Psychology* **21**, (1991) 49–74.
54. S. Galam, Rational group decision making: A random field Ising model at $T = 0$. *Physica A* **238**, (1997) 66–80.
55. S. Galam and J.D. Zucker, From individual choice to group decision-making. *Physica A* **287**, (2000) 644–659.
56. U. Garibaldi, E. Scalas and P. Viarengo, Statistical equilibrium in simple exchange games II. The redistribution game. *Eur. Phys. Jour. B* **60**(2) (2007) 241–246.
57. J.C. Gonzalez–Avella, V.M. Eguiluz, M. Marsili, F. Vega–Redondo and M. San Miguel, Threshold learning dynamics in social networks. *PLoS ONE* **6** (5) (2011) e20207.
58. J. Grilli, M. Romano, F. Bassetti and M. Cosentino Lagomarsino, Cross-species gene-family fluctuations reveal the dynamics of horizontal transfers. *Nucleic Acids Research* **42** (2014) 6850–60.
59. A.K. Gupta, Models of wealth distributions: a perspective. In *Econophysics and sociophysics: trends and perspectives* B.K. Chakrabarti, A. Chakraborti, A. Chatterjee (Eds.) Wiley VHC, Weinheim (2006) 161–190.
60. S-Y. Ha, E. Jeong, J-H. Kang and K. Kang, Emergence of multi-cluster configurations from attractive and repulsive interactions. *Math. Models Methods Appl. Sci.* **22** (2012) 1250013 (42 pages).
61. S-Y. Ha, M-J. Kang and B. Kwon, A hydrodynamic model for the interaction of Cucker-Smale particles and incompressible fluid. *Math. Models Methods Appl. Sci.* **24** (2014) 2311–2359.
62. S-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models* **1** (2008) 415–435.
63. B. Hayes, Follow the money. *American Scientist* **90** (5), (2002) 400–405.
64. P. Herd, K. Holden and Y.T. Su, The links between early-life cognition and schooling and late-life financial knowledge. *The Journal of Consumer Affairs*, Fall 2012, (2012) 411–435
65. M. A. Huynen and E. van Nimwegen, The frequency distribution of gene family size in complete genomes. *Mol. Biol. Evol.* **15** (1998) 583–589.
66. I. A. Ibragimov and M. ÃÂ. Rradavichyus, On large deviation probabilities for max-

imum likelihood estimators. (Russian) *Dokl. Akad. Nauk SSSR* **5** (1981), 1048–1052
67. S. Ispolatov, P.L. Krapivsky and S. Redner, Wealth distributions in asset exchange models. *Eur. Phys. Jour. B* **2**, (1998) 267–276.
68. O. Johnson and A. Barron, Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields* **129** (3) (2004) 391–409.
69. G. P. Karev, Y. I. Wolf, A. Y. Rzhetsky, F. S. Berezovskaya and E. V. Koonin, Birth and death of protein domains: A simple model of evolution explains power law behaviour. *BMC Evol. Biol.* **2** (2002) 1–26.
70. E. Kashdan and L. Pareschi, Mean field mutation dynamics and the continuous Luria-Delbrück distribution. *Math. Biosci.* **240** (2012) 223–230.
71. C. A. Klaassen, On an Inequality of Chernoff. *Ann. Probability* **13** (3) (1985) 966–974.
72. S. Kullback, *Information Theory and Statistics.* (John Wiley, 1959).
73. J. Lin, Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theo.* **37**, (1991) 145–151.
74. D.E. Lea and C.A. Coulson. The distribution of the numbers of mutants in bacterial populations. *Journal of genetics* **49** (1949) 264–285.
75. C. Le Bris and P.L. Lions Existence and Uniqueness of Solutions to Fokker–ÂPlanck Type Equations with Irregular Coefficients, *Communications in Partial Differential Equations* **33** (7) (2008) 1272–1317.
76. S.E. Luria and M. Delbrück, Interference between inactivated bacterial virus and active virus of the same strain and of a different strain. *Arch. Biochem* (1942) 207–218.
77. S.E. Luria and M. Delbrück, Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* (1943) 491-511.
78. D. Maldarella and L. Pareschi, Kinetic models for socio–economic dynamics of speculative markets, *Physica A*, **391** (2012) 715–730 .
79. A.C.R. Martins and S. Galam, Building up of individual inflexibility in opinion dynamics. *Phys. Rev. E*, **87** (2013) 042807.
80. D. Matthes, A. Juengel and G.Toscani, Convex Sobolev inequalities derived from entropy dissipation. *Arch. Rat. Mech. Anal.* **199** (2) (2011) 563–596.
81. D. Matthes and G. Toscani, On steady distributions of kinetic models of conservative economies. *J. Stat. Phys.* **130** (2008) 1087–1117.
82. G. Naldi, L. Pareschi and G. Toscani, *Mathematical modeling of collective behavior in socio-economic and life sciences.* (Springer Verlag, Heidelberg, 2010).
83. L. Pareschi, G. Toscani, Wealth distribution and collective knowledge. A Boltzmann approach, *Phil. Trans. R. Soc. A* **372**, 20130396, 6 October (2014).
84. L. Pareschi and G. Toscani, *Interacting multiagent systems. Kinetic equations & Monte Carlo methods.* (Oxford University Press, Oxford, 2013).
85. V. Pareto, *Cours d'Économie Politique.* Lausanne and Paris (1897).
86. A. Pulvirenti, G. Toscani, Asymptotic properties of the inelastic Kac model, *J. Statist. Phys.*, **114** (2004) 1453–1480
87. R. Rudnicki and J. Tiuryn, Size distribution of the gene families in a genome. *Math. Models Methods Appl. Sci.* **24** (2014) 697–717.
88. P. Sánchez Moreno, A. Zarzo and J. S. Dehesa, Jensen divergence based on Fisher's information. *J. Phys. A: Math. Theor.* **45** (12) (2012) 125305
89. E. Scalas, U. Garibaldi and S. Donadio, Statistical equilibrium in the simple exchange games I. Methods of solution and application to the Bennati–Dragulescu–Yakovenko (BDY) game. *Eur. Phys. J. B* **53** (2006) 267–272.
90. F. Slanina, Inelastically scattering particles and wealth distribution in an open eco-

nomy. *Phys. Rev. E* **69**, (2004) 046102.
91. K. Sznajd–Weron and J. Sznajd, Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11**, (2000) 1157–1165.
92. R.C. Teevan and R.C. Birney, *Readings for Introductory Psychology* Harcourt, Brace & World, New York, (1965).
93. F. Topsøe, Some Inequalities for Information Divergence and Related Measures of Discrimination. *IEEE Trans. Inf. Theory* **46** (4) (2000) 1602–1609.
94. M. Torregrossa and G. Toscani, On a Fokker–Planck equation for wealth distribution, (preprint) (2016)
95. G. Toscani, Entropy dissipation and the rate of convergence to equilibrium for the Fokker-Planck equation, *Quart. Appl. Math.*, **LVII** (1999), 521–541
96. G. Toscani, Kinetic models of opinion formation. *Commun. Math. Sci.* **4** (2006) 481–496.
97. G. Toscani, An information-theoretic proof of Nash's inequality, *Rend. Lincei Mat. Appl.*, **24** (2013) 83–93
98. G. Toscani, C. Brugna and S. Demichelis, Kinetic models for the trading of goods. *J. Stat. Phys.* **151** 549–566 (2013).
99. G. Toscani, A kinetic description of mutation processes in bacteria. *Kinet. Relat. Models* **6** (2013) 1043–1055.
100. C. Villani, *Contribution à l'étude mathématique des équations de Boltzmann et de Landau en théorie cinétique des gaz et des plasmas.* PhD Thesis, Univ. Paris-Dauphine (1998).
101. C. Villani, On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rational Mech. Anal*, **143** (1998) 273–307.
102. C. Villani, A Review of Mathematical Topics in Collisional Kinetic Theory, in S. Friedlander and D. Serre (eds), *Handbook of Mathematical Fluid Dynamics*, Vol. 1, New York: Elsevier (2002)
103. V. M. Zolotarev, Properties and relations of certain types of metrics. (Russian. English summary) Studies in mathematical statistics, 3. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **87** (1979), 18–35, 206, 212.