



**HAL**  
open science

# Mean Field Games: Convergence of a Finite Difference Method

Yves Achdou, Fabio Camilli, Italo Capuzzo-Dolcetta

► **To cite this version:**

Yves Achdou, Fabio Camilli, Italo Capuzzo-Dolcetta. Mean Field Games: Convergence of a Finite Difference Method. SIAM Journal on Numerical Analysis, 2013, 51 (5), pp.2585 - 2612. 10.1137/120882421 . hal-01456506

**HAL Id: hal-01456506**

**<https://hal.science/hal-01456506>**

Submitted on 5 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MEAN FIELD GAMES: CONVERGENCE OF A FINITE DIFFERENCE METHOD

YVES ACHDOU <sup>\*</sup>, FABIO CAMILLI <sup>†</sup>, AND ITALO CAPUZZO-DOLCETTA <sup>‡</sup>

**Abstract.** Mean field type models describing the limiting behavior of stochastic differential games as the number of players tends to  $+\infty$ , have been recently introduced by J-M. Lasry and P-L. Lions. Numerical methods for the approximation of the stationary and evolutive versions of such models have been proposed by the authors in previous works. Here, convergence theorems for these methods are proved under various assumptions on the coupling operator.

**Key words.** Mean field games, finite difference schemes, convergence.

**AMS subject classifications.** 65M06,65M012,9108,91A23,49L25

**1. Introduction.** Mean field type models describing the asymptotic behavior of stochastic differential games (Nash equilibria) as the number of players tends to  $+\infty$  have recently been introduced by J-M. Lasry and P-L. Lions [19, 20, 21], and termed *mean field games* by the same authors. For brevity, the acronym *MFG* will sometimes be used for *mean field games*. Examples of MFG models with applications in economics and social sciences are proposed in [16]. Many important aspects of the mathematical theory developed by J-M. Lasry and P-L. Lions on MFG are not published in journals or books, but can be found in the videos of the lectures of P-L. Lions (in French) at Collège de France: see [23]. A very good introduction is also given in the notes by P. Cardaliaguet, [8], with a special emphasis on the deterministic case, i.e.  $\nu = 0$  in (1.1)-(1.2) below. Related ideas have been developed independently in the engineering literature by Huang-Caines-Malhamé, see for example [17].

MFG may lead to systems of evolutive partial differential equations involving two unknown scalar functions: the density of the agents in a given state  $x \in \mathbb{R}^d$ , namely  $m = m(t, x)$  and the value function  $u = u(t, x)$ . The present work is devoted to finite difference schemes for the systems of partial differential equations. Although the methods and the theoretical results obtained below can be easily generalized, the present work focuses on the two-dimensional case for the following reasons: 1) the one dimensional case is easier and allows too special arguments; 2) in dimension two, the description of the discrete methods discussed below remain fairly simple. Besides, several important applications of the mean field games theory are two-dimensional, in particular those related to crowd motion, see [3].

In the state-periodic setting, typical MFG model comprises the following system of partial differential equations in  $(0, T) \times \mathbb{T}^2$

$$\frac{\partial u}{\partial t}(t, x) - \nu \Delta u(t, x) + H(x, \nabla u(t, x)) = \Phi[m(t, \cdot)](x), \quad (1.1)$$

$$\frac{\partial m}{\partial t}(t, x) + \nu \Delta m(t, x) + \operatorname{div} \left( m(t, \cdot) \frac{\partial H}{\partial p}(\cdot, \nabla u(t, \cdot)) \right) (x) = 0, \quad (1.2)$$

with the initial and terminal conditions

$$u(0, x) = u_0(x), \quad m(T, x) = m_T(x), \quad \text{in } \mathbb{T}^2, \quad (1.3)$$

---

<sup>\*</sup> Univ. Paris Diderot, Sorbonne Paris Cité, Laboratoire Jacques-Louis Lions, UMR 7598, UPMC, CNRS, F-75205 Paris, France. achdou@ljl-univ-paris-diderot.fr

<sup>†</sup> "Sapienza", Università di Roma, Dipartimento di Scienze di Base e Applicate per l'Ingegneria, 00161 Roma, Italy, camilli@dmmm.uniroma1.it

<sup>‡</sup>"Sapienza", Università di Roma, Dipartimento di Matematica, 00185 Roma, Italy, capuzzo@mat.uniroma1.it

given a cost function  $u_0$  and a probability density  $m_T$ . Here, we denote by  $\mathbb{T}^2 = [0, 1]^2$  the 2-dimensional unit torus, and by  $\Delta$ ,  $\nabla$  and  $\text{div}$ , respectively, the Laplace, the gradient and the divergence operator acting on the state variable  $x$ . The parameter  $\nu$  is the diffusion coefficient. Hereafter, we will always assume that  $\nu > 0$ . The system also involves the scalar Hamiltonian  $H(x, p)$ , which is assumed to be convex with respect to  $p$  and  $\mathcal{C}^1$  regular w.r.t.  $x$  and  $p$ . The notation  $\frac{\partial H}{\partial p}(x, q)$  is used for the gradient of  $p \mapsto H(x, p)$  at  $p = q$ . Finally, in the term  $\Phi[m(t, \cdot)](x)$ ,  $\Phi$  may be

- either a local operator, i.e.  $\Phi[m(t, \cdot)](x) = F(m(t, x))$  where  $F$  is a  $\mathcal{C}^1$  regular function defined on  $\mathbb{R}_+$ . In this case, there are existence theorems of either classical (see [9]) or weak solutions (see [20]), under suitable assumptions on the data,  $H$  and  $F$ .
- or a non local operator which continuously maps the set of probability measures on  $\mathbb{T}^2$  (endowed with the weak \* topology) to a bounded subset of  $Lip(\mathbb{T}^2)$ , the Lipschitz functions on  $\mathbb{T}^2$ , and for example maps continuously  $C^{k, \alpha}(\mathbb{T}^2)$  to  $C^{k+1, \alpha}(\mathbb{T}^2)$ , for all  $k \in \mathbb{N}$  and  $0 \leq \alpha < 1$ . In this case, classical solutions of (1.1)-(1.3) are shown to exist under natural assumptions on the data and some technical assumptions on  $H$ , see [20, 21].

System (1.1)-(1.2) consists then of a forward Bellman equation coupled with a backward Fokker-Planck equation. The forward-backward structure is an important feature of this system, which makes it necessary to design new strategies for its mathematical analysis (see [20, 21]) and for numerical approximation.

If the Hamiltonian is of the form

$$H(x, \nabla u) = \sup_{\gamma} [\gamma \cdot \nabla u - L(x, \gamma)], \quad (1.4)$$

and if  $u$  and  $m$  solve (1.1)-(1.3), then Dynamic Programming arguments, see Bardi-Capuzzo Dolcetta [5], Fleming- Soner [12], show that  $u$  is the value function of an optimal control problem for the controlled dynamics defined on  $\mathbb{T}^2$  by

$$dX_s = -\gamma_s ds + \sqrt{2\nu} dW_s,$$

(  $(W_s)$  is a Brownian motion and the time variable  $s$  is linked to the variable  $t$  in (1.1)-(1.3) by the relation  $t = T - s$ , i.e.  $t$  is the time to the horizon), running cost density  $L(X_s, \gamma_s) + \Phi[m(s, \cdot)](X_s)$  depending on the position  $X_s$ , the control  $\gamma_s$  and the probability density  $m(s, \cdot)$ . The choice of a nonlocal operator  $\Phi$  is justified if a given agent in the state  $x$  is not only sensitive to  $m(x)$ , but also to a global information on the density in a given neighborhood of  $x$ , for example the number of agents in the latter. On the other hand, (1.2) is a backward Fokker-Planck equation with velocity field  $\frac{\partial H}{\partial p}(x, \nabla u)$  depending on the value function itself. Since  $t$  in (1.1)-(1.3) is the remaining time to the horizon,  $u_0$  corresponds to the final cost or incitation of the optimal control problem whereas  $m_T$  is the density of the agents at the beginning of the game. The only reason for us to use the variable  $t$  instead of  $s$  in the partial differential equations is to keep the formalism of the pioneering articles [19, 20, 21].

We have chosen to focus on the case when the cost  $u_0$  depends directly on  $x$ . In some realistic situations, the final cost may depend on the density of the players, i.e.  $u|_{t=0} = \Phi_0[m|_{t=0}](x)$ , where  $\Phi_0$  is an operator acting on probability densities, which may be local or not. This case can be handled by the methods presented below with no additional difficulty, and the same kind of convergence results as those given below can be obtained with additional assumptions on  $\Phi_0$  such as monotonicity (see (1.5) below). Yet, we will not tackle this aspect, in order to keep the discussion as

simple as possible.

By working on the torus  $\mathbb{T}^2$ , we avoid the discussion of the boundary conditions, but other boundary value problems can be considered, for example

- Neumann conditions on  $u$  and  $m$ : for example, in the context of crowd motion, see [3], such conditions are relevant if the domain where the crowd moves is limited by walls. The results of the present paper hold when there are homogeneous Neumann boundary conditions instead of periodic boundary conditions
- Dirichlet conditions on  $u$  and  $m$ : in the example of crowd motion, Dirichlet conditions are relevant if the crowd exits a room and heads toward a much larger space: at the exit, it is sensible to assume that the density  $m$  vanishes and that the value function is given (exit cost, which may be zero)
- In the deterministic case  $\nu = 0$ , other boundary conditions can be used, for example when the state of the players is constrained to remain in the domain. In this work, we focus on  $\nu > 0$ , because the analysis of the deterministic case requires quite different arguments.

A very important feature of the mean field models above is that uniqueness and stability may be obtained under reasonable assumptions, see [19, 20, 21], in contrast with the Nash system describing the individual behavior of each player, for which uniqueness hardly occurs. To be more precise, a sufficient condition for the uniqueness of a solution of (1.1)-(1.3) is that  $\Phi$  is monotone in the sense that for all probability measures  $m$  and  $\tilde{m}$  on  $\mathbb{T}^2$ ,

$$\int (\Phi[m](x) - \Phi[\tilde{m}](x))(dm(x) - d\tilde{m}(x)) \leq 0 \Rightarrow m = \tilde{m}. \quad (1.5)$$

In fact, a weaker condition than (1.5) is sufficient for uniqueness, namely that

$$\int (\Phi[m](x) - \Phi[\tilde{m}](x))(dm(x) - d\tilde{m}(x)) \leq 0 \Rightarrow \Phi[m] = \Phi[\tilde{m}], \quad (1.6)$$

but we will not make this assumption to avoid additional technical details in the proof of Theorem 4.2 below.

**Remark 1.** *Note that in the special case when*

- $H$  is given by (1.4) and  $L$  is strictly convex
- $\Phi[m(t, \cdot)](x) = F(m(t, x))$  and  $F$  is the derivative of a function  $W : \mathbb{R}_+ \rightarrow \mathbb{R}$  which is  $\mathcal{C}^2$  and strictly convex

*the system (1.1)-(1.3) may be obtained (at least formally) as the optimality conditions of the following optimal control problem driven by a transport equation: Minimize*

$$J(m, b) = \int_0^T \int_{\mathbb{T}^2} (m(t, x)L(x, b(t, x)) + W(m(t, x))) dx dt + \int_{\mathbb{T}^2} u_0(x)m(x, 0) dx$$

*subject to the constraints*

$$\begin{cases} \frac{\partial m}{\partial t} + \nu \Delta m + \operatorname{div}(m b) & = 0, & \text{in } (0, T) \times \mathbb{T}^2, \\ m(T, x) & = m_T(x) & \text{in } \mathbb{T}^2. \end{cases}$$

*For more details on this, see [21]. Note that obtaining (1.1)-(1.3) as the optimality condition of a global optimization problem is not possible in general.*

An important research activity is currently going on about approximation procedures of different types of mean field games models, see [18] for a numerical method

based on the reformulation of the model as the optimal control problem discussed in Remark 1, with an application in economics and [13] for a work on discrete time, finite state space mean field games. We also refer to [14, 15] for a specific constructive approach for quadratic Hamiltonians. Finally, a semi-discrete approximation for a first order mean field games problem has been studied in [7].

In [2], the authors have proposed and studied finite difference methods basically relying on monotone approximations of the Hamiltonian and on a suitable weak formulation of the Fokker-Planck equation, both for infinite and finite horizon mean field games. These schemes were shown to have several important features:

- existence and uniqueness for the discretized problems can be obtained by similar arguments as those used in the continuous case,
- they are robust when  $\nu \rightarrow 0$  (the deterministic limit of the models),
- bounds on the solutions (especially on the Lipschitz norm of  $u(t, \cdot)$ ), which are uniform in the grid step, may be proved under reasonable assumptions on the data.

In [1], the previously mentioned finite difference method has been extended to planning problems with MFG. This article contains results on existence and uniqueness for the systems of nonlinear equations and on the convergence of a penalty method. Fast algorithms for solving the discrete nonlinear systems arising in [2] have been proposed in [4].

In the present paper, we discuss the convergence of the schemes proposed in [2] in the reference case when  $H(x, p)$  is of the form

$$H(x, p) = \mathcal{H}(x) + |p|^\beta, \quad (1.7)$$

where  $\beta > 1$  is a real number and  $\mathcal{H}$  is a periodic continuous function, under suitable monotonicity assumptions on the operator  $\Phi$ . All these assumptions lead to uniqueness for the continuous and discrete systems, and also to a priori and stability estimates under further assumptions.

We have chosen to focus on the Hamiltonians in (1.7) because

- their fairly simple form (i.e. separate dependency on  $x$  and  $p$  and isotropy w.r.t.  $p$ ) will ease the algebra in the proofs of convergence
- the corresponding discrete Hamiltonians remain simple, see (2.3)-(2.4) below
- even if the Hamiltonians are simple, they lead to various situations depending on  $\beta$ : we will see below that the analysis of the numerical scheme requires different arguments for  $1 < \beta < 2$  and  $\beta \geq 2$ . This is not surprising since in the literature related to Hamilton-Jacobi-Bellman equations, the distinction between subquadratic and superquadratic Hamiltonians is often made because the techniques of analysis differ.

Note that the Hamiltonians in (1.7) are of the form (1.4) with  $L(x, \gamma) = -\mathcal{H}(x) + c|\gamma|^{\beta'}$  for  $\beta' = \beta/(\beta-1)$  and a suitable constant  $c > 0$ , in which there is separate dependency on  $x$  and  $\gamma$  and isotropy w.r.t.  $\gamma$ . Subquadratic Hamiltonians correspond to superquadratic Lagrangians and vice-versa.

This work is organized as follows: in Section 2, we recall the finite difference schemes proposed in [2]. Section 3 is devoted to basic facts concerning the discrete Hamiltonians and to the fundamental identity which leads to uniqueness and stability. When  $\Phi$  is a local operator, a consequence of this key identity is the a priori estimates on the solutions of the discrete MFG system that is presented in Section 5.1. Convergence theorems in the case when  $\Phi$  is a nonlocal smoothing operator are discussed in Section 4. Finally, Section 5 contains convergence theorems in the case when  $\Phi$  is a local operator.

**2. Finite difference schemes.** In the present paragraph, we discuss the finite difference method originally proposed in [2]. Before that, we stress the fact it has been obtained directly from the boundary value problem (1.1)-(1.3), and that it is not restricted to the case described in Remark 1. Yet, in the latter case, the discrete Bellman and Fokker-Planck equations discussed below can also be found as the optimality conditions of a global discrete optimization problem, as shown in [1] in the context of a planning problem with MFG.

Let  $N_T$  be a positive integer and  $\Delta t = T/N_T$ ,  $t_n = n\Delta t$ ,  $n = 0, \dots, N_T$ . Let  $\mathbb{T}_h^2$  be a uniform grid on the torus with mesh step  $h$ , (assuming that  $1/h$  is an integer  $N_h$ ), and  $x_{ij}$  denote a generic point in  $\mathbb{T}_h^2$ . The values of  $u$  and  $m$  at  $(x_{i,j}, t_n)$  are respectively approximated by  $u_{i,j}^n$  and  $m_{i,j}^n$ . Let  $u^n$  (resp.  $m^n$ ) be the vector containing the values  $u_{i,j}^n$  (resp.  $m_{i,j}^n$ ), for  $0 \leq i, j < N_h$  indexed in the lexicographic order. Hereafter, such vectors will be termed *grid functions on  $\mathbb{T}_h^2$*  or simply *grid functions*. For all grid function  $z$ , all  $i$  and  $j$ , we agree that  $z_{i,j} = z_{(i \bmod N_h), (j \bmod N_h)}$ .

*Elementary finite difference operators.* Let us introduce the elementary finite difference operators

$$(D_1^+ u)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{h} \quad \text{and} \quad (D_2^+ u)_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{h}, \quad (2.1)$$

and define  $[D_h u]_{i,j}$  as the collection of the four possible one sided finite differences at  $x_{i,j}$ :

$$[D_h u]_{i,j} = \left( (D_1^+ u)_{i,j}, (D_1^+ u)_{i-1,j}, (D_2^+ u)_{i,j}, (D_2^+ u)_{i,j-1} \right) \in \mathbb{R}^4. \quad (2.2)$$

We will also need the standard five point discrete Laplace operator

$$(\Delta_h u)_{i,j} = -\frac{1}{h^2} (4u_{i,j} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1}).$$

*Numerical Hamiltonian.* In order to approximate the term  $H(x, \nabla u)$  in (1.1) or (5.5), we consider a numerical Hamiltonian  $g : \mathbb{T}^2 \times \mathbb{R}^4 \rightarrow \mathbb{R}$ ,  $(x, q_1, q_2, q_3, q_4) \mapsto g(x, q_1, q_2, q_3, q_4)$ . Hereafter we will often assume that the following conditions hold:

(**g1**) *monotonicity:*  $g$  is nonincreasing with respect to  $q_1$  and  $q_3$  and nondecreasing with respect to  $q_2$  and  $q_4$ .

(**g2**) *consistency:*  $g(x, q_1, q_1, q_2, q_2) = H(x, q)$ ,  $\forall x \in \mathbb{T}^2, \forall q = (q_1, q_2) \in \mathbb{R}^2$ .

(**g3**) *differentiability:*  $g$  is of class  $\mathcal{C}^1$ .

(**g4**) *convexity:*  $(q_1, q_2, q_3, q_4) \mapsto g(x, q_1, q_2, q_3, q_4)$  is convex.

We will approximate  $H(\cdot, \nabla u)(x_{i,j})$  by  $g(x_{i,j}, [D_h u]_{i,j})$ .

Standard examples of numerical Hamiltonians fulfilling these requirements are provided by Lax-Friedrichs or upwind schemes, see [2]. In this work, we focus on Hamiltonians of the form  $H(x, p) = \mathcal{H}(x) + |p|^\beta$ ,  $\beta \in (1, \infty)$ , for which we choose

$$g(x, q) = \mathcal{H}(x) + G(q_1^-, q_2^+, q_3^-, q_4^+), \quad (2.3)$$

where, for a real number  $r$ ,  $r^+ = \max(r, 0)$  and  $r^- = \max(-r, 0)$  and where  $G : (\mathbb{R}_+)^4 \rightarrow \mathbb{R}_+$  is given by

$$G(p) = |p|^\beta = (p_1^2 + p_2^2 + p_3^2 + p_4^2)^{\frac{\beta}{2}}. \quad (2.4)$$

*Discrete version of the cost term  $\Phi[m(t, \cdot)](x)$ .* We introduce the compact and convex set

$$\mathcal{K}_h = \left\{ (m_{i,j})_{0 \leq i, j < N_h} : h^2 \sum_{i,j} m_{i,j} = 1; \quad m_{i,j} \geq 0 \right\} \quad (2.5)$$

which can be viewed as the set of the discrete probability measures.

We make the following assumptions,  $\Phi_h$  being local or not:

( $\Phi_{h1}$ )  $\Phi_h$  is continuous on  $\mathcal{K}_h$ .

( $\Phi_{h2}$ ) The numerical cost  $\Phi_h$  is monotone in the following sense:

$$(\Phi_h[m] - \Phi_h[\tilde{m}], m - \tilde{m})_2 \leq 0 \Rightarrow \Phi_h[m] = \Phi_h[\tilde{m}], \quad (2.6)$$

where  $(u, v)_2 = \sum_{0 \leq i, j < N_h} u_{i,j} v_{i,j}$ . This assumption and ( $g_4$ ) will be a sufficient condition for the discrete MFG system to have at most a solution,  $\Phi_h$  being local or not.

If  $\Phi$  is a local operator, i.e.  $\Phi[m](x) = F(m(x))$ ,  $F$  being a continuous function from  $\mathbb{R}^+$  to  $\mathbb{R}$ , then the discrete version of  $\Phi$  is naturally given by  $(\Phi_h[m])_{i,j} = F(m_{i,j})$ . In this case, the operator  $\Phi_h$  is continuous on the set of nonnegative grid functions. If  $\Phi$  is a nonlocal operator, then we assume that the discrete operator  $\Phi_h$  has the following additional properties:

( $\Phi_{h3}$ ) There exists a constant  $C$  independent of  $h$  such that for all grid function  $m \in \mathcal{K}_h$ ,

$$\|\Phi_h[m]\|_\infty \leq C \quad (2.7)$$

and

$$|(\Phi_h[m])_{i,j} - (\Phi_h[m])_{k,\ell}| \leq C d_{\mathbb{T}}(x_{i,j}, x_{k,\ell}) \quad (2.8)$$

where  $d_{\mathbb{T}}(x, y)$  is the distance between the two points  $x$  and  $y$  in the torus  $\mathbb{T}^2$ .

( $\Phi_{h4}$ ) Define  $\mathcal{K}$  as the set of probability densities, i.e. nonnegative integrable functions  $m$  on  $\mathbb{T}^2$  such that  $\int_{\mathbb{T}^2} m(x) dx = 1$ . For a grid function  $m_h \in \mathcal{K}_h$ , let  $\mathcal{J}_h m_h$  be the piecewise bilinear interpolation of  $m_h$  at the grid nodes: it is clear that  $\mathcal{J}_h m_h \in \mathcal{K}$ . There exists a continuous and bounded function  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $\omega(0) = 0$  and for all  $m \in \mathcal{K}$ , for all sequences  $(m_h)_h$ ,  $m_h \in \mathcal{K}_h$ ,

$$\|\Phi[m] - \Phi_h[m_h]\|_{L^\infty(\mathbb{T}_h^2)} \leq \omega(\|m - \mathcal{J}_h m_h\|_{L^1(\mathbb{T}^2)}). \quad (2.9)$$

Let  $\mathcal{I}_h m$  be the grid function whose value at  $x_{i,j}$  is

$$\frac{1}{h^2} \int_{|x - x_{i,j}|_\infty \leq h/2} m(x) dx.$$

It is clear that if  $m \in \mathcal{K}$  then  $\mathcal{I}_h m \in \mathcal{K}_h$  and that (2.9) implies that

$$\limsup_{h \rightarrow 0} \sup_{m \in \mathcal{K}} \|\Phi[m] - \Phi_h[\mathcal{I}_h m]\|_{L^\infty(\mathbb{T}_h^2)} = 0. \quad (2.10)$$

For example, if  $\Phi[m]$  is defined as the solution  $w$  of the equation  $\Delta^2 w + w = m$  in  $\mathbb{T}^2$ , ( $\Delta^2$  being the bilaplacian), then one can define  $\Phi_h[m_h]$  as the solution  $w_h$  of  $\Delta_h^2 w_h + w_h = m_h$  in  $\mathbb{T}_h^2$ . It is possible to check that all the above properties are true. Another example is the following: if  $\zeta$  is a nonnegative regularizing kernel,  $\Phi[m] = \zeta \star (\zeta \star m)$  ( $\star$  stands for the convolution) and  $\Phi_h[m_h] = \zeta \star \zeta \star \mathcal{J}_h m_h$  (assuming that the convolution can be computed exactly). The monotonicity assumptions on  $\Phi$  and  $\Phi_h$  are natural even for non local operators, in situations where a given agent has access to some global and partial information on the density around and has aversion to large densities.

*Discrete Bellman equation.* The discrete version of the Bellman equation is obtained by applying a semi-implicit Euler scheme to (1.1),

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} - \nu(\Delta_h u^{n+1})_{i,j} + g(x_{i,j}, [D_h u^{n+1}]_{i,j}) = (\Phi_h[m^n])_{i,j}, \quad (2.11)$$

for all points in  $\mathbb{T}_h^2$  and all  $n$ ,  $0 \leq n < N_T$ , where all the discrete operators have been introduced above. Given  $(m^n)_{n=0, \dots, N_T-1}$ , (2.11) and the initial condition  $u_{i,j}^0 = u_0(x_{i,j})$  for all  $(i, j)$  completely characterizes  $(u^n)_{0 \leq n \leq N_T}$ .

*Discrete Fokker-Planck equation.* In order to approximate equation (1.2), it is convenient to consider its weak formulation which involves in particular the term

$$\int_{\mathbb{T}^2} \operatorname{div} \left( m \frac{\partial H}{\partial p}(\cdot, \nabla u) \right) (x) w(x) dx.$$

By periodicity,

$$\int_{\mathbb{T}^2} \operatorname{div} \left( m \frac{\partial H}{\partial p}(\cdot, \nabla u) \right) (x) w(x) dx = - \int_{\mathbb{T}^2} m(x) \frac{\partial H}{\partial p}(x, \nabla u(x)) \cdot \nabla w(x) dx$$

holds for any test function  $w$ . The right hand side in the identity above will be approximated by

$$-h^2 \sum_{i,j} m_{i,j} \nabla_q g(x_{i,j}, [D_h u]_{i,j}) \cdot [D_h w]_{i,j} = h^2 \sum_{i,j} \mathcal{T}_{i,j}(u, m) w_{i,j},$$

where the transport operator  $\mathcal{T}$  is defined as follows:

$$\mathcal{T}_{i,j}(u, m) = \frac{1}{h} \left( \begin{array}{l} \left( \begin{array}{l} m_{i,j} \frac{\partial g}{\partial q_1}(x_{i,j}, [D_h u]_{i,j}) - m_{i-1,j} \frac{\partial g}{\partial q_1}(x_{i-1,j}, [D_h u]_{i-1,j}) \\ + m_{i+1,j} \frac{\partial g}{\partial q_2}(x_{i+1,j}, [D_h u]_{i+1,j}) - m_{i,j} \frac{\partial g}{\partial q_2}(x_{i,j}, [D_h u]_{i,j}) \end{array} \right) \\ + \\ \left( \begin{array}{l} m_{i,j} \frac{\partial g}{\partial q_3}(x_{i,j}, [D_h u]_{i,j}) - m_{i,j-1} \frac{\partial g}{\partial q_3}(x_{i,j-1}, [D_h u]_{i,j-1}) \\ + m_{i,j+1} \frac{\partial g}{\partial q_4}(x_{i,j+1}, [D_h u]_{i,j+1}) - m_{i,j} \frac{\partial g}{\partial q_4}(x_{i,j}, [D_h u]_{i,j}) \end{array} \right) \end{array} \right). \quad (2.12)$$

The discrete version of equation (1.2) is chosen as follows:

$$\frac{m_{i,j}^{n+1} - m_{i,j}^n}{\Delta t} + \nu(\Delta_h m^n)_{i,j} + \mathcal{T}_{i,j}(u^{n+1}, m^n) = 0. \quad (2.13)$$

This scheme is implicit w.r.t. to  $m$  and explicit w.r.t.  $u$  because the considered Fokker-Planck equation is backward. Given  $u$  this is a system of linear equations for  $m$ . It is easy to see that if  $m^n$  satisfies (2.13) for  $0 \leq n < N_T$  and if  $m^{N_T} \in \mathcal{K}_h$ , then  $m^n \in \mathcal{K}_h$  for all  $n$ ,  $0 \leq n < N_T$ .

**Remark 2.** An important property of  $\mathcal{T}$  is that the operator  $m \mapsto (-\nu(\Delta_h m)_{i,j} - \mathcal{T}_{i,j}(u, m))_{i,j}$  is the adjoint of the linearized version of the operator  $u \mapsto (-\nu(\Delta_h u)_{i,j} + g(x_{i,j}, [D_h u]_{i,j}))_{i,j}$ .

This property implies that the structure of (1.1)-(1.2) is preserved in the discrete version (2.11)-(2.13). In particular, it implies the fundamental identity given in § 3.3 and then the uniqueness result stated in Theorem 2.2 below.

*Summary.* The fully discrete scheme for system (1.1),(1.2),(1.3) is therefore the following: for all  $0 \leq i, j < N_h$  and  $0 \leq n < N_T$

$$\begin{cases} \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} - \nu(\Delta_h u^{n+1})_{i,j} + g(x_{i,j}, [D_h u^{n+1}]_{i,j}) = (\Phi_h[m^n])_{i,j}, \\ \frac{m_{i,j}^{n+1} - m_{i,j}^n}{\Delta t} + \nu(\Delta_h m^n)_{i,j} + \mathcal{T}_{i,j}(u^{n+1}, m^n) = 0, \end{cases} \quad (2.14)$$

with the initial and terminal conditions

$$u_{i,j}^0 = u_0(x_{i,j}), \quad m_{i,j}^{N_T} = \frac{1}{h^2} \int_{|x-x_{i,j}|_\infty \leq h/2} m_T(x) dx, \quad 0 \leq i, j < N_h. \quad (2.15)$$



The following theorem was proved in [2] (using essentially a Brouwer fixed point argument and estimates on the solutions of the discrete Bellman equation):

**THEOREM 2.1.** *Assume that  $(g_1)$ – $(g_3)$  and  $(\Phi_{h1})$  hold, that  $u_0$  is a continuous function on  $\mathbb{T}^2$  and that  $m_T \in \mathcal{K}$ : then (2.14)–(2.15) has a solution such that  $m^n \in \mathcal{K}_h, \forall n$ .*

If furthermore

- $(\Phi_{h3})$  holds
- there exists a constant  $C$  such that

$$\left| \frac{\partial g}{\partial x}(x, (q_1, q_2, q_3, q_4)) \right| \leq C(1 + |q_1| + |q_2| + |q_3| + |q_4|) \quad \forall x \in \mathbb{T}^2, \forall q_1, q_2, q_3, q_4$$

- $u_0$  is Lipschitz continuous

then  $\max_{0 \leq n \leq N_T} (\|u^n\|_\infty + \|D_h u^n\|_\infty) \leq c$  for a constant  $c$  independent of  $h$  and  $\Delta t$ .

**Remark 3.** *Theorem 2.1 states in particular that  $m^n \in \mathcal{K}_h$  for all  $n$ : this means that the finite method preserves the positivity and the total mass of the density  $m$ . This is of course an essential feature of the discrete scheme.*

**Remark 4.** *Note that the technical condition on  $\frac{\partial g}{\partial x}$  is automatically true if  $g$  is given by (2.3)–(2.4) and  $\mathcal{H}$  is  $\mathcal{C}^1$ .*

**Remark 5.** *A priori estimates in the case when  $\Phi$  is a local operator will be given in § 5.1. Since (2.14)–(2.15) has exactly the same structure as the continuous problem (1.1)–(1.3), uniqueness has been obtained in [2] with the same arguments as in [20]:*

**THEOREM 2.2.** *Assume that  $(g_1)$ – $(g_4)$  and  $(\Phi_{h1})$ – $(\Phi_{h2})$  hold, then (2.14)–(2.15) has a unique solution.*

**Remark 6.** *Efficient algorithms for solving system (2.14)–(2.15) require special efforts, essentially because of the forward-backward structure already discussed in §1. We refer to [2] and [4] for the description of possible algorithms and numerical results.*

**3. Basic facts for numerical Hamiltonians of the form (2.3)–(2.4).** We focus on numerical Hamiltonians in the form (2.3)–(2.4) but obviously, what follows holds for  $g(x, q) = \mathcal{H}(x) + cG(q_1^-, q_2^+, q_3^-, q_4^+)$ , where  $G$  is given by (2.4) and  $c$  is a positive constant.

We use the following notations: for  $q \in \mathbb{R}^4$ ,  $|q|$  is the Euclidean norm given by

$$|q| = \sqrt{\sum_{i=1}^4 q_i^2}, \text{ and } |q|_\infty \text{ is the max norm, } |q|_\infty = \max_{i=1, \dots, 4} |q_i|.$$

For a sufficiently regular function  $\psi : \mathbb{R}^4 \rightarrow \mathbb{R}$ ,  $p \mapsto \psi(p)$ ,  $\psi_p(p) \in \mathbb{R}^4$  (resp.  $\psi_{pp}(p) \in \mathbb{R}^{4 \times 4}$ ) stands for the gradient of  $\psi$  (resp. the Hessian of  $\psi$ ). For a sufficiently regular function  $\psi : \mathbb{T}^2 \times \mathbb{R}^4 \rightarrow \mathbb{R}$ ,  $(x, p) \mapsto \psi(x, p) \in \mathbb{R}$ ,  $\psi_p(x, p) \in \mathbb{R}^4$  is the gradient of  $p \mapsto \psi(x, p)$  and  $\psi_{pp}(x, p) \in \mathbb{R}^{4 \times 4}$  is the Hessian of  $p \mapsto \psi(x, p)$ .

Let us define the map  $\Theta : \mathbb{R}^4 \rightarrow \mathbb{R}_+^4$  by

$$p = \Theta(q) \Leftrightarrow \begin{cases} p_1 = q_1^-, \\ p_2 = q_2^+, \\ p_3 = q_3^-, \\ p_4 = q_4^+. \end{cases} \quad (3.1)$$

**3.1. Basic lemmas.** The following lemmas about  $g$  and  $G$  will be useful. Their proofs are given in Appendix A.

**LEMMA 3.1.**

1. (a) For all  $p \in (\mathbb{R}_+)^4$ ,  $G_p(p) = \beta|p|^{\beta-2}p$  and

$$G_{pp}(p) = \beta|p|^{\beta-2}I_4 + \beta(\beta-2)|p|^{\beta-4}p \otimes p. \quad (3.2)$$

(b) If  $\beta \geq 2$ , then  $G_{pp}(p) - \beta|p|^{\beta-2}I_4$  is a positive semi-definite matrix, which may be written

$$G_{pp}(p) \geq \beta|p|^{\beta-2}I_4. \quad (3.3)$$

(c) If  $1 < \beta < 2$ , then

$$G_{pp}(p) \geq \beta(\beta - 1)|p|^{\beta-2}I_4. \quad (3.4)$$

2. For all  $q, \tilde{q} \in \mathbb{R}^4$ ,

$$g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \geq G(\tilde{p}) - G(p) - G_p(p) \cdot (\tilde{p} - p), \quad (3.5)$$

where  $p = \Theta(q)$ ,  $\tilde{p} = \Theta(\tilde{q})$  and  $\Theta$  is defined in (3.1).

LEMMA 3.2.

1. (a) If  $\beta \geq 2$ , then for all  $q, \tilde{q} \in \mathbb{R}^4$ ,

$$g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \geq \frac{1}{\beta - 1} \max(|p|^{\beta-2}, |\tilde{p}|^{\beta-2})|p - \tilde{p}|^2 \quad (3.6)$$

$$\geq \frac{|p - \tilde{p}|^\beta}{2^{\beta-2}(\beta - 1)}, \quad (3.7)$$

where  $p = \Theta(q)$ ,  $\tilde{p} = \Theta(\tilde{q})$  and  $\Theta$  is defined in (3.1).

1. (b) If  $1 < \beta < 2$ , then for all  $q, \tilde{q} \in \mathbb{R}^4$  such that  $\Theta(q) + \Theta(\tilde{q}) \neq 0$ ,

$$g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \geq 2^{\beta-3}\beta(\beta - 1) \min(|p|_\infty^{\beta-2}, |\tilde{p}|_\infty^{\beta-2})|p - \tilde{p}|^2, \quad (3.8)$$

where  $p = \Theta(q)$ ,  $\tilde{p} = \Theta(\tilde{q})$ .

2. If  $\beta \geq 2$ , there exists a positive constant  $c$  such that for all  $q, \tilde{q}, r \in \mathbb{R}^4$ , for all  $\eta > 0$ ,

$$|(g_q(x, \tilde{q}) - g_q(x, q)) \cdot r| \leq \max(|p|^{\beta-2}, |\tilde{p}|^{\beta-2}) \left( \frac{c}{\eta} |p - \tilde{p}|^2 + \eta |r|^2 \right), \quad (3.9)$$

where  $p = \Theta(q)$ ,  $\tilde{p} = \Theta(\tilde{q})$ .

**3.2. A nonlinear functional  $\mathcal{G}(M, U, \tilde{U})$ .** Let us define the nonlinear functional  $\mathcal{G}$  by

$$\mathcal{G}(M, U, \tilde{U}) = \sum_{n=1}^{N_T} \sum_{i,j} m_{i,j}^{n-1} R_{i,j}^n, \quad (3.10)$$

where  $M = (m^n)_{0 \leq n \leq N_T}$ ,  $U = (u^n)_{0 \leq n \leq N_T}$ ,  $\tilde{U} = (\tilde{u}^n)_{0 \leq n \leq N_T}$ , and  $m^n$ ,  $u^n$  and  $\tilde{u}^n$  are grid functions, i.e. for example  $m^n = (m_{i,j}^n)_{0 \leq i,j < N_h}$ , and where  $R_{i,j}^n$  is defined by

$$R_{i,j}^n = g(x_{i,j}, [D\tilde{u}^n]_{i,j}) - g(x_{i,j}, [Du^n]_{i,j}) - g_q(x_{i,j}, [Du^n]_{i,j}) \cdot ([D\tilde{u}^n]_{i,j} - [Du^n]_{i,j}). \quad (3.11)$$

Under Assumption  $(g_4)$ , it is clear that if  $M \geq 0$  (meaning that  $m^n$  is a nonnegative grid function for all  $n$ ,  $0 \leq n \leq N_T$ ), then  $\mathcal{G}(M, U, \tilde{U}) \geq 0$ . If  $g$  is of the form (2.3)-(2.4), we have a more precise estimate:

LEMMA 3.3. If  $M \geq 0$  in the sense defined above and if  $g$  is of the form (2.3)-(2.4) with  $\beta \geq 2$ , then

$$\begin{aligned} \mathcal{G}(M, U, \tilde{U}) &\geq \frac{1}{\beta - 1} \sum_{n=1}^{N_T} \sum_{i,j} m_{i,j}^{n-1} \max(|p_{i,j}^n|^{\beta-2}, |\tilde{p}_{i,j}^n|^{\beta-2}) |p_{i,j}^n - \tilde{p}_{i,j}^n|^2 \\ &\geq \frac{1}{2^{\beta-2}(\beta - 1)} \sum_{n=1}^{N_T} \sum_{i,j} m_{i,j}^{n-1} |p_{i,j}^n - \tilde{p}_{i,j}^n|^\beta. \end{aligned} \quad (3.12)$$

where  $p_{i,j}^n$  and  $\tilde{p}_{i,j}^n$  are the four dimensional vectors

$$\begin{aligned} p_{i,j}^n &= \left( (D_1^+ u^n)_{i,j}^-, ((D_1^+ u^n)_{i-1,j})^+, ((D_2^+ u^n)_{i,j})^-, ((D_2^+ u^n)_{i,j-1})^+ \right), \\ \tilde{p}_{i,j}^n &= \left( (D_1^+ \tilde{u}^n)_{i,j}^-, ((D_1^+ \tilde{u}^n)_{i-1,j})^+, ((D_2^+ \tilde{u}^n)_{i,j})^-, ((D_2^+ \tilde{u}^n)_{i,j-1})^+ \right). \end{aligned} \quad (3.13)$$

Furthermore, if  $M$  is bounded from below by  $\underline{m}$ , which means that  $m_{i,j}^n \geq \underline{m}$  for all  $n, i, j$ , then

$$\mathcal{G}(M, U, \tilde{U}) \geq \frac{\underline{m}}{2^{2\beta-3}(\beta-1)} \sum_{n=1}^{N_T} \sum_{i,j} \left| [D\tilde{u}^n]_{i,j} - [Du^n]_{i,j} \right|^\beta. \quad (3.14)$$

*Proof.* We deduce (3.12) from (3.6) and (3.7). If  $M$  is bounded from below by  $\underline{m} > 0$ , we deduce that

$$\begin{aligned} \mathcal{G}(M, U, \tilde{U}) &\geq \frac{\underline{m}}{2^{\beta-2}(\beta-1)} \sum_{n=1}^{N_T} \sum_{i,j} |p_{i,j}^n - \tilde{p}_{i,j}^n|^\beta \\ &\geq \frac{\underline{m}}{2^{\beta-2}(\beta-1)} \sum_{n=1}^{N_T} \sum_{i,j} \sum_{k=1}^4 |(p_{i,j}^n)_k - (\tilde{p}_{i,j}^n)_k|^\beta. \end{aligned}$$

Since for each  $0 \leq i, j < N_h$ , each  $1 \leq k \leq 4$ , the quantity  $\left| ([D\tilde{u}^n]_{i,j})_k - ([Du^n]_{i,j})_k \right|^\beta$  appears at least once in the sum above, we deduce that

$$\begin{aligned} \mathcal{G}(M, U, \tilde{U}) &\geq \frac{\underline{m}}{2^{\beta-1}(\beta-1)} \sum_{n=1}^{N_T} \sum_{i,j} \sum_{k=1}^4 \left| ([D\tilde{u}^n]_{i,j})_k - ([Du^n]_{i,j})_k \right|^\beta \\ &\geq \frac{\underline{m}}{2^{2\beta-3}(\beta-1)} \sum_{n=1}^{N_T} \sum_{i,j} \left| [D\tilde{u}^n]_{i,j} - [Du^n]_{i,j} \right|^\beta \end{aligned}$$

where the second inequality comes from the fact that for four positive numbers  $(a_k)_{k=1,\dots,4}$   $\sum_{k=1}^4 a_k^\beta \geq \frac{1}{4^{\beta/2-1}} (\sum_{k=1}^4 a_k^2)^\beta$ .  $\square$

**Remark 7.** The same kind of argument shows that if  $1 < \beta < 2$ , and  $m_{i,j}^n \geq \underline{m}$ , then

$$\begin{aligned} \mathcal{G}(M, 0, U) &\geq 2^{\beta-3}\beta(\beta-1)\underline{m} \sum_{n=1}^{N_T} \sum_{i,j} |p_{i,j}^n|_{\infty}^{\beta-2} |p_{i,j}^n|^2 \geq 2^{\beta-3}\beta(\beta-1)\underline{m} \sum_{n=1}^{N_T} \sum_{i,j} |p_{i,j}^n|^\beta \\ &\geq 2^{2\beta-5}\beta(\beta-1)\underline{m} \sum_{n=1}^{N_T} \sum_{i,j} \sum_{k=1}^4 |(p_{i,j}^n)_k|^\beta \\ &\geq 2^{2\beta-6}\beta(\beta-1)\underline{m} \sum_{n=1}^{N_T} \sum_{i,j} \sum_{k=1}^4 |([Du^n]_{i,j})_k|^\beta \\ &\geq 2^{2\beta-6}\beta(\beta-1)\underline{m} \sum_{n=1}^{N_T} \sum_{i,j} |[Du^n]_{i,j}|^\beta, \end{aligned}$$

where  $p_{i,j}^n$  is given by (3.13).

**3.3. A fundamental identity .** In this paragraph, we discuss a key identity which leads to the stability of the finite difference scheme under additional assumptions.

Since the arguments below are similar to those used for studying the system of partial differential equations (1.1)-(1.2), let us first sketch them in the continuous context: consider a solution  $(u, m)$  of (1.1)-(1.2), and a solution  $(\tilde{u}, \tilde{m})$  of the perturbed system in  $(0, T) \times \mathbb{T}^2$

$$\frac{\partial \tilde{u}}{\partial t}(t, x) - \nu \Delta \tilde{u}(t, x) + H(x, \nabla \tilde{u}(t, x)) = \Phi[\tilde{m}(t, \cdot)](x) + a(t, x), \quad (3.15)$$

$$\frac{\partial \tilde{m}}{\partial t}(t, x) + \nu \Delta \tilde{m}(t, x) + \operatorname{div} \left( \tilde{m}(t, \cdot) \frac{\partial H}{\partial p}(\cdot, \nabla \tilde{u}(t, \cdot)) \right) (x) = b(t, x), \quad (3.16)$$

where  $a$  and  $b$  are  $L^\infty$  functions. By

1. subtracting (3.15) to (1.1), multiplying the result by  $(m - \tilde{m})$ , and integrating on  $(0, T) \times \mathbb{T}^2$
2. subtracting (3.16) to (1.2), multiplying the result by  $(u - \tilde{u})$ , and integrating on  $(0, T) \times \mathbb{T}^2$
3. subtracting the two identities obtained in the first two steps and performing suitable integrations by parts taking advantage of the periodicity,

one gets the following important identity:

$$\begin{aligned} & \int_0^T \int_{\mathbb{T}^2} m(t, x) R(t, x) + \int_0^T \int_{\mathbb{T}^2} \tilde{m}(t, x) \tilde{R}(t, x) \\ & + \int_0^T \int_{\mathbb{T}^2} (\Phi[m] - \Phi[\tilde{m}])(m(t, x) - \tilde{m}(t, x)) \\ & + \int_{\mathbb{T}^2} (u(0, x) - \tilde{u}(0, x))(m(0, x) - \tilde{m}(0, x)) \\ & - \int_{\mathbb{T}^2} (u(T, x) - \tilde{u}(T, x))(m(T, x) - \tilde{m}(T, x)) \\ & = \int_0^T \int_{\mathbb{T}^2} a(t, x)(m(t, x) - \tilde{m}(t, x)) + \int_0^T \int_{\mathbb{T}^2} b(t, x)(u(t, x) - \tilde{u}(t, x)). \end{aligned}$$

where

$$\begin{aligned} R(t, x) &= H(x, \nabla u(t, x)) - H(x, \nabla \tilde{u}(t, x)) - \frac{\partial H}{\partial p}(x, \nabla u(t, x)) \cdot \nabla(u(t, x) - \tilde{u}(t, x)), \\ \tilde{R}(t, x) &= H(x, \nabla \tilde{u}(t, x)) - H(x, \nabla u(t, x)) - \frac{\partial H}{\partial p}(x, \nabla \tilde{u}(t, x)) \cdot \nabla(\tilde{u}(t, x) - u(t, x)). \end{aligned}$$

The same arguments can be used at the discrete level: consider a perturbed system:

$$\begin{cases} \frac{\tilde{u}_{i,j}^{n+1} - \tilde{u}_{i,j}^n}{\Delta t} - \nu(\Delta_h \tilde{u}^{n+1})_{i,j} + g(x_{i,j}, [D_h \tilde{u}^{n+1}]_{i,j}) &= (\Phi_h[\tilde{m}^n])_{i,j} + a_{i,j}^n, \\ \frac{\tilde{m}_{i,j}^{n+1} - \tilde{m}_{i,j}^n}{\Delta t} + \nu(\Delta_h \tilde{m}^n)_{i,j} + \mathcal{T}_{i,j}(\tilde{u}^{n+1}, \tilde{m}^n) &= b_{i,j}^n. \end{cases} \quad (3.17)$$

Multiplying the first equations in (3.17) and (2.14) by  $m_{i,j}^n - \tilde{m}_{i,j}^n$  and subtracting, then summing the results for all  $n = 0, \dots, N_T - 1$  and all  $(i, j)$ , we obtain

$$\begin{aligned} & \sum_{n=0}^{N_T-1} \frac{1}{\Delta t} ((u^{n+1} - \tilde{u}^{n+1}) - (u^n - \tilde{u}^n), (m^n - \tilde{m}^n))_2 \\ & - \sum_{n=0}^{N_T-1} \nu(\Delta_h(u^{n+1} - \tilde{u}^{n+1}), m^n - \tilde{m}^n)_2 \\ & + \sum_{n=0}^{N_T-1} \sum_{i,j} (g(x_{i,j}, [D_h u^{n+1}]_{i,j}) - g(x_{i,j}, [D_h \tilde{u}^{n+1}]_{i,j}))(m_{i,j}^n - \tilde{m}_{i,j}^n) \\ & = \sum_{n=0}^{N_T-1} (\Phi_h[m^n] - \Phi_h[\tilde{m}^n], m^n - \tilde{m}^n)_2 - \sum_{n=0}^{N_T-1} (a^n, m^n - \tilde{m}^n)_2, \end{aligned} \quad (3.18)$$

where  $(X, Y)_2 = \sum_{i,j} X_{i,j} Y_{i,j}$ . Similarly, subtracting the second equation in (3.17) from the second equation in (2.13), multiplying the result by  $u_{i,j}^{n+1} - \tilde{u}_{i,j}^{n+1}$  and summing for all  $n = 0, \dots, N_T - 1$  and all  $(i, j)$  leads to

$$\begin{aligned}
& \sum_{n=0}^{N_T-1} \frac{1}{\Delta t} ((m^{n+1} - m^n) - (\tilde{m}^{n+1} - \tilde{m}^n), (u^{n+1} - \tilde{u}^{n+1}))_2 \\
& + \nu((m^n - \tilde{m}^n), \Delta_h(u^{n+1} - \tilde{u}^{n+1}))_2 \\
& - \sum_{n=0}^{N_T-1} \sum_{i,j} m_{i,j}^n [D_h(u^{n+1} - \tilde{u}^{n+1})]_{i,j} \cdot g_q(x_{i,j}, [D_h u^{n+1}]_{i,j}) \\
& + \sum_{n=0}^{N_T-1} \sum_{i,j} \tilde{m}_{i,j}^n [D_h(u^{n+1} - \tilde{u}^{n+1})]_{i,j} \cdot g_q(x_{i,j}, [D_h \tilde{u}^{n+1}]_{i,j}) = - \sum_{n=1}^{N_T} (b^{n-1}, u^n - \tilde{u}^n)_2.
\end{aligned} \tag{3.19}$$

Adding (3.18) and (3.19) leads to the fundamental identity

$$\begin{aligned}
& - \frac{1}{\Delta t} (m^{N_T} - \tilde{m}^{N_T}, u^{N_T} - \tilde{u}^{N_T})_2 + \frac{1}{\Delta t} (m^0 - \tilde{m}^0, u^0 - \tilde{u}^0)_2 \\
& + \mathcal{G}(M, U, \tilde{U}) + \mathcal{G}(\tilde{M}, \tilde{U}, U) + \sum_{n=0}^{N_T-1} (\Phi_h[m^n] - \Phi_h[\tilde{m}^n], m^n - \tilde{m}^n)_2 \\
& = \sum_{n=0}^{N_T-1} (a^n, m^n - \tilde{m}^n)_2 + \sum_{n=1}^{N_T} (b^{n-1}, u^n - \tilde{u}^n)_2
\end{aligned} \tag{3.20}$$

where  $M = (m^n)_{0 \leq n \leq N_T}$ ,  $\tilde{M} = (\tilde{m}^n)_{0 \leq n \leq N_T}$ ,  $U = (u^n)_{0 \leq n \leq N_T}$ ,  $\tilde{U} = (\tilde{u}^n)_{0 \leq n \leq N_T}$ . Note that under assumptions  $(g_4)$  and  $(\Phi_{h2})$ , the second line of (3.20) is made of three nonnegative terms. This is the key observation leading to uniqueness for (2.14)-(2.15), but it may also lead to a priori estimates or stability estimates under additional assumptions including for example the assumptions of Lemma 3.3.

**4. Study of the convergence in the case when  $\Phi$  is a nonlocal smoothing operator.** Hereafter the Hamiltonian is of the form (1.7). In all the following convergence results, we assume that system (1.1)-(1.3) has a unique classical solution. This is always the case if  $\Phi$  is monotone in the sense of (1.5) and continuously maps the set of probability measures on  $\mathbb{T}^2$  (endowed with the weak \* topology) to a bounded subset of  $Lip(\mathbb{T}^2)$ , the Lipschitz functions on  $\mathbb{T}^2$ , and for example maps continuously  $C^{k,\alpha}(\mathbb{T}^2)$  to  $C^{k+1,\alpha}(\mathbb{T}^2)$ , for all  $k \in \mathbb{N}$  and  $0 \leq \alpha < 1$ .

We summarize the assumptions made in § 4 as follows:

*Standing assumptions (in Section 4).*

- The Hamiltonian is of the form (1.7) and the function  $x \rightarrow \mathcal{H}(x)$  is  $\mathcal{C}^1$  on  $\mathbb{T}^2$
- The functions  $u_0$  and  $m_T$  are smooth, and  $m_T \in \mathcal{K}$  is bounded from below by a positive number
- The operator  $\Phi$  is monotone in the sense of (1.5), nonlocal and smoothing, so that there is a unique classical solution  $(u, m)$  of (1.1)-(1.3) such that  $m > 0$
- The numerical Hamiltonian given by (2.3)-(2.4) and  $(\Phi_{h1})$ ,  $(\Phi_{h2})$ ,  $(\Phi_{h3})$ , and  $(\Phi_{h4})$  hold for the numerical cost function  $\Phi_h$ .

**Remark 8.** *The assumption on the monotonicity of  $\Phi$  will be made in all the convergence theorems that follow. Without such an assumption, one may only expect the convergence of subsequences; the techniques of proofs would differ much from the ones proposed below, the main difficulty being to prove compactness in order to pass to the limit in the nonlinearities.*

#### 4.1. The case when $\beta \geq 2$ .

**THEOREM 4.1.** *With the standing assumptions stated above, we choose  $\beta \geq 2$ . Let  $u_h$  (resp.  $m_h$ ) be the piecewise trilinear function in  $\mathcal{C}([0, T] \times \mathbb{T}^2)$  obtained by interpolating the values  $u_{i,j}^n$  (resp.  $m_{i,j}^n$ ) at the nodes of the space-time grid. The functions  $u_h$  converge uniformly and in  $L^\beta(0, T; W^{1,\beta}(\mathbb{T}^2))$  to  $u$  as  $h$  and  $\Delta t$  tend to 0. The functions  $m_h$  converge to  $m$  in  $C^0([0, T]; L^2(\mathbb{T}^2)) \cap L^2(0, T; H^1(\mathbb{T}^2))$  as  $h$  and  $\Delta t$  tend to 0.*

*Proof.* Note that  $m_h(t, \cdot) \in \mathcal{K}$  for any  $t \in [0, T]$ .

Hereafter, we use the following notations:

1.  $\tilde{u}^n = (\tilde{u}_{i,j}^n)$  is the grid function defined by  $\tilde{u}_{i,j}^n = u(t_n, x_{i,j})$ , where  $u$  is defined by (1.1)-(1.3)
2.  $\tilde{U} = (\tilde{u}^n)_{0 \leq n \leq N_T}$
3.  $\tilde{m}^n = \mathcal{I}_h(m(t_n, \cdot))$  with the notation introduced in § 2 and  $m$  defined by (1.1)-(1.3). This means that  $\tilde{m}_{i,j}^n = \int_{|x-x_{i,j}|_\infty < h/2} m(t_n, x) dx$
4.  $\tilde{M} = (\tilde{m}^n)_{0 \leq n \leq N_T}$
5.  $U = (u^n)_{0 \leq n \leq N_T}$  and  $M = (m^n)_{0 \leq n \leq N_T}$  where  $(u^n, m^n)_n$  is the solution of (2.14)-(2.15)

The grid functions  $\tilde{u}^n$  and  $\tilde{m}^n$  satisfy (3.17) where  $a_{i,j}^n$  and  $b_{i,j}^n$  are consistency errors. From (2.10), the consistency of the discrete scheme and since  $(u, m)$  is a classical solution of (1.1)-(1.3), we know that  $\max_{0 \leq n < N_T} (\|a^n\|_{L^\infty(\mathbb{T}_h^2)} + \|b^n\|_{L^\infty(\mathbb{T}_h^2)})$  goes to zero as  $h$  and  $\Delta t$  go to zero.

**Remark 9.** *In order to give more precise information on the rate of convergence of  $a^n$  to 0, we would need to make further assumptions on  $\Phi_h$ . In any case, the convergence to 0 of  $\max_{0 \leq n < N_T} (\|a^n\|_{L^\infty(\mathbb{T}_h^2)} + \|b^n\|_{L^\infty(\mathbb{T}_h^2)})$  is not better than linear, because the discrete scheme for  $H(x, \nabla u)$  is first order.*

*Strategy.* The proof is done in three steps:

1. in the first step, we prove the convergence to zero of a discrete semi-norm of the error on  $u$ , more precisely that  $\lim_{h, \Delta t \rightarrow 0} h^2 \Delta t \sum_{n=1}^{N_T} \sum_{i,j} \left| [D_h \tilde{u}^n]_{i,j} - [D_h u^n]_{i,j} \right|^\beta = 0$ . For that, we essentially use the fundamental identity (3.20) and the fact that the function  $m$  in (1.1)-(1.3) is bounded from below by a positive constant.
2. The second step consists proving the convergence of  $m_h$  to  $m$  by working only on the discrete Fokker-Planck equations and taking advantage of the results obtained in the first step.
3. The third step consists of focusing on the discrete Bellman equation and deducing the convergence of  $u_h$  to  $u$  from the results of the second step.

*Step 1: convergence to zero of a discrete semi-norm of the error on  $u$ .* As a consequence of the previous observations, the fundamental identity (3.20) holds, and from (2.15), can be written as follows:

$$\begin{aligned}
 & h^2 \Delta t \mathcal{G}(M, M, \tilde{U}) + h^2 \Delta t \mathcal{G}(\tilde{M}, \tilde{U}, U) + h^2 \Delta t \sum_{n=0}^{N_T-1} (\Phi_h[m^n] - \Phi_h[\tilde{m}^n], m^n - \tilde{m}^n)_2 \\
 & = h^2 \Delta t \sum_{n=0}^{N_T-1} (a^n, m^n - \tilde{m}^n)_2 + h^2 \Delta t \sum_{n=1}^{N_T} (b^{n-1}, u^n - \tilde{u}^n)_2.
 \end{aligned} \tag{4.1}$$

The a priori estimate on the discrete Lipschitz norm of  $u^n$  given at the end of Theorem 2.1 implies that  $\lim_{h, \Delta t \rightarrow 0} h^2 \max_n |(b^{n-1}, u^n - \tilde{u}^n)_2| = 0$ . From the fact that  $m^n \in \mathcal{K}_h$ , we also get that  $\lim_{h, \Delta t \rightarrow 0} h^2 \max_n |(a^n, m^n - \tilde{m}^n)_2| = 0$ . We can

then use (3.12) to deduce that

$$\begin{aligned} \Delta t h^2 \sum_{n=0}^{N_T-1} \sum_{i,j} \tilde{m}_{i,j}^n \max(|p_{i,j}^{n+1}|^{\beta-2}, |\tilde{p}_{i,j}^{n+1}|^{\beta-2}) |p_{i,j}^{n+1} - \tilde{p}_{i,j}^{n+1}|^2 &= o(1), \\ \Delta t h^2 \sum_{n=0}^{N_T-1} \sum_{i,j} m_{i,j}^n \max(|p_{i,j}^{n+1}|^{\beta-2}, |\tilde{p}_{i,j}^{n+1}|^{\beta-2}) |p_{i,j}^{n+1} - \tilde{p}_{i,j}^{n+1}|^2 &= o(1), \end{aligned} \quad (4.2)$$

where  $p_{i,j}^n$  and  $\tilde{p}_{i,j}^n$  are given by (3.13). Here and in all the remaining part of the paper,  $o(1)$  stands for a quantity that tends to zero as the parameters  $h$  and  $\Delta t$  tend to zero, see Remark 9. Since  $\tilde{m}_{i,j}^n$  is bounded from below by a positive constant, we also deduce from (3.14) that

$$h^2 \Delta t \sum_{n=1}^{N_T} \sum_{i,j} \left| [D_h \tilde{u}^n]_{i,j} - [D_h u^n]_{i,j} \right|^\beta = o(1). \quad (4.3)$$

*Step 2: convergence of  $m_h$  to  $m$ .* We define the grid function  $e^\ell$  by  $e^\ell = m^\ell - \tilde{m}^\ell$ . Subtracting the second equation in (3.17) from the second equation in (2.14), multiplying the result by  $e_{i,j}^\ell$  and summing for all  $\ell = 0, \dots, n-1$  and all  $(i, j)$  leads to

$$\begin{aligned} & h^2 \sum_{\ell=0}^{n-1} (e^{\ell+1} - e^\ell, e^\ell)_2 + \Delta t h^2 \sum_{\ell=0}^{n-1} \nu(\Delta_h e^\ell, e^\ell)_2 \\ & - \Delta t h^2 \sum_{\ell=0}^{n-1} \sum_{i,j} [D_h e^\ell]_{i,j} \cdot (m_{i,j}^\ell g_q(x_{i,j}, [D_h u^{\ell+1}]_{i,j}) - \tilde{m}_{i,j}^\ell g_q(x_{i,j}, [D_h \tilde{u}^{\ell+1}]_{i,j}^{\ell+1})) \\ & = -h^2 \Delta t \sum_{\ell=0}^{n-1} (b^\ell, e^\ell)_2. \end{aligned}$$

This implies that

$$\begin{aligned} & \Delta t h^2 \sum_{\ell=0}^{n-1} \left( \left( \frac{e^{\ell+1} - e^\ell}{\Delta t}, e^\ell \right)_2 + \nu(\Delta_h e^\ell, e^\ell)_2 - \sum_{i,j} e_{i,j}^\ell [D_h e^\ell]_{i,j} \cdot g_q(x_{i,j}, [D_h u^{\ell+1}]_{i,j}) \right) \\ & = -h^2 \Delta t \sum_{\ell=0}^{n-1} (b^\ell, e^\ell)_2 \\ & - \Delta t h^2 \sum_{\ell=0}^{n-1} \sum_{i,j} \tilde{m}_{i,j}^\ell [D_h e^\ell]_{i,j} \cdot (g_q(x_{i,j}, [D_h \tilde{u}^{\ell+1}]_{i,j}) - g_q(x_{i,j}, [D_h u^{\ell+1}]_{i,j})). \end{aligned} \quad (4.4)$$

It is clear that

$$h^2 \Delta t \left| \sum_{\ell=0}^{n-1} (b^\ell, e^\ell)_2 \right| = o(1) \left( h^2 \Delta t \sum_{\ell=0}^{n-1} \|e^\ell\|_2^2 \right)^{\frac{1}{2}}.$$

From (3.9), we know that there exists an absolute constant  $c$  such that for all  $\eta > 0$ ,

$$\begin{aligned} & \left| \tilde{m}_{i,j}^\ell [D_h e^\ell]_{i,j} \cdot (g_q(x_{i,j}, [D_h \tilde{u}^{\ell+1}]_{i,j}) - g_q(x_{i,j}, [D_h u^{\ell+1}]_{i,j})) \right| \\ & \leq \tilde{m}_{i,j}^\ell \max(|p_{i,j}^{\ell+1}|^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|^{\beta-2}) \left( \frac{c}{\eta} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 + \eta |[D_h e^\ell]_{i,j}|^2 \right) \end{aligned}$$

where  $p_{i,j}^{\ell+1}, \tilde{p}_{i,j}^{\ell+1} \in (\mathbb{R}_+)^4$  are given by (3.13). Using the  $L^\infty$  bound on  $[D_h u^{\ell+1}]$  uniform w.r.t.  $\ell, h$  and  $\Delta t$  given in Theorem 2.1, we obtain that there exists a constant  $c$  such that for all  $\eta > 0$ ,

$$\begin{aligned} & \left| \tilde{m}_{i,j}^\ell [D_h e^\ell]_{i,j} \cdot (g_q(x_{i,j}, [D_h \tilde{u}^{\ell+1}]_{i,j}) - g_q(x_{i,j}, [D_h u^{\ell+1}]_{i,j})) \right| \\ & \leq \tilde{m}_{i,j}^\ell \left( \frac{c}{\eta} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^\beta + \eta |[D_h e^\ell]_{i,j}|^2 \right) \\ & \leq \tilde{m}_{i,j}^\ell \left( \frac{c}{\eta} |[D_h u^{\ell+1}]_{i,j} - [D_h \tilde{u}^{\ell+1}]_{i,j}|^\beta + \eta |[D_h e^\ell]_{i,j}|^2 \right). \end{aligned} \quad (4.5)$$

We shall also use the standard estimate:

$$\| [D_h e^\ell] \|_2^2 \leq -C(\Delta_h e^\ell, e^\ell)_2, \quad (4.6)$$

for a positive constant  $C$  independent of  $h$ . Finally, from (4.4), a very classical argument making use of (4.5), (4.6) and (4.3), the  $L^\infty$  bound on  $[D_h u^{\ell+1}]$  uniform w.r.t.  $\ell, h$  and  $\Delta t$ , and the fact that  $h \|e^{N_T}\|_2 = o(1)$ , leads to the estimate:

$$\max_{0 \leq \ell < N_T} h^2 \|e^\ell\|_2^2 + \Delta t h^2 \sum_{\ell=0}^{N_T-1} \| [D_h e^\ell] \|_2^2 = o(1). \quad (4.7)$$

We easily deduce from (4.7) the claim on the convergence of  $m_h$  to  $m$ .

*Step 3: convergence of  $u_h$  to  $u$ .* We have found that  $\max_{0 \leq \ell < N_T} h^2 \|m^\ell - \mathcal{I}_h(m(t_\ell, \cdot))\|_2^2 = o(1)$ . From  $(\Phi_{h4})$ , this implies that

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t} - \nu(\Delta_h u^{n+1})_{i,j} + g(x_{i,j}, [D_h u^{n+1}]_{i,j}) = (\Phi[m(t_{n+1}, \cdot)])(x_{i,j}) + o(1).$$

The uniform convergence of the piecewise linear functions  $u_h$  (defined by interpolating the values  $u_{i,j}^n$ ) to  $u$  is obtained from classical results on the approximation of Bellman equations by consistent and monotone schemes: one may proceed by induction on  $n$  and at each step use comparison arguments for  $u^{n+1}$  and  $\tilde{u}^{n+1}$ , thanks to the monotonicity of  $g$ , see for example [2], proof of Theorem 8. No relation between  $h$  and  $\Delta t$  is needed because the scheme is implicit w.r.t.  $u$ . From this and (4.3), we also deduce the convergence of  $u_h$  to  $u$  in  $L^\beta(0, T; W^{1,\beta}(\mathbb{T}^2))$ .  $\square$

**Remark 10.** *In the present case, additional assumptions on the order of the discrete scheme should lead to error estimates. We will not discuss this matter.*

## 4.2. The case when $1 < \beta < 2$ .

**THEOREM 4.2.** *With the standing assumptions stated at the beginning of § 4, we choose  $\beta$  such that  $1 < \beta < 2$ .*

*Let  $u_h$  (resp.  $m_h$ ) be the piecewise trilinear function in  $\mathcal{C}([0, T] \times \mathbb{T}^2)$  obtained by interpolating the values  $u_{i,j}^n$  (resp.  $m_{i,j}^n$ ) at the nodes of the space-time grid. The functions  $u_h$  converge uniformly and in  $L^2(0, T; W^{1,2}(\mathbb{T}^2))$  to  $u$  as  $h$  and  $\Delta t$  tend to 0. The functions  $m_h$  converge to  $m$  in  $L^2((0, T) \times \mathbb{T}^2)$  as  $h$  and  $\Delta t$  tend to 0.*

*Proof.*

*Strategy.* The proof is organized in three steps

1. The first step consists of proving the convergence to zero of a discrete seminorm of the error on  $u$  thanks to the fundamental identity (4.1)
2. In the present case, it is no longer possible to use (3.9); hence, the convergence of  $m_h$  to  $m$  will not be obtained by subtracting the discrete Fokker-Planck equations in (2.14) and in (3.17) from each other. Instead, Step 2 is devoted to a priori estimates on  $m_h$



3. The third step consists of showing the convergence of  $m_h$  to  $m$  by using compactness properties stemming from the above mentioned a priori estimates and again the fundamental identity (4.1). The convergence of  $u_h$  follows.

*Step 1: convergence to zero of a discrete semi-norm of the error on  $u$ .* We start from (4.1) where  $a$  and  $b$  are the same consistency errors (with the same bounds) as in the previous paragraph; using (3.8), this implies that

$$\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} \tilde{m}_{i,j}^\ell \mathbf{1}_{\{p_{i,j}^{\ell+1} + \tilde{p}_{i,j}^{\ell+1} \neq 0\}} \min(|p_{i,j}^{\ell+1}|_\infty^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|_\infty^{\beta-2}) |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 = o(1), \quad (4.8)$$

where  $p_{i,j}^{\ell+1}, \tilde{p}_{i,j}^{\ell+1} \in (\mathbb{R}_+)^4$  are given by (3.13). From the a priori estimates on  $\| [Du^{\ell+1}] \|_\infty$  and  $\| [D\tilde{u}^{\ell+1}] \|_\infty$ , the inequalities  $1 < \beta < 2$  and the bound from below on  $m$ , there exists a positive constant  $c$  such that

$$\tilde{m}_{i,j}^\ell \min(|p_{i,j}^{\ell+1}|_\infty^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|_\infty^{\beta-2}) \geq c.$$

Hence

$$\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 = o(1). \quad (4.9)$$

It is easy to see from the periodicity that (4.9) implies

$$\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} |[D_h u^{\ell+1}]_{i,j} - [D_h \tilde{u}^{\ell+1}]_{i,j}|^2 = o(1). \quad (4.10)$$

*Step 2: a priori estimates on  $m_h$ .* Multiplying the second equation in (2.14) by  $m_{i,j}^n$ , summing with respect to the indices  $i$  and  $j$ , and using the identity  $z(z-y) = \frac{1}{2}(z^2 - y^2 + (z-y)^2)$ , we obtain that

$$\begin{aligned} & \frac{1}{2} (\|m^n\|_2^2 - \|m^{n+1}\|_2^2 + \|m^n - m^{n+1}\|_2^2) + \nu \Delta t \|\nabla_h m^n\|_2^2 \\ &= -\Delta t \sum_{i,j} m_{i,j}^n g_q(x_{i,j}, [D_h u^{n+1}]_{i,j}) \cdot [D_h m^n]_{i,j}. \end{aligned}$$

Using the continuity of  $g_q$  and the a priori bound on  $\|D_h u^{n+1}\|_{L^\infty(\mathbb{T}_h^2)}$  leads to the existence of a positive constant  $c$  (which depends on  $\nu$  and on the uniform bound on  $\|D_h u^{n+1}\|_{L^\infty(\mathbb{T}_h^2)}$ ) such that

$$\|m^n\|_2^2 + \nu \Delta t \|\nabla_h m^n\|_2^2 \leq \|m^{n+1}\|_2^2 + c \Delta t \|m^n\|_2^2.$$

A discrete Gronwall estimate then leads to the existence of a positive constant  $C$  such that

$$\max_n \|m^n\|_2^2 + \Delta t \sum_{n=0}^{N_T-1} \|[D_h m^n]\|_2^2 \leq C \|m^{N_T}\|_2^2. \quad (4.11)$$

Similarly, from (2.14), we see that for all grid function  $(r_{i,j})$  on  $\mathbb{T}_h$ ,

$$\begin{aligned} & \sum_{n=0}^{N_T-1} \sum_{i,j} r_{i,j}^n \frac{m_{i,j}^{n+1} - m_{i,j}^n}{\Delta t} \\ &= \nu \sum_{n=0}^{N_T-1} \sum_{i,j} (\nabla_h m^n)_{i,j} \cdot (\nabla_h r^n)_{i,j} + \sum_{n=0}^{N_T-1} \sum_{i,j} m_{i,j}^n g_q(x_{i,j}, [D_h u^{n+1}]_{i,j}) \cdot [D_h r^n]_{i,j}. \end{aligned} \quad (4.12)$$

From the a priori bound on  $\|D_h u^{n+1}\|_{L^\infty(\mathbb{T}_h^2)}$  and (4.11), we infer that

$$\sup_{(r^n)_n} \frac{\sum_{n=0}^{N_T-1} \sum_{i,j} r_{i,j}^n \frac{m_{i,j}^{n+1} - m_{i,j}^n}{\Delta t}}{\sqrt{\sum_{n=0}^{N_T-1} \sum_{i,j} |[D_h r^n]_{i,j}|^2}} \leq C.$$

*Step 3: convergence of  $m_h$  and  $u_h$ .* From (4.11), we see that the family of functions  $(m_h)$  is bounded in  $L^2(0, T; H^1(\mathbb{T}^2))$ . Moreover, from (4.12) and the a priori bound on  $\|D_h u^{n+1}\|_{L^\infty(\mathbb{T}_h^2)}$ , we can use the same arguments as in e.g. [10] pages 855-858 and prove that the family of functions  $(m_h)$  has the following property: there exists a constant  $C$  such that

$$\|m_h(\cdot - \tau, \cdot) - m_h(\cdot, \cdot)\|_{L^2((\tau, T) \times \mathbb{T}^2)}^2 \leq C\tau h^2 \Delta t \sum_{n=0}^{N_T-1} \left( \sum_{i,j} |[D_h m^n]_{i,j}|^2 + \sum_{i,j} (m_{i,j}^n)^2 \right), \quad (4.13)$$

for all  $\tau \in (0, T)$ . Since the right hand side is bounded by  $C\tau h^2 \|m^{N_T}\|_2^2$ , Kolmogorov's theorem (see e.g. [6], [11], [10] page 833) implies that the family of functions  $(m_h)$  is relatively compact in  $L^2((0, T) \times \mathbb{T}^2)$ : we can extract a subsequence of parameters  $h$  and  $\Delta t$  tending to 0 such that  $m_h$  converges to  $\bar{m}$  strongly in  $L^2((0, T) \times \mathbb{T}^2)$ , and (4.13) holds for  $\bar{m}$ .

Therefore, from  $(\Phi_{h4})$ ,

$$\lim_{h, \Delta t \rightarrow 0} \Delta t \sum_{n=0}^{N_T} \|\Phi[\bar{m}(t_n, \cdot)] - \Phi_h[m^n]\|_{L^\infty(\mathbb{T}_h^2)}^2 = 0.$$

On the other hand, from (4.1), we see that

$$\Delta t h^2 \sum_{n=0}^{N_T-1} (\Phi_h[m^n] - \Phi_h[\tilde{m}^n], m^n - \tilde{m}^n)_2 = o(1).$$

Then, using (2.10), we deduce from the previous two formulas that

$$\lim_{\Delta t \rightarrow 0} \Delta t \sum_{n=0}^{N_T} \int_{\mathbb{T}_h^2} (\Phi[m(t_n, \cdot)](x) - \Phi[\bar{m}(t_n, \cdot)](x)) (m(t_n, x) - \bar{m}(t_n, x)) dx = 0,$$

which implies that

$$\int_0^T \int_{\mathbb{T}_h^2} (\Phi[m(t, \cdot)](x) - \Phi[\bar{m}(t, \cdot)](x)) (m(t, x) - \bar{m}(t, x)) dx dt = 0,$$

The monotonicity of  $\Phi$ , see (1.5), then implies that  $m = \bar{m}$ . From the uniqueness of the limit  $\bar{m}$ , we have proven that the whole family  $m_h$  converges to  $m$  in  $L^2((0, T) \times \mathbb{T}^2)$  as  $h$  and  $\Delta t$  go to zero.

We conclude as in Step 3 of the proof of Theorem 4.1 for the convergence of  $u_h$  to  $u$  in  $C^0[0, T] \times \mathbb{T}^2$  and in  $L^2(0, T; W^{1,2}(\mathbb{T}^2))$ .  $\square$

## 5. Study of the convergence in the case when $\Phi$ is a local operator.

**5.1. A priori estimates for (2.14)-(2.15) with local operators  $\Phi_h$ .** We have a result similar to Theorem 2.7 in [21]:

LEMMA 5.1. *Assume that  $0 \leq m_T(x) \leq \bar{m}_T$  and that  $u_0$  is a continuous function. If  $g$  is given by (2.3)-(2.4),  $(\Phi_h[m])_{i,j} = F(m_{i,j})$ , where  $(F_1)$   $F$  is a  $C^0$  function defined on  $[0, \infty)$*

(F<sub>2</sub>) there exist three constants  $\delta > 0$  and  $\gamma > 1$  and  $C_1 \geq 0$  such that

$$mF(m) \geq \delta|F(m)|^\gamma - C_1, \quad \forall m \geq 0,$$

then there exists two constants  $c$  and  $C > 0$  such that

- $u_{i,j}^n \geq c$ , for all  $n, i$  and  $j$
- 

$$h^2 \Delta t \sum_{n=1}^{N_T} \sum_{i,j} \left| [D_h u^n]_{i,j} \right|^\beta + h^2 \Delta t \sum_{n=0}^{N_T-1} \sum_{i,j} |F(m_{i,j}^n)|^\gamma \leq C \quad (5.1)$$

- 

$$\max_{0 \leq n \leq N_T} h^2 \sum_{i,j} |u_{i,j}^n| \leq C \quad (5.2)$$

- Finally, let us call  $Z^n$  the sum  $z^n = h^2 \sum_{i,j} u_{i,j}^n$  and  $z_h$  the piecewise linear function obtained by interpolating the values  $z^n$  at the points  $(t_n)$ : the family of functions  $(z_h)$  is bounded in  $W^{1,1}(0, T)$  by a constant independent of  $h$  and  $\Delta t$ .

*Proof.* From the two assumptions on  $F$ , we deduce that  $\underline{F} \equiv \inf_{m \in \mathbb{R}_+} F(m)$  is a real number and that  $\underline{F} = \min_{m \geq 0} F(m)$ . Note that  $\underline{F} = F(0)$  if  $F$  is nondecreasing. A standard comparison argument shows that

$$u_{i,j}^n \geq \min_{x \in \mathbb{T}^2} u_0(x) + \left( \underline{F} - \max_{x \in \mathbb{T}^2} \mathcal{H}(x) \right) t_n \geq \min_{x \in \mathbb{T}^2} u_0(x) - T \left( \underline{F} - \max_{x \in \mathbb{T}^2} \mathcal{H}(x) \right)^-,$$

so  $u_{i,j}^n$  is bounded from below by a constant independent of  $h$  and  $\Delta t$ . Consider  $\tilde{u}_{i,j}^n = n \Delta t F(\bar{m}_T)$  and  $\tilde{m}_{i,j}^n = \bar{m}_T$  for all  $i, j, n$ . We have

$$\begin{cases} \frac{\tilde{u}_{i,j}^{n+1} - \tilde{u}_{i,j}^n}{\Delta t} - \nu(\Delta_h \tilde{u}^{n+1})_{i,j} + g(x_{i,j}, [D_h \tilde{u}^{n+1}]_{i,j}) & = F(\bar{m}_T) + \mathcal{H}(x_{i,j}), \\ \frac{\tilde{m}_{i,j}^{n+1} - \tilde{m}_{i,j}^n}{\Delta t} + \nu(\Delta_h \tilde{m}^n)_{i,j} + \mathcal{I}_{i,j}(\tilde{u}^{n+1}, \tilde{m}^n) & = 0. \end{cases}$$

Identity (3.20) becomes

$$\begin{aligned} & h^2 \Delta t \left( \mathcal{G}(M, U, \tilde{U}) + \mathcal{G}(\tilde{M}, \tilde{U}, U) + \sum_{n=0}^{N_T-1} \sum_{i,j} (F(m_{i,j}^n) - F(\bar{m}_T))(m_{i,j}^n - \bar{m}_T) \right) \\ & = h^2 \Delta t \sum_{n=0}^{N_T-1} \sum_{i,j} \mathcal{H}(x_{i,j})(m_{i,j}^n - \tilde{m}_{i,j}^n) \\ & \quad + h^2 (m^{N_T} - \bar{m}_T, u^{N_T} - T F(\bar{m}_T))_2 - h^2 (m^0 - \bar{m}_T, u^0)_2. \end{aligned} \quad (5.3)$$

Note that

$$\mathcal{G}(M, U, \tilde{U}) = \mathcal{G}(M, U, 0) \quad \text{and} \quad \mathcal{G}(\tilde{M}, \tilde{U}, U) = \mathcal{G}(\tilde{M}, 0, U),$$

because  $\tilde{u}_{i,j}^n$  does not depend on  $i, j$ . On the other hand,

1. Since the function  $\mathcal{H}$  is bounded, and  $m^n$  is a discrete probability density, there exists a constant  $C$  such that

$$\left| h^2 \Delta t \sum_{n=0}^{N_T-1} \sum_{i,j} \mathcal{H}(x_{i,j})(m_{i,j}^n - \tilde{m}_{i,j}^n) \right| \leq C.$$

2. Since  $m^{N_T} - \bar{m}_T$  is nonpositive with a bounded mass, and since  $u^n$  is bounded from below by a constant, there exists a constant  $C$  such that

$$h^2(m^{N_T} - \bar{m}_T, u^{N_T} - T F(\bar{m}_T))_2 \leq C.$$

3. Since  $u_0$  is continuous on  $\mathbb{T}^2$  and  $m_0$  is a discrete probability density, there exists a constant  $C$  such that

$$-h^2(m^0 - \bar{m}_T, u^0)_2 \leq C.$$

4. Finally, we know that

$$\begin{aligned} & (F(m_{i,j}^n) - F(\bar{m}_T))(m_{i,j}^n - \bar{m}_T) \\ & \geq \delta |F(m_{i,j}^n)|^\gamma - C_1 - \bar{m}_T F(m_{i,j}^n) - m_{i,j}^n F(\bar{m}_T) + \bar{m}_T F(\bar{m}_T). \end{aligned}$$

Moreover, since  $\gamma > 1$ , there exists two constants  $c = \frac{\delta}{2}$  and  $C$  such that  $\delta |F(m_{i,j}^n)|^\gamma - \bar{m}_T F(m_{i,j}^n) \geq c |F(m_{i,j}^n)|^\gamma - C$ . Since  $m^n \in \mathcal{K}_h$ , summing yields that for a possibly different constant  $C$ ,

$$h^2 \sum_{i,j} (F(m_{i,j}^n) - F(\bar{m}_T))(m_{i,j}^n - \bar{m}_T) \geq ch^2 \sum_{i,j} |F(m_{i,j}^n)|^\gamma - C.$$

In the case  $\beta \geq 2$ , we get (5.1) from (5.3), from (3.14) and from the four points above. In the case  $1 < \beta < 2$ , we get (5.1) from (5.3), from Remark 7 and from the four points above.

Finally, summing the first equation in (2.14) for all  $i, j$ ,  $0 \leq \ell < n$  one gets that

$$h^2 \sum_{i,j} u_{i,j}^n + h^2 \Delta t \sum_{\ell=0}^{q-1} \sum_{i,j} g(x_{i,j}, [D_h u^{\ell+1}]_{i,j}) = h^2 \Delta t \sum_{\ell=0}^{n-1} \sum_{i,j} F(m_{i,j}^\ell) + h^2 \sum_{i,j} u_{i,j}^0.$$

Using (5.1), we get that there exists a constant  $C$  such that

$$h^2 \sum_{i,j} u_{i,j}^n \leq C,$$

and since  $u_{i,j}^n$  is bounded from below by a constant, we get (5.2).

Finally, let us estimate  $z^n = h^2 \sum_{i,j} u_{i,j}^n$ ; summing the first equations in (2.14) for all  $i, j$ , we obtain that

$$\frac{z^{n+1} - z^n}{\Delta t} = -h^2 \sum_{i,j} (g(x_{i,j}, [D_h u^{n+1}]_{i,j}) + F(m_{i,j}^n)).$$

The a priori estimate (5.1) implies that  $\Delta t \sum_{n=0}^{N_T-1} |z^{n+1} - z^n|$  is bounded by a constant. This implies that the piecewise linear function  $z_h$  obtained by interpolating the values  $z^n$  at the points  $(t_n)$  is bounded in  $W^{1,1}(0, T)$  by a constant independent of  $h$  and  $\Delta t$ .  $\square$

**5.2. A convergence theorem.** The case when  $\Phi$  is a local operator, i.e.  $\Phi[m](x) = F(m(x))$  brings additional difficulties, because there is no a priori Lipschitz estimates on  $u_h$ : such estimates were used several times in the proofs of Theorems 4.1 and 4.2.

For simplicity, we are going to make the assumption that the continuous problem has a classical solution: existence of a classical solution can be true for local operators  $\Phi$ : for example, it has been proved in [9] that if  $\beta = 2$ , and  $F$  is  $C^1$  and bounded from below, and if the functions  $u_0$  and  $m_T$  are  $C^2$  then there is a classical solution.

Standing assumptions (in § 5.2 and § 5.3).

- The Hamiltonian is of the form (1.7) and the function  $x \rightarrow \mathcal{H}(x)$  is  $\mathcal{C}^1$  on  $\mathbb{T}^2$
- The functions  $u_0$  and  $m_T$  are smooth, and  $m_T \in \mathcal{K}$  is bounded from below by a positive number
- $(F_1)$  and  $(F_2)$  hold and there exist three positive constants  $\underline{\delta}$ ,  $\eta_1 > 0$  and  $0 < \eta_2 < 1$  such that  $F'(m) \geq \underline{\delta} \min(m^{\eta_1}, m^{-\eta_2})$
- The numerical Hamiltonian is given by (2.3)-(2.4)

THEOREM 5.2. *We make the standing assumptions stated above and we assume furthermore that there is a unique classical solution  $(u, m)$  of (1.1)-(1.3) such that  $m > 0$ .*

Let  $u_h$  (resp.  $m_h$ ) be the piecewise trilinear function in  $\mathcal{C}([0, T] \times \mathbb{T}^2)$  obtained by interpolating the values  $u_{i,j}^n$  (resp  $m_{i,j}^n$ ) at the nodes of the space-time grid. The functions  $u_h$  converge in  $L^\beta(0, T; W^{1,\beta}(\mathbb{T}^2))$  to  $u$  as  $h$  and  $\Delta t$  tend to 0. The functions  $m_h$  converge to  $m$  in  $L^{2-\eta_2}((0, T) \times \mathbb{T}^2)$  as  $h$  and  $\Delta t$  go to 0.

*Proof.* Call  $\bar{m} = \max m(t, x)$  and  $0 < \underline{m} = \min m(t, x)$ .

*Strategy.* The proof is organized in three steps

1. The first step consists of proving (4.3). This will more tricky if  $1 < \beta < 2$ .
2. The second step consists of proving the convergence of  $m_h$  to  $m$  by taking advantage on the assumptions on  $F$  and  $F'$ .
3. The third step consists of proving the convergence of  $u_h$  to  $u$ , by passing to the limit in the Bellman equation: in particular, Vitali's theorem will be used for passing to the limit in the nonlinear term  $F(m)$ .

*Step 1: convergence to zero of a discrete semi-norm of the error on  $u$ .* We start from (4.1) where  $a$  and  $b$  are the same consistency errors (with the same bounds) as in § 4. From Lemma 5.1, the a priori bound (5.2) holds. This implies that  $\lim_{h, \Delta t \rightarrow 0} h^2 \max_n |(b^{n-1}, u^n - \tilde{u}^n)_2| = 0$ . From the fact that  $m^n \in \mathcal{K}_h$ , we also get that  $\lim_{h, \Delta t \rightarrow 0} h^2 \max_n |(a^n, m^n - \tilde{m}^n)_2| = 0$ .

Therefore, if  $\beta \geq 2$ , we obtain (4.2) and (4.3).

If  $1 < \beta < 2$ , we are going to prove that (4.3) also holds: we have

$$\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} \tilde{m}_{i,j}^\ell \mathbf{1}_{p_{i,j}^{\ell+1} + \tilde{p}_{i,j}^{\ell+1} \neq 0} \min(|p_{i,j}^{\ell+1}|_\infty^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|_\infty^{\beta-2}) |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 = o(1),$$

where  $p_{i,j}^{\ell+1}, \tilde{p}_{i,j}^{\ell+1} \in (\mathbb{R}_+)^4$  are given by (3.13). Let us define for brevity

$$V_{i,j}^{\ell+1} = 2^{\beta-3} \beta(\beta-1) \mathbf{1}_{p_{i,j}^{\ell+1} + \tilde{p}_{i,j}^{\ell+1} \neq 0} \min(|p_{i,j}^{\ell+1}|_\infty^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|_\infty^{\beta-2}) |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2.$$

Assume that  $\tilde{p}_{i,j}^{\ell+1} \neq 0$ . We have

$$V_{i,j}^{\ell+1} \geq 2^{\beta-3} \beta(\beta-1) \min(|p_{i,j}^{\ell+1}|_\infty^{\beta-2}, |\tilde{p}_{i,j}^{\ell+1}|_\infty^{\beta-2}) |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2.$$

- If  $|p_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|$ , then

$$V_{i,j}^{\ell+1} \geq 2^{\beta-3} \beta(\beta-1) |\tilde{p}_{i,j}^{\ell+1}|^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2.$$

- If  $|\tilde{p}_{i,j}^{\ell+1}| \leq |p_{i,j}^{\ell+1}|$  and  $|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|/2$ , then

$$\begin{aligned} V_{i,j}^{\ell+1} &\geq 2^{\beta-3} \beta(\beta-1) |p_{i,j}^{\ell+1}|^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &\geq 2^{\beta-3} \beta(\beta-1) (3|\tilde{p}_{i,j}^{\ell+1}|/2)^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &= \frac{3^{\beta-2}}{2} \beta(\beta-1) |\tilde{p}_{i,j}^{\ell+1}|^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2. \end{aligned}$$

- If  $|\tilde{p}_{i,j}^{\ell+1}| \leq |p_{i,j}^{\ell+1}|$  and  $|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|/2$ , then

$$\begin{aligned}
V_{i,j}^{\ell+1} &\geq 2^{\beta-3}\beta(\beta-1)|p_{i,j}^{\ell+1}|^{\beta-2}|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\
&\geq 2^{\beta-3}\beta(\beta-1)(|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| + |\tilde{p}_{i,j}^{\ell+1}|)^{\beta-2}|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\
&\geq 2^{\beta-3}\beta(\beta-1)(3|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|)^{\beta-2}|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\
&= \frac{6^{\beta-2}}{2}\beta(\beta-1)|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^\beta.
\end{aligned}$$

If  $\tilde{p}_{i,j}^{\ell+1} = 0$  and  $p_{i,j}^{\ell+1} \neq 0$ , then  $V_{i,j}^{\ell+1} \geq 2^{\beta-3}\beta(\beta-1)|p_{i,j}^{\ell+1}|^\beta$ .

From the observation above, and since  $\tilde{m}_{i,j}^\ell \geq \underline{m}$ ,

$$\begin{aligned}
\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} &\left( \begin{aligned} &1_{\{|p_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|\}} |\tilde{p}_{i,j}^{\ell+1}|^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &+ 1_{\{|p_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|\}} 1_{\{|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|/2\}} |\tilde{p}_{i,j}^{\ell+1}|^{\beta-2} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &+ 1_{\{|p_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|\}} 1_{\{|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|/2\}} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^\beta. \end{aligned} \right) \\
&= o(1).
\end{aligned}$$

Furthermore, looking at the definition of  $\tilde{p}_{i,j}^{\ell+1}$  in (3.13), we see that  $|\tilde{p}_{i,j}^{\ell+1}|$  is smaller than an absolute constant times the  $\mathcal{C}^1$  norm of  $u$  (recall that  $(u, m)$  is the classical solution of (1.1)-(1.3)). Hence, since  $1 < \beta < 2$ ,  $|\tilde{p}_{i,j}^{\ell+1}|^{\beta-2}$  is bounded from below by a constant independent of  $h, \Delta t, i, j, \ell$ .

Thus

$$\Delta t h^2 \sum_{\ell=0}^{N_T-1} \sum_{i,j} \left( \begin{aligned} &1_{\{|p_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|\}} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &+ 1_{\{|p_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|\}} 1_{\{|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \leq |\tilde{p}_{i,j}^{\ell+1}|/2\}} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^2 \\ &+ 1_{\{|p_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|\}} 1_{\{|p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}| \geq |\tilde{p}_{i,j}^{\ell+1}|/2\}} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^\beta. \end{aligned} \right) = o(1).$$

Using a Hölder inequality, we deduce that

$$h^2 \Delta t \sum_{\ell=1}^{N_T} \sum_{i,j} |p_{i,j}^{\ell+1} - \tilde{p}_{i,j}^{\ell+1}|^\beta = o(1),$$

and finally (4.3).

*Step 2: convergence of  $m_h$  to  $m$ .* We also obtain from (4.1) that

$$h^2 \Delta t \sum_{n=0}^{N_T-1} \sum_{i,j} (F(m_{i,j}^n) - F(\tilde{m}_{i,j}^n))(m_{i,j}^n - \tilde{m}_{i,j}^n) = o(1). \quad (5.4)$$

We split the sum w.r.t.  $(i, j)$  in the left hand side of (5.4) into

$$\begin{aligned}
S_1^n &= \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^-)^2 \int_0^1 F'(\tilde{m}_{i,j}^n + t(m_{i,j}^n - \tilde{m}_{i,j}^n)) dt, \\
S_2^n &= \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^+)^2 \int_0^1 F'(\tilde{m}_{i,j}^n + t(m_{i,j}^n - \tilde{m}_{i,j}^n)) dt.
\end{aligned}$$

Call  $\bar{m} = \max m(t, x)$  and  $\underline{m} = \min m(t, x) > 0$ ; there exists a positive number  $c$  depending on  $\underline{m}$  and  $\bar{m}$  but independent of  $h$  and  $\Delta t$ , and  $(i, j, n)$  such that

$$\begin{aligned}
S_1^n &\geq c \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^-)^2 \int_0^1 (\tilde{m}_{i,j}^n + t(m_{i,j}^n - \tilde{m}_{i,j}^n))^{\eta_1} dt \\
&= \frac{c}{\eta_1 + 1} \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^-)((\tilde{m}_{i,j}^n)^{\eta_1+1} - (m_{i,j}^n)^{\eta_1+1}) \\
&\geq \frac{c}{\eta_1 + 1} \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^-)^2 (\tilde{m}_{i,j}^n)^{\eta_1}.
\end{aligned}$$

The latter inequality comes from the nondecreasing character of the function  $\chi : [0, y] \mapsto \mathbb{R}$ ,  $\chi(z) = \frac{y^{\eta_1+1} - z^{\eta_1+1}}{y-z}$ . Thus,  $\chi(z) \geq \chi(0) = y^{\eta_1}$ . Hence, there exists a constant  $c$  depending on the bounds on the density  $m$  solution of (1.1)-(1.3) but not on  $h$  and  $\Delta t$ , and  $(i, j, n)$  such that

$$S_1^n \geq c \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^-)^2.$$

On the other hand

$$\begin{aligned} S_2^n &\geq c \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^+)^2 \int_0^1 (\tilde{m}_{i,j}^n + t(m_{i,j}^n - \tilde{m}_{i,j}^n))^{-\eta_2} dt \\ &= \frac{c}{1-\eta_2} \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^+) ((m_{i,j}^n)^{1-\eta_2} - (\tilde{m}_{i,j}^n)^{1-\eta_2}) \\ &\geq c \sum_{i,j} ((m_{i,j}^n - \tilde{m}_{i,j}^n)^+)^2 (m_{i,j}^n)^{-\eta_2}. \end{aligned}$$

But there exists a constant  $c$  such that for all  $y \in [\underline{m}, \bar{m}]$ : if  $z \geq y + 1$

$$(z-y)^2 z^{-\eta_2} \geq (z-y)^{2-\eta_2} \inf_{z \geq y+1} \frac{(z-y)^{\eta_2}}{z^{\eta_2}} \geq c(z-y)^{2-\eta_2},$$

and if  $y \leq z \leq y + 1$ ,

$$(z-y)^2 z^{-\eta_2} \geq c(z-y)^2.$$

Therefore there exists a constant  $c$  such that

$$\begin{aligned} &S_1^n + S_2^n \\ &\geq c \left( \sum_{i,j} (m_{i,j}^n - \tilde{m}_{i,j}^n)^2 \mathbf{1}_{\{m_{i,j}^n \leq \tilde{m}_{i,j}^n + 1\}} + \sum_{i,j} (m_{i,j}^n - \tilde{m}_{i,j}^n)^{2-\eta_2} \mathbf{1}_{\{m_{i,j}^n \geq \tilde{m}_{i,j}^n + 1\}} \right). \end{aligned}$$

Then (5.4) implies that

$$\lim_{h, \Delta t \rightarrow 0} h^2 \Delta t \sum_{n=0}^{N_T-1} \left( \sum_{i,j} (m_{i,j}^n - \tilde{m}_{i,j}^n)^2 \mathbf{1}_{\{m_{i,j}^n \leq \tilde{m}_{i,j}^n + 1\}} + \sum_{i,j} (m_{i,j}^n - \tilde{m}_{i,j}^n)^{2-\eta_2} \mathbf{1}_{\{m_{i,j}^n \geq \tilde{m}_{i,j}^n + 1\}} \right) = 0.$$

Then, a Hölder inequality leads to

$$\lim_{h, \Delta t \rightarrow 0} h^2 \Delta t \sum_{n=0}^{N_T-1} \sum_{i,j} |m_{i,j}^n - \tilde{m}_{i,j}^n|^{2-\eta_2} = 0.$$

*Step 3: convergence of  $u_h$  to  $u$ .* From the previous two steps, up to an extraction of a sequence,  $m_h \rightarrow m$  in  $L^{2-\eta_2}((0, T) \times \mathbb{T}^2)$  and almost everywhere in  $(0, T) \times \mathbb{T}^2$ ,  $\nabla u_h$  converges to  $\nabla u$  strongly in  $L^\beta((0, T) \times \mathbb{T}^2)$ . Moreover, from the last point in Lemma 5.1, the sequence of piecewise linear functions  $(z_h)$  on  $[0, T]$  obtained by interpolating the values  $z^n = h^2 \sum_{i,j} u_{i,j}^n$  at the points  $(t_n)$  is bounded in  $W^{1,1}(0, T)$ , so up to a further extraction of a subsequence, it converges to some function  $z$  in  $L^\beta(0, T)$ . As a result, there exists a function  $\psi$  of the variable  $t$  such that  $u_h \rightarrow u + \psi$  in  $L^\beta(0, T; W^{1,\beta}(\mathbb{T}^2))$ . We want to prove that  $\psi = 0$ .

From the a priori estimate (5.1), the sequence  $(F(m_h))$  is bounded in  $L^\gamma((0, T) \times \mathbb{T}^2)$  for some  $\gamma > 1$ , which implies that it is uniformly integrable on  $(0, T) \times \mathbb{T}^2$ . On the

other hand,  $F(m_h)$  converges almost everywhere to  $F(m)$ . Therefore, from Vitali's theorem, see e.g. [24],  $F(m_h)$  converges to  $F(m)$  in  $L^1((0, T) \times \mathbb{T}^2)$ , (in fact, it is also possible to show that  $F(m_h)$  converges to  $F(m)$  in  $L^q((0, T) \times \mathbb{T}^2)$  for all  $q \in [1, \gamma)$ ).

It is then possible to pass to the limit in the discrete Bellman equation, which yields that  $\frac{\partial \psi}{\partial t} = 0$  in the sense of distributions in  $(0, T)$ . Hence  $\psi$  is a constant.

We are left with proving that  $\psi$  is indeed 0. For that, we split  $\frac{\partial u_h}{\partial t}$  into the sum  $\mu_h + \eta_h$ , where

- $\mu_h|_{t \in (t_n, t_{n+1}]}$  is constant w.r.t.  $t$  and piecewise linear w.r.t.  $x$ , and takes the value  $\nu(\Delta_h u^{n+1})_{i,j}$  at the node  $\xi_{i,j}$
- $\eta_h$  is the remainder, see (2.11). This term is constructed by interpolating the values  $F(m_{i,j}^n) - g(x_{i,j}, [D_h u^{n+1}]_{i,j})$  at the grid nodes.

From the observations above,  $(\eta_h)$  converges in  $L^1((0, T) \times \mathbb{T}^2)$ , (because of the strong convergence of  $\nabla u_h$  and of  $F(m_h)$ ). On the other hand, from (4.3), it is not difficult to see that  $(\mu_h)$  is a Cauchy sequence in  $L^\beta(0, T; (W^{s, \beta/(\beta-1)}(\mathbb{T}^2))')$  for  $s$  large enough, (here  $(W^{s, \beta/(\beta-1)}(\mathbb{T}^2))'$  is the topological dual of  $W^{s, \beta/(\beta-1)}(\mathbb{T}^2)$ ). Hence,  $(\frac{\partial u_h}{\partial t})$  converges in  $L^1(0, T; (W^{s, \beta/(\beta-1)}(\mathbb{T}^2))')$ . Therefore,  $u_h$  converges in  $C^0([0, T]; (W^{s, \beta/(\beta-1)}(\mathbb{T}^2))')$ ; since  $(u_h(t=0))$  converges to  $u_0$ , we see that  $\psi = 0$ . This implies that the extracted sequence  $u_h$  converges to  $u$  in  $L^\beta(0, T; W^{1, \beta}(\mathbb{T}^2))$ . Since the limit is unique, the whole family  $(u_h)$  converges to  $u$  in  $L^\beta(0, T; W^{1, \beta}(\mathbb{T}^2))$  as  $h$  and  $\Delta t$  go to 0.  $\square$

**5.3. The stationary case.** The following steady state version of (1.1)-(1.3) arises when mean field games with infinite horizon are considered (ergodic problem):

$$-\nu \Delta u(x) + H(x, \nabla u(x)) + \lambda = \Phi[m(\cdot)](x), \quad \text{in } \mathbb{T}^2, \quad (5.5)$$

$$-\nu \Delta m(x) - \operatorname{div} \left( m \frac{\partial H}{\partial p}(\cdot, \nabla u) \right) (x) = 0, \quad \text{in } \mathbb{T}^2, \quad (5.6)$$

$$\int_{\mathbb{T}^2} m = 1 \quad \text{and} \quad m \geq 0 \quad \text{in } \mathbb{T}^2. \quad (5.7)$$

The unknowns in (5.5)-(5.7) are the density  $m$ , the function  $u$  and the scalar  $\lambda$ . The scalar  $\lambda$  is the limit of the ergodic costs of infinite horizon Nash equilibria with  $N$  agents when  $N$  tends to  $\infty$ , see [19, 21]. Clearly, if  $(u, m, \lambda)$  solves (5.5)-(5.7), then for all real number  $c \neq 0$ ,  $(u + c, m, \lambda)$  is another solution. Hence, it is possible to impose an additional normalization on  $u$ , for example  $\int_{\mathbb{T}^2} u = 0$ .

It can be proved that, if  $(F_2)$  holds with  $\gamma > 2$  (2 is the space dimension) and  $F$  is nondecreasing, then (5.5)-(5.7) has a classical solution for any  $\beta > 1$ , by using the weak Bernstein method studied in [22].

We give the counterpart of Theorem 5.2 in the stationary case. We omit the proof because it is quite similar to that of Theorem 5.2.

**THEOREM 5.3.** *Let us make the standing assumptions stated at the beginning of § 5.2 and the further assumption that there is unique classical solution  $(u, m, \lambda)$  of (5.5)-(5.7) such that  $m > 0$  and  $\int_{\mathbb{T}^2} u(x) dx = 0$ .*

*Let  $u_h$  (resp.  $m_h$ ) be the piecewise bilinear function in  $\mathcal{C}(\mathbb{T}^2)$  obtained by interpolating the values  $u_{i,j}$  (resp  $m_{i,j}$ ) at the nodes of  $\mathbb{T}_h^2$ , where  $((u_{i,j}), (m_{i,j}), \lambda_h)$  is the unique solution of the following system:*

*for all  $0 \leq i, j < N_h$ ,  $m_{i,j} \geq 0$ ,*

$$\begin{cases} -\nu(\Delta_h u)_{i,j} + g(x_{i,j}, [D_h u]_{i,j}) + \lambda_h & = F(m_{i,j}), \\ -\nu(\Delta_h m)_{i,j} - \mathcal{T}_{i,j}(u, m) & = 0, \\ h^2 \sum_{i,j} u_{i,j} = 0, \quad h^2 \sum_{i,j} m_{i,j} & = 1. \end{cases} \quad (5.8)$$

*As  $h$  goes to 0, the functions  $u_h$  converge in  $W^{1, \beta}(\mathbb{T}^2)$  to  $u$ , the functions  $m_h$  converge to  $m$  in  $L^{2-\eta_2}(\mathbb{T}^2)$ , and  $\lambda_h$  tends to  $\lambda$ .*



### Appendix A. Proofs of some technical lemmas.

*Proof of Lemma 3.1.* We focus on point 2, since point 1 is obtained by straightforward calculus.

For all  $r \in \mathbb{R}^4$ , we have

$$\begin{aligned} g_q(x, q) \cdot r &= G_p(p) \cdot (-1_{q_1 < 0} r_1, 1_{q_2 > 0} r_2, -1_{q_3 < 0} r_3, 1_{q_4 > 0} r_4) \\ &= \beta |p|^{\beta-2} (-p_1 r_1 + p_2 r_2 - p_3 r_3 + p_4 r_4) \end{aligned}$$

Hence,

$$\begin{aligned} &-g_q(x, q) \cdot (\tilde{q} - q) \\ &= -G_p(p) \cdot \left( -1_{q_1 < 0} (\tilde{q}_1 - q_1), 1_{q_2 > 0} (\tilde{q}_2 - q_2), -1_{q_3 < 0} (\tilde{q}_3 - q_3), 1_{q_4 > 0} (\tilde{q}_4 - q_4) \right) \\ &= -\beta |p|^{\beta-2} \left( -p_1 (\tilde{q}_1 - q_1) + p_2 (\tilde{q}_2 - q_2) - p_3 (\tilde{q}_3 - q_3) + p_4 (\tilde{q}_4 - q_4) \right). \end{aligned}$$

But

$$\begin{aligned} -p_1 (\tilde{q}_1 - q_1) &= p_1 (\tilde{p}_1 - p_1) - p_1 \tilde{q}_1^+ \leq p_1 (\tilde{p}_1 - p_1), \\ p_2 (\tilde{q}_2 - q_2) &= p_2 (\tilde{p}_2 - p_2) - p_2 \tilde{q}_2^- \leq p_2 (\tilde{p}_2 - p_2), \\ -p_3 (\tilde{q}_3 - q_3) &= p_3 (\tilde{p}_3 - p_3) - p_3 \tilde{q}_3^+ \leq p_3 (\tilde{p}_3 - p_3), \\ p_4 (\tilde{q}_4 - q_4) &= p_4 (\tilde{p}_4 - p_4) - p_4 \tilde{q}_4^- \leq p_4 (\tilde{p}_4 - p_4). \end{aligned}$$

Therefore

$$-g_q(x, q) \cdot (\tilde{q} - q) \geq -\beta |p|^{\beta-2} p \cdot (\tilde{p} - p) = -G_p(p) \cdot (\tilde{p} - p), \quad (\text{A.1})$$

and (3.5) follows immediately.

*Proof of Lemma 3.2.*

1.(a) If  $\beta \geq 2$ , then

$$\begin{aligned} &g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \\ &\geq G(\tilde{p}) - G(p) - G_p(p) \cdot (\tilde{p} - p) \\ &= \int_0^1 (1-s) G_{pp}(s\tilde{p} + (1-s)p) (\tilde{p} - p) \cdot (\tilde{p} - p) ds \\ &\geq \beta |p - \tilde{p}|^2 \int_0^1 (1-s) |s\tilde{p} + (1-s)p|^{\beta-2} ds, \end{aligned}$$

where the first (resp. second) inequality comes from from point 2 (resp. point 1) in Lemma 3.1. Hence,

$$\begin{aligned} g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) &\geq \beta |p|^{\beta-2} |p - \tilde{p}|^2 \int_0^1 (1-s)^{\beta-1} ds \\ &= |p|^{\beta-2} |p - \tilde{p}|^2. \end{aligned}$$

On the other hand,

$$\begin{aligned} g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) &\geq \beta |\tilde{p}|^{\beta-2} |p - \tilde{p}|^2 \int_0^1 (1-s) s^{\beta-2} ds \\ &= \frac{1}{\beta-1} |\tilde{p}|^{\beta-2} |p - \tilde{p}|^2. \end{aligned}$$

The last two estimates yield (3.6) because  $1 \geq \frac{1}{\beta-1}$ , then (3.7).

1.(b) If  $1 < \beta < 2$  and  $p + \tilde{p} \neq 0$ , then

$$\begin{aligned} &g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \\ &\geq \int_0^1 (1-s) G_{pp}(s\tilde{p} + (1-s)p) (\tilde{p} - p) \cdot (\tilde{p} - p) ds \\ &\geq \beta(\beta-1) |p - \tilde{p}|^2 \int_0^1 (1-s) |s\tilde{p} + (1-s)p|^{\beta-2} ds, \end{aligned}$$

where the first (resp. second) inequality comes from point 2 (resp. point 1) in Lemma 3.1. But  $|s\tilde{p} + (1-s)p| \leq 2 \max(|p|_\infty, |\tilde{p}|_\infty)$  and since  $\beta < 2$ ,  $|s\tilde{p} + (1-s)p|^{\beta-2} \geq 2^{\beta-2} \min(|p|_\infty^{\beta-2}, |\tilde{p}|_\infty^{\beta-2})$ . Hence,

$$g(x, \tilde{q}) - g(x, q) - g_q(x, q) \cdot (\tilde{q} - q) \geq 2^{\beta-2} \beta (\beta - 1) \min(|p|_\infty^{\beta-2}, |\tilde{p}|_\infty^{\beta-2}) \int_0^1 (1-s) ds,$$

which yields (3.8).

2. We have that

$$\begin{aligned} & (g_q(x, \tilde{q}) - g_q(x, q)) \cdot r \\ &= \beta (|\tilde{p}|^{\beta-2} (-\tilde{p}_1 r_1 + \tilde{p}_2 r_2 - \tilde{p}_3 r_3 + \tilde{p}_4 r_4) - |p|^{\beta-2} (-p_1 r_1 + p_2 r_2 - p_3 r_3 + p_4 r_4)). \end{aligned}$$

Call  $\ell$  the function defined on  $(\mathbb{R}_+)^4$  by

$$\ell(p) = \beta |p|^{\beta-2} p \cdot \Xi r,$$

where  $\Xi = \text{Diag}(-1, 1, -1, 1)$  stands for the diagonal matrix in  $\mathbb{R}^{4 \times 4}$  whose diagonal is  $(-1, 1, -1, 1)$ . We have

$$\begin{aligned} (g_q(x, \tilde{q}) - g_q(x, q)) \cdot r &= \ell(\tilde{p}) - \ell(p) = \int_0^1 \ell_p(s\tilde{p} + (1-s)p) \cdot (\tilde{p} - p) ds & \text{if } p \neq -\tilde{p}, \\ (g_q(x, \tilde{q}) - g_q(x, q)) \cdot r &= 0 & \text{if } p = -\tilde{p}. \end{aligned}$$

But

$$\ell_p(p) = \beta(\beta - 2) |p|^{\beta-4} (p \cdot \Xi r) p + \beta |p|^{\beta-2} \Xi r, \quad \forall p \neq 0.$$

Hence, if  $p + \tilde{p} \neq 0$ ,

$$\begin{aligned} & (g_q(x, \tilde{q}) - g_q(x, q)) \cdot r \\ &= \beta(\beta - 2) \int_0^1 |s\tilde{p} + (1-s)p|^{\beta-4} \left( (s\tilde{p} + (1-s)p) \cdot \Xi r \right) \left( (s\tilde{p} + (1-s)p) \cdot (\tilde{p} - p) \right) \\ & \quad + \beta \int_0^1 |s\tilde{p} + (1-s)p|^{\beta-2} \left( (\tilde{p} - p) \cdot \Xi r \right) ds. \end{aligned} \tag{A.2}$$

Call  $I$  (respectively  $II$ ) the first (respectively second) integral in (A.2). It is clear that

$$|I| \leq |p - \tilde{p}| |r| \int_0^1 |s\tilde{p} + (1-s)p|^{\beta-2} ds \leq \max(|p|^{\beta-2}, |\tilde{p}|^{\beta-2}) |p - \tilde{p}| |r|.$$

We also have

$$|II| \leq \max(|p|^{\beta-2}, |\tilde{p}|^{\beta-2}) |p - \tilde{p}| |r|,$$

and (3.9) follows from the last two estimates.

*Acknowledgement.* The first author would like to thank O. Guéant for helpful discussions. The first author was partially funded by the ANR projects ANR-12-MONU-0013 and ANR-12-BS01-0008-01.

#### REFERENCES

- [1] Y. Achdou, F. Camilli, and I. Capuzzo Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM J. Control Optim.*, 50(1):77–109, 2012.
- [2] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.

- [3] Y. Achdou and J-M. Lasry. On the use of mean field games for modeling of crowd motion. in preparation.
- [4] Yves Achdou and Victor Perez. Iterative strategies for solving linearized discrete mean field games systems. *Netw. Heterog. Media*, 7(2):197–217, 2012.
- [5] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [6] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master’s Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [7] F. Camilli and F. Silva. A semi-discrete approximation for a first order mean field game problem. *Netw. Heterog. Media*, 7(2):263–277, 2012.
- [8] P. Cardaliaguet. Notes on mean field games. preprint, 2011.
- [9] P. Cardaliaguet, , J-M. Lasry, P-L. Lions, and A. Porretta. Long time average of mean field games. *Networks and heterogeneous media*, 7(2):279–301, 2012.
- [10] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.
- [11] R. Eymard, T. Gallouët, D. Hilhorst, and Y. Naït Slimane. Finite volumes and nonlinear diffusion equations. *RAIRO Modél. Math. Anal. Numér.*, 32(6):747–761, 1998.
- [12] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.
- [13] D. A. Gomes, J. Mohr, and R. R. Souza. Discrete time, finite state space mean field games. *J. Math. Pures Appl. (9)*, 93(3):308–328, 2010.
- [14] O. Guéant. Mean field games equations with quadratic Hamiltonian: a specific approach. *Math. Models Methods Appl. Sci.*, 22(9):1250022, 37, 2012.
- [15] O. Guéant. New numerical methods for mean field games with quadratic costs. *Netw. Heterog. Media*, 7(2):315–336, 2012.
- [16] O. Guéant, J-M. Lasry, and P-L. Lions. Mean field games and applications. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 205–266. Springer, Berlin, 2011.
- [17] M. Huang, P.E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria. *IEEE Trans. Automat. Control*, 52(9):1560–1571, 2007.
- [18] A. Lachapelle, J. Salomon, and G. Turinici. Computation of mean field equilibria in economics. *Math. Models Methods Appl. Sci.*, 20(4):567–588, 2010.
- [19] J-M. Lasry and P-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.
- [20] J-M. Lasry and P-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.
- [21] J-M. Lasry and P-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [22] P.-L. Lions. Quelques remarques sur les problèmes elliptiques quasilinéaires du second ordre. *J. Analyse Math.*, 45:234–254, 1985.
- [23] P-L. Lions. Cours du Collège de France. <http://www.college-de-france.fr/default/EN/all/equ-der/>, 2007-2011.
- [24] W Rudin. *Real and complex analysis. 3rd ed.* New York, NY: McGraw-Hill. xiv, 416 p. \$ 44.95 , 1987.