Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Anders Hoel

The Master Equation in Mean Field Games

Master's thesis in Industrial Mathematics
Supervisor: Espen R. Jakobsen
July 2022

NTNU
Norwegian University of Science and Technology

Anders Hoel

The Master Equation in Mean Field Games

Master's thesis in Industrial Mathematics
Supervisor: Espen R. Jakobsen
July 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Abstract

In this master's thesis we prove the well-posedness of the Master Equation, which is a second order partial differential equation on the space of probability measures. The proofs utilise a "method of characteristics", where we employ well-posed solutions of the closely related Mean Field Game systems as the characteristics. Furthermore, we sketch and discuss the proof of the convergence of the Nash system, which is a system of N strongly coupled Hamilton-Jacobi-Bellman equations, as N tends to infinity. For both well-posedness and convergence, we derive our results on the domain $\mathbb{R}^d$, which diverges from the primary source material, where analysis is performed on $\mathbb{T}^d$.

Sammendrag

I denne masteroppgaven beviser vi velstiltheten til Master Equation, som er en annenordens partiell differensiallikning definert for sannsynlighetsmål. Bevisene benytter en "karakteristikkmetode", der vi bruker velstilte løsninger av de nært beslektede Mean Field Game-systemene som karakteristikker. Videre skisserer og diskuterer vi beviset for konvergensen til Nash-systemet, som er et system av N sterkt koblede Hamilton-Jacobi-Bellman-likninger, når N går mot uendelig. For både velstilthet og konvergens utleder vi våre resultater på domenet $\mathbb{R}^d$, som avviker fra det primære kildematerialet, hvor analysen utføres på $\mathbb{T}^d$.

Preface

This thesis is the result of my work in the subject *TMA4900 - Industrial Mathematics, Master's Thesis* finished in my last year as student at the study program of Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU). I would like to thank my supervisor, Espen R. Jakobsen, for his patience, directness, and him providing an interesting and truly challenging masters topic. I would also like to thank Artur Rutkowski for enlightening discussions and good advice. Lastly, I would like to thank the citizens of Matteland and people of Nabla for making the last 5 years truly memorable.

Anders Hoel

July 2022

Trondheim

Table of Contents

1 Introduction

Game theory has since its introduction by John von Neumann and Oskar Morgenstern in 1944 [25] been a large and active branch of mathematics. Being the study of mathematical models of strategic interaction, the field has been especially fruitful for a wide array of applications, but has been particularly successful in social science and economics, where game theorists have won the Nobel Memorial Prize in Economic Sciences 15 times. Further development in the field was done by John Forbes Nash Jr., who introduced the equilibrium solutions that now bear his name as the natural concept of solution for non-cooperative games [24]. Indeed, we say that a set of strategies in a game is in Nash equilibrium if none of the players have anything to gain from changing their strategy.

A modern branch of game theory is the study of Mean Field Game systems, where strategic interaction in very large populations of rational agents are modelled. Each of the modelled agents has negligible impact upon the total system, which gives rise to the key idea of the Mean Field theory, namely that instead of viewing each agent of the system seperately, we view the system as an averaged statistical distribution of agents. Mean Field Game theory was introduced using tools from stochastic control theory by Jean-Michel Lasry and Pierre-Louis Lions in 2006 [20]. The same concept, under a different name and using a stochastic approach, was discovered independently at about the same time by M. Huang, P. Caines, and R. Malhamé. Since its introduction, Mean Field Game theory has been applied to a wide range of fields, some examples being crowd dynamics, trading of financial securities and power grid planning.

From the control theoretic interpretation of Lasry and Lions arises the Mean Field Game system

$$\begin{cases} -\partial_t u - \varepsilon\Delta u + H(x, Du) = F(x, m(t)) & \text{in } [t_0, T) \times \mathbb{R}^d, \\ u(T) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ \partial_t m - \varepsilon\Delta m - \operatorname{div}(m D_p H(x, Du)) = 0 & \text{in } (t_0, T] \times \mathbb{R}^d, \\ m(0) = m_0 & \text{in } \mathbb{R}^d, \end{cases} \tag{1}$$

which is a coupled system of a Hamilton-Jacobi-Bellman equation, governing the control problem, and a Fokker-Planck equation, governing how the agents in the system move. Under sufficiently regularising coupling terms F, G, (1) has unique classical solutions, as has been proven for an explicit choice of $H(x,p) = \frac{1}{2}|p|^2$ in the present authors project thesis [14], and on the simplified domain $\mathbb{T}^d$ for a general H in [1].

A core topic of study in Mean Field Game theory is how differential game systems with a finite number of players N converge to a mean field formulation as $N \to \infty$. As with other non-cooperative games, the concept of solution for the N-player game is the Nash equilibrium, which can be shown [7] is equivalent to the value functions $v^{N,i}$ of the corresponding control problem satisfying the following strongly coupled system of Hamilton-Jacobi-Bellman equations that we will call the Nash system

$$\begin{cases} -\partial_t v^{N,i}(t, \boldsymbol{x}) - \varepsilon \sum_{j=1}^{N} \Delta_{x_j} v^{N,i}(t, \boldsymbol{x}) + H(x_i, D_{x_i} v^{N,i}(t, \boldsymbol{x})) \\ \quad + \sum_{j \neq i} D_p H(x_j, D_{x_j} v^{N,j}(t, \boldsymbol{x})) \cdot D_{x_j} v^{N,i}(t, \boldsymbol{x}) \\ = F^{N,i}(\boldsymbol{x}) & \text{in } [0,T] \times \left(\mathbb{R}^d\right)^N, \\ v^{N,i}(T, \boldsymbol{x}) = G^{N,i}(\boldsymbol{x}) & \text{in } \left(\mathbb{R}^d\right)^N, \end{cases}$$

Our hope is that this system simplifies as $N \to \infty$. However, passing the Nash system to the limit using standard PDE techniques is very difficult as the strongly coupled nature of the systems makes it seemingly impossible to establish the compactness results necessary for convergence. To solve this problem we introduce the main object of study in this thesis, the Master Equation, which is a non-local PDE, closely related to (1), taking the following form

$$\begin{cases} -\partial_t U(t,x,m) - \varepsilon \Delta_x U(t,x,m) + H(x, D_x U(x,t,m)) \\ \quad -\varepsilon \int_{\mathbb{R}^d} \operatorname{div}_y [D_m U](t,x,m,y) m(dy) \\ \quad + \int_{\mathbb{R}^d} D_m U(t,x,m,y) \cdot D_p H(y, D_x U(t,y,m)) m(dy) \\ = F(x,m) \qquad\qquad\qquad \text{in } [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d), \\ U(T,x,m) = G(x,m) \qquad\qquad \text{in } \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d). \end{cases} \tag{2}$$

To the unfamiliar eye, this equation might look fierce and formidable, as it takes in probability measures $m \in \mathcal{P}_1(\mathbb{R}^d)$ as variables, and contains a new type of derivative, namely the Lions derivative $D_m U$. The Lions derivative is taken in the space of measures and related to the measure equivalent of a Frechet derivative.

The Master Equation bridges the gap of understanding between the finite dimensional systems and the Mean Field formulation, and is indeed what is needed for establishing some sort of convergence for the N-player Nash system. Indeed, given an empirical measure $m_{\boldsymbol{x}}^{N,i} := \frac{1}{N-1} \sum_{j \neq i}$ we have that

$$\sup_{i \in \{1,\dots,N\}} \left| v^{N,i}(t_0, \boldsymbol{x}) - U(t_0, x_i, m_{\boldsymbol{x}}^{N,i}) \right| \leq \frac{C}{N},$$

which shows that the solution U to the Master Equation (2) is to some extent the natural limit to the Nash systems. This convergence result has been proven on the compact domain $\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d$ with periodic boundaries in [7], while the full result on $\mathbb{R}^d$ has been conjectured by Lasry and Lions in the original paper from 2006 [20].

The main topic and core of this thesis is proving existence and uniqueness of classical solutions of the Master Equation on $\mathbb{R}^d$. To this end we will follow an approach set out in the monograph [7] where a similar proof is performed for the simpler case $\mathbb{T}^d$. The existence proof is based on a "method of characteristics", using the Mean Field Game system (1) as the characteristics. Let (u, m) be the unique solution of (1) with respect to the initial time and measure $(t_0, m_0) \in [0,T] \times \mathcal{P}_1(\mathbb{R}^d)$, and define the following ansatz

$$U(t_0, x, m_0) := u(t_0, x).$$

With this choice of ansatz, much of the analysis boils down to estimates upon the Mean Field Game system (1) and how its solution changes as the initial distibution m_0 changes, all in order to show that the "characteristic" defined above satisfies (2). By performing the analysis on the whole space $\mathbb{R}^d$ as opposed to the torus $\mathbb{T}^d$ as is done in the source material, the proofs often get quite a bit more technical and difficult. This is due to many compactness and boundedness results that are almost immediate for $\mathbb{T}^d$ are either elusive or impossible for $\mathbb{R}^d$. For example continuous functions on $\mathbb{T}^d$ are always bounded by the extreme value theorem. However, the price we pay with difficult analysis is payed back by the fact that the Master Equation for domain $\mathbb{R}^d$ is more useful for applications.

The theory of Mean Field Games is at the meeting point between several large fields of mathematics, including, but not limited to PDE analysis, stochastic processes, measure theory, and optimal control theory. A consequence of this is that the amount of preliminaries needed to rigorously study these systems is quite large. If this text was to be self contained, it would have been way too long for an enjoyable read. We therefore presuppose the readers familiarity with measure and integration theory, as well as some PDE theory and stochastic analysis.

The thesis is structured as follows. Chapter 2 introduces the relevant background material needed to perform the analysis. Particular focus is placed upon derivatives in the space of measures and the relation between differential games and optimal control theory. Chapter 3 Contains a treatment of the Mean Field Game system (1), and contains some well-posedness and regularity estimates employed in the proofs for the Master Equation. Chapter 4 is dedicated in its entirety to the well-posedness of the Master Equation, and most of the exposition is in establishing theory and technical results leading up to the well-posedness proof. Finally, chapter 5 covers the Nash system and the manner of which it converges to the Master Equation.

2 Background Material

Before we head into the analysis and tackle the main topics of this thesis, we will perform a brief survey of the required background material. We will keep the introduction to a minimum, focusing the exposition on the theory that will be directly applied in the sequel, while giving precise references and recommendations to the sources where further explanation and exposition can be found.

2.1 Spaces and Notation

We start by defining the key spaces, norms, and metrics in which our analysis will be performed. Our primary domain of analysis will be the standard d-dimensional Euclidian space $\mathbb{R}^d$ equipped with the Euclidian norm $|\cdot|$. Another important domain, that we will not use ourselves, but is quite popular in the analysis on Mean Field Games due to its simplifying nature is the d-dimensional torus $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. This domain is compact and has a periodic boundary.

Another very central space is $\mathcal{P}(\mathbb{R}^d)$, the set of Borel probability measures on $\mathbb{R}^d$. We note that $\mathcal{P}(\mathbb{R}^d)$ is a subset of $rba(\mathbb{R}^d)$, the space of all regular bounded finitely additive Borel measures on $\mathbb{R}^d$. Next, we introduce one of the core concepts of establishing compactness, and hence convergence in these spaces, namely tightness.

Definition 2.1. We say that a subset $\mathcal{K} \subseteq rba(X)$ is tight, if for any $\epsilon > 0$ there exists a compact subset $K \subseteq X$ such that

$$\sup_{m\in\mathcal{K}} m(X\backslash K) \leq \epsilon.$$

An important theorem that we will not use directly, but is the basis for many of the results we will employ is the famous Prokhorov's theorem, which equates tightness with compactness.

Theorem 2.2. *(Prokhorov) A subset $\mathcal{K} \subseteq rba(\mathbb{R}^d)$ is sequentially precompact in weak-* if and only if it is tight.*

Proof. The original proof is given in [27] □

Here the weak-* convergence is defined in the usual way given that $rba(\mathbb{R}^d)$ is the dual of the space of bounded continuous functions $C_b(\mathbb{R}^d)$ (Theorem IV.6.2 in [10]).

As it turns out, $\mathcal{P}(\mathbb{R}^d)$ is too general for the approach taken in the analysis in this thesis, and a more natural choice is the space of measures with finite first moment $\mathcal{P}_1(\mathbb{R}^d)$. We define $\mathcal{P}_1(\mathbb{R}^d)$ as the subset of $\mathcal{P}(\mathbb{R}^d)$ where all measures m satisfy

$$\int_{\mathbb{R}^d} |x| m(dx) < \infty.$$

Simply speaking, this is the set of probability distributions with a well defined statistical mean. As a metric, we choose the Monge-Kantorovich distance which is defined in the following way, for $m, m' \in \mathcal{P}_1(\mathbb{R}^d)$

$$d_1(m, m') := \sup \phi \in 1 - \text{Lip} \int_{\mathbb{R}^d} \phi(x)(m' - m)(dx),$$

where 1-Lip is the space of the Lipschitz functions ϕ with Lipschitz constant 1, that is, functions satisfying $|\phi(x) - \phi(y)| \leq |x - y|$. Furthermore, we have from [1] that $\mathcal{P}_1(\mathbb{R}^d)$ is a complete metric space with $d_1(\cdot, \cdot)$.

We will also need to introduce the notation $C([t_0, T], \mathcal{P}_1(\mathbb{R}^d))$ m, which is the space of continuously time indexed families of probability measures. The function m solving the Mean Field Game system (1) will typically be a function of this space.

Having finished the exposition on probability measures, we move on to the function spaces. We define the spaces of bounded continuously differentiable functions $C_b^n(\mathbb{R}^d)$ for each $n \in \mathbb{N}$ as the linear space of maps $\phi : \mathbb{R}^d \to \mathbb{R}$ such that $\|\phi\|_{n+\alpha} < \infty$ where

$$\|\phi\|_{C_b^n} \coloneqq \sum_{|\ell| \le n} \sup_{x \in \mathbb{R}^d} |D^\ell \phi(x)|, \tag{3}$$

where ℓ denotes a multiindex of differentiation with respect to $x \in \mathbb{R}^d$.

Another, and an in many ways more refined, notion of continously differentiable functions comes though the theory of Hölder continuity. For some $\alpha \in (0,1)$, we define the Hölder spaces $C^{n+\alpha}(\mathbb{R}^d)$ as the space of functions $\phi \in C_b^n(\mathbb{R}^d)$ that are in addition bounded in the following norm

$$\|\phi\|_{n+\alpha} \coloneqq \|\phi\|_{C_b^n} + \sum_{|\ell|=n} \sup_{x \neq x'} \frac{|D^\ell(x) - D^\ell(x')|}{|x - x'|^\alpha}. \tag{4}$$

Furthermore, as a key part of the analysis of coupled linear systems in the sequel, we have to introduce the dual spaces of the aformentioned Hölder spaces $C^{n+\alpha}(\mathbb{R}^d)$, which we denote by $C^{-(n+\alpha)}(\mathbb{R}^d)$ with norm defined in the usual way for functionals

$$\|\rho\|_{-(n+\alpha)} \coloneqq \sup_{\|\phi\|_{n+\alpha} \le 1} \langle \rho, \phi \rangle_{n+\alpha}, \tag{5}$$

where $\langle \cdot, \cdot \rangle_{n+\alpha}$ denotes the action of a functional in the left hand slot upon a $C^{n+\alpha}$-function in the right hand slot. We also have the following lemma regarding inclusion in the relevant spaces.

Lemma 2.3. *For natural numbers $n \le m$, $\alpha \in (0,1)$, and considering fuctions spaces over $\mathbb{R}^d$, we have that $C^{m+\alpha} \subseteq C_b^m \subseteq C^{n+\alpha} \subseteq C_b^n$. Furthermore, for the corresponding dual spaces, we have $C_b^{-n} \subseteq C^{-(n+\alpha)} \subseteq C_b^{-m} \subseteq C^{-(m+\alpha)}$.*

Proof. By induction, we can choose $m = n + 1$ without loss of generality. The first chain of inclusions follows from observing that, by the definition of the norms (3) and (4), $\|\phi\|_{C_b^n} \le \|\phi\|_{n+\alpha}$ and $\|\phi\|_{C_b^m} \le \|\phi\|_{m+\alpha}$, as well as noting that if $\phi \in C_b^m$ we have by the Mean Value Theorem

$$\sum_{|\ell|=n} \sup_{x \neq x'} \frac{|D^\ell \phi(x) - D^\ell \phi(x')|}{|x - x'|} \le \sum_{|\ell|=n+1} \sup_{x \in \mathbb{R}^d} |D^\ell \phi(x)| \le \|\phi\|_{C_b^m}.$$

Now, if $|x - x'| \le 1$ we have

$$\frac{|D^\ell \phi(x) - D^\ell \phi(x')|}{|x - x'|^\alpha} \le \frac{|D^\ell \phi(x) - D^\ell \phi(x')|}{|x - x'|} |x - x'|^{1-\alpha} \le \|\phi\|_{C_b^m} |x - x'|^{1-\alpha} \le \|\phi\|_{C_b^m},$$

and equivalently if $|x - x'| > 1$ then by the triangle inequality

$$\frac{|D^\ell \phi(x) - D^\ell \phi(x')|}{|x - x'|^\alpha} \le \frac{|D^\ell \phi(x)| + |D^\ell \phi(x')|}{|x - x'|^\alpha} \le 2\|\phi\|_{C_b^m}.$$

Consequently, $\|\phi\|_{n+\alpha} \le \|\phi\|_{C_b^m}$ and the inclusion chain holds.

For the inclusion chain in the dual spaces, we note that since $\|\phi\|_{n+\alpha} \le \|\phi\|_{C_b^m}$

$$\|\rho\|_{-m} = \sup_{\|\phi\|_{C_b^m} \le 1} \langle \rho, \phi \rangle_{n+\alpha} \le \sup_{\|\phi\|_{n+\alpha} \le 1} \langle \rho, \phi \rangle_{n+\alpha} = \|\rho\|_{-(n+\alpha)},$$

since we are taking the supremum over a larger class of functions. Thus $C^{-(n+\alpha)} \subseteq C_b^{-m}$. The same argument holds for the other two inclusions. □

We will also need to work with functions of two space variables in $\mathbb{R}^d$, often denoted x and y, and to control derivatives with respect to the two variables simultaneously. To this end we introduce notation for a paired Hölder norm, for $\phi = \phi(x, y)$ we set

$$\|\phi\|_{m,n} \coloneqq \sum_{|\ell| \le m, |\ell'| \le n} \|D_x^\ell D_y^{\ell'} \phi\|_\infty,$$

and

$$\|\phi\|_{m+\alpha,n+\alpha} := \|\phi\|_{m,n} + \sum_{|\ell|=m,|\ell'|=n} \sup_{(x,y\neq(x',y'))} \frac{|D_x^\ell D_y^{\ell'}\phi(x,y) - D_x^\ell D_y^{\ell'}\phi(x',y')|}{|x-x'|^\alpha + |y-y'|^\alpha}.$$

Finally, we have the parabolic Hölder spaces, which are spaces of functions $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with Hölder-type regularity in time and space. For compactness of notation, we will sometimes denote $Q_T := [0,T] \times \mathbb{R}^d$. Let

$$[u]_{\alpha/2,\alpha;Q_T} := \sup_{\substack{(t_1,x_1)\neq(t_2,x_2)\\(t_i,x_i)\in Q_T}} \frac{|u(t_1,x_1) - u(t_2,x_2)|}{|x_1-x_2|^\alpha + |t_1-t_2|^{\frac{\alpha}{2}}}$$

$$[u]_{1+\alpha/2,2+\alpha;Q_T} := [u_t]_{\alpha/2,\alpha;Q_T} + \sum_{i,j=1}^d [u_{x_ix_j}]_{\alpha/2,\alpha;Q_T}.$$

We define the parabolic Hölder spaces $C^{\alpha/2,\alpha}(Q_T), C^{1+\alpha/2,2+\alpha}(Q_T)$ where inclusion is decided by boundedness in the following norms

$$\|u\|_{C^{\alpha/2,\alpha}(Q_T)} = \|u\|_{\infty;Q_T} + [u]_{\alpha/2,\alpha;Q_T},$$

$$\|u\|_{C^{1+\alpha/2,2+\alpha}(Q_T)} = \|u\|_{\infty;Q_T} + \|Du\|_{\infty;Q_T} + \|u_t\|_{\infty;Q_T} + \sum_{i,j=1}^d \|\frac{\partial^2 u}{\partial_{x_i}\partial_{x_j}}\|_{\infty;Q_T} + [u]_{1+\alpha/2,2+\alpha;Q_T}.$$

The concept of Hölder spaces and their relation to parabolic PDE is treated in a thorough manner in [17] and [21].

An important type of continuous function which we will apply numerous times in our analysis is the mollifiers.

Definition 2.4. We define the standard mollifier $\eta \in C_c^\infty(\mathbb{R}^d)$ in the following way

$$\eta := \begin{cases} C_d \exp\left(-\frac{1}{1-|x|^2}\right) & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where C_d is a dimensionally dependent constant chosen so that $\int_{\mathbb{R}^d} \eta \; dx = 1$. Furthermore, for each $\epsilon > 0$ we denote

$$\eta_\epsilon(x) := \frac{1}{\epsilon^d}\eta(\frac{x}{\epsilon}).$$

The concept of mollification will be used quite heavily in the sequel, primary in order to build smooth approximations to given functions. For a locally integrable function $f : \mathbb{R}^d \to \mathbb{R}$ we define the mollification of f as the convolution against a scaled standard mollifier

$$f^\epsilon := \eta_\epsilon * f.$$

We state some important properties of mollification in the following proposition taken from [12]

Proposition 2.5. *Let η be the standard mollifier, and let $f \in L^1_{loc}(\mathbb{R}^d)$. Then the following statements hold*

(i) $\int_{\mathbb{R}^d} \eta \; dx = 1$ *and* $supp(\eta_\epsilon) \subset B(0,\epsilon)$.

(ii) $f^\epsilon \in C^\infty(\mathbb{R}^d)$.

(iii) $f^\epsilon \to f$ *a.e as* $\epsilon \to 0$.

(iv) If f is continuous, then $f^\epsilon \to f$ uniformly on compacts.

As the last topic in this subsection, we have to mention the Dirac distribution. For our purposes the Dirac delta function δ_{x_0} is a generalised function in the dual of $C_b(\mathbb{R}^d)$ defined by the following action: For $x_0 \in \mathbb{R}^d$ and any $\phi \in C_b(\mathbb{R}^d)$

$$\langle \delta_{x_0}, \phi \rangle := \phi(x_0).$$

Simply spoken, this is expressing the concept of point evaluation as a functional. We are furthermore able to define derivatives. For $|\ell| \leq n$ we define the distributional derivative of the Dirac distribution $D^\ell \delta_{x_0} \in C_b^{-n}(\mathbb{R}^d)$ as the following action upon the function $\phi \in C_b^n(\mathbb{R}^d)$

$$\langle D^\ell \delta_{x_0}, \phi \rangle := (-1)^{|\ell|} \langle \delta_{x_0}, D^\ell \phi \rangle = (-1)^{|\ell|} D^\ell \phi(x_0),$$

where we have, analogously to weak derivatives, moved the derivatives over to a test function ϕ.

The Dirac delta function is an example of what are called distributions, a huge and fascinating topic in and of itself that we cannot go into here. The curious reader is recommended to look up the nice and comprehensive presentation of distributions theory given in [13].

2.2 Derivatives in a Measure Variable

In order to study the Master Equation, we have to make sense of the calculus of how changes in initial distribution $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ influences the Mean Field Game system. We formalise this properly by defining the notion of a derivative with respect to a measure.

The following definition is taken from [7], and adjusted to the case in the whole space $\mathbb{R}^d$.

Definition 2.6. A function $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is said to be $\mathcal{C}^1$ if there exists a continuous mapping $\frac{\delta U}{\delta m} : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ such that for all $m, m' \in \mathcal{P}(\mathbb{R}^d)$

$$\lim_{h \to 0^+} \frac{U((1-h)m + hm') - U(m)}{h} = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, y)(m' - m)(dy)$$

Furthermore, in order to ensure uniqueness, we set the following normalisation convention.

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, y) m(dy) := 0$$

This way of defining a derivative in a function space is similar to the Gateaux derivative, a treatment of which can be found in Chapter 2.6 of [28]. However, since $\mathcal{P}_1(\mathbb{R}^d)$ is not a vector space, we define the measure derivative using a convex combination.

We continue by introducing a property mirroring the fundamental theorem of calculus

Lemma 2.7. *Let U be $\mathcal{C}^1$, then for all $m, m' \in \mathcal{P}(\mathbb{R}^d)$, we have*

$$U(m') - U(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}((1-s)m + sm', y)(m' - m)(dy) ds.$$

Proof. We consider $U((1-s)m+sm')$ as a function of $s \in \mathbb{R}$ and attempt to compute the derivative with respect to this variable.

$$\begin{aligned}
&\frac{d}{ds}\left(U((1-s)m + sm')\right) \\
&= \lim_{h \to 0^+} \frac{U((1-(s+h))m + (s+h)m') - U((1-s)m + sm')}{h} \\
&= \lim_{h \to 0^+} \frac{U(((1-s)m + sm')(1-h) + h((1+s)m' - sm)) - U((1-s)m + sm')}{h}.
\end{aligned}$$

Assuming that $s \in [0,1]$, we can apply definition 2.6 for the measure derivative, and we get

$$\begin{aligned}\frac{d}{ds}\left(U((1-s)m+sm')\right) &= \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}((1-s)m+sm',y)(((1-s)m+sm')-(1-s)m+sm')(dy) \\ &= \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}((1-s)m+sm',y)(m'-m)(dy)\end{aligned}$$

Now, applying the fundamental theorem of calculus, we get

$$U(m')-U(m) = \int_0^1 \frac{d}{ds}\left(U((1-s)m+sm')\right)ds = \int_0^1\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}((1-s)m+sm',y)(m'-m)(dy)ds$$

□

With this property, we can perform what we in the sequel will call the Lipschitz in y trick. If $\sup_{m\in\mathcal{P}_1(\mathbb{R}^d)} \|D_y \frac{\delta U}{\delta m}(m,\cdot)\|_\infty = L < \infty$, then we know that $\frac{\delta U}{\delta m}(m,\cdot)$ is Lipschitz continuous in y with Lipschitz constant smaller or equal to L. By the definition of d_1, we then have for any $m, m' \in \mathcal{P}(\mathbb{R}^d)$

$$|U(m')-U(m)| = \left|\int_0^1\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}((1-s)m+sm',y)(m'-m)(dy)ds\right| \le \|D_y\frac{\delta U}{\delta m}(m,\cdot)\|_\infty d_1(m',m).$$

This cross derivative in y and m is quite important to the study of the Master Equation, so important in fact that we give it its own name and notation. We call the following derivative the Lions derivative, named after the Fields medallist Pierre-Louis Lions who co-introduced the field of Mean Field Games

$$D_mU(m,y) \coloneqq D_y\frac{\delta U}{\delta m}(m,y).$$

We also need to introduce the concept of a pushforward measure

$$\phi\#m(\phi^{-1}(A)),$$

for any Borel set $A \subseteq \mathbb{R}^d$. This is in some ways the generalisation of the change of variables method from integration theory. We also include a final useful result.

Proposition 2.8. *Assume U to be $\mathcal{C}^1$, with $\frac{\delta U}{\delta m}$ being continuously differentiable in y and D_mU is continuous in m and y. Furthermore, let $\phi \in L^1(m,\mathbb{R}^d)$. Then,*

$$\lim_{h\to 0}\frac{U((id+h\phi)\#m)-U(m)}{h} = \int_{\mathbb{R}^d} D_mU(m,y)\cdot\phi(y)m(dy).$$

Proof. We follow the proof for Proposition 2.2.3. in [7]. By the fundamental theorem of calculus for measures from Lemma 2.7, setting $m_{h,s} \coloneqq s(\mathrm{id}+h\phi)\#m + (1-s)m$,

$$\begin{aligned}U((\mathrm{id}+h\phi)\#m)-U(m) &= \int_0^1\int_{\mathbb{R}^d}\frac{\delta U}{\delta m}(m_{h,s},y)((\mathrm{id}+h\phi)\#m-m)(dy)ds \qquad (6)\\ &= \int_0^1\int_{\mathbb{R}^d}\big(\frac{\delta U}{\delta m}(m_{h,s},y+h\phi(y))-\frac{\delta U}{\delta m}(m_{h,s},y)\big)m(dy)ds \\ &= h\int_0^1\int_{\mathbb{R}^d}\int_0^1 D_mU(m_{h,s},y+th\phi(y))\cdot\phi(y)dtm(dy)ds.\end{aligned}$$

From the definition of the d_1-metric, we check

$$d_1((\mathrm{id}+h\phi)\#m,m) \le \int_{\mathbb{R}^d}|y+h\phi-y|m(dy) = h\|\phi\|_{L^1(m)},$$

and hence $m_{h,s}$ tends to m as $h\to 0$. We divide both sides of (6) by h and let $h\to 0$. By the continuity of D_m we are left with the desired result. □

This introduction to measure valued derivatives has been kept at a bare minimum of what is necessary for the analysis of the Master Equation. For a more thorough introduction, see the probabilistic introduction of Carmona and Delarue [8] or the embedding strategy from the lectures of Lions [22].

2.3 Short Introduction in Game Theory

Game theory is the mathematical study of models of strategic interaction, and has since its introduction in the 1950s by Morgenstern and von Neumann evolved into a vast field with countless applications in fields like social science, biology, and economics. In this thesis, we consider the subfield of noncooperative games, where two or more players choose strategies to optimise for individual success.

The concept of the equilibrium as the solution to a non-cooperative game was introduced by John Forbes Nash Jr. in [24]. In a Nash equilibrium, none of the agents in the non-cooperative game has anything to gain from changing only ones own strategy. We phrase this more mathematically by considering a non cooperative game of two agents, agent A and agent B. They both choose a strategy, α_A and α_B respectively, from a set of admissible strategies $\mathcal{A}$. Each agent want to minimise a given cost function $J_i : A \times A \to \mathbb{R}$ for $i \in \{A, B\}$ by choosing the best strategy in response to the admissible strategies of the other. We can then state the Nash equilibrium mathematically by noting that if (α_A^*, α_B^*) is the equilibrium strategies, the following inequalities hold

$$\begin{aligned} J_A(\alpha_A^*, \alpha_B^*) &\le J_A(\alpha_A, \alpha_B^*), \quad \forall \alpha_A \in \mathcal{A}, \\ J_B(\alpha_A^*, \alpha_B^*) &\le J_A(\alpha_A^*, \alpha_B), \quad \forall \alpha_B \in \mathcal{A}. \end{aligned}$$

A famous example of a Nash equilibrium is the prisoners' dilemma: Two people, person A and person B, are charged with a crime and are given the option to either confess or stay silent. If both stay silent, they both go to prison for 1 year. If one confesses and the other stays silent, the one who confesses goes free while the other gets 20 years. Lastly, if both confess, both get 5 years. The crux of this scenario is that, with no cooperation, you are always better off by confessing, even though both confessing will lead to a worse result for both people.

Game theory problems with continuous time and space variables are called differential games, and their dynamic and optimality are governed by differential equations. We introduce the N-player differential game, which indeed can be cast as an optimal control problem, each player $i \in \{1, ..., N\}$ controls their own state $\{X_{i,t}\}_{t\in[0,T]}$ using their control function $\{\alpha_{i,t}\}_{t\in[0,T]}$ which is governed by the following dynamic

$$\begin{cases} dX_{i,t} = \alpha_{i,t} dt + \sqrt{2\varepsilon} dB_t^i, \quad t \in (t_0, T] \\ X_{i,t_0} = x_{i,0}, \end{cases} \tag{7}$$

where $\{B_t^i\}_{t\in[0,T]}$ is the standard d-dimensional Brownian motion as described in [26], and $\boldsymbol{x}_0 = (x_{1,0}, ..., x_{N,0}) \in \left(\mathbb{R}^d\right)^N$ is the initial condition of the whole system at time t_0. To complete the control problem formulation, we need a cost functional, and define

$$J_i^N(t_0, \boldsymbol{x}_0, \{\alpha_{j,\cdot}\}_{j\in\{1,...,N\}}) = E\left[\int_{t_0}^T \Big(L\big(X_{i,s}, \alpha_{i,s}\big) + F^{N,i}\big(\boldsymbol{X}_s\big)\Big) ds + G^{N,i}\big(\boldsymbol{X}_T\big)\right], \tag{8}$$

where $\boldsymbol{X}_t = (X_{1,t}, ..., X_{N,t})$, and $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $F^{N,i} : \mathbb{R}^{Nd} \to \mathbb{R}$, and $G^{N,i} : \mathbb{R}^{Nd} \to \mathbb{R}$ are Borel measurable functions. Here, L corresponds to the running cost of the agents own position and strategy, $F^{N,i}$ is the running cost given by the position of the other $N-1$ agents in the system, and $G^{N,i}$ is a terminal cost associated with the position of all the agents at time T. Note that the cost functional includes an expectation, as the dynamic is stochastic.

We now seek an optimal solution, which in the realm of non-cooperative games is the Nash equilibrium. For an optimal set of strategies $\{\alpha_{i,\cdot}^*\}_{i\in\{1,...,N\}}$ we define the value functions $\{v^{N,i}\}_{i\in\{1,...,N\}}$

as the cost functionals evaluated at equilibrium,

$$v^{N,i}(t_0, \boldsymbol{x}_0) \coloneqq J_i^N(t_0, \boldsymbol{x}_0, \{\alpha^*_{i,\cdot}\}_{i\in\{1,\dots,N\}}),$$

which by definition satisfies the inequality of the Nash equilibrium

$$v^{N,i}(t_0, \boldsymbol{x}_0) \leq J_i^N(t_0, \boldsymbol{x}_0, \alpha_{i,\cdot}, \{\alpha^*_{i,\cdot}\}_{i\neq j}),$$

for each admissible strategy $\{\alpha_{i,t}\}_{t\in[0,T]}$. If all interaction between the agents of the comes through the cost function due to the changing of strategies, it can be shown that the set of value functions satisfy the following parabolic partial differential equation.

$$\begin{cases} -\partial_t v^{N,i}(t,\boldsymbol{x}) - \varepsilon \sum_{j=1}^{N} \Delta_{x_j} v^{N,i}(t,\boldsymbol{x}) + H(x_i, D_{x_i} v^{N,i}(t,\boldsymbol{x})) \\ \quad + \sum_{j\neq i} D_p H(x_j, D_{x_j} v^{N,j}(t,\boldsymbol{x})) \cdot D_{x_j} v^{N,i}(t,\boldsymbol{x}) \\ = F^{N,i}(\boldsymbol{x}) & \text{in } [0,T]\times \left(\mathbb{R}^d\right)^N, \\ v^{N,i}(T,\boldsymbol{x}) = G^{N,i}(\boldsymbol{x}) & \text{in } \left(\mathbb{R}^d\right)^N, \end{cases}$$

where the function $H(x,p) = \sup_{\alpha\in\mathbb{R}^d}\{-\alpha\cdot p - Lx, \alpha\}$ is called the Hamitonian. We call this system the N-dimensional Nash system, and part of the analysis of this paper is concerned with its behaviour as $N\to\infty$. Furthermore, we also have that the optimal strategies can be expressed using solutions of the Nash system in the following manner:

$$\alpha^*_i(t,\boldsymbol{x}) \coloneqq -D_p H\big(x_i, D_{x_i} v^{N,i}(t,\boldsymbol{x})\big), \quad i\in\{1,...,N\},$$

which we can combine with the dynamic (7) to express the set of so called optimal trajectories of the problem

$$\begin{cases} dX_{i,t} = -D_p H\big(x_i, D_{x_i} v^{N,i}(t,\boldsymbol{X}_t)\big)dt + \sqrt{2\varepsilon}dB^i_t, & t\in(t_0,T] \\ X_{i,t_0} = x_{i,0}. \end{cases}$$

Solving the Nash system for finite N is quite difficult, as the PDE system is strongly coupled, however there is a simplification to be made for large N in going to a Mean Field formulation. Neumann and Morgenstern remarked that [25], in borrowing intuition from physics and statistics, system of very large size where each agent is negligible are often easier to handle than small or medium sized ones. Further development was made by Aumann, who introduced the idea of simplification to infinitesimal actors [2] to systems of game theory in financial applications. Finally, the jump to a Mean Field formulation was done, in the PDE sense, by Lasry and Lions in 2006 [20]. They proved that, roughly speaking, for a system of infinitely many agents following the same dynamic (7) and the same cost functional (8), where the cost terms for each agent is only dependent upon its own position and the cumulative distribution of the other agents, and not upon the exact position of every other agent, the distribution of agents and their cost is governed by the Mean Field Game system (1). Thus we only have to compute the solution of a single system of parabolic PDE in stead of having to deal with a N-dimensional strongly coupled system. For a more rigorous and thorough description of the origin of the Mean Field Game system, see [1] or [14].

However, even though an approximation by letting $N\to\infty$ is convenient, this is easier said than done for the Nash system. Due to the strongly coupled nature it is very difficult to establish compactness estimates upon the system in and of itself, and thus many tools and techniques for convergence fall short. The solution is to look to the Master Equation (2), a PDE closely related to (1). If the Master Equation is well-posed, we can use it to approximate Nash systems under certain symmetry assumptions, and use the approximation to obtain a notion of convergence. The well-posedness of the Master Equation will be treated in Section 3 and the approximation and convergence in Section 4.

A more detailed introduction of differential games this context can be found in [7], from which the approach in this subsection was inspired.

3 The Mean Field Game System

Before we can undertake the main problem and purpose of this thesis, the well-posedness of the Master Equation, we first need to pay the topic of Mean Field Game systems a mathematical visit. Indeed, as mentioned in the introduction, the core idea for proving existence of the Master Equation is to represent its solution using solutions of its corresponding Mean Field Game system, which in a general setting takes the following form

$$\begin{cases} -\partial_t u - \varepsilon\Delta u + H(x, Du) = F(x, m(t)) & \text{in } [t_0, T) \times \mathbb{R}^d, \\ u(T) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ \partial_t m - \varepsilon\Delta m - \operatorname{div}(mD_pH(x, Du)) = 0 & \text{in } (t_0, T] \times \mathbb{R}^d, \\ m(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases} \tag{9}$$

Remark 3.1. The ε in the system is a viscosity parameter in $(0, \infty)$, related to the stochastic term $\sqrt{2\varepsilon}dB_t$ in the optimal control problem from which the Mean Field Game arises. Some texts, as [7], prefer to set $\varepsilon = 1$, while others like [1] prefer to keep it as a parameter. The two approaches are qualitatively equivalent as one can scale the spatial domain, and the terms F, G, H, by a factor $\sqrt{\varepsilon}$, setting $x^* = \sqrt{\varepsilon}x \in \mathbb{R}^d$, and the parameter cancels. This approach is described, in a somewhat applied manner, in [19]. One of the advantages of keeping the parameter ε in the equation is that one can keep track of what estimates on (9) depend on ε so one can later pass $\varepsilon \to 0$ through a vanishing viscosity argument, a method described clearly in Chapter 10.1 of [12]. This allows us to establish solutions of the so called First Order Mean Field Game System.

$$\begin{cases} -\partial_t u + H(x, Du) = F(x, m(t)) & \text{in } [t_0, T) \times \mathbb{R}^d, \\ u(T) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ \partial_t m - \operatorname{div}(mD_pH(x, Du)) = 0 & \text{in } (t_0, T] \times \mathbb{R}^d, \\ m(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Vanishing viscosity arguments for proving existence of solution for the first order system have been performed in [5] and in the present author's project thesis [14]. We will not give the first order system, nor the method of vanishing viscosity, any further treatment in the sequel.

To simplify the analysis, and to make the presentation of the concepts in this thesis less repetitive and proof-technical, we will work with a specific choice of Hamiltonian $H(x, p)$, namely a quadratic one $H(x, p) = \frac{1}{2}|p|^2$. This simplification results in the following Mean Field Game system, to which we will devote our attention

$$\begin{cases} -\partial_t u - \epsilon\Delta u + \frac{1}{2}|Du|^2 = F(x, m(t)) & \text{in } [t_0, T) \times \mathbb{R}^d, \\ u(T) = G(x, m(T)) & \text{in } \mathbb{R}^d, \\ \partial_t m - \varepsilon\Delta m - \operatorname{div}(mDu) = 0 & \text{in } (t_0, T] \times \mathbb{R}^d, \\ m(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases} \tag{10}$$

In this section we will recall, reproduce and enhance some existence and uniqueness results of this system. Furthermore, we will introduce the core sufficient assumptions upon the coupling terms F, G that will underpin the analysis of both Master, and Mean Field Game equation. Before we can state well-posedness, we have to define precisely the concept of solution we are working with.

Definition 3.2. A pair (u, m) is a classical solution of the system (10) if

(i) $m \in C([0, T], \mathcal{P}_1(\mathbb{R}^d))$ and $m(t_0) = m_0$; $u \in C([t_0, T] \times \mathbb{R}^d)$ and $u(T, x) = G(x, m(T))$, $\forall x \in \mathbb{R}^d$

(ii) u is a continuous function in $(0, T) \times \mathbb{R}^d$, of class C^2 in space and C^1 in time, and the first equation of (10) is satisfied pointwise for $x \in \mathbb{R}^d$ and $t \in (0, T)$. Furthermore, m is a solution of the distributional formulation of the second equation.

The notion of a distributional formulation of the system's second equation will be introduced in the subsection on the Fokker-Planck equation. We continue by treating each of the constituent equations of the system in their own subsections.

3.1 The Hamilton-Jacobi-Bellman Equation

The first equation in the Mean Field game system is called the Hamilton-Jacobi-Bellman equation, and will be introduced in a very brief manner below. This equation in full generality takes the following form

$$\begin{cases} -\partial_t u + \tilde{H}(t,x,Du,D^2u) = 0 \quad \text{ in } \mathbb{R}^d \times (0,T), \\ u(T) = g \quad \text{ in } \mathbb{R}^d. \end{cases}$$

The function $\tilde{H} : [t_0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \to \mathbb{R}^d$ is called the Hamiltonian and is deeply tied to the theory of optimal control. Indeed, considering a control problem with a dynamic

$$\begin{cases} d\boldsymbol{x}(s) & = b(\boldsymbol{x}(\tau), \alpha_\tau)d\tau, \quad s \in (t,T], \\ \boldsymbol{x}(t) & = x, \end{cases}$$

and cost function

$$J(t,x,\alpha) = \int_t^T L(s, \boldsymbol{x}(s), \alpha_s)ds + G(\boldsymbol{x}(T)),$$

the Hamiltonian arises through the so called Legendre transformation of the problem

$$\tilde{H}(t,x,p,M) := \sup_{a\in A} \left[-L(t,x,a,m) - p\cdot b(x,a,m) - \varepsilon \mathrm{Tr}(M) \right].$$

Furthermore, the solution of the Hamilton-Jacobi-Bellman equation itself serves as the so called value function

$$u(t,x) := \inf_{\alpha\in\mathcal{A}} J(t,x,\alpha),$$

of the optimal control problem.

In this thesis, we consider Hamiltonians of the form $\tilde{H}(x,p,M,m) = -\varepsilon \mathrm{Tr} M + \frac{1}{2}|p|^2 - F(t,x)$, resulting in the following parabolic PDE

$$\begin{cases} -\partial_t u - \varepsilon\Delta u + \frac{1}{2}|Du|^2 = F(t,x) \quad \text{ in } [t_0,T) \times \mathbb{R}^d, \\ u(T) = G(x) \quad \text{ in } \mathbb{R}^d. \end{cases} \tag{11}$$

For producing smooth classical solutions for this equation, we rely on the regularity of the coupling terms $F(t,x)$ and $G(x)$, and utilise existence theory for parabolic PDE.

Proposition 3.3. *Let $F(t,x) \in C^{\alpha/2,\alpha}([0,T] \times \mathbb{R}^d)$ and $G(x) \in C^{2+\alpha}([0,T]\times\mathbb{R}^d)$ Then the Hamilton-Jacobi-Bellman equation* (11) *has a unique solution $u \in C^{1+\alpha/2,2+\alpha}([0,T]\times\mathbb{R}^d)$.*

Proof. In order to apply parabolic existence results, we first transform (11) to a linear equation using the Cole-Hopf transform, a method described in Chapter 4.4 of [12]. In line with the transformation, we define $w := \exp(u/2\varepsilon)$, which yields the linear parabolic PDE

$$\begin{cases} -\partial_t w - \varepsilon\Delta w = \frac{1}{2\varepsilon} wF(t,x) \quad \text{ in } [0,T) \times \mathbb{R}^d, \\ w(x,T) = e^{G(x)/2\varepsilon} \quad \text{ in } \mathbb{R}^d. \end{cases}$$

We can now use the global parabolic existence result Theorem 5.1 in [18], which states that since $w(x,T) \in C^{2+\alpha}([0,T]\times\mathbb{R}^d))$ and $F(x,m(t)) \in C^{\alpha/2,\alpha}([0,T]\times\mathbb{R}^d)$, then the transformed equation has a unique solution $w \in C^{1+\alpha/2,2+\alpha}([0,T] \times \mathbb{R}^d)$. It is readily shown using the comparison principle that $w(t,x) > 0$ for all $(t,x) \in ([0,T] \times \mathbb{R}^d)$ implying that the inverse transformation $u = 2\varepsilon \ln(w)$ is well defined, and hence that $u \in C^{1+\alpha/2,2+\alpha}([0,T] \times \mathbb{R}^d)$ is the unique solution of (11). $\square$

After this surface-level treatment, we remark that the Hamilton-Jacobi-Bellman equation is introduced in a thorough and eloquent manner in chapter 10.3 of [12], while a compact summary written with Mean Field Games in mind can be found in the present author's project thesis [14].

3.2 The Fokker-Planck Equation

The second equation in our Mean Field Game system is called the Fokker-Planck equation, which for our use cases takes the following form

$$\begin{cases} \partial_t m - \varepsilon \Delta m - \operatorname{div}(mb(t,x)) = 0 \quad \text{in } (t_0, T] \times \mathbb{R}^d, \\ m(t_0) = m_0 \quad \text{in } \mathbb{R}^d. \end{cases} \tag{12}$$

where $b : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is continuous in time and uniformly Lipschitz continuous in space.

Due to the regularising term $\varepsilon \Delta m$ we also consider this equation to have a parabolic nature. Indeed, if m_0 and b are sufficiently smooth, say $C^{2+\alpha}$ and $C^{\alpha/2,1+\alpha}$ respectively, we can obtain classical solutions in $C^{1+\alpha/2,2+\alpha}$ using the same strategy as in Proposition 3.3, that is, applying existence results from [18]. However, while the parabolic smoothness of the Hamilton-Jacobi-Bellman equation (11) is always assured through the smoothness inducing coupling terms $F(x, m(t))$ and $G(x, m(T))$, the ability of m_0 to be any measure in $\mathcal{P}_1(\mathbb{R}^d)$ makes such a parabolic approach impossible, and we have to consider entirely different techniques.

It turns out that, since we are dealing with general functions of probability measures, a specific type of weakened formulation of the Fokker-Planck equations suits our purpose rather nicely. We define what is called a distributional, or very-weak, solution of (12).

Definition 3.4. We say that $m \in C([0,T], \mathcal{P}_1(\mathbb{R}^d))$ is a distribution solution of the equation (12) if for any test function $\varphi \in C_c^\infty([t_0,T] \times \mathbb{R}^d)$ we have

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(t,x) m(t,dx) = \int_{\mathbb{R}^d} & \varphi(t_0,x) m_0(dx) \\ & + \int_{t_0}^t \int_{\mathbb{R}^d} [\varphi_t(s,x) + D\varphi(s,x) \cdot b(s,x) + \epsilon \Delta \varphi(s,x)] m(s,dx) ds \end{aligned} \tag{13}$$

We collect the necessary information considering well-posedness of these equations in the following proposition.

Proposition 3.5. *Assume that $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ and $b \in C_b^2((t_0,T) \times \mathbb{R}^d)$. Then the Fokker-Planck equation (12) has a unique distributional solution $m \in C([0,T], \mathcal{P}_1(\mathbb{R}^d))$.*

Furthermore, if additionally $m_0 \in C^{2+\alpha}(\mathbb{R}^d)$, then (12) has a unique classical solution where we can consider $m(t,\cdot)$ as a probability density function with

$$\sup_{t \in [t_0,T]} \Big(\|m(t,\cdot)\|_\infty + \|Dm(t,\cdot)\|_\infty + \|D^2 m(t,\cdot)\|_\infty + \|\partial_t m(t,\cdot)\|_\infty \Big) \le C,$$

and C is a constant dependent upon $T,\ d,\ \|b\|_\infty,\ \|Db\|_\infty$, and $\|D^2 b\|_\infty$.

Remark 3.6. A rigorous proof of the first part of this proposition uses quite a bit of advanced probability theory and outside the scope of, and not of primary concern of this thesis. A comprehensive article on the subject of weak elliptic and parabolic equations for measures is [3], a paper which [7] claims is sufficient to cite to obtain the existence and uniqueness of $m \in C([0,T], \mathcal{P}_1(\mathbb{R}^d))$ above. The article performs the analysis on the whole space $\mathbb{R}^d$, and not just $\mathbb{T}^d$, and is hence applicable in our case.

The classical existence is treated and proven using heat kernel estimates in Proposition 6.8 of [11] for fractional diffusion operators $\mathcal{L}$ in the place of our Laplacian term $\varepsilon\Delta m$. These operators are of the form

$$\mathcal{L}u(x) = \int_{\mathbb{R}^d} u(x+z) - u(x) - Du(x) \cdot z\mathbf{1}_{|z|<1}\mu(dz),$$

for some nonnegative Borel measure satisfying the Lévy-condition $\int_{\mathbb{R}^d} 1 \wedge |z|^2 \mu(dz) < \infty$, and lead to a non-local formulation of the Fokker-Planck and hence the Mean Field Game system. However, and fortunately, the proof performed in [11] works for the Laplacian as well.

Another important property of the distribution solution of the Fokker-Planck is its deep relation to the solution of an SDE

$$\begin{cases} dX_t = b(X_t, t)dt + \sqrt{2\varepsilon}dB_t \quad t \in (t_0, T], \\ X_0 = Z_0. \end{cases}$$

where $\{B_t\}_{t\in[0,T]}$ denotes standard Brownian motion given some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $Z_0 \in L^1(\Omega)$ is a random variable independent of $\{B_t\}_{t\in[0,T]}$.

Their relation, and an important technical tool for the temporal continuity of the solutions, described in the following lemma.

Lemma 3.7. *If $\mathcal{L}(Z_0) = m_0$, then $m(t) := \mathcal{L}(X_t)$ is a distribution solution of* (12). *Furthermore, the solution satisfies*

$$d_1(m(t), m(s)) \leq \|b\|_\infty |t-s| + \sqrt{2\varepsilon|t-s|} \quad \forall s, t \in [0, T]$$

Proof. A full proof of this statement is found in the present author's project thesis [14], which in turn is an adaption and enhancement of a proof performed in [1]. □

Now, having established some exposition for the constituent equations of the Mean Field Game system, we are able to undertake the problem of establishing, and expanding upon, well-posedness results for the system itself.

3.3 Assumptions

In this subsection, we will present the assumptions needed in order to perform the proof for the well-posedness of the Master Equation. We note that the following assumptions are sufficient but strictly stronger than necessary. As this is a masters thesis, we would rather overshoot our assumptions in order to gain clarity of presentation instead of obtaining optimal results. Several of the assumptions can be significantly weakened, however, some come at the cost of having to perform technical estimates beyond the scope of this paper.

To start off, we define the coupling terms F, G. Let the maps $F, G : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$ be globally Lipschitz in both variables, that is, for any $(x_1, m_1), (x_2, m_2) \in \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$ we have, for some constant L

$$|F(x_1, m_1) - F(x_2, m_2)| \leq L\left(|x_1 - x_2| + d_1(m_1, m_2)\right),$$
$$|G(x_1, m_1) - G(x_2, m_2)| \leq L\left(|x_1 - x_2| + d_1(m_1, m_2)\right).$$

Furthermore, for a function $U : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ which is C^1 in the measure variable, we define the following Lipschitz-type criterion

$$\mathrm{Lip}_n(\frac{\delta U}{\delta m}) := \sup_{m_1 \neq m_2} (d_1(m_1, m_2))^{-1} \left\| \frac{\delta U(\cdot, m_1, \cdot)}{\delta m} - \frac{\delta U(\cdot, m_2, \cdot)}{\delta m} \right\|_{(n+\alpha, n+\alpha)}$$

We are now able to define the main assumptions of this paper.

$$\sup_{m\in\mathcal{P}_1(\mathbb{R}^d)}\left(\|F(\cdot,m)\|_{3+\alpha}+\left\|\frac{\delta F(\cdot,m,\cdot)}{\delta m}\right\|_{(3+\alpha,3+\alpha)}+\mathrm{Lip}_3(\frac{\delta F}{\delta m})\right)<\infty, \tag{F}$$

$$\sup_{m\in\mathcal{P}_1(\mathbb{R}^d)}\left(\|G(\cdot,m)\|_{4+\alpha}+\left\|\frac{\delta G(\cdot,m,\cdot)}{\delta m}\right\|_{(4+\alpha,4+\alpha)}+\mathrm{Lip}_4(\frac{\delta G}{\delta m})\right)<\infty. \tag{G}$$

We again remark that the aforementioned assumptions are not optimal. For instance, one does not in general need the same order of regularity in the x and y variable for $\frac{\delta F(\cdot,m,\cdot)}{\delta m}$. However, since the bounds (F), (G) will be applied in the majority of results in the analysis of the Master Equation, we choose the same order to increase the legibility of the text. When these bounds are invoked, we will also assume that F, G have the two following monotonicity properties.

The first monotonicity property is called The Lasry-Lions monotonicity condition, and states that for any $m, m' \in \mathcal{P}_1(\mathbb{R}^d)$

$$\begin{cases}\int_{\mathbb{R}^d}(F(x,m)-F(x,m'))(m-m')(dx) & \geq 0,\\ \int_{\mathbb{R}^d}(G(x,m)-G(x,m'))(m-m')(dx) & \geq 0.\end{cases} \tag{14}$$

This condition is precisely what is necessary in order to gain uniqueness of solutions to the Mean Field Game system (10). The second criterion is of a functional kind, and is the following: for every $\rho \in C^{-(3+\alpha)}(\mathbb{R}^d)$ and $m \in \mathcal{P}_1(\mathbb{R}^d)$

$$\left\langle\langle\frac{\delta F(\cdot,m,\cdot)}{\delta m},\rho\rangle_x,\rho\right\rangle_y\geq 0,$$
$$\left\langle\langle\frac{\delta G(\cdot,m,\cdot)}{\delta m},\rho\rangle_x,\rho\right\rangle_y\geq 0,$$

where the functional actions $\langle\cdot,\cdot\rangle_x, \langle\cdot,\cdot\rangle_y$ are applied in the x and y variable respectively. The second monotonicity criterion is of a more technical nature, and will not be applied directly, but rather serve as an important assumptions for the application for the upcoming technical Lemma 4.8, whose proof is beyond the scope of this text.

Remark 3.8. If we are not satisfied with choosing the specific Hamiltonian $H(x,p) = \frac{1}{2}|p^2|$, and want to treat the system in more generality, we would need to introduce some further assumptions upon H in order to get well-posedness for the more general Mean Field Game system (9). Sufficient restrictions on $H(x,p)$, as stated in [1] and [7], are global Lipschitz continuity in both variables, and that there exists a constant C such that

$$\frac{1}{C}I_d \leq D^2_{pp}H(x,p)\leq CI_d \qquad \text{for } (x,p)\in\mathbb{R}^d\times\mathbb{R}^d$$

3.4 Existence and Uniqueness - Removing Moment Assumptions

As mentioned in the introduction, in the project thesis [14], a proof for the existence and uniqueness of the Mean Field Game system (10) was performed under some regularity and moment assumptions on the initial data. The core assumption for these results was assuming $m_0 \in \mathcal{P}_1(\mathbb{R}^d)\cap C^{2+\alpha}(\mathbb{R}^d)$, as well as requiring that

$$\int_{\mathbb{R}^d}\psi(x)m_0(dx)<\infty \tag{15}$$

for some radially increasing superlinear function ψ. This is done to induce tightness, and hence compactness in the existence proof. In the project thesis [14], the explicit choice of $\psi(x) = \sqrt{1+|x|^2}\log\sqrt{1+|x|^2}$ was made to satisfy this particular assumption, as it was a simple example of a smooth superlinear function. However, using methods from [9], we can infer the existence of a ψ for any $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$, and hence that (15) actually comes for free. Moment assumptions are only an issue in Mean Field Game systems defined upon $\mathbb{R}^d$, and not for the $\mathbb{T}^d$-cases, as presented in [1] and [7]. This is due to (15) always being satisfied on the torus for any $\psi \in C(\mathbb{T}^d)$ since $\mathbb{T}^d$ is compact, and hence the extreme value theorem holds yielding $\|\psi\|_\infty < \infty$.

Before developing this argument further, we give introduce the restricted form of the existence theorem. By observing that the assumptions (F) and (G) upon F, G are strictly stronger than in the project thesis, we can combine the existence and uniqueness Theorems 4.2 and 4.3 from [14] into the following result.

Theorem 3.9. *Let the assumptions* (F) *and* (G) *be satisfied. Furthermore, let* $m_0 \in \mathcal{P}_1(\mathbb{R}^d) \cap C^{2+\alpha}(\mathbb{R}^d)$ *and assume that* $\int_{\mathbb{R}^d} \psi(x) m_0(dx) < \infty$*, where* $\psi(x) = \sqrt{1+|x|^2}\log\sqrt{1+|x|^2}$*. Then there exists a unique classical solution* (u, m) *of the MFG-system* (10)*. Moreover, we additionally have that* $u, m \in C^{1+\alpha/2, 2+\alpha}([t_0, T] \times \mathbb{R}^d)$*.*

However, in order to use the "method of characteristics" $U(t_0, x, m_0) := u(t_0, x)$ mentioned in the introduction to derive solutions for the Master Equation, we need to be able to obtain unique solutions for the Mean Field Game system (10) for any $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. The purpose of this subsection is thus to provide and prove the necessary results in this generality.

The first step towards this easing of assumptions is removing the integrability condition with respect to ψ, for which we employ the following lemma:

Lemma 3.10. *If* $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$*, then there exists a radial function* $\psi \in C^2(\mathbb{R}^d)$*, with* $|D\psi(x)\|, |D^2\psi(x)| \leq C(1+|x|)$*, for some constant* $C > 0$*, such that*

$$\int_{\mathbb{R}^d} \psi(x) m_0(dx) \leq 1$$

and

$$\lim_{|x|\to\infty} \frac{\psi(x)}{|x|} = \infty$$

Proof. We use Lemma 4.9 from [9], which equates tightness of a set of measures $\mathcal{K}$ with the existence of an increasing radial function $V \in C^2(\mathbb{R}^d)$, where $\|DV|_\infty, \|D^2V\|_\infty \leq 1$ and

$$\sup_{m\in\mathcal{K}} \int_{\mathbb{R}^d} V(x) m(dx) \leq 1$$

To utilise this result, we first show that the singleton set of any $m \in \mathcal{P}(\mathbb{R}^d)$, $\{m\}$, is tight. In other words, for each $\epsilon > 0$, we need to be able to find a compact set $K_\epsilon \subset \mathbb{R}^d$ such that

$$m(\mathbb{R}^d \backslash K_\epsilon) < \epsilon. \tag{16}$$

We define for each $n \in \mathbb{N}$ the closed balls $B_n = \{x \in \mathbb{R}^d : |x| \leq n\}$. By the rules of complements

$$\mathbb{R}^d = \bigcup_{n=1}^{\infty} B_n \implies \emptyset = \left(\mathbb{R}^d\right)^c = \left(\bigcup_{n=1}^{\infty} B_n\right)^c = \bigcap_{n=1}^{\infty} B_n^c$$

Thus, by observing that $B_{n+1}^c \subset B_n^c$ and $m(B_n^c) \leq 1$ for all $n \in \mathbb{N}$, we can use the continuity property of measures of intersections

$$\lim_{n\to\infty} m(B_n^c) = m(\bigcap_{n=1}^{\infty} B_n^c) = m(\emptyset) = 0$$

By the definition of convergence of sequences in $\mathbb{R}^d$, we have that for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $m(\mathbb{R}^d \backslash B_n) = m(B_n^c) < \epsilon$ for all $n \geq N$. This means that we can choose $K_\epsilon = B_N$, which satisfies (16), proving that $\{m\}$ is tight.

For the next step, we first observe that since $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \sqrt{1+|x|^2} m_0(dx) \leq \int_{\mathbb{R}^d} \big(1+|x|\big) m_0(dx) < \infty$$

We can thus define a probability measure in the following way for any Borel set $A \in \mathbb{R}^d$

$$\nu_0(A) := \frac{1}{C_0} \int_A \sqrt{1+|x|^2} m_0(dx),$$

where, for the sake of normalisation

$$C_0 = \int_{\mathbb{R}^d} \sqrt{1+|x|^2} m_0(dx).$$

Since $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$, $\{\nu_0\}$ is tight, and by Lemma 4.9 in [9], we have the existence of a radially non-decreasing function $V \in C^2(\mathbb{R}^d)$ on the form with $\|DV\|_\infty, \|D^2V\|_\infty \leq 1$ where

$$\lim_{|x|\to\infty} V(x) = \infty$$

and

$$\int_{\mathbb{R}^d} V(x)\nu_0(dx) \leq 1$$

Rewriting this integral with respect to m_0, we have

$$\int_{\mathbb{R}^d} V(x) \frac{1}{C_0} \sqrt{1+|x|^2} m_0(dx) \leq 1,$$

from which we can define the desired function $\psi(x) := \frac{1}{C_0}\sqrt{1+|x|^2}V(x)$. We see that

$$\frac{\psi(x)}{|x|} = \frac{1}{C_0} \frac{\sqrt{1+|x|^2}V(x)}{|x|} \leq \frac{1}{C_0} V(x) \xrightarrow{|x|\to\infty} \infty,$$

as desired. What remains of the proof is to establish the properties of ψ. ψ is obviously of class $C^2(\mathbb{R}^d)$, as it is a product of a C^2 and a C^∞ function. Furthermore,

$$D\psi(x) = \frac{1}{C_0}\left(\sqrt{1+|x|^2}DV(x) + V(x)\frac{x}{\sqrt{1+|x|^2}}\right),$$
$$D^2\psi(x) = \frac{1}{C_0}\left(V(x)D^2(\sqrt{1+|x|^2}) + \frac{x^T}{\sqrt{1+|x|^2}}DV + DV^T\frac{x}{\sqrt{1+|x|^2}} + \sqrt{1+|x|^2}D^2V\right),$$

When combining with the estimates $\|DV\|_\infty, \|D^2V\|_\infty \leq 1$, we get for some $C > 0$

$$|D\psi(x)| \leq C(1+|x|),$$
$$|D^2\psi(x)| \leq C(1+|x|),$$

which completes the proof. □

Having shown existence of such a "tightness inducing function" ψ for each initial condition $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$, we can show that a similar integral bound also holds for the solution of the Fokker-Planck equation with m_0 as initial distribution.

Lemma 3.11. *Let $m(t) = \mathcal{L}(X_t)$ be the solution to the Fokker-Planck equation (13) with $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ and $\|b\|_\infty < \infty$. Furthermore, let $\psi \in C^2(\mathbb{R}^d)$ be a non-negative function such that $|D\psi(x)|, |D^2\psi(x)| \leq C_0(1+\psi(x))$, for some constant $C_0 > 0$. Then*

$$\int_{\mathbb{R}^d} \psi(x)m(t)(dx) \leq \left(\int_{\mathbb{R}^d} \psi(x)m_0(dx) + CT(\epsilon + \|b\|_\infty)\right) e^{CT(\epsilon+\|b\|_\infty)}$$

Proof. In order to circumvent the need for integrability of the function ψ with respect to the family of measures $\{m(t)\}_{t\in[0,T]}$ we introduce a smooth cutoff function by convolving with a mollifier $\eta \in C_c^\infty(\mathbb{R}^d)$ as defined in Definition 2.4.

$$\chi_{B_k,1} := \mathbf{1}_{B_k} * \eta(x) = \int_{\mathbb{R}^d} \mathbf{1}_{B_k}(x-y)\eta(y)dy$$

We then define $\psi_k(x) := \psi(x)\chi_{B_k,1}$, which is twice continuously differentiable in space and constant in time, and we can apply Itô's lemma:

$$\psi_k(X_t) = \psi_k(Z_0) + \int_0^t [D\psi(X_s)\cdot b(X_s) + \epsilon\Delta\psi_k(X_s)]ds + \int_0^t D\psi_k(X_s)\cdot dB_s$$

Furthermore, since $\chi_{B_k,1} \in C_c^\infty(\mathbb{R}^d)$, the bounds on the derivatives of ψ yields the following estimates

$$\begin{aligned} D\psi_k(x)\cdot b(x) &\leq C\|b\|_\infty(\psi(x)+1) \\ \epsilon\Delta\psi_k(x) &\leq \epsilon C(\psi(x)+1) \\ \psi_k(x) &\leq \psi(x) \end{aligned}$$

for some common constant C. This yields

$$\psi_k(X_t) \leq \psi(Z_0) + \int_0^t C(\epsilon + \|b\|_\infty)(\psi_k(X_s)+1)ds + \int_0^t D\psi_k(X_s)\cdot dB_s$$

Taking the expectation of the inequality, where the integral with respect to the Brownian motion will vanish, and using Tonelli's theorem to swap the order of the integrals, we get

$$\int_{\mathbb{R}^d} \psi_k(x)m(t,dx) \leq \int_{\mathbb{R}^d} \psi(x)m_0(dx) + C(\epsilon + \|b\|_\infty)\int_0^t \left(1 + \int_{\mathbb{R}^d} \psi_k(x)m(s,dx)\right) ds$$

Grönwall's inequality applied to the map $t \to \int_{\mathbb{R}^d} \psi_k(x)m(t)(dx)$, which is bounded for each $k \in \mathbb{N}$, yields

$$\int_{\mathbb{R}^d} \psi_k(x)m(t,dx) \leq \left(\int_{\mathbb{R}^d} \psi(x)m_0(dx) + CT(\epsilon + \|b\|_\infty)\right) e^{CT(\epsilon+\|b\|_\infty)}$$

Finally, we apply Fatou's Lemma (Lemma 4.1 in [4]), noting that $\psi(x) = \lim_{k\to\infty} \psi_k(x)$

$$\int_{\mathbb{R}^d} \psi(x)m(t,dx) \leq \liminf_k \int_{\mathbb{R}^d} \psi_k(x)m(t,dx) \leq \left(\int_{\mathbb{R}^d} \psi(x)m_0(dx) + CT(\epsilon + \|b\|_\infty)\right) e^{CT(\epsilon+\|b\|_\infty)}$$

which is the estimate we wanted.

□

An immediate consequence of Lemma 3.11 is that solutions of the Fokker-Planck equation with a bounded drift term have a first moment.

Corollary 3.12. *Let $m(t) = \mathcal{L}(X_t)$ be the solution to the Fokker-Planck equation* (13) *with $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ and $\|b\|_\infty < \infty$. Then $m(t) \in \mathcal{P}_1(\mathbb{R}^d)$ for all $t \in [0,T]$.*

Proof. Choose $\psi(x) = \sqrt{1+|x|^2}$ in Lemma 3.11.

$$\begin{aligned} \int_{\mathbb{R}^d} |x|m(t)(dx) &\leq \int_{\mathbb{R}^d} \psi(x)m(t)(dx) \leq \left(\int_{\mathbb{R}^d} \sqrt{1+|x|^2}m_0(dx) + CT(\epsilon + \|b\|_\infty)\right) e^{CT(\epsilon+\|b\|_\infty)} \\ &\leq \left(\int_{\mathbb{R}^d} (1+|x|)m_0(dx) + CT(\epsilon + \|b\|_\infty)\right) e^{CT(\epsilon+\|b\|_\infty)} < \infty \end{aligned}$$

where the last inequality follows from $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. □

Having established boundedness of the integrals of the Fokker-Planck solutions against a tightness inducing function ψ in Lemma 3.11, we set out to convert this into compactness in $C([0,T],\mathcal{P}_1(\mathbb{R}^d))$ for a family of solutions, which is exactly what is needed to amend the moment assumption (15). To that end, we need the following technical lemma, which is proven in Lemma 2.5 in the project thesis [14].

Lemma 3.13. *Let $\mathcal{K} \subseteq \mathcal{P}_1(\mathbb{R}^d)$. If there exists a positive continuous function $\psi : \mathbb{R}^d \to \mathbb{R}$ where $\lim_{|x|\to\infty} \psi(x)/|x| = \infty$ and*

$$\sup_{m\in\mathcal{K}} \int_{\mathbb{R}^d} \psi(x)m(dx) = C < \infty$$

for some constant C, then $\mathcal{K}$ is precompact in $\mathcal{P}_1(\mathbb{R}^d)$ with respect to the d_1 metric.

Combining Lemmas 3.10, 3.11, and 3.13, we can finally construct the required compactness result.

Lemma 3.14. *Let $\{m_i\}_{i\in\mathcal{I}} \subset C([0,T],\mathcal{P}_1(\mathbb{R}^d))$ be a set of solutions to the Fokker-Planck equation (13) with respect to $\{b_i\}_{i\in\mathcal{I}}$, all sharing initial condition $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. Assume there exists a $K > 0$ such that $\sup_{i\in\mathcal{I}} \|b_i\|_\infty \leq K$. Then $\{m_i\}_{i\in\mathcal{I}}$ is precompact in $C([0,T],\mathcal{P}_1(\mathbb{R}^d))$*

Proof. We proceed by applying the version of the Arzelà-Ascoli Theorem found in [23] and [16], which requires the precompactness of $\{m_i(t)\}_{i\in\mathcal{I}}$ for each $t \in [0,T]$, as well as for every m_i to be equicontinous in t. For the precompactness for each $t \in [0,T]$ we note, using Lemma 3.10 and Lemma 3.11, that there exists a radially increasing function $\psi \in C^2(\mathbb{R}^d)$ with

$$\lim_{|x|\to\infty} \frac{\psi(x)}{|x|} = \infty$$

such that

$$\sup_{i\in\mathcal{I}} \int_{\mathbb{R}^d} \psi(x)m_i(t)(dx) \leq \left(\int_{\mathbb{R}^d} \psi(x)m_0(dx) + CT(\epsilon + K) \right) e^{CT(\epsilon+K)} < \infty$$

By Lemma 3.13, the set $\{m_i(t)\}_{i\in\mathcal{I}}$ is precompact in $\mathcal{P}_1(\mathbb{R}^d)$ for each $t \in [0,T]$. Furthermore, for the equicontinuity, we invoke Lemma 3.7 and get

$$d_1(m_i(t), m_i(s)) \leq K|t-s| + \sqrt{2\epsilon|t-s|} \quad \forall s,t \in [0,T]$$

Applying the Arzelà-Ascoli theorem, we conclude that $\{m_i\}_{i\in\mathcal{I}}$ is precompact in $C([0,T],\mathcal{P}_1(\mathbb{R}^d))$. ☐

Remark 3.15. Lemma 3.14 might not seem significant at first glance. However, when one inspects the proof of existence in Theorem 3.9, one can observe that the set stated as compact can serve as an "upgrade" in a central part in the proof. The core idea in establishing existence of solutions is an application of the Schauder fixed point theorem, which can be found in Chapter 9.2 of [12]. This theorem states that if X is a real Banach space, $K \subset X$ is nonempty, compact, and convex, and a map $A : K \to K$ is continuous, then A has a fixed point. To utilise this method, we recast the system (10) as a fixed point iteration in the following way. For some $\mu \in K \subset C([0,T],\mathcal{P}_1(\mathbb{R}^d))$ we denote by u_μ the solution of

$$\begin{cases} -\partial_t u_\mu - \varepsilon\Delta u_\mu + \frac{1}{2}|Du_\mu|^2 = F(x,\mu(t)) & \text{in } [t_0,T]\times\mathbb{R}^d, \\ u_\mu(T) = G(x,\mu(T)) & \text{in } \mathbb{R}^d. \end{cases}$$

Then we set $m := A(\mu)$ as the solution to

$$\begin{cases} \partial_t m - \varepsilon\Delta m - \operatorname{div}(mDu_\mu) = 0 & \text{in } [t_0,T]\times\mathbb{R}^d, \\ m(0) = m_0 & \text{in } \mathbb{R}^d. \end{cases}$$

Note that the existence of a fixed point $m = A(\mu) = \mu$ implies the existence of a solution to the system (10). By choosing K as the set $\{m_i\}_{i\in\mathcal{I}} \subset C([0,T],\mathcal{P}_1(\mathbb{R}^d))$ from Lemma 3.14, instead of the $\psi(x) = \sqrt{1+|x|^2}\log\sqrt{1+|x|^2}$ dependent set from the proof in [14], we rid ourselves of the moment assumption (15). The rest of the proof is identical.

Next, we have to remove the assumption $m_0 \in C^{2+\alpha}(\mathbb{R}^d)$. To this end, for any $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ we construct a family of regularisers $m_0^\delta \in C_b^\infty(\mathbb{R}^d)$ converging to m_0. We solve the MFG system for each m^δ and pass the solutions to the limit using compactness.

Theorem 3.16. *Assume that* (F) *and* (G) *holds. Then there exists a unique classical solution* (u, m) *of the MFG-system* (10) *for any initial condition* $(t_0, m_0) \in [0, T) \times \mathcal{P}_1(\mathbb{R}^d)$*. Moreover,* $(u, m) \in C^{1+\alpha/2,2+\alpha}([t_0, T] \times \mathbb{R}^d)) \times C([t_0, T], \mathcal{P}_1(\mathbb{R}^d))$*. Furthermore, if* $m_0 \in C^{2+\alpha}(\mathbb{R}^d)$*, then* $u, m \in C^{1+\alpha/2,2+\alpha}([t_0, T] \times \mathbb{R}^d)$*.*

Proof. We introduce the following δ-indexed family of mollifiers $\{\eta_\delta\}$ in accordance with Definition 2.4, and construct a family of mollified measures

$$m_0^\delta(x) := m_0 * \eta_\delta = \int_{\mathbb{R}^d} \eta_\delta(x - y) m_0(dy).$$

We can readily check that $m_0^\delta \in \mathcal{P}_1(\mathbb{R}^d) \cap C_b^\infty(\mathbb{R}^d)$. By interpreting m_0 as a measure and applying Tonelli's theorem (Theorem 4.4 in [4])

$$m_0^\delta(\mathbb{R}^d) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \eta_\delta(x - y) m_0(dy) \right) dx = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \eta_\delta(x - y) dx \right) m_0(dy) = \int_{\mathbb{R}^d} m_0(dy) = 1,$$

and for any $n \in \mathbb{N}$ and any multi index ℓ such that $|\ell| = n$ we have

$$\sup_{x \in \mathbb{R}^d} |D^\ell m_0^\delta(x)| \leq \int_{\mathbb{R}^d} \|D^\ell \eta_\delta\|_\infty m_0(dy) < \infty.$$

Since $m_0^\delta \in \mathcal{P}_1(\mathbb{R}^d) \cap C^{2+\alpha}(\mathbb{R}^d)$, we have by Theorem 3.9 and Remark 3.15 a unique solution $(u^\delta, m^\delta) \in C^{1+\frac{\alpha}{2},2+\alpha}([t_0, T] \times \mathbb{R}^d) \times C^{1+\frac{\alpha}{2},2+\alpha}([t_0, T] \times \mathbb{R}^d) \cap C([0, T], \mathcal{P}_1(\mathbb{R}^d))$ to the system

$$\begin{cases} -\partial_t u^\delta - \epsilon \Delta u^\delta + \frac{1}{2}|Du^\delta|^2 = F(x, m^\delta(t)), \\ u^\delta(T) = G(x, m^\delta(T)), \\ \partial_t m^\delta - \epsilon \Delta m^\delta - \operatorname{div}(m^\delta D u^\delta) = 0, \\ m^\delta(0) = m_0^\delta. \end{cases} \tag{17}$$

Since $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ we have by Lemma 3.10 a radial function $\psi \in C^2(\mathbb{R}^d)$, with $|D\psi(x)|, |D^2\psi(x)| \leq C(1 + |x|)$, for some constant $C > 0$, such that

$$\int_{\mathbb{R}^d} \psi(x) m_0(dx) \leq 1. \tag{18}$$

Furthermore, by Lemma 3.11 we get

$$\int_{\mathbb{R}^d} \psi(x) m^\delta(t)(dx) \leq \left(\int_{\mathbb{R}^d} \psi(x) m_0^\delta(dx) + CT(\epsilon + \|Du^\delta\|_\infty) \right) e^{CT(\epsilon + \|Du^\delta\|_\infty)}. \tag{19}$$

By (F) and (G), we have for some constant C_0

$$\sup_{m \in \mathcal{P}_1(\mathbb{R}^d)} (\|F(\cdot, m)\|_{3+\alpha} + \|G(\cdot, m)\|_{4+\alpha}) = C_0 < \infty,$$

which states that the Hölder norms of the coupling terms of (17), and hence also the bounding constant C_0, are m^δ-independent. We can fix an arbitrary $h \in \mathbb{R}^d$ and introduce the spatial shift $u_h^\delta(x) := u^\delta(x - h)$, which satisfies the equation

$$\begin{cases} -\partial_t u_h^\delta - \epsilon \Delta u_h^\delta + \frac{1}{2}|Du_h^\delta|^2 = F(x - h, m^\delta(t)), \\ u_h^\delta(T) = G(x - h, m^\delta(T)). \end{cases}$$

We can then choose super/subsolutions $w^{\pm}(t,x) \coloneqq u_h^\delta(t,x) \pm C_0(T-t+1)|h|$ for u^δ, which we can verify using the shifted equation above. Thus the comparison principle yields

$$\|u^\delta(t,x) - u^\delta(t,x-h)\|_\infty \leq C_0(1+T)|h|,$$

and by dividing both sides by $|h|$ and passing $|h| \to 0$ we can conclude that

$$\sup_\delta \|Du^\delta\|_\infty \leq C_0(1+T). \tag{20}$$

Moreover, again using Tonelli's theorem (Theorem 4.4 in [4])

$$\begin{aligned}\int_{\mathbb{R}^d} \psi(x) m_0^\delta(dx) &= \int_{\mathbb{R}^d} \psi(x) \int_{\mathbb{R}^d} \eta_\delta(x-y) m_0(dy) dx = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \psi(x)\eta_\delta(x-y)dx \right) m_0(dy) \\ &= \int_{\mathbb{R}^d} (\psi * \eta_\delta)\, m_0(dy).\end{aligned}$$

By noting the support of the mollifier η_δ, we have

$$\sup_{x\in B_r(0)} \psi * \eta_\delta(x) = \sup_{x \in B_r(0)} \int_{B_{r+\delta}(0)} \psi(y) * \eta_\delta(x-y) dy \leq \sup_{x\in B_{r+\delta}(0)} \psi(x).$$

Since ψ is radial and increasing, we can use the interpretation $\psi(x) = \overline{\psi}(|x|)$, where $\overline{\psi} \in C^2(\mathbb{R})$ and $\frac{d}{dt}\overline{\psi}(t) \leq C(1+t)$, which yields that

$$\psi * \eta_\delta(x) \leq \overline{\psi}(|x| + \delta).$$

This makes us able to bound the convolution product using the growth condition on the derivative of $\overline{\psi}$

$$\begin{aligned}\int_{\mathbb{R}^d} (\psi * \eta_\delta)\, m_0(dy) &\leq \int_{\mathbb{R}^d} \overline{\psi}(|y|+\delta) m_0(dy) = \int_{\mathbb{R}^d} \left(\overline{\psi}(|y|) + \int_{|y|}^{|y|+\delta} \frac{d}{dt}\overline{\psi}(s) ds \right) m_0(dy) \\ &\leq \int_{\mathbb{R}^d} \left(\psi(y) + \int_{|y|}^{|y|+\delta} C(1+s)ds \right) m_0(dy) \\ &= \int_{\mathbb{R}^d} \left(\psi(y) + C(\delta + |y|\delta + \frac{1}{2}\delta^2) \right) m_0(dy) \leq C_1(1+\delta+\delta^2),\end{aligned}$$

where the last inequality employs (18) and that $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$, and $C_1 > 0$ is a suitably chosen common constant. We desire to pass δ to 0 later in the proof, so we can consequently assume $\delta \leq 1$ without lack of generality. We combine the estimate with (19), and get

$$\sup_\delta \int_{\mathbb{R}^d} \psi(x) m^\delta(t)(dx) \leq (3C_1 + CT(\epsilon + K))\, e^{CT(\epsilon+K)}.$$

Thus, the set $\{m^\delta(t)\}_{\delta\in(0,1)}$ is precompact by Lemma 3.13. Moreover, from Lemma 3.7

$$d_1(m^\delta(t), m^\delta(s)) \leq \|Du^\delta\|_\infty |t-s| + \sqrt{2\epsilon|t-s|} \quad \forall s,t \in [0,T].$$

Thus, by (20), the set $\{m^\delta\}$ is equicontinous in $[0,T]$. By the Arzelà-Ascoli theorem [16, 23], $\{m^\delta\}_{\delta\in(0,1)}$ is precompact in $C([0,T], \mathcal{P}_1(\mathbb{R}^d))$.

Choose a decreasing sequence $\delta_n \to 0$, then by the definition of the metric $d_1(\cdot,\cdot)$

$$d_1(m_0^{\delta_n}, m_0) = \sup_{\phi \in 1-\mathrm{Lip}} \int_{\mathbb{R}^d} \phi(x)(m_0^{\delta_n} - m_0)(dx) = \sup_{\phi\in 1-\mathrm{Lip}} \int_{\mathbb{R}^d} (\phi * \eta_{\delta_n}(x) - \phi(x)) m_0(dx) \xrightarrow{\delta_n \to 0} 0.$$

By compactness, there exist a subsequence $\{m^{\delta_{n'}}\}$ which converges in $C([0,T], \mathcal{P}_1(\mathbb{R}^d))$ to some $\bar{m} \in C([0,T], \mathcal{P}_1(\mathbb{R}^d))$. Let $\bar{u}$ be the solution to the Hamilton-Jacobi-Bellman equation in (17) with respect to $\bar{m}$

$$\begin{cases} -\partial_t \bar{u} - \epsilon\Delta\bar{u} + \frac{1}{2}|D\bar{u}|^2 = F(x, \bar{m}(t)), \\ \bar{u}(T) = G(x, \bar{m}(T)). \end{cases} \tag{21}$$

Since F, G only requires $\bar{m}$ to be of class $C([0,T], \mathcal{P}_1(\mathbb{R}^d))$, and the equicontinuity estimate above also holds for$\bar{m}$, $\bar{u}$ is of class $C^{1+\alpha/2, 2+\alpha}(\mathbb{R}^d)$.

Next, in order to show that the sequence$\{u^{\delta_{n'}}\}$ converges to $\bar{u}$, we construct super/subsolutions for (21). Let $w^{\pm} = u^{\delta} \pm \big(\|F(\cdot, \bar{m}(t)) - F(\cdot, m^{\delta}(t))\|_{\infty}(T-t) + \|G(\cdot, \bar{m}(T)) - G(\cdot, m^{\delta}(T))\|_{\infty}\big)$. By the comparison principle, we have

$$\|\bar{u}(t) - u^{\delta}(t)\|_{\infty} \leq \|F(\cdot, \bar{m}(t)) - F(\cdot, m^{\delta}(t))\|_{\infty}(T-t) + \|G(\cdot, \bar{m}(T)) - G(\cdot, m^{\delta}(T))\|_{\infty},$$

which by the Lipschitz continuity of F, G in $d_1(\cdot, \cdot)$ implies

$$\sup_{t \in [0,T]} \|\bar{u}(t) - u^{\delta}(t)\|_{\infty} \leq C \sup_{t \in [0,T]} d_1(\bar{m}(t), m^{\delta}(t)).$$

Consequently, $\{u^{\delta_{n'}}\}$ converges uniformly to $\bar{u}$. Furthermore, since $u^{\delta_{n'}} \in C^{2+\alpha, 1+\frac{\alpha}{2}}(\mathbb{R}^d)$ with Hölder-norm independent of $\delta_{n'}$, we have

$$|Du^{\delta_{n'}}(t_1, x_1) - Du^{\delta_{n'}}(t_2, x_2)| \leq C(|t_1 - t_2|^{\frac{\alpha}{2}} + |x_1 - x_2|^{\alpha}).$$

Along with the bound (20), this implies through a version of the Arzelà-Ascoli theorem found in Appendix C.8 of [12] , which implies that there exists a further subsequence $\{Du^{\delta_{n''}}\} \subseteq \{Du^{\delta_{n'}}\}$ such that $\{Du^{\delta_{n''}}\}$ converges uniformly towards some object $\widetilde{Du} \in C(Q_T)$. We can identify $\widetilde{Du}$ with $D\bar{u}$ by fixing an arbitrary box $B := \prod_{i=1}^{d}[a_i, b_i] \subset \mathbb{R}^d$ containing the points $x_0 = (x_{01}, ..., x_{0d})$and $x = (x_{01}, ..., x_i, ..., x_{0d})$, with $x_{0i} < x_i$, and considering the sequence of difference quotients

$$\frac{u^{\delta_{n''}}(t,x) - u^{\delta_{n''}}(t, x_0)}{x_i - x_{0i}} = \partial_{x_i} u^{\delta_{n''}}(t, z).$$

For the equality, we have applied the mean value theorem with $z := (x_{01}, ..., z_i, ..., x_{0d})$ where $z_i \in (x_i, x_{0i})$. By the uniform convergence of the right hand side, the difference quotient converges to $\widetilde{Du}(t, x_0)$ as $x_i \to x_{0i}$, so the convergence of $\{u^{\delta_{n''}}\}$ implies that

$$\partial_{x_i} \bar{u}(t, x_0) = \lim_{x_i \to x_{0i}} \frac{\bar{u}(t,x) - \bar{u}(t, x_0)}{x_i - x_{0i}} = (\widetilde{Du})_i(t, x_0),$$

which shows that $\{Du^{\delta_{n''}}\}$ does indeed converge locally uniformly to $D\bar{u}$.

We complete the proof by showing that $\bar{m} \in C([0,T], \mathcal{P}_1(\mathbb{R}^d))$ is a distribution solution of the Fokker-Planck equation by passing to the limit in the equation (13)

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(t,x) m^{\delta_{n''}}(t)(dx) = &\int_{\mathbb{R}^d} \varphi(0,x) m_0^{\delta_{n''}}(dx) \\ &+ \int_0^t \int_{\mathbb{R}^d} [\varphi_t(s,x) + D\varphi(s,x) \cdot Du^{\delta_{n''}}(s,x) + \epsilon \Delta \varphi(s,x)] m^{\delta_{n''}}(s, dx) ds.\end{aligned}$$

For the terms involving only $m^{\delta_{n''}}$ and any derivative of the test function $\varphi \in C_c^{\infty}(\overline{Q_T})$ this is quite straightforward

$$\left| \int_{\mathbb{R}^d} \varphi(t,x) m^{\delta_{n''}}(t)(dx) - \int_{\mathbb{R}^d} \varphi(t,x) \bar{m}(t)(dx) \right| \leq \|D\varphi(t, \cdot)\|_{\infty} d_1(m^{\delta_{n''}}(t), \bar{m}(t)) \xrightarrow{\delta_{n''} \to 0} 0.$$

This procedure is the same for every term, except for the one containing $Du^{\delta_{n''}}$, which requires a splitting

$$\begin{aligned}&\left| \int_0^t \int_{\mathbb{R}^d} D\varphi(s,x) \cdot Du^{\delta_{n''}}(s,x) m^{\delta_{n''}}(s,dx) ds - \int_0^t \int_{\mathbb{R}^d} D\varphi(s,x) \cdot D\bar{u}(s,x) \bar{m}(s,dx) ds \right| \\ \leq &\left| \int_0^t \int_{\mathbb{R}^d} D\varphi(s,x) \cdot Du^{\delta_{n''}}(s,x)(m^{\delta_{n''}}(s,dx) - \bar{m}(s,dx)) ds \right| \\ + &\left| \int_0^t \int_{\mathbb{R}^d} D\varphi(s,x) \cdot (Du^{\delta_{n''}}(s,x) - D\bar{u}(s,x)) \bar{m}(s,dx) ds \right|.\end{aligned}$$

For the second term on the left hand side we utilise that $D\varphi$ has compact support to employ the locally compact convergence of $\{Du^{\delta_{n''}}\}$

$$\begin{aligned}&\left|\int_0^t\int_{\mathbb{R}^d} D\varphi(s,x)\cdot(Du^{\delta_{n''}}(s,x)-D\bar{u}(s,x))\bar{m}(s,dx)ds\right|\\ \leq&\|D\varphi\|_\infty\int_0^t\int_{\mathbb{R}^d}\bar{m}(s,dx)ds\sup_{(t,x)\in\mathrm{supp}\{D\varphi\}}\|Du^{\delta_{n''}}-D\bar{u}\|_\infty\xrightarrow{\delta_{n''}\to 0}0.\end{aligned}$$

For the other term we use that the spatial derivative of $D\varphi\cdot Du^{\delta_{n''}}$ is bounded independently of $\delta_{n''}$, and again bound by the definition of $d_1(\cdot,\cdot)$ and the Lipschitz constant of $D\varphi\cdot Du^{\delta_{n''}}$

$$\begin{aligned}&\left|\int_0^t\int_{\mathbb{R}^d} D\varphi(s,x)\cdot Du^{\delta_{n''}}(s,x)(m^{\delta_{n''}}(s,dx)-\bar{m}(s,dx))ds\right|\\ \leq&\|D(D\varphi\cdot Du^{\delta_{n''}})\|_\infty d_1(m^{\delta_{n''}}(t),\bar{m}(t))\xrightarrow{\delta_{n''}\to 0}0.\end{aligned}$$

Thus we can conclude that $\bar{m}$ is a distribution solution to the Fokker-Planck equation, and that $(\bar{u},\bar{m})$ is a classical solution to the Mean Field Game system.

The uniqueness of the solution is an immediate consequence of the Lasry-Lions monotonicity argument, which is treated in the next subsection. □

3.5 Monotonicity

A key discovery of Lasry and Lions, made while developing the early theory of Mean Field Games, is the monotonicity argument. This technique is a source to many properties of the solutions of mean field games, among others uniqueness, and is precisely what necessitates the monotonicity assumptions (14) upon the coupling terms F and G. Our version of the monotonicity argument, featured in the following lemma, is tailored for our choice of specific Hamiltonian $H(x,Du)=\frac{1}{2}|Du|^2$.

Lemma 3.17 (The Lasry-Lions monotonicity argument)**.** *Let (u^1,m^1) and (u^2,m^2) be solutions of (10) with initial conditions $m_0^1,m_0^2\in\mathcal{P}_1(\mathbb{R}^d)$. Then*

$$\int_{t_0}^T\int_{\mathbb{R}^d}|Du^1-Du^2|^2(m^1+m^2)(t,dx)dt\leq -2\int_{\mathbb{R}^d}(u^1(t_0,x)-u^2(t_0,x))(m_0^1(dx)-m_0^2(dx)) \tag{22}$$

Proof. The general version of this proof has become quite standard in MFG literature and can be found in [1],[5], and [7], all of which prove the more general statement

$$\begin{aligned}&\int_{t_0}^T\int_{\mathbb{R}^d}\left(H(x,Du^2)-H(x,Du^1)-D_pH(x,Du^1)\cdot(Du^2-Du^1)\right)m^1(t,dx)dt\\ &+\int_{t_0}^T\int_{\mathbb{R}^d}\left(H(x,Du^1)-H(x,Du^2)-D_pH(x,Du^2)\cdot(Du^1-Du^2)\right)m^2(t,dx)dt\\ &\leq-\int_{\mathbb{R}^d}(u^1(t_0,x)-u^2(t_0,x))(m_0^1(dx)-m_0^2(dx)).\end{aligned} \tag{23}$$

By inserting our specific Hamiltonian, and performing elementary computations we get that

$$\begin{aligned}&\left(H(x,Du^2)-H(x,Du^1)-D_pH(x,Du^1)\cdot(Du^2-Du^1)\right)m^1\\ &+\left(H(x,Du^1)-H(x,Du^2)-D_pH(x,Du^2)\cdot(Du^1-Du^2)\right)m^2\\ &=\frac{m^1-m^2}{2}\left(|Du^2|^2-|Du^1|^2\right)+\left(Du^1-Du^2\right)(m^1Du^1-m^2Du^2)\\ &=\frac{m^1+m^2}{2}\left|Du^1-Du^2\right|^2,\end{aligned}$$

which by insertion into (23) precisely yields the desired inequality. □

An immediate result of this inequality is the uniqueness of the solution of (10).

Proof of Theorem 3.16, uniqueness. Let (u^1, m^1) and (u^2, m^2) be solutions of (10) with the same initial condition $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. Then by Lemma 3.17, we have

$$\int_{t_0}^{T} \int_{\mathbb{R}^d} |Du^1 - Du^2|^2 (m^1 + m^2)(t, dx)dt \leq 0.$$

With m^1, m^2 being elements of $\mathcal{P}_1(\mathbb{R}^d)$, and hence nonnegative everywhere, this requires $|Du^1(t,x) - Du^2(t,x)|^2 = 0$ for $m^1 + m^2$- almost all $(t,x) \in [t_0, T] \times \mathbb{R}^d$. Recalling the distribution formulation of the Fokker-Planck (13), given some $\varphi \in C_c^\infty([t_0, T] \times \mathbb{R}^d)$ we have

$$\begin{aligned} &\left| \int_{t_0}^{t} \int_{\mathbb{R}^d} D\varphi(s,x) \cdot Du^1(s,x) m^1(s,dx)ds - \int_{t_0}^{t} \int_{\mathbb{R}^d} D\varphi(s,x) \cdot Du^2(s,x) m^1(s,dx)ds \right| \\ &\leq \|D\varphi\|_\infty \int_{t_0}^{T} \int_{\mathbb{R}^d} |Du^1 - Du^2| m^1(t,dx)dt = 0. \end{aligned}$$

This in turn implies that m^1 and m^2 solve the same Fokker-Planck equation, and are hence equal by the uniqueness of the equation. Similarly, since $m^1 = m^2$, u^1 and u^2 solve the same Hamilton-Jacobi-Bellman equation, and by uniqueness of that equation, we can conclude that $(u^1, m^1) = (u^2, m^2)$, and hence that the solution of (10) is unique. □

3.6 Higher Order Regularity of u

As a final subsection in our treatment of the Mean Field Game system, we demonstrate how to obtain higher order regularity for the solution u of (10) utilising the generous assumptions upon the coupling terms F, G. The technique employed will be differentiating the Hamilton-Jacobi-Bellman equation, and applying existence results for parabolic PDE to the resulting linear equation.

Lemma 3.18. *Let $(u, m) \in C^{1+\frac{\alpha}{2}, 2+\alpha}([t_0, T] \times \mathbb{R}^d)) \times C([t_0, T], \mathcal{P}_1(\mathbb{R}^d))$ be classical solutions of the system* (10) *with F, G obeying assumptions* (F) *and* (G) *respectively. Then there exists a (t_0, m_0)-independent constant $C_1 > 0$ such that*

$$\sum_{|\ell| \leq 2} \|D^\ell u\|_{1+\alpha/2, 2+\alpha} \leq C_1$$

Proof. To prove this statement, we adapt a proof for a part of Proposition 3.1.1. in [7]. For brevity, we skip some details, and the curious reader can consult how this methodology is implemented in [7] and in Theorem 1.7 in [1]. Let $v := Du \cdot e$ where $e \in \mathbb{R}^d$ is an arbitrary vector with $|e| = 1$. We compute the temporal derivative of v, and we use the equation to get

$$\begin{aligned} \partial_t v &= D\big(-\varepsilon \Delta u + \frac{1}{2}|Du|^2 - F(x, m(t))\big) \cdot e \\ &= -\varepsilon \Delta v + Du \cdot Dv - D_x F(x, m(t)) \cdot e. \end{aligned}$$

For the terminal condition, we have

$$v(T, x) = D_x u(T, x) \cdot e = D_x G(x, m(T)) \cdot e,$$

and hence, v satisfies the following parabolic partial differential equation

$$\begin{cases} -\partial_t v - \varepsilon \Delta v + Du \cdot Dv = D_x F(x, m(t)) \cdot e & \text{in } \mathbb{R}^d \times (0, T), \\ v(T) = D_x G(x, m(T)) \cdot e & \text{in } \mathbb{R}^d. \end{cases}$$

By the strong assumptions (F) and (G), and that by Theorem we have 3.16 $u \in C^{1+\alpha/2, 2+\alpha}([t_0, T] \times \mathbb{R}^d)$, the coefficients of the PDE are sufficiently Hölder continuous, so we can conclude by Theorem 5.1 in [18] that $v = Du \cdot e \in C^{1+\alpha/2, 2+\alpha}([t_0, T] \times \mathbb{R}^d)$. We can bootstrap this regularity once more, now using $v := D^2 ue \cdot e$, and by the precise same estimation method, we can conclude that $D^2 ue \cdot e \in C^{1+\alpha/2, 2+\alpha}([t_0, T] \times \mathbb{R}^d)$, and since e was chosen arbitrarily, the bound in the Lemma statement holds. □

This finalises out technical treatment of the Mean Field Game system, and we can continue to the primary focus of this paper, the well-posedness of the Master Equation.

4 The Master Equation

In this section we will perform analysis for existence and uniqueness of the Master Equation, which in our case for our explicit quadratic Hamiltonian $H(x,p) = \frac{1}{2}|p|^2$ takes the following form

$$\begin{cases} -\partial_t U(t,x,m) - \varepsilon \Delta_x U(t,x,m) + \frac{1}{2}|D_x U(x,t,m)|^2 \\ \quad -\varepsilon \int_{\mathbb{R}^d} \mathrm{div}_y [D_m U](t,x,m,y) m(dy) \\ \quad + \int_{\mathbb{R}^d} D_m U(t,x,m,y) \cdot D_x U(t,y,m) m(dy) \\ = F(x,m) & \text{in } [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d), \\ U(T,x,m) = G(x,m) & \text{in } \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d). \end{cases} \tag{24}$$

Here, as before, $D_m U \coloneqq \frac{\delta U}{\delta m}$ denotes the Lions derivative.

Before we start with the analysis, let us familiarise ourselves with the equation. We observe that, except for the integral terms containing the measure derivative D_m, the equation shares all terms with the Hamilton-Jacobi-Bellman equation from the Mean Field Game System (10). As the $D_m U$-terms correspond to a change in measure, one can intuitively think that the Master Equation correspond to a "nudging" of sorts of the Hamilton-Jacobi-Bellman in the measure variable.

If we assume that a solution $U(t,x,m)$ of (24) is constant with respect to the measure variable, we get by Definition 2.6 of the measure derivative that for any $m', m \in \mathcal{P}_1(\mathbb{R}^d)$

$$0 = \lim_{h \to 0^+} \frac{U(t,x,(1-h)m + hm') - U(t,x,m)}{h} = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t,x,m,y)(m' - m)(dy),$$

which by the normalisation of the measure derivative yields

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t,x,m,y) m'(dy) = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t,x,m,y) m(dy) = 0,$$

which can only be the case if $\frac{\delta U}{\delta m}(t,x,m,y) = 0$, $\forall m \in \mathcal{P}_1(\mathbb{R}^d)$, and hence that $D_m U(t,x,m,u) = 0$. In this case, the system reduces to the following form

$$\begin{cases} -\partial_t U(t,x,m) - \varepsilon \Delta_x U(t,x,m) + \frac{1}{2}|D_x U(x,t,m)|^2 = F(x,m) & \text{in } [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d), \\ U(T,x,m) = G(x,m) & \text{in } \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d). \end{cases} \tag{25}$$

This equation is really close to the Hamilton-Jacobi-Bellman equation, and if we let (u,m) be the solution to (10) with initial data $(t_0, m_0) \in [0,T] \times \mathcal{P}_1(\mathbb{R}^d)$, then $U(t_0, x, m_0) = u(t_0, x)$ solves the equation (25) in the point (t_0, x, m_0).

This heuristic computation motivates the following definition for an ansatz solution inspired by a method of characteristics.

Definition 4.1. Let $t_0 \in [0,T]$, $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$, and denote by (u,m) the solution to the Mean Field Game System (10) with initial condition (t_0, m_0). We define the Master Characteristic as the function $U : [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$ such that for all $x \in \mathbb{R}^d$

$$U(t_0, x, m_0) \coloneqq u(t_0, x) \tag{26}$$

We will prove in this chapter that the solutions springing from Master Characteristic are the classical solutions for the Master Equation. However, before we can start proving existence and uniqueness of classical solutions, we need to define them precisely.

Definition 4.2. A map $U : [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$ is a classical solution to the Master Equation if

- U is bounded and continuous in all arguments, specifically in the d_1-metric on $\mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$, and so is $\partial_t U,\ D_x U,\ D_x^2 U$, and $D_x^3 U$.
- U is $\mathcal{C}^1$ in m as described in Definition 2.6, and $\frac{\delta U}{\delta m}(t,x,m,y)$, $D_y D_x \frac{\delta U}{\delta m}(t,x,m,y)$, and $D_y^2 \frac{\delta U}{\delta m}(t,x,m,y)$ are bounded and continuous in all arguments in every $(t,x,m,y) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \times \mathbb{R}^d$.
- U satisfies the Master Equation (24).

Remark 4.3. Some might find the $D_x^3 U$ in the classical solution somewhat strange as the third derivative does not appear in the Master Equation. However, it is required in order to establish uniqueness of solutions, as the proof for uniqueness relies on the uniqueness of solutions of the Fokker-Planck equation. In order to apply Proposition 3.5, which supplies the uniqueness of the Fokker-Planck equation (12), we need the existence of $D_x^3 U$. If uniqueness for the Fokker-Planck with weaker assumptions are found, we are able to weaken Definition 4.2

With the classical solution well defined, we can finally state the main theorem of this thesis, namely the well-posedness of the Master Equation. The rest of this chapter will be dedicated to establishing the results leading up to its proof.

Theorem 4.4 (Well-posedness of the Master Equation)**.** *Assume that the boundedness and monotonicity assumptions* (F) *and* (G) *upon F and G hold for some $\alpha \in (0,1)$ and $\beta \in (0,\alpha)$. Then the Master Characteristic defined in* (26) *is the unique classical solution to the Master Equation* (24)*. Moreover, we have that for any $(t,m) \in [0,T] \times \mathcal{P}_1(\mathbb{R}^d)$, $U(t,\cdot,m)$ is bounded in $C^{4+\alpha}(\mathbb{R}^d)$ and $\frac{\delta U}{\delta m}(t,\cdot,m,\cdot)$ is bounded in $C^{4+\alpha}(\mathbb{R}^d) \times C^{2+\alpha-\beta}(\mathbb{R}^d)$, both independently of (t,m).*

Furthermore, for $m_1, m_2 \in \mathcal{P}_1(\mathbb{R}^d)$ and a constant C independent of m_1, m_2,

$$\|U(t,\cdot,m^1) - U(t,\cdot,m^2)\|_{4+\alpha} \leq C d_1(m^1,m^2). \tag{27}$$

For the increased legibility of the following results, we fix an arbitrary $\alpha \in (0,1)$ for the remainder of this chapter.

4.1 Lipschitz continuity in m_0

One of the main goals of this chapter is to be able to apply a method of characteristics with respect to the initial time t_0 and the initial measure variable m_0 in order to prove existence of solution of the Master Equation. In order to achieve this, we need to make sure that the characteristic solutions $U(t_0,x,m_0)$ change in a well behaved manner when the initial measure m_0 changes.

To this end we establish that the solution of the Mean Field Game system (10) depends Lipschitz continuously upon the initial measure $m_0 \in \mathbb{R}^d$. To help us establishing such a result, we turn to the following technical lemma.

Lemma 4.5. *Assume $V(t,x) \in C([0,T], C_b^{k-1}(\mathbb{R}^d))$, $f \in C([0,T], C_b^{k-1}(\mathbb{R}^d))$, and $g \in C_b^{k+\alpha}(\mathbb{R}^d)$ for some natural number $k \geq 2$. Then*

$$\begin{cases} -\partial_t z - \epsilon \Delta z + V(t,x) \cdot Dz = f(t,x) & \text{in } (0,\mathrm{T}) \times \mathbb{R}^\mathrm{d}, \\ z(T,x) = g(x) & \text{in } \mathbb{R}^\mathrm{d}. \end{cases}$$

has a unique classical solution z, where

$$\sup_{t \in [0,T]} \|z(t,\cdot)\|_{C^{k+\alpha}} \leq C(\|g\|_{C^{k+\alpha}} + \sup_{t \in [0,T]} \|f(t,\cdot)\|_{C^{k-1}})$$

Remark 4.6. This Lemma is part of Lemma 3.2.2. in [7], where the domain of the torus $\mathbb{T}^d$ has been swapped for our domain of the whole space $\mathbb{R}^d$. The result comes from the classical theory of linear PDE, and is well known according to [7], who cites a theorem from [18] for existence and uniqueness. The proof relies on a heat kernel argument, and can be proven for the whole space using the Duhamel formula combined with the Banach fixed point theorem. Treating these techniques properly is outside the scope of this thesis. A full proof for $\mathbb{R}^d$ and operators more general than the Laplacian is given in an upcoming paper by Espen R. Jakobsen and Artur Rutkowski at.

With this technical lemma in hand, we can establish that the solution of the Mean Field Game system is Lipschitz continuous with respect to its initial measure. We structure this insight in a lemma, which will be applied in the proofs of many of the upcoming results required for the well-posedness of the Master Equation.

Lemma 4.7. *Assume that* (F) *and* (G) *hold. Furthermore, let* (u^1, m^1), (u^2, m^2) *be solutions of the system* (10) *with respect to initial conditions* (t_0, m_0^1), (t_0, m_0^2), *where* $t_0 \in [0, T]$ *and* $m_0^1, m_0^2 \in \mathcal{P}_1(\mathbb{R}^d)$. *Then*

$$\sup_{t\in[0,T]} \left\{ d_1(m^1(t), m^2(t)) + \|u^1(t,\cdot) - u^2(t,\cdot)\|_{4+\alpha} \right\} \leq C d_1(m_0^1, m_0^2),$$

where C is a constant independent of t_0, m_0^1, and m_0^2. By Definition 4.1, we also get

$$\|U(t_0, \cdot, m_0^1) - U(t_0, \cdot, m_0^2)\|_{4+\alpha} \leq C d_1(m_0^1, m_0^2).$$

Proof. Using the monotone stability (22) and the definition of the metric d_1, we get

$$\begin{aligned} \int_{t_0}^{T} \int_{\mathbb{R}^d} |Du^2 - Du^1|^2 (m^1 + m^2)(t, dx)dt &\leq 2 \int_{\mathbb{R}^d} (u^1(t_0, x) - u^2(t_0, x))(m_0^1(dx) - m_0^2(dx)) \\ &\leq 2\|(Du^1 - Du^2)(t_0, \cdot)\|_\infty d_1(m_0^1, m_0^2). \end{aligned} \tag{28}$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a standard probability space and let X_0^1, X_0^2 be random variables with law m_0^1, m_0^2 respectively, such that $E[|X_0^1 - X_0^2|] = d_1(m_0^1, m_0^2)$. We also define the stochastic processes X_t^1, X_t^2 as the solution to the SDEs

$$\begin{cases} dX_t^i = -Du^i(t, X_t^i)dt + \sqrt{2\epsilon}dB_t & t \in (t_0, T], \\ X_t^i = X_0^i & t = t_0, \end{cases}$$

where $i = 1, 2$ and $\{B_t\}_{t\in[t_0,T]}$ is a standard d-dimensional Brownian motion. By the properties of the Fokker-Planck equation established earlier in Lemma 3.7, we have that m_t^i is the law of the process X_t^i for all $t \in [t_0, T]$.

Integrating the SDEs and taking the difference yields

$$|X_t^1 - X_t^2| = \left| (X_0^1 - X_0^2) + \int_{t_0}^{t} (Du^2(t, X_s^2) - Du^1(t, X_s^1))ds \right|.$$

Next, we take the expectation of the difference and apply the triangle inequality.

$$\begin{aligned} E\left[|X_t^1 - X_t^2|\right] &\leq E\left[|X_0^1 - X_0^2|\right] \\ &\quad + E\left[\int_{t_0}^{t} (|Du^1(t, X_s^1) - Du^1(t, X_s^2))| + |Du^1(t, X_s^2) - Du^2(t, X_s^2))|)ds\right]. \end{aligned}$$

Since $u^1, u^2 \in C^{1+\alpha/2,2+\alpha}([t_0, T] \times \mathbb{R}^d)$, with Hölder norm independent m_0^i, we have by Tonelli's theorem (Theorem 4.4 in [4])

$$
\begin{aligned}
E\left[|X_t^1-X_t^2|\right] &\le E\left[|X_0^1-X_0^2|\right]+\|D^2u^1\|_\infty\int_{t_0}^t E[|X_s^1-X_s^2|]ds\\
&\quad+\int_{t_0}^t\int_{\mathbb{R}^d}|Du^1(s,x)-Du^2(s,x)|m^2(s,dx)ds\\
&\le d_1(m_0^1,m_0^2)+\|D^2u^1\|_\infty\int_{t_0}^t(E[|X_s^1-X_s^2|]ds\\
&\quad+T\left(\int_{t_0}^t\int_{\mathbb{R}^d}|Du^1(s,x)-Du^2(s,x)|^2m^2(s,dx)ds\right)^{\frac12},
\end{aligned}
$$

where the last inequality is given by Jensen's inequality and the fact that since the map $x\mapsto x$ is trivially 1-Lipschitz, we have by the definition of d_1

$$
\begin{aligned}
E\left[X_0^1-X_0^2\right]&=\int_{\mathbb{R}^d}x(m_0^1-m_0^2)(dx)\le\sup_{\phi\in 1-\mathrm{Lip}}\int_{\mathbb{R}^d}\phi(x)(m_0^1-m_0^2)(dx)=d_1(m_0^1,m_0^2),\\
E\left[X_0^2-X_0^1\right]&=\int_{\mathbb{R}^d}x(m_0^2-m_0^1)(dx)\le\sup_{\phi\in 1-\mathrm{Lip}}\int_{\mathbb{R}^d}\phi(x)(m_0^2-m_0^1)(dx)=d_1(m_0^1,m_0^2).
\end{aligned}
$$

We combine this estimate with (28) and apply Grönwall's inequality.

$$
E\left[|X_t^1-X_t^2|\right]\le C\left[d_1(m_0^1,m_0^2)+\|(Du^1-Du^2)(0,\cdot)\|_\infty^{\frac12}d_1(m_0^1,m_0^2)^{\frac12}\right].
$$

Additionally, since d_1 is defined as a supremum with respect to 1−Lipschitz functions ϕ

$$
d_1(m^1(t),m^2(t))=\sup_\phi(\int_{\mathbb{R}^d}\phi(x)(m^1-m^2)(t,dx)=\sup_\phi E[\phi(X_t^1)-\phi(X_t^2)]\le E\left[|X_t^1-X_t^2|\right].
$$

so we can conclude

$$
\sup_{t\in[t_0,T]}d_1(m^1(t),m^2(t))\le C\left[d_1(m_0^1,m_0^2)+\|(Du^1-Du^2)(0,\cdot)\|_\infty^{\frac12}d_1(m_0^1,m_0^2)^{\frac12}\right]. \tag{29}
$$

Finally, we produce a bound on of the norm of the difference u^1-u^2. We introduce $z:=u^1-u^2$ and observe that it satisfies a linear PDE of the form

$$
\begin{cases}
-\partial_t z-\epsilon\Delta z+V(t,x)\cdot Dz=f(t,x) & \text{in } (0,T)\times\mathbb{R}^d,\\
z(T,x)=g(x) & \text{in } \mathbb{R}^d,
\end{cases}
$$

where $V(t,x)=\frac12(Du^1+Du^2)$ and by the measure version of the fundamental theorem of calculus from Lemma 2.7

$$
\begin{aligned}
f(t,x)&=\int_0^1\int_{\mathbb{R}^d}\frac{\delta F}{\delta m}(x,sm^1(t)+(1-s)m^2(t),y)(m^1(t,dy)-m^2(t,dy))ds,\\
g(x)&=\int_0^1\int_{\mathbb{R}^d}\frac{\delta G}{\delta m}(x,sm^1(T)+(1-s)m^2(T),y)(m^1(T,dy)-m^2(T,dy))ds.
\end{aligned}
$$

We have for $|\ell|\le 3$

$$
\begin{aligned}
D^\ell f(t,x)&=\int_0^1\int_{\mathbb{R}^d}D_x^\ell\frac{\delta F}{\delta m}(x,sm^1(t)+(1-s)m^2(t),y)(m^1(t,dy)-m^2(t,dy))ds\\
&\le\int_0^1\left\|\frac{\delta F}{\delta m}(\cdot,sm^1(t)+(1-s)m^2(t),\cdot)\right\|_{(3+\alpha,1)}d_1(m^1(t,dy),m^2(t,dy))ds,
\end{aligned}
$$

and thus

$$\|f(t,\cdot)\|_3 \leq Cd_1(m^1(t), m^2(t)),$$

where the constant C is independent of the choice of m^1, m^2 by the assumption (F). By the same argument, we also have

$$\|g(\cdot)\|_{4+\alpha} \leq Cd_1(m^1(T), m^2(T)).$$

Consequently, by Lemma 4.5 with $k = 4$, we have

$$\sup_{t\in[t_0,T]} \|(u^1 - u^2)(t,\cdot)\|_{4+\alpha} \leq C(\|g\|_{4+\alpha} + \sup_{t\in[t_0,T]} \|f(t,\cdot)\|_3) \leq C \sup_{t\in[t_0,T]} d_1(m^1(t), m^2(t)),$$

which we combine with (29) and rearrange to get

$$\sup_{t\in[t_0,T]} \|(u^1 - u^2)(t,\cdot)\|_{3+\alpha} \leq Cd_1(m_0^1, m_0^2).$$

Finally, we insert this back into (29), which yields

$$\begin{aligned}
\sup_{t\in[t_0,T]} d_1(m^1(t), m^2(t)) &\leq C\left[d_1(m_0^1, m_0^2) + \|(Du^1 - Du^2)(0,\cdot)\|_\infty^{\frac{1}{2}} d_1(m_0^1, m_0^2)^{\frac{1}{2}}\right] \\
&\leq C\left[d_1(m_0^1, m_0^2) + \left(\sup_{t\in[t_0,T]} \|(u^1 - u^2)(t,\cdot)\|_{3+\alpha}\right)^{\frac{1}{2}} d_1(m_0^1, m_0^2)^{\frac{1}{2}}\right] \\
&\leq C\left[d_1(m_0^1, m_0^2) + \left(d_1(m_0^1, m_0^2)\right)^{\frac{1}{2}} d_1(m_0^1, m_0^2)^{\frac{1}{2}}\right] \\
&\leq Cd_1(m_0^1, m_0^2),
\end{aligned}$$

which is the estimate we sought to prove.

□

With Lipschitz continuity in hand, we next set out to establish existence of the measure derivative of the Master Characteristic.

4.2 A Linearised System

The first step on the way to constructing the measure derivative $\frac{\delta U}{\delta m}$ is the introduction of the main technical tool of this chapter, the well-posedness of a quite general forward-backward system of linear equations. It takes the following form

$$\begin{cases}
-\partial_t z - \varepsilon\Delta z + V(t,x)\cdot Dz = \langle \frac{\delta F}{\delta m}(x, m(t)), \rho(t)\rangle + b(t,x) & \text{in } [t_0,T]\times\mathbb{R}^d, \\
z(T,x) = \langle \frac{\delta G}{\delta m}(x, m(T)), \rho(T)\rangle + z_T(x) & \text{in } \mathbb{R}^d, \\
\partial_t\rho - \varepsilon\Delta\rho - \operatorname{div}(\rho V) - \operatorname{div}(mDz + c) = 0 & \text{in } [t_0,T]\times\mathbb{R}^d, \\
\rho(t_0) = \rho_0 & \text{in } \mathbb{R}^d.
\end{cases} \tag{30}$$

Here, $V : [t_0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ is a vector field and $c : [t_0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ is a map. Furthermore, assume that $m \in C([t_0,T], \mathcal{P}_1(\mathbb{R}^d))$ where

$$d_1(m(t), m(s)) \leq C|t-s|^{\frac{1}{2}}, \quad \forall t,s \in [t_0,T]. \tag{31}$$

This system might seem a bit unruly at first glance. However, what it lacks in clarity, it makes up for in applicability, and will be at the core of every proof from now up until we can prove Theorem 4.4. For the well-posedness of this system we have the following comprehensive and complicated result

Lemma 4.8. *First, let $k \in \{1,2\}$, $\sigma \in [0,1)$ and $\beta \in (0,\sigma)$. Let $b \in C([t_0,T], C^{k+1+\sigma}(\mathbb{R}^d))$, $z_T \in C^{k+2+\sigma}(\mathbb{R}^d)$, and $\rho_0 \in C^{-(k+1+\sigma-\beta)}(\mathbb{R}^d)$. Furthermore, assume that* (31) *holds, let $V \in C([t_0,T], C_b^{k+1}(\mathbb{R}^d))$, $c \in C([t_0,T], C^{-(k+\sigma-\beta)}(\mathbb{R}^d))^d$, and assume that the family $\{c(t) * \eta_\gamma : t \in [t_0,T]\}$ is tight for any $\gamma > 0$. Lastly, assume that F, G obeys* (F) *and* (G)*. Then the system* (30) *has a unique solution $(z,\rho) \in C([t_0,T], C^{k+2+\sigma-\beta}(\mathbb{R}^d) \times C^{-(k+1+\sigma+\beta)}(\mathbb{R}^d))$ where*

$$\sup_{t\in[t_0,T]} \left(\|z(t,\cdot)\|_{k+2+\sigma} + \|\rho(t)\|_{-(k+1+\sigma)}\right) \le C_k M,$$

where C_k depends on $k, T, \beta, \sup_{t\in[0,T]} \|V(t,\cdot)\|_{k+1+\sigma}$ and M is defined by

$$M := \|z_T\|_{k+2+\sigma} + \|\rho_0\|_{-(k+1+\sigma)} + \sup_{t\in[t_0,T]} \|b(t,\cdot)\|_{k+1+\sigma} + \sup_{t\in[t_0,T]} \|c(t,\cdot)\|_{-(k+\sigma)}.$$

The solution of the second equation has to be interpreted in the distributional sense, that is, for each function $\phi : [t_0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that for every $t \in [t_0,T]$ $\phi(t_0), \phi(t) \in C^{k+1+\sigma}(\mathbb{R}^d)$ and $\partial_t \phi + \varepsilon\Delta\phi + V \cdot D\phi \in C([t_0,t], C^{k+1+\sigma}(\mathbb{R}^d))$ then

$$\begin{aligned}\langle \rho(t), \phi(t)\rangle - \langle \rho_0, \phi(t_0)\rangle = &\int_{t_0}^t \langle \rho(s), (\partial_t\phi + \varepsilon\Delta\phi + V\cdot D\phi)(s)\rangle ds \\ &- \int_{t_0}^t \int_{\mathbb{R}^d} D\phi Dz(s,x) m(s,dx) ds - \int_{t_0}^t \langle c(s), D\phi(s)\rangle ds.\end{aligned} \tag{32}$$

Remark 4.9. This workhorse of a Lemma is the $\mathbb{R}^d$ equivalent of Lemma 3.3.1. from [7]. However, since we lose the luxury of a compact domain, Lemma 4.8 is quite a bit more complicated, both in statement and in proof, than its $\mathbb{T}^d$-based sibling. The proof of the lemma is dependent on approximation of functionals by smooth functions, which is the reason for the pathological-looking losses of arbirtrarily small orders of regularity $\beta \in (0,\sigma)$. The proof of this Lemma is outside of scope for this thesis, but the core of the proof relies on an application of the Leray–Schauder theorem in order to get simultaneous solutions for both equations. A full proof of the result is given in an upcoming paper by Espen R. Jakobsen and Artur Rutkowski.

4.3 Differentiability of U With Respect to m

In this subsection, we derive all necessary results for the establishment of existence, and construction of, the measure derivative $\frac{\delta U}{\delta m}$. Having established the technical Lemma 4.8, we straight away apply it to a system that in many ways can be considered to be the linearisation of the Mean Field Game system (10) with respect to the measure variable.

$$\begin{cases} -\partial_t v - \varepsilon\Delta v + Du\cdot Dv = \langle \frac{\delta F}{\delta m}(x, m(t)), \mu(t)\rangle & \text{in } [t_0,T]\times\mathbb{R}^d, \\ v(T,x) = \langle \frac{\delta G}{\delta m}(x, m(T)), \mu(T)\rangle & \text{in } \mathbb{R}^d, \\ \partial_t\mu - \varepsilon\Delta\mu - \operatorname{div}(\mu Du) - \operatorname{div}(mDv) = 0 & \text{in } [t_0,T]\times\mathbb{R}^d, \\ \mu(t_0) = \mu_0 & \text{in } \mathbb{R}^d. \end{cases} \tag{33}$$

We fix an arbitrary $\beta \in (0,\alpha)$ which, as can be seen in the statement of Lemma 4.8, encodes the infinitesimal loss of regularity associated with the application of the technical result. The first step on the way to finding the measure derivative is by solving, and establishing the desired regularity of the solutions of, system (33). We structure this in the following result, where we bootstrap regularity in the v and μ variable by applying Lemma 4.8 twice.

Proposition 4.10. *Let assumption* (F) *and* (G) *hold. If $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ and $\mu_0 \in C^{-(3+\alpha-\beta)}$ for some $\beta \in (0,\alpha/2)$, we have a unique solution $(v,\mu) \in C([t_0,T], C^{4+\alpha-\beta}(\mathbb{R}^d) \times C^{-(3+\alpha+\beta)}(\mathbb{R}^d))$ of* (33) *and*

$$\sup_{t\in[t_0,T]} \left(\|v(t,\cdot)\|_{4+\alpha} + \|\mu(t)\|_{-(3+\alpha)}\right) \le C\|\mu_0\|_{-(3+\alpha)}.$$

Furthermore, if $\mu_0 \in C^{-(2+\alpha-2\beta)}$*, we have* $(v,\mu) \in C([t_0,T], C^{4+\alpha-\beta}(\mathbb{R}^d) \times C^{-(2+\alpha)}(\mathbb{R}^d))$*, where*

$$\sup_{t\in[t_0,T]} \big(\|v(t,\cdot)\|_{4+\alpha} + \|\mu(t)\|_{-(2+\alpha-\beta)}\big) \le C\|\mu_0\|_{-(2+\alpha-\beta)}.$$

The constant $C > 0$ *is independent of* (t_0, m_0)*.*

Proof. For the proof we apply Lemma 4.8 twice in a straightforward manner, using the same $\rho_0 = \mu_0 \in C^{-(2+\alpha-2\beta)}$. Firstly, we choose $k = 1$ and $\sigma = \alpha - \beta$ and set $z_T = b = c = 0$. We note that $V = Du \in C^{1+\alpha/2,2+\alpha}(\mathbb{R}^d) \subset C([0,T], C_b^2(\mathbb{R}^d))$ by Lemma 3.18. $\{c(t) * \eta_\gamma : t \in [t_0,T]\} = \{0\}$ is a singleton set, and is thus trivially tight. This results in a unique solution $(v,\mu) \in C([t_0,T], C^{3+\alpha-2\beta}(\mathbb{R}^d) \times C^{-(2+\alpha)}(\mathbb{R}^d))$ with

$$\sup_{t\in[t_0,T]} \big(\|v(t,\cdot)\|_{3+\alpha-\beta} + \|\mu(t)\|_{-(2+\alpha-\beta)}\big) \le C\|\mu_0\|_{-(2+\alpha-\beta)}.$$

Next we apply the same Lemma 4.8 again, now with $k = 2$ and $\sigma = \alpha$ with the same coefficients z_T, b, c, ρ_0, noting that $\rho_0 = \mu_0 \in C^{-(2+\alpha-2\beta)} \subset C^{-(3+\alpha-\beta)}$ by Lemma 2.3. $z_T = b = c = 0$. Again, $V = Du \in C^{1+\alpha/2,3+\alpha}(\mathbb{R}^d) \subset C([0,T], C_b^3(\mathbb{R}^d))$ by Lemma 3.18. This once more results in a unique solution $v, \mu \in C([t_0,T], C^{4+\alpha-\beta}(\mathbb{R}^d) \times C^{-(3+\alpha)}(\mathbb{R}^d))$ with

$$\sup_{t\in[t_0,T]} \big(\|v(t,\cdot)\|_{4+\alpha} + \|\mu(t)\|_{-(3+\alpha)}\big) \le C\|\mu_0\|_{-(3+\alpha)} \le C\|\mu_0\|_{-(2+\alpha-\beta)}.$$

By the linearity of the equation (33), and the uniqueness from Lemma 4.8 we conclude that the pair (v,μ) obtained by the application of the lemma with $k = 2$ is the same as the pair obtained with $k = 1$. □

If (33) indeed is the linearisation with respect to the measure variable, we would expect the following relation to hold in some sense

$$v(t_0,x) = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0,x,m_0,y)\mu_0(y)dy,$$

for some suitable choice of μ_0. Having established existence and regularity of the linearised system in Proposition 4.10, we can utilise the flexibility brought by being able to choose any $\mu_0 \in C^{-(2+\alpha-\beta)}$ to construct a candidate function $K(t_0,x,m_0,y)$ that should satisfy the properties required of it from Definitions 2.6 and 4.2 of the measure derivative and the classical solution of the Master Equation respectively. We formalise this choice of candidate in the following proposition.

Proposition 4.11. *With the same assumptions as Proposition 4.10, for each* (t_0,m_0) *we have a* $C^{4+\alpha-\beta} \times C^{2+\alpha-\beta}$ *map* $(x,y) \mapsto K(t_0,x,m_0,y)$ *such that if* $\mu_0 \in C^{-(2+\alpha-2\beta)}(\mathbb{R}^d)$ *is a finite signed measure, the* v*-component of the solution of* (33) *is given by*

$$v(t_0,x) = \langle \mu_0, K(t_0,x,m_0,\cdot)\rangle.$$

Furthermore K *satisfies*

$$\|K(t_0,\cdot,m_0,\cdot)\|_{(4+\alpha,3+\alpha)} \le C,$$

with constant $C > 0$ *independent of* (t_0,m_0)*, and has derivatives in* (x,y) *of order* $4+\alpha-\beta$ *and* $2+\alpha-\beta$ *which are continuous on* $[0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}) \times \mathbb{R}^d$*, and Lipschitz on* $\mathcal{P}_1(\mathbb{R})$*.*

Proof. Let $\ell \in \mathbb{N}^d$ be a multi index such that $|\ell| \le 3$ and $y \in \mathbb{R}^d$. We denote by $(v^{(\ell)}(\cdot,\cdot,y), \mu^{(\ell)}(\cdot,\cdot,y))$ the solution of (33) with respect to initial condition $\mu_0 = D^\ell \delta_y$, the ℓ-th distributional derivative of the Dirac distribution. We can readily check that $\mu_0 \in C^{-(3+\alpha-\beta)}$, by applying the definition of the dual norm (5). Let ϕ be an arbitrary function in $C^{3+\alpha-\beta}(\mathbb{R}^d)$ with $\|\phi\|_{3+\alpha-\beta} \le 1$ then

$$|\langle D^\ell \delta_y, \phi\rangle| = |(-1)^{|\ell|}\langle \delta_y, D^\ell \phi\rangle| = |D^\ell \phi(y)| \le \|\phi\|_{3+\alpha-\beta} \le 1.$$

Next, let $K(t_0, x, m_0, y) := v^{(0)}(t_0, x, y)$. We will check that $\partial_{y_i} K(t_0, x, m_0, y) = -v^{(e_i)}$, where $e_i = (0, ..., 0, 1, 0, ..., 0)$ is the i-th unit vector in $\mathbb{R}^d$. We first note that since

$$\begin{aligned}\|\frac{1}{\epsilon}(\delta_{y+\epsilon e_i} - \delta_y) + D^{(e_i)}\delta_y\|_{-(3+\alpha-\beta)} &= \sup_{\phi \in C^{3+\alpha-\beta}, \|\phi\| \leq 1} |\langle \frac{1}{\epsilon}(\delta_{y+\epsilon e_i} - \delta_y) + D^{(e_i)}\delta_y, \phi\rangle| \\ &= \sup_{\phi \in C^{3+\alpha-\beta}, \|\phi\| \leq 1} |\frac{1}{\epsilon}(\phi(y + \epsilon e_i) - \phi(y)) - D^{(e_i)}\phi| \xrightarrow{\epsilon \to 0} 0,\end{aligned}$$

$\frac{1}{\epsilon}(\delta_{y+\epsilon e_i} - \delta_y)$ converges to $-D^{(e_i)}$ in $C^{-(3+\alpha-\beta)}(\mathbb{R}^d)$. By the bound given in Proposition 4.10, and the linearity of the system (33) in v, μ, we can conclude that the map $\mu_0 \mapsto (v, \mu)$ is linear and continuous from $C^{-(2+\alpha-\beta)}$ to $C([0, T], C^{4+\alpha-\beta} \times C^{-(3+\alpha+\beta)})$. Thus, by the linearity in the v-term of the map, we have

$$\frac{1}{\epsilon}(\delta_{y+\epsilon e_i} - \delta_y) \mapsto \frac{1}{\epsilon}(K(\cdot, \cdot, m_0, y + \epsilon e_i) - K(\cdot, \cdot, m_0, y)).$$

Due to the map being continuous, we can pass to the limit and conclude that

$$-v^{(e_i)} = \partial_{y_i} K(\cdot, \cdot, m_0, y).$$

Repeating this process for derivatives up to $|\ell| \leq 3$ yields that

$$D_y^\ell K(t_0, x, m_0, y) = (-1)^{|\ell|} v^{(\ell)}(t_0, x, y). \tag{34}$$

This implies through the bound in Proposition 4.10 that for each $y \in \mathbb{R}^d$

$$\|D_y^\ell K(t_0, \cdot, m_0, y)\|_{4+\alpha-\beta} = \|v^{(\ell)}(t_0, \cdot, y)\|_{4+\alpha-\beta} \leq C.$$

Next, we check the Hölder continuity in y by applying the linearity of the system

$$\begin{aligned}&\|D_y^\ell K(t_0, \cdot, m_0, y) - D_y^\ell K(t_0, \cdot, m_0, y')\|_{4+\alpha} \\ &\leq C\|D^\ell \delta_y - D^\ell \delta_{y'}\|_{-(3+\alpha)} = C\left(\sup_{\|\phi\|_{3+\alpha} \leq 1} |\langle \delta_y - \delta_{y'}, D^\ell \phi\rangle|\right) \\ &\leq C\left(\sup_{\|\phi\|_{3+\alpha} \leq 1} |D^\ell \phi(y) - D^\ell \phi(y')|\right) \leq C|y - y'|^\alpha.\end{aligned}$$

Consequently, we can conclude that $K(t_0, \cdot, m_0, \cdot) \in C^{4+\alpha-\beta} \times C^{3+\alpha}$. This immediately implies the continuity of K and its derivatives in x and y.

For the continuity of the derivatives in (x, y) on $[0, T] \times \mathcal{P}_1(\mathbb{R})$ we let m_0^1, m_0^2 be two different initial measures in $\mathcal{P}_1(\mathbb{R}^d)$. Let $(u^1, m^1), (u^2, m^2)$ be solutions to the Mean Field Game system (10), with initial conditions (t_0, m_0^1) and (t_0, m_0^2) respectively, and let $(v^1, \mu^1), (v^2, \mu^2)$ be the respective solutions of the linear system (33), both with $\mu_0 = D^\ell \delta_y$. Here, we only have $|\ell| \leq 2$, since we need to have $\mu_0 \in C^{-(2+\alpha-2\beta)}$to apply both inequalities in Proposition 4.10. In order to estimate the difference, we let $(z, \rho) := (v^1 - v^2, \mu^1 - \mu^2)$, which gives rise to another linearised system

$$\begin{cases} -\partial_t z - \varepsilon\Delta z + Du^1 \cdot Dz = \langle \frac{\delta F}{\delta m}(x, m^1(t)), \rho(t)\rangle + b(t, x) & \text{in } [t_0, T] \times \mathbb{R}^d, \\ z(T, x) = \langle \frac{\delta G}{\delta m}(x, m^1(T)), \rho(T)\rangle + z_T & \text{in } \mathbb{R}^d, \\ \partial_t \rho - \varepsilon\Delta\rho - \operatorname{div}(\rho Du^1) - \operatorname{div}(m^1 Dz + c) = 0 & \text{in } [t_0, T] \times \mathbb{R}^d, \\ \rho(t_0) = 0 & \text{in } \mathbb{R}^d. \end{cases}$$

where

$$\begin{aligned}b(t, x) &= \langle \frac{\delta F}{\delta m}(x, m^1(t)) - \frac{\delta F}{\delta m}(x, m^2(t)), \mu^2(t)\rangle + (Du^2 - Du^1)Dv^2, \quad (35)\\ z_T(x) &= \langle \frac{\delta G}{\delta m}(x, m^1(T)) - \frac{\delta G}{\delta m}(x, m^2(T)), \mu^2(T)\rangle, \\ c(t, x) &= (Du^1 - Du^2)\mu^2 + (m^1 - m^2)Dv^2.\end{aligned}$$

Again we want to use Lemma 4.8 with $k = 2$ and $\sigma \coloneqq \alpha$. We proceed to check the regularity of the various terms. Again $V = Du^1 \in C^{1+\alpha/2,3+\alpha}(\mathbb{R}^d) \subset C([0,T], C_b^3(\mathbb{R}^d))$ by Lemma 3.18. For b, we apply the linearity and continuity of the duality brackets, as well as the definition of the operator norm, to get

$$\begin{aligned}\sup_{t\in[t_0,T]} \|b(t,\cdot)\|_{3+\alpha} &\leq \sup_{t\in[t_0,T]} \Big(\| \frac{\delta F}{\delta m}(\cdot, m^1(t), \cdot) - \frac{\delta F}{\delta m}(\cdot, m^2(t), \cdot) \|_{(3+\alpha,3+\alpha)} \|\mu^2\|_{-(3+\alpha)} \\ &\quad + C\|Du^2 - Du^1\|_{3+\alpha} \|Dv^2\|_{3+\alpha} \Big) \\ &\leq C\|\mu_0\|_{-(2+\alpha-\beta)} \sup_{t\in[t_0,T]} \Big(d_1(m^1(t), m^2(t)) + \|u^1(t,\cdot) - u^2(t,\cdot)\|_{4+\alpha} \Big) \\ &\leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)}.\end{aligned}$$

Here, the second inequality used the Lipschitz continuity in $\mathcal{P}(\mathbb{R}^d)$ from assumption (F) as well as the regularity on v^2, μ^2 from Proposition 4.10, while the last inequality utilised the Lipschitz continuity from Lemma 4.7. Furthermore, applying the triangle inequality and the second term of (F), we get that $\sup_{t\in[t_0,T]} \|b(t,\cdot)\|_{3+\alpha} < \infty$.

By the exact same estimation, instead using (G), we also get

$$\|z_T\|_{4+\alpha} \leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)}.$$

Finally, we check $c(t)$:

$$\sup_{t\in[t_0,T]} \|c(t,\cdot)\|_{-(2+\alpha-\beta)} = \sup_{t\in[t_0,T]} \Big(\sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \langle (Du^1 - Du^2)\mu^2, \phi\rangle + \langle (m^1(t) - m^2(t)) Dv^2, \phi\rangle \Big),$$

where again from the estimates of Proposition 4.10 and Lemma 4.7

$$\begin{aligned}\sup_{t\in[t_0,T]} \Big(\sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \langle (Du^1 - Du^2)\mu^2, \phi\rangle \Big) &\leq \sup_{t\in[t_0,T]} \Big(\sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \|\mu^2(t)\|_{-(2+\alpha-\beta)} \|Du^1 - Du^2\|_{2+\alpha} \Big) \\ &\leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)},\end{aligned}$$

and, using the definition of the d_1-norm

$$\begin{aligned}\sup_{t\in[t_0,T]} \Big(\sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \langle (m^1(t) - m^2(t)) Dv^2, \phi\rangle \Big) &\leq \sup_{t\in[t_0,T]} \Big(\sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \int_{\mathbb{R}^d} Dv^2 \phi (m^1(t) - m^2(t))(dx) \Big) \\ &\leq C\|v^2\|_{2+\alpha} \sup_{t\in[t_0,T]} d_1(m^1(t), m^2(t)) \\ &\leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)}.\end{aligned}$$

We can thus conclude that

$$\sup_{t\in[t_0,T]} \|c(t,\cdot)\|_{-(2+\alpha-\beta)} \leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)}.$$

By noting that $\{c(t) * \eta_\gamma : t \in [t_0, T]\}$ is tight for any $\gamma > 0$, we have by application of Lemma 4.8 that

$$\sup_{t\in[t_0,T]} \|z(t,\cdot)\|_{(4+\alpha)} \leq C d_1(m_0^1, m_0^2) \|\mu_0\|_{-(2+\alpha-\beta)}.$$

Combining this with the fact that

$$\|\mu_0\|_{-(2+\alpha-\beta)} = \sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} |\langle D^\ell \delta_y, \phi\rangle| \leq \sup_{\|\phi\|_{2+\alpha-\beta}\leq 1} \|\phi\|_{2+\alpha-\beta} \leq 1,$$

we have through (34) that for every $|\ell| \leq 2$

$$\|D_y^\ell K(t_0, \cdot, m_0^1, y) - D_y^\ell K(t_0, \cdot, m_0^2, y)\|_{(4+\alpha)} \leq C d_1(m_0^1, m_0^2).$$

Lastly, in order to show the Hölder continuity, we redefine $\mu_0 = D^\ell \delta_y - D^\ell \delta_{y'}$, a change which alters none of the estimates above, and compute

$$\|\mu_0\|_{-(2+\alpha-\beta)} = \sup_{\|\phi\|_{2+\alpha-\beta}\le 1} |\langle D^\ell \delta_y - D^\ell \delta_{y'}, \phi\rangle| \le |y-y'|^{2+\alpha-\beta},$$

which by the linearity of (33) implies

$$\begin{aligned}&\|\big(D_y^\ell K(t_0,\cdot,m_0^1,y) - D_y^\ell K(t_0,\cdot,m_0^1,y')\big) - \big(D_y^\ell K(t_0,\cdot,m_0^2,y) - D_y^\ell K(t_0,\cdot,m_0^2,y')\big)\|_{(4+\alpha)}\\ &\le C d_1(m_0^1,m_0^2)|y-y'|^{2+\alpha-\beta}.\end{aligned}$$

And we have proven the Lipschitz continuity of K and its derivatives in $\mathcal{P}_1(\mathbb{R}^d)$.

For the temporal variable, we apply a similar trick. We fix initial times $0 \le t_0^1 < t_0^2 \le T$ let $(u^1,m^1),(u^2,m^2)$ be solutions to the Mean Field Game system (10), with initial conditions (t_0^1,m_0) and (t_0^2,m_0) respectively, and let $(v_1^{(\ell)},\mu_1^{(\ell)}),(v_2^{(\ell)},\mu_2^{(\ell)})$ be the respective solutions of the linear system (33), both with $\mu_0 = D^\ell\delta_y$. In order to establish continuity, we have to circumvent the fact that $v_2^{(\ell)}$ is undefined for $t\in[0,t_0^1)$. We do this by rewriting in the following way

$$\begin{aligned}D_y^\ell K(t_0^1,x,m_0,y) - D_y^\ell K(t_0^2,x,m_0,y) &= v_1^{(\ell)}(t_0^1,x,y) - v_2^{(\ell)}(t_0^2,x,y)\\ &= v_1^{(\ell)}(t_0^1,x,y) - v_1^{(\ell)}(t_0^2,x,y) + v_1^{(\ell)}(t_0^2,x,y) - v_2^{(\ell)}(t_0^2,x,y).\end{aligned}$$

The continuity of the first term, $v_1^{(\ell)}(t_0^1,x,y) - v_1^{(\ell)}(t_0^2,x,y)$ is evident from the fact that $v_1^{(\ell)} \in C([t_0^1,T],C^{4+\alpha-\beta}(\mathbb{R}^d))$. For the other term we perform the same technique as in the continuity in measure: We let $(z,\rho) \coloneqq (v_1^{(\ell)} - v_2^{(\ell)}, \mu_1^{(\ell)} - \mu_2^{(\ell)})$, which yields the similar linear system

$$\begin{cases}-\partial_t z - \varepsilon\Delta z + Du^1\cdot Dz = \langle \frac{\delta F}{\delta m}(x,m^1(t)),\rho(t)\rangle + b(t,x) & \text{in } [t_0^2,T]\times\mathbb{R}^d,\\ z(T,x) = \langle \frac{\delta G}{\delta m}(x,m^1(T)),\rho(T)\rangle + z_T & \text{in } \mathbb{R}^d,\\ \partial_t\rho - \varepsilon\Delta\rho - \operatorname{div}(\rho Du^1) - \operatorname{div}(m^1 Dz + c) = 0 & \text{in } [t_0^2,T]\times\mathbb{R}^d,\\ \rho(t_0) = \mu_1^{(\ell)}(t_0^2) - D^\ell\delta_y & \text{in } \mathbb{R}^d,\end{cases}$$

with coefficients b, z_T, c same as in (35). We estimate using Proposition 4.10 and assumptions (F), (G)

$$\begin{aligned}\sup_{t\in[t_0^2,T]} \|b(t,\cdot)\|_{3+\alpha} &\le C\Big(\sup_{t\in[t_0^2,T]}\|u^1-u^2\|_{4+\alpha} + \sup_{t\in[t_0^2,T]} d_1(m^1(t),m^2(t))\Big),\\ \sup_{t\in[t_0^2,T]} \|c(t)\|_{-(2+\alpha-\beta)} &\le C\Big(\sup_{t\in[t_0^2,T]}\|u^1-u^2\|_{3+\alpha} + \sup_{t\in[t_0^2,T]} d_1(m^1(t),m^2(t))\Big),\\ \|z_T\|_{4+\alpha} &\le C d_1(m^1(T),m^2(T)),\\ \|\rho_0\|_{-(2+\alpha-\beta)} &= \|\mu_1^{(\ell)}(t_0^2) - \mu_1^{(\ell)}(t_0^1)\|_{-(2+\alpha-\beta)}.\end{aligned}$$

next, let $(\overline{u},\overline{m})$ solve the Mean Field Game System (10) with initial condition $(t_0^2, m^1(t_0^2))$. Thus, by the uniqueness of the MFG system $(u^1(t,x),m^1(t,x)) = (\overline{u}(t,x),\overline{m}(t,x))$ for all $(t,x)\in[t_0^2,T]\times\mathbb{R}^d$. Since u_2 and $\overline{u}$ have the same initial time, we can use Lemma 4.7 to estimate their difference

$$\begin{aligned}\sup_{t\in[t_0^2,T]}\|u^1-u^2\|_{4+\alpha} &\le \sup_{t\in[t_0^2,T]}\Big(\|u^1(t,\cdot)-\overline{u}(t,\cdot)\|_{4+\alpha} + \|\overline{u}(t,\cdot)-u^2(t,\cdot)\|_{4+\alpha}\Big)\\ &\le \sup_{t\in[t_0^2,T]}\Big(0 + Cd_1(m^1(t_0^2),m_0)\Big),\end{aligned}$$

and similarly

$$\sup_{t\in[t_0^2,T]} d_1(m^1(t),m^2(t)) \le \sup_{t\in[t_0^2,T]} \Big(d_1(m^1(t),\overline{m}(t)) + d_1(\overline{m}(t),m^2(t))\Big)$$
$$\le \sup_{t\in[t_0^2,T]} \Big(0 + Cd_1(m^1(t_0^2),m_0)\Big).$$

We apply Lemma 4.8 with $k=2$ and $\sigma := \alpha$, and get

$$\sup_{t\in[t_0^2,T]} \|v_1^{(\ell)}(t,\cdot) - v_2^{(\ell)}(t,\cdot)\|_{4+\alpha} \le C\Big(d_1(m^1(t_0^2),m_0) + \|\mu_1^{(\ell)}(t_0^2) - \mu_1^{(\ell)}(t_0^1)\|_{-(2+\alpha-\beta)}\Big).$$

Since $m^1 \in C([t_0^1,T],\mathcal{P}_1(\mathbb{R}^d))$ and $\mu_1^{(\ell)} \in C([t_0^1,T],C^{-(2+\alpha)}(\mathbb{R}^d))$, we can conclude by passing $t_0^2 \to t_0^1$ that $v_1^{(\ell)}(t_0^2,x,y) - v_2^{(\ell)}(t_0^2,x,y)$ and its derivatives tend to zero, and consequently that $D_y^\ell K(t_0,x,m_0,y)$ is continuous on $[0,T]$.

Finally, we have to verify that if $\mu_0 \in \mathcal{M}(\mathbb{R}^d)$ then $v(t_0,x) = \langle \mu_0, K(t_0,x,m_0,\cdot)\rangle$ solves (33). Recall that we denote by $(v^{(0)}(t,x,y),\mu^{(0)}(t,y))$ the solution to (33) with initial condition δ_y and that $K(t_0,x,m_0,\cdot) := v^{(0)}(t_0,x,y)$. We define for some $\phi \in C^{2+\alpha-\beta}(\mathbb{R}^d)$

$$v(t,x) := \langle \mu_0, v^{(0)}(t,x)\rangle = \int_{\mathbb{R}^d} v^{(0)}(t,x,y)\mu_0(dy),$$
$$\langle \mu(t),\phi\rangle := \langle \mu_0, \langle \mu^{(0)}(t,y),\phi\rangle\rangle = \int_{\mathbb{R}^d} \langle \mu^{(0)}(t,y),\phi\rangle \mu_0(dy).$$

We will show that (v,μ) satisfies the first equation of (33) pointwise, and the second in the distributional sense as in (32).

$$\begin{aligned}
-\partial_t v(t,x) &= -\lim_{h\to 0} \frac{v(t+h,x)-v(t,x)}{h}\\
&= -\lim_{h\to 0}\int_{\mathbb{R}^d} \frac{v^{(0)}(t+h,x,y)-v^{(0)}(t,x,y)}{h}\mu_0(dy)\\
&= -\int_{\mathbb{R}^d} \lim_{h\to 0} \frac{v^{(0)}(t+h,x,y)-v^{(0)}(t,x,y)}{h}\mu_0(dy)\\
&= \int_{\mathbb{R}^d} -\partial_t v^{(0)}(t,x,y)\mu_0(dy)\\
&= \int_{\mathbb{R}^d} \varepsilon\Delta v^{(0)}(t,x,y) - Du\cdot Dv^{(0)}(t,x,y) + \langle \frac{\delta F}{\delta m}(x,m(t)),\mu^{(0)}(t)\rangle \mu_0(dy)\\
&= \varepsilon\Delta v(t,x) - Du\cdot Dv(t,x) + \int_{\mathbb{R}^d} \langle \frac{\delta F}{\delta m}(x,m(t)),\mu^{(0)}(t)\rangle \mu_0(dy)\\
&= \varepsilon\Delta v(t,x) - Du\cdot Dv(t,x) + \langle \mu(t), \frac{\delta F}{\delta m}(x,m(t))\rangle.
\end{aligned}$$

For the third equality above we applied the Dominated Convergence Theorem, which is applicable here since $\sup_{t\in[t_0,T]}\|v^{(0)}(t,\cdot)\|_{4+\alpha} < \infty$ and μ_0 is a bounded measure on $\mathbb{R}^d$. The application of the convergence theorem is repeated for the derivative terms $\varepsilon\Delta v$ and $Du\cdot Dv$ in the sixth equality. In the last equality we apply the definition of $\langle\mu(t),\phi\rangle$ noting that by assumption (F) $\frac{\delta F}{\delta m}(x,m) \in C^{2+\alpha-\beta}(\mathbb{R}^d)$ for any $\mathcal{P}(\mathbb{R}^d)$.

Likewise, we check the terminal condition

$$v(T,x) = \int_{\mathbb{R}^d} v^{(0)}(T,x,y)\mu_0(dy) = \int_{\mathbb{R}^d} \langle \frac{\delta G}{\delta m}(x,m(T)),\mu^{(0)}(T)\rangle\mu_0(dy) = \langle \frac{\delta G}{\delta m}(x,m(T)),\mu(T)\rangle,$$

and conclude that v satisfies the first equation of (33) pointwise. For the distributional equation we choose a function ϕ such that $\phi(t_0),\phi(t) \in C^{k+1+\sigma}(\mathbb{R}^d)$ and $\partial_t\phi + \varepsilon\Delta\phi + V\cdot D\phi \in C([t_0,t],C^{k+1+\sigma}(\mathbb{R}^d))$, and compute

$$
\begin{aligned}
\langle \mu(t), \phi(t)\rangle - \langle \mu_0, \phi(t_0)\rangle &= \int_{\mathbb{R}^d} \langle \mu^{(0)}(t,y), \phi(t)\rangle \mu_0(dy) - \langle \mu_0, \phi(t_0)\rangle \\
&= \int_{\mathbb{R}^d} \Big(\langle \delta_y, \phi(t_0)\rangle + \int_{t_0}^t \langle \mu^{(0)}(s), (\partial_t \phi + \varepsilon \Delta \phi + Du \cdot D\phi)(s)\rangle ds \\
&- \int_{t_0}^t \int_{\mathbb{R}^d} D\phi \cdot Dv^{(0)}(s,x,y) m(s,dx) ds \Big) \mu_0(dy) - \langle \mu_0, \phi(t_0)\rangle \\
&= \langle \mu_0, \phi(t_0)\rangle + \int_{t_0}^t \int_{\mathbb{R}^d} \langle \mu^{(0)}(s), (\partial_t \phi + \varepsilon \Delta \phi + Du \cdot D\phi)(s)\rangle \mu_0(dy) ds \\
&- \int_{t_0}^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} D\phi \cdot Dv^{(0)}(s,x,y) \mu_0(dy) m(s,dx) ds - \langle \mu_0, \phi(t_0)\rangle \\
&= \int_{t_0}^t \langle \mu(s), (\partial_t \phi + \varepsilon \Delta \phi + Du \cdot D\phi)(s)\rangle ds \\
&- \int_{t_0}^t \int_{\mathbb{R}^d} D\phi \cdot Dv(s,x) m(s,dx) ds.
\end{aligned}
$$

For the second equality, we used Fubini's Theorem (Theorem 4.5 in [4]) to swap the order of integration since all terms are bounded and the measure μ_0 is finite. And for the final equality, we used the same Dominated Convergence Theorem trick as in the pointwise equation to handle the term Dv. Finally, we check the initial condition

$$
\langle \mu(t_0), \phi\rangle = \int_{\mathbb{R}^d} \langle \delta_y, \phi\rangle \mu_0(dy) = \int_{\mathbb{R}^d} \phi(y) \mu_0(dy) = \langle \mu_0, \phi\rangle,
$$

which shows that μ is a distibution solution to the second equation of (33). □

Remark 4.12. Recall that it was mentioned in the subsection on assumptions that the choice of (F) and (G) was sufficient, but not optimal. The proof of Proposition 4.11 is the primary reason we overshoot assumptions. This is due to the proof requiring Lemma 4.8 to be applied several times over the course of establishing the properties of $K(t_0, x, m_0, y)$, each relying on the assumptions (F) and (G) in different ways for different estimates. As this proof is the major bottleneck of the chapter when it comes to regularity, immediate improvement with respect to assuming less of (F) and (G) can be easily achieved by observing precisely what is required of the coupling terms in order to complete this proof.

Having established $K(t_0, x, m_0, y)$ as the clear candidate for our measure derivative, we prove one last technical proposition which will enable us to apply Definition 2.6. The proof of this result also has Lemma 4.8 at its core.

Proposition 4.13. *Let assumption* (F) *and* (G) *hold. Assume $t_0 \in [0,T]$ and $m_0, \hat{m}_0 \in \mathcal{P}_1(\mathbb{R}^d)$ to be fixed. Let (u,m) and $(\hat{u}, \hat{m})$ be solutions of* (10) *with respect to initial conditions (t_0, m_0) and $(t_0, \hat{m}_0)$, and let (v, μ) solve* (33) *with initial value $(t_0, \hat{m}_0 - m_0)$. Then the following estimate holds*

$$
\sup_{t \in [t_0, T]} \left(\|\hat{u}(t,\cdot) - u(t,\cdot) - v(t,\cdot)\|_{4+\alpha} + \|\hat{m}(t,\cdot) - m(t,\cdot) - \mu(t,\cdot)\|_{-(3+\alpha)} \right) \leq C d_1^2(m_0, \hat{m}_0) \quad (36)
$$

Proof. We yet again want to apply Lemma 4.8 to obtain an estimate for a linearised system. Let $z := \hat{u} - u - v$ and $\rho := \hat{m} - m - \mu$. After summing the systems and applying the measure version of the fundamental theory of analysis from Lemma 2.7 we are left with the following linearised system.

$$\begin{cases} -\partial_t z - \varepsilon\Delta z + Du\cdot Dz = \langle \frac{\delta F}{\delta m}(x, m(t)), \rho(t)\rangle + b(t,x) & \text{in } [t_0,T]\times\mathbb{R}^d,\\ z(T,x) = \langle \frac{\delta G}{\delta m}(x, m(t)), \rho(t)\rangle + z_T & \text{in } \mathbb{R}^d,\\ \partial_t\rho - \varepsilon\Delta\rho - \operatorname{div}(\rho Du) - \operatorname{div}(mDz + c) = 0 & \text{in } [t_0,T]\times\mathbb{R}^d,\\ \rho(t_0) = 0 & \text{in } \mathbb{R}^d. \end{cases}$$

where

$$\begin{aligned} b(t,x) &= \int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta F}{\delta m}(x,(1-s)m+s\hat m,y) - \frac{\delta F}{\delta m}(x,m,y)\right)(\hat m - m)(dy)ds - \frac{1}{2}|D\hat u - Du|^2,\\ z_T(x) &= \int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta G}{\delta m}(x,(1-s)m(T)+s\hat m(T),y) - \frac{\delta G}{\delta m}(x,m(T),y)\right)(\hat m(T) - m(T))(dy)ds,\\ c(t,x) &= (\hat m - m)(D\hat u - Du). \end{aligned}$$

From Lemma 4.7, we get that

$$\sup_{t\in[t_0,T]}\|\frac{1}{2}|D\hat u - Du|^2\|_{3+\alpha} \le C\sup_{t\in[t_0,T]}\|\hat u(t,\cdot) - u(t,\cdot)\|_{4+\alpha}^2 \le Cd_1^2(\hat m_0, m_0).$$

Furthermore, for any $|\ell|\le 3$

$$\begin{aligned} &\Big|D_x^\ell\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta F}{\delta m}(x,(1-s)m+s\hat m,y) - \frac{\delta F}{\delta m}(x,m,y)\right)(\hat m - m)(dy)ds\Big|\\ &= \Big|\int_0^1\int_{\mathbb{R}^d}D_x^\ell\Big(\frac{\delta F}{\delta m}(x,(1-s)m+s\hat m,y) - \frac{\delta F}{\delta m}(x,m,y)\Big)(\hat m - m)(dy)ds\Big|\\ &\le \int_0^1\|\frac{\delta F}{\delta m}(\cdot,(1-s)m+s\hat m,\cdot) - \frac{\delta F}{\delta m}(\cdot,m,\cdot)\|_{(3+\alpha,1)}d_1(\hat m_0,m_0)ds\\ &\le C\int_0^1 d_1((1-s)m+s\hat m,m)d_1(\hat m_0,m_0)ds = C\int_0^1 sd_1(\hat m(t),m(t))d_1(\hat m_0,m_0)ds\\ &\le Cd_1(\hat m(t),m(t))d_1(\hat m_0,m_0). \end{aligned}$$

Where we have applied the usual Lipschitz in y trick for the first inequality, the assumption (F) for the second, and the definition of d_1 for the last equality. By applying the same steps for the Hölder quotient of the derivatives with $|\ell| = 3$, we can conclude that

$$\left\|\int_0^1\int_{\mathbb{R}^d}\left(\frac{\delta F}{\delta m}(\cdot,(1-s)m+s\hat m,y) - \frac{\delta F}{\delta m}(\cdot,m,y)\right)(\hat m - m)(dy)ds\right\|_{3+\alpha} \le Cd_1(\hat m(t),m(t))d_1(\hat m_0,m_0).$$

Through another application of Lemma 4.7 we can conclude that

$$\sup_{t\in[t_0,T]}\|b(t,\cdot)\|_{3+\alpha} \le Cd_1^2(\hat m_0,m_0).$$

For the estimate on c we observe that it can be interpreted as a signed measure, and consequently its functional action takes the form of integration with respect to the measure

$$
\begin{aligned}
\sup_{t\in[t_0,T]} \|c(t)\|_{-(2+\alpha-\beta)} &= \sup_{t\in[t_0,T]} \sup_{\|\phi\|_{2+\alpha-\beta}\le 1} |\langle(\hat{m}-m)(D\hat{u}-Du),\phi\rangle| \\
&= \sup_{t\in[t_0,T]} \sup_{\|\phi\|_{2+\alpha-\beta}\le 1} \Big| \int_{\mathbb{R}^d} (D\hat{u}-Du)\phi(x)(\hat{m}(t)-m(t))(dx)\Big| \\
&\le \sup_{t\in[t_0,T]} \sup_{\|\phi\|_{2+\alpha-\beta}\le 1} \|D((D\hat{u}-Du)\phi)\|_\infty d_1(\hat{m}(t),m(t)) \\
&\le \sup_{t\in[t_0,T]} \sup_{\|\phi\|_{2+\alpha-\beta}\le 1} \|\hat{u}(t,\cdot)-u(t,\cdot)\|_{3+\alpha}\|\phi\|_{2+\alpha-\beta} d_1(\hat{m}(t),m(t)) \\
&\le C d_1^2(\hat{m}_0,m_0).
\end{aligned}
$$

In estimating z_T using the same method as with b, and applying Lemma 4.8, we are left with

$$
\begin{aligned}
\sup_{t\in[t_0,T]} \Big(\|z(t,\cdot)\|_{4+\alpha} + \|\rho\|_{-(3+\alpha)}\Big) &\le \sup_{t\in[t_0,T]} \Big(\|b(t,\cdot)\|_{3+\alpha} + \|c(t)\|_{-(2+\alpha-\beta)}\Big) + \|z_T\|_{4+\alpha} \\
&\le C d_1^2(\hat{m}_0,m_0).
\end{aligned}
$$

which is precisely the estimate we sought to prove. □

This result immediately implies that the Master Characteristic $U(t_0,x,m_0) \coloneqq u(t_0,x)$ from Definition 4.1 is $\mathcal{C}^1$ in the measure variable in the sense of Definition 2.6.

Corollary 4.14. *With the same assumptions as Proposition 4.13, and with $K(t_0,x,m_0,y)$ as given in Proposition 4.11, then U is $\mathcal{C}^1$ in measure and*

$$
\frac{\delta U}{\delta m}(t_0,x,m_0,y) = K(t_0,x,m_0,y).
$$

Hence,

$$
\|\frac{\delta U}{\delta m}(t_0,\cdot,m_0,\cdot)\|_{(4+\alpha,3+\alpha)} \le C,
$$

with constant $C>0$ independent of (t_0,m_0), and $\frac{\delta U}{\delta m}$ has derivatives in (x,y) of order $4+\alpha-\beta$ and $2+\alpha-\beta$ which are continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R})\times\mathbb{R}^d$.

Furthermore,

$$
\|U(t_0,\cdot,\hat{m}_0) - U(t_0,\cdot,m_0) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0,\cdot,m_0,y)(\hat{m}_0-m_0)(dy)\|_{4+\alpha} \le C d_1^2(m_0,\hat{m}_0).
$$

Proof. We will demonstrate that $K(t_0,x,m_0,y)$ satisfies the definition of a derivative in the space of measures. As Proposition 4.11 states that $v(t_0,x) = \langle \hat{m}_0, K(t_0,x,m_0,\cdot)\rangle$ we have through (36) that

$$
|U(t_0,\cdot,\hat{m}_0) - U(t_0,\cdot,m_0) - \int_{\mathbb{R}^d} K(t_0,x,m_0,y)(\hat{m}_0-m_0)(dy)| \le C d_1^2(m_0,\hat{m}_0)
$$

Let m,m' be arbitrary measures in $\mathcal{P}_1(\mathbb{R}^d)$. We choose $\hat{m}_0 = (1-h)m + hm'$ and $m_0 = m$ for some $h\in[0,1]$, then dividing both sides of the inequality with h and rearranging yields

$$
\begin{aligned}
\left| \frac{U(t_0,\cdot,(1-h)m+hm') - U(t_0,\cdot,m)}{h} - \int_{\mathbb{R}^d} K(t_0,x,m,y)(m'-m)(dy)\right| &\le C\frac{1}{h} d_1^2(m_0,\hat{m}_0) \\
&\le \frac{h^2}{h} d_1(m,m'),
\end{aligned}
$$

which tends to zero for any choice of $m, m \in \mathcal{P}_1(\mathbb{R}^d)$as $h \to 0$.

Next we check the normalisation, that is

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0, x, m_0, y) m_0(dy) = 0.$$

To confirm this we observe that this integral corresponds to $v(t_0, x) = \langle m_0, K(t_0, x, m_0, \cdot)\rangle$, which implies that $v(t, x)$ is a solution to the linear system (33) with $\mu_0 = m_0$. Since F, G are by assumption $\mathcal{C}^1$ in measure, normalisation of the measure derivative yields that

$$\langle \frac{\delta F}{\delta m}(t, m(t)), m(t)\rangle = \langle \frac{\delta G}{\delta m}(t, m(T)), m(T)\rangle = 0.$$

This in turn means that $(0, m)$ is a solution, which by Proposition 4.10 is a unique solution, to (33), and we can conclude

$$\int_{\mathbb{R}^d} K(t_0, x, m_0, y) m_0(dy) = v(t_0, x) = 0.$$

U is then $\mathcal{C}^1$ in measure, and the inequality in the corollary statement follows immediately from (36). ☐

4.4 Existence and Uniqueness of The Master Equation

We are finally ready to perform the proofs for Theorem 4.4, establishing existence and uniqueness for the Master Equation. For the existence proof there are some different approaches one can take. In this text we will do an approximation argument, where we first assume a smooth initial measure $m_0 \in C_b^\infty(\mathbb{R}^d)$, show existence, and then pass to the limit using mollifiers as was done in the proof for Theorem 3.16. This approach is performed in the existence proofs in [7] and [6]. Another approach uses the distributional formulation of the Fokker-Planck directly, cleverly using $\frac{\delta U}{\delta m}$ as a test function, which circumvents the need to use integration by parts. This second idea will feature in the uniqueness proof.

Proof of Theorem 4.4 (existence). We proceed by an approximation argument, first showing the result for $m_0 \in C_b^\infty(\mathbb{R}^d)$ before passing to any $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ using Lemma 4.7

Step 1: Solution for a smooth measure. First let $m_0 \in C_b^\infty(\mathbb{R}^d) \cap \mathcal{P}_1(\mathbb{R}^d)$ let (u, m) be the corresponding solution of the Mean Field Game system (10). By Theorem 3.16 $(u, m) \in C^{1+\frac{\alpha}{2}, 2+\alpha}([t_0, T] \times \mathbb{R}^d) \times C^{1+\alpha/2, 2+\alpha}([t_0, T] \times \mathbb{R}^d) \cap C([0, T], \mathcal{P}_1(\mathbb{R}^d))$, which implies that $m \in L^1(\mathbb{R}^d)$. We will show that the Master Characteristic U solves the Master Equation 24 for the case of a smooth initial measure.

Firstly, consider $\partial_t U$ by applying the definition of the derivative. Fix an arbitrary $t_0 \in [0, T]$ and let $h > 0$. Instead of tackling the difference quotient directly, we split it up into two more manageable pieces

$$\begin{aligned}\frac{U(t_0+h, x, m_0) - U(t_0, x, m_0)}{h} &= \frac{U(t_0+h, x, m_0) - U(t_0+h, x, m(t_0+h))}{h} \\ &+ \frac{U(t_0+h, x, m(t_0+h)) - U(t_0, x, m_0)}{h}.\end{aligned} \tag{37}$$

For the first quotient we use that U is $\mathcal{C}^1$ in measure from Corollary 4.14, and apply the measure version of the fundamental theorem of calculus from Lemma 2.7, denoting $m_s = (1-s)m_0 + sm(t_0+h)$

$$U(t_0+h,x,m(t_0+h)) - U(t_0+h,x,m_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(t_0+h,x,m_s,y)(m(t_0+h)-m_0)(dy)ds$$

$$= \int_0^1 \int_{\mathbb{R}^d} \int_{t_0}^{t_0+h} \frac{\delta U}{\delta m}(t_0+h,x,m_s,y)\partial_t m(t,y)dtdyds.$$

Since $m \in C^{1+\frac{\alpha}{2},2+\alpha}([t_0,T]\times\mathbb{R}^d)$ we can use the strong solution of the PDE (10) to swap out $\partial_t m$, and since $m \in L^1(\mathbb{R}^d)$ and $\|\frac{\delta U}{\delta m}(t_0+h,x,m_s,y)\| \le C$ from Corollary 4.14 we can integrate by parts yielding

$$\int_0^1 \int_{\mathbb{R}^d} \int_{t_0}^{t_0+h} \frac{\delta U}{\delta m}(t_0+h,x,m_s,y)\Big(\varepsilon\Delta_y m(t,y) + \operatorname{div}\big(m(t,y)Du(t,y)\big)\Big)dtdyds$$
$$= \int_0^1 \int_{\mathbb{R}^d} \int_{t_0}^{t_0+h} \Big(\varepsilon\Delta_y \frac{\delta U}{\delta m}(t_0+h,x,m_s,y) - D_y\frac{\delta U}{\delta m}(t_0+h,x,m_s,y)\cdot Du(t,y)\Big)m(t,y)dtdyds.$$

By appealing to the continuity of $\Delta_y\frac{\delta U}{\delta m}$ and $D_y\frac{\delta U}{\delta m}$ in all the variables, and by using the dominated convergence theorem, we can divide by h and pass to the limit, getting

$$\lim_{h\to 0}\frac{U(t_0+h,x,m(t_0+h)) - U(t_0+h,x,m_0)}{h}$$
$$= \int_0^1 \int_{\mathbb{R}^d} \Big(\varepsilon\Delta_y \frac{\delta U}{\delta m}(t_0,x,m_0,y) - D_y\frac{\delta U}{\delta m}(t_0,x,m_0,y)\cdot Du(t_0,y)\Big)m(t_0,y)dyds$$
$$= \int_{\mathbb{R}^d} \Big(\varepsilon\operatorname{div}_y[D_mU](t_0,x,m_0,y) - D_mU(t_0,x,m_0,y)\cdot Du(t_0,y)\Big)m(t_0,y)dy,$$

where we have inserted the definition of the Lions derivative $D_m \coloneqq D_y\frac{\delta U}{\delta m}$.

For the second difference quotient, we use the definition of the Master Characteristic U. $U(t_0+h,x,m(t_0+h))$ is defined as the initial value of the solution of the Hamilton-Jacobi-Bellman equation of (10) started at time t_0+h with the initial measure $m(t_0+h)$. This is equivalent to the function u, which is the solution of (10) started in (t_0,m_0), evaluated at time t_0+h, since they are governed uniquely by the same Mean Field Game system with solutions agreeing at time t_0+h. From the regularity of u we then have

$$\lim_{h\to 0}\frac{U(t_0+h,x,m(t_0+h)) - U(t_0,x,m_0)}{h} = \lim_{h\to 0}\frac{u(t_0+h,x)-u(t_0,x)}{h} = \partial_t u(t_0,x)$$
$$= -\varepsilon\Delta u(t_0,x) + \frac{1}{2}|Du(t_0,x)|^2 - F(x,m(t_0))$$
$$= -\varepsilon\Delta_x U(t_0,x,m_0) + \frac{1}{2}|D_xU(t_0,x,m_0)|^2 - F(x,m(t_0)).$$

Combining the two difference quotients from (37), we get

$$\begin{aligned}\partial_t U(t_0,x,m_0) &= \lim_{h\to 0}\frac{U(t_0+h,x,m(t_0+h)) - U(t_0,x,m_0)}{h} \\ &= -\varepsilon\Delta_x U(t_0,x,m_0) + \frac{1}{2}|D_xU(t_0,x,m_0)|^2 - F(x,m_0) \\ &- \int_{\mathbb{R}^d}\Big(\varepsilon\operatorname{div}_y[D_mU](t_0,x,m_0,y) - D_mU(t_0,x,m_0,y)\cdot D_xU(t_0,y,m_0)\Big)m_0(y)dy,\end{aligned} \tag{38}$$

which is precisely the Master Equation (24). Furthermore, from the convergence of the right hand side above, we get that U is C^1 in time for any $(t_0,x,m_0) \in [0,T)\times\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R}^d)\cap C_b^\infty(\mathbb{R}^d)$. To check the terminal value for the equation we observe that starting the MFG system in (T,m_0)

yields a somewhat degenerate solution for the simultaneous initial and terminal point (T,x) where $(u(T,\cdot),m(T,\cdot)) = (G(\cdot,m_0),m_0)$, and the terminal value of the master equation is thus

$$U(T,x,m_0) = u(T,x) = G(x,m_0),$$

which is exactly what was desired.

Step 2: Approximation of a general measure. For this approximation part we follow a similar approach as was done in the proof of Theorem 3.16 and regularise an arbitrary $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ using a δ-indexed family of standard mollifiers, as introduced in Definition 2.4, and pass to the limit in the master equation in a controlled manner. For $\delta > 0$ let

$$m_0^\delta(x) := m_0 * \eta_\delta = \int_{\mathbb{R}^d} \eta_\delta(x-y)m_0(dy).$$

As was shown in the proof of Theorem 3.16, we have that $m_0^\delta \in \mathcal{P}_1(\mathbb{R}^d) \cap C_b^\infty(\mathbb{R}^d)$, and that for any decreasing sequence $\{\delta_n\}$ tending to zero

$$\lim_{n\to\infty} d_1(m_0^{\delta_n}, m_0) = 0. \tag{39}$$

Now, let (u^δ, m^δ) be the solutions of the Mean Field game system (10) starting in (t_0, m_0^δ), and let (u,m) be the solution starting in (t_0,m_0). Again, by applying the definition of the Master Characteristic, we have $U(t,x,m_0^{\delta_n}) = u^{\delta_n}(t_0,x)$ and $U(t,\cdot,m_0) = u(t_0,x)$, and we want to show by passing $\{\delta_n\}$ to zero that the latter solves the Master Equation.

Next, we argue that the right hand side of (38) is continuous, and hence that we are able to pass to the limit. For the terms containing U we apply Lemma 4.7 and the limit (39) and get

$$\|U(t,\cdot,m_0^{\delta_n}) - U(t,\cdot,m_0)\|_{4+\alpha} \leq C d_1(m_0^{\delta_n}, m_0) \xrightarrow{n\to\infty} 0. \tag{40}$$

Likewise, for the terms involving D_mU we appeal to the continuity of the derivatives from Corollary 4.14

$$\lim_{n\to\infty} \|\frac{\delta U}{\delta m}(t_0,\cdot,m_0,\cdot) - \frac{\delta U}{\delta m}(t_0,\cdot,m_0^{\delta_n},\cdot)\|_{(4+\alpha-\beta,2+\alpha-\beta)} = 0.$$

Lastly, from the assumptions (F),(G) on F and G

$$\|F(\cdot,m_0^{\delta_n}) - F(\cdot,m_0)\|_\infty + \|G(\cdot,m_0^{\delta_n}) - G(\cdot,m_0)\|_\infty \leq C d_1(m_0^{\delta_n}, m_0) \xrightarrow{n\to\infty} 0, \tag{41}$$

and we can finally begin to estimate the terms of (38). The convergence of the non-integral part $-\varepsilon\Delta_x U(t_0,x,m_0) + \frac{1}{2}|D_xU(t_0,x,m_0)|^2 - F(x,m_0)$ is immediate from the limits (40) and (41). For the two terms containing integrals we can estimate

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d}\Big(\varepsilon \mathrm{div}_y[D_mU](t_0,x,m_0^{\delta_n},y)\Big)m_0^{\delta_n}(dy) - \int_{\mathbb{R}^d}\Big(\varepsilon \mathrm{div}_y[D_mU](t_0,x,m_0,y)\Big)m_0(dy)\Big| \\
&\leq \varepsilon\Big|\int_{\mathbb{R}^d}\Big(\mathrm{div}_y[D_mU](t_0,x,m_0^{\delta_n},y)\Big)(m_0^{\delta_n}-m_0)(dy)\Big| \\
&+ \varepsilon\Big|\int_{\mathbb{R}^d}\Big(\mathrm{div}_y[D_mU](t_0,x,m_0^{\delta_n},y) - \mathrm{div}_y[D_mU](t_0,x,m_0,y)\Big)m_0(dy)\Big| \\
&\leq \varepsilon\|\frac{\delta U}{\delta m}(t_0,\cdot,m_0^{\delta_n},\cdot)\|_{(4+\alpha,3+\alpha)} d_1(m_0^{\delta_n},m_0) + \varepsilon\|\frac{\delta U}{\delta m}(t_0,\cdot,m_0,\cdot) - \frac{\delta U}{\delta m}(t_0,\cdot,m_0^{\delta_n},\cdot)\|_{(4+\alpha-\beta,2+\alpha-\beta)},
\end{aligned}$$

where we have used Lipschitz in y trick for the first term, and the regularity and boundedness of $\frac{\delta U}{\delta m}$ and that m_0 is a finite measure, for the second. By (39) and (40), the integral estimate above converges to zero, implying the convergence of the integrals.

The estimate on, and convergence of, the integral containing $D_mU(t_0,x,m_0,y) \cdot D_xU(t_0,y,m_0)$ follows using the same techniques and bounds in the same way, and we can thus conclude that the right hand side of (38) converges to the right hand side of the Master Equation with respect to a general $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. Furthermore, with the right hand side being continuous, we can conclude that $\partial_t U(t_0,x,m_0)$ exists, and is continuous, for any $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$. □

Remark 4.15. Note that the previous proof only computed the right derivative. The proof for the left derivative uses similar techniques, but is quite a bit more technical and is often ignored in the literature, like in [7] and [6]. The system is not as well-behaved backwards in time. The proof for the left derivative gets full treatment in an upcoming publication by Jakobsen and Rutkowski.

As with many uniqueness proofs the core consists of assuming existence of another solution to the Master Equation, and demonstrating that it has to coincide with the solution constructed from the Master Characteristic U. We apply some ideas from the uniqueness proof of Theorem 2.4.2 in [7], but instead of using smooth approximation and integration by parts for a classical solution of the Fokker-Planck, we use the distributional formulation directly.

Proof of Theorem 4.4 (uniqueness). Let U be the classical solution to the Master Equation derived using the Master Characteristic, and let V be any other classical solution. By Definition 4.2 of a classical solution, $D_x D_y V$ is globally bounded in $[0,T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \times \mathbb{R}^d$, and hence by our standard Lipschitz in y trick we have that D_x is globally Lipschitz continuous in the m-variable.

Let $m_0 \in \mathcal{P}_1(\mathbb{R}^d)$ and fix a $t_0 \in [0,T]$. We introduce the following Fokker-Planck equation

$$\begin{cases} \partial_t \tilde{m} - \varepsilon \Delta \tilde{m} - \operatorname{div}\big(\tilde{m} D_x V(t,x,\tilde{m})\big) = 0 & \text{in } [t_0,T] \times \mathbb{R}^d, \\ \tilde{m}(t_0) = m_0 & \text{in } \mathbb{R}^d. \end{cases}$$

By Proposition 3.5 (which utilises results from [3]), since $D_x V, D_x^2 V, D_x^3 V \in C_b\big([0,T] \times \mathbb{R}^d\big)$ this Fokker-Planck equation has a unique distributional solution. Thus for any $\varphi \in C_c^\infty([t_0,T] \times \mathbb{R}^d)$ we have

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(t,x) \tilde{m}(t,dx) = \int_{\mathbb{R}^d} \varphi(t_0,x) m_0(dx) \\ + \int_{t_0}^{t} \int_{\mathbb{R}^d} [\varphi_t(s,x) - D\varphi(s,x) \cdot D_x V(s,x,\tilde{m}) + \varepsilon \Delta \varphi(s,x)] \tilde{m}(s,dx) ds. \end{aligned} \tag{42}$$

Actually the above formulation holds for any function $\varphi \in C_b^{1,2}([t_0,T] \times \mathbb{R}^d)$, as can be shown by regularising and truncating φ by cutoff functions, as described in [5].

As an ansatz, we set $\tilde{u}(t,x) = V(t,x,\tilde{m}(t))$ and use the regularity of V and the measure version of the fundamental theorem of calculus from Lemma 2.7 to get the time derivative at the initial time t_0

$$\begin{aligned} \partial_t \tilde{u}(t_0,x) &= \lim_{h \to 0} \frac{V(t_0+h,x,\tilde{m}(t_0+h)) - V(t_0,x,\tilde{m}(t_0))}{h} \\ &= \lim_{h \to 0} \Big(\frac{V(t_0+h,x,\tilde{m}(t_0+h)) - V(t_0,x,\tilde{m}(t_0+h))}{h} + \frac{V(t_0,x,\tilde{m}(t_0+h)) - V(t_0,x,\tilde{m}(t_0))}{h} \Big) \\ &= \partial_t V(t_0,x,m_0) + \lim_{h \to 0} \frac{1}{h} \int_0^1 \int_{\mathbb{R}^d} \frac{\delta V}{\delta m}(t_0,x,\tilde{m}_s,y)(\tilde{m}(t_0+h,dy) - \tilde{m}(t_0,dy)) ds, \end{aligned} \tag{43}$$

where $\tilde{m}_s = s\tilde{m}(t_0+h) + (1-s)\tilde{m}(t_0)$, which tends to $\tilde{m}(t_0)$ as $h \to 0$. We can then utilise our distributional formulation, choosing $\varphi(t,y) = \frac{\delta V}{\delta m}(t_0,x,\tilde{m}_s(t_0),y)$. Note that $\varphi \in C_b^{1,2}([t_0,T] \times \mathbb{R}^d)$, $\varphi(t_0+h,y) = \varphi(t_0,y)$ and hence $\partial_t \varphi(t,y) = 0$. We can then use the distributional formulation (42) in (43) to get

$$\begin{aligned} \partial_t \tilde{u}(t_0,x) = \partial_t V(t_0,x,m_0) + \lim_{h \to 0} \frac{1}{h} \int_0^1 \int_{t_0}^{t_0+h} \int_{\mathbb{R}^d} \Big(\varepsilon \Delta_y \frac{\delta V}{\delta m}(t_0,x,\tilde{m}_s,y) \\ - D_x V(t,x,\tilde{m}(t)) \cdot D_y \frac{\delta V}{\delta m}(t_0,x,\tilde{m}_s,y) \Big) \tilde{m}(t,dy) dt ds \\ = \partial_t V(t_0,x,m_0) + \int_{\mathbb{R}^d} \Big(\varepsilon \operatorname{div}_y [D_m V](t_0,x,m_0,y) - D_x V \cdot D_m V(t_0,x,m_0,y) \Big) m_0(dy), \end{aligned}$$

where the second equality utilises Fubini's Theorem (Theorem 4.5 in [4]), and that by Definition 4.2 of the classical solution, $D_x \frac{\delta V}{\delta m}, D_y \frac{\delta V}{\delta m}$ and $D_y^2 \frac{\delta V}{\delta m}$ are continuous in all variables. We furthermore use that V by assumption solves the Master Equation, and after comparing terms get

$$\begin{aligned} \partial_t \tilde{u}(t_0, x) &= -\varepsilon \Delta V(t_0, x, m_0) + \frac{1}{2}|DV(t, x, m_0)|^2 - F(x, m_0) \\ &= -\varepsilon \Delta \tilde{u}(t_0, x) + \frac{1}{2}|D\tilde{u}(t_0, x)|^2 - F(x, m_0). \end{aligned} \tag{44}$$

We recognise this as the Hamilton-Jacobi-Bellman equation from the Mean Field Game system (10) started in point (t_0, m_0), evaluated at time t_0. By Theorem 3.16, (44) has a unique solution $\tilde{u}$, and by the definition of the Master Characteristic

$$U(t_0, x, m_0) = \tilde{u}(t_0, x) = V(t_0, x, m_0), \quad \forall x \in \mathbb{R}^d.$$

Since the initial time and inital measure $(t_0, m_0) \in [0, T] \times \mathcal{P}_1(\mathbb{R}^d)$ were chosen arbitrarily, we conclude that U and V agree everywhere on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \times \mathbb{R}^d$, and hence that the classical solution to the Master Equation is unique. □

For the later convergence of the master equation to the N-dimensional Nash system, we will need the Lipschitz continuity of $\frac{\delta U}{\delta m}$ with respect to m. This result follows immediately from properties previously established.

Proposition 4.16. *Assume that* (F) *and* (G) *holds. Then*

$$\sup_{t \in [0,T]} \sup_{m_1 \neq m_2 \in \mathcal{P}_1(\mathbb{R}^d)} \big(d_1(m_1, m_2)\big)^{-1} \Big\| \frac{\delta U}{\delta m}(t, \cdot, m_1, \cdot) - \frac{\delta U}{\delta m}(t, \cdot, m_2, \cdot) \Big\|_{(4+\alpha, 2+\alpha-\beta)} \leq C,$$

where C is dependent upon F, G, and T.

Proof. From the Lipschitz continuity in measure in Proposition 4.11, we have

$$\|K(t_0, \cdot, m_0^1, \cdot) - K(t_0, \cdot, m_0^2, \cdot)\|_{(4+\alpha, 2+\alpha-\beta)} \leq C d_1(m_0^1, m_0^2),$$

where $C > 0$ is independent of t_0, m_0^1, and m_0^2. By the identification between K and $\frac{\delta U}{\delta m}$ made in Corollary 4.14, we get

$$\|\frac{\delta U}{\delta m}(t_0, \cdot, m_0^1, \cdot) - \frac{\delta U}{\delta m}(t_0, \cdot, m_0^2, \cdot)\|_{(4+\alpha, 2+\alpha-\beta)} \leq C d_1(m_0^1, m_0^2),$$

which holds for any $t_0 \in [0, T]$, and the proposition is proven.

□

5 The Convergence Problem

We wrap up this thesis with a final section on the convergence problem. First we describe the necessary further assumptions to be made upon the Nash system, then we approximate the system using the well-posed Master Equation, and finally we sketch the convergence properties of the systems.

5.1 The N-dimensional Nash system

As we discussed in the background section of game theory, a N-player differential game has a Nash equilibrium taking the form of the following system.

$$\begin{cases} -\partial_t v^{N,i}(t,\boldsymbol{x}) - \varepsilon \sum_{j=1}^{N} \Delta_{x_j} v^{N,i}(t,\boldsymbol{x}) + \frac{1}{2}|D_{x_i} v^{N,i}(t,\boldsymbol{x})|^2 \\ \quad + \sum_{j\neq i} D_{x_j} v^{N,j}(t,\boldsymbol{x}) \cdot D_{x_j} v^{N,i}(t,\boldsymbol{x}) \\ = F^{N,i}(\boldsymbol{x}) & \text{in } [0,T] \times \left(\mathbb{R}^d\right)^N, \\ v^{N,i}(T,\boldsymbol{x}) = G^{N,i}(\boldsymbol{x}) & \text{in } \left(\mathbb{R}^d\right)^N, \end{cases}$$

where $\boldsymbol{x} = (x_1, ..., x_n) \in \left(\mathbb{R}^d\right)^N$, and we have inserted our explicit choice of Hamiltonian $H(x,p) = \frac{1}{2}|p|^2$. This system is still to general to apply Mean Field Game techniques, and we need to make some slight assumptions upon the cost terms F, G. Two core ideas of the Mean Field formulation are 1. That each agent is identical and 2. That each agent chooses strategies based on the distribution of the other agents, and not every other agents exact position. We satisfy these two requirements by making the following assumptions

$$F^{N,i}(\boldsymbol{x}) = F(x_i, m_{\boldsymbol{x}}^{N,i}),$$
$$G^{N,i}(\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}),$$

where F, G obeys the assumptions (F),(G), and

$$m_{\boldsymbol{x}}^{N,i} := \frac{1}{N-1} \sum_{j\neq i}^{N} \delta_{x_j}$$

is called the empirical measure. For any finite N, $m_{\boldsymbol{x}}^{N,i} \in \mathcal{P}_1(\mathbb{R}^d)$, and integration with respect to the empirical measure takes the form.

$$\int_{\mathbb{R}^d} f(y) m_{\boldsymbol{x}}^{N,i}(dy) = \frac{1}{N-1} \sum_{j\neq i} f(x_j).$$

Inserting these coupling terms into the Nash system yields

$$\begin{cases} -\partial_t v^{N,i}(t,\boldsymbol{x}) - \varepsilon \sum_{j=1}^{N} \Delta_{x_j} v^{N,i}(t,\boldsymbol{x}) + \frac{1}{2}|D_{x_i} v^{N,i}(t,\boldsymbol{x})|^2 \\ \quad + \sum_{j\neq i} D_{x_j} v^{N,j}(t,\boldsymbol{x}) \cdot D_{x_j} v^{N,i}(t,\boldsymbol{x}) \\ = F(x_i, m_{\boldsymbol{x}}^{N,i}) & \text{in } [0,T] \times \left(\mathbb{R}^d\right)^N, \\ v^{N,i}(T,\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}) & \text{in } \left(\mathbb{R}^d\right)^N. \end{cases} \tag{45}$$

We observe that under our new assumptions, the system is quite symmetric, and the solutions $v^{N,i}(t,\boldsymbol{x})$ are identical, save for the starting point $\boldsymbol{x}$. This is sufficient for an approximation using the Master Equation.

5.2 Approximation Using the Master Equation

Having described the properties of the N-dimensional Nash system, we seek to demonstrate how the well-posedness of the master equation provides us with the means to describe the asymptotic behaviour of the equilibrium system (45) as $N \to \infty$.

The key idea for proving convergence is to evaluate the solution of the Master Equation U in the empirical measure $m_{\boldsymbol{x}}^{N,i}$, and to show that this yields a system approximately equal to (45). We again follow an approach set out in [7], which as before is performed on $\mathbb{T}^d$, and the proofs will have to be amended in order to reflect our non-compact domain of $\mathbb{R}^d$. To implement this idea, we introduce what in the literature are called the finite dimensional projections of U.

Definition 5.1. Let U be the solutions of the Master Equation (24), and let $\boldsymbol{x} = (x_1, ..., x_N) \in (\mathbb{R}^d)^N$. We then define the set of finite dimensional projections $\{u^{N,i}(t, \boldsymbol{x})\}_{i\in\{1,...,N\}}$ as

$$u^{N,i}(t, \boldsymbol{x}) := U(t, x_i, m_{\boldsymbol{x}}^{N,i}). \tag{46}$$

Remark 5.2. For the definition of the projections to be well defined, we need that $m_{\boldsymbol{x}}^{N,i} \in \mathcal{P}_1(\mathbb{R}^d)$ in order for the solution of the Master Equation to be well defined. This is satisfied as the measures $m_{\boldsymbol{x}}^{N,i}$ are finite sums of Dirac measures

$$\sup_{i\in\{1,...,N\}} \int_{\mathbb{R}^d} |y| m_{\boldsymbol{x}}^{N,i}(dy) = \sup_{i\in\{1,...,N\}} \left(\frac{1}{N-1} \sum_{j\neq i}^{N} |x_j| \right) < \infty.$$

We also remark that the stronger condition $\sup_{N\in\{2,3,...\}} \sup_{i\in\{1,...,N\}} \int_{\mathbb{R}^d} |y| m_{\boldsymbol{x}}^{N,i}(dy)$ satisfied if $\{x_1, ..., x_n\}_{i\in\{1,...,N\}}$ are independent and identically distributed (i.i.d.) samples of some distribution $m \in \mathcal{P}_1(\mathbb{R}^d)$ by the Strong Law of Large Numbers (Theorem 20.1 in [15]). This might be a key assumption in applications using a $\mathcal{P}_1(\mathbb{R}^d)$-based Mean Field Game approach in order to show stronger forms of convergence.

The next step in establishing that $u^{N,i}(t, \boldsymbol{x})$ almost solves (45) is computing the relevant derivatives. The regularity of these projections is, perhaps as expected, directly tied to the regularity of the classical solution of the Master Equation as given in Theorem 4.4. Consequently, we immediately get the existence of

$$D_{x_i} u^{N,i}(t, \boldsymbol{x}) = D_x U(t, x_i, m_{\boldsymbol{x}}^{N,i}),$$

as the variable x_i only appears in the x-slot of U. For $D_{x_j} u^{N,i}(t, \boldsymbol{x})$ with $j \neq i$, computation is a bit more intricate due to the dependence of x_j through the measure $m_{\boldsymbol{x}}^{N,i}$. To handle this, we introduce the following technical proposition.

Proposition 5.3. *Assume that $U : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$ is $\mathcal{C}^1$ measure, that for each $m, m' \in \mathcal{P}_1(\mathbb{R}^d)$, $U(\cdot, m) \in C^{2+\alpha}(\mathbb{R}^d)$ and $U(\cdot, m, \cdot) \in C^{2+\alpha}(\mathbb{R}^d) \times C_b^2(\mathbb{R}^d)$ and*

$$\left\| U(\cdot, m') - U(\cdot, m) - \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(\cdot, m, y)(m' - m)(dy) \right\|_{2+\alpha} \leq C d_1^2(m, m').$$

Then for a fixed $m \in \mathcal{P}_1(\mathbb{R}^d)$ and a vector field $\phi \in L^2(m, \mathbb{R}^d)$

$$\left\| U(\cdot, (id + \phi)\#m) - U(\cdot, m) - \int_{\mathbb{R}^d} D_m U(\cdot, m, y) \cdot \phi(y) m(dy) \right\|_{2+\alpha} \leq C' \|\phi\|_{L^2(m)}^2.$$

where C' only depends on the constant C.

Proof. This result is a quantitative version of Proposition 2.8, and is stated and proven for the torus in Proposition A.2.1 in the appendix of 4.4. The proof in $\mathbb{R}^d$ is identical. □

With this technical estimate in hand, we can construct a proposition restating the derivatives of the projection $u^{N,i}$ as the derivatives of the solution of the Master Equation.

Lemma 5.4. *Assume* (F) *and* (G) *holds. Then for any $N \geq 2$, $i \in \{1, ..., N\}$, $u^{N,i} \in C_b^2\big((\mathbb{R}^d)^N\big)$ and for $j \neq i$*

$$\begin{aligned}
D_{x_j} u^{N,i}(t, \boldsymbol{x}) &= \frac{1}{N-1} D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j), \\
D_{x_i, x_j}^2 u^{N,i}(t, \boldsymbol{x}) &= \frac{1}{N-1} D_x D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j), \\
\left| D_{x_j, x_j}^2 u^{N,i}(t, \boldsymbol{x}) - \frac{1}{N-1} D_y [D_m U](t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j) \right| &\leq \frac{C}{N},
\end{aligned}$$

where C depends on the Lipschitz constant from Proposition 4.16.

Remark 5.5. This lemma is an adaption of Remark 6.1.2. in [7], and an analogue to Proposition 6.1.2. for the case for the Master Equation with common noise.

Proof. Assume without lack of generality that $\boldsymbol{x} = \{x_j\}_{j\in\{1,...,N\}}$ where $x_j \neq x_k$ for $j \neq k$ and let $\epsilon = \min_{j\neq k} |x_j - x_k|$. For the tuple $\boldsymbol{v} = \{v_j\}_{j\in\{1,...,N\}} \in (\mathbb{R}^d)^N$ where $v_i = 0$ for a fixed i, we define a smooth function ϕ such that

$$\phi(x) = v_j \qquad \text{if } x \in B_{\epsilon/4}(x_j),$$

and decaying quickly and smoothly to zero otherwise.

By the definition for the pushforward measure, we get

$$(\mathrm{id} + \phi)\#m_{\boldsymbol{x}}^{N,i} = \frac{1}{N-1}\sum_{j\neq i}\delta_{x_j+\phi(x_j)} = \frac{1}{N-1}\sum_{j\neq i}\delta_{x_j+v_j} = m_{\boldsymbol{x}+\boldsymbol{v}}^{N,i}$$

Inserting this into our projected solution yields

$$\begin{aligned} u^{N,i}(t,\boldsymbol{x}+\boldsymbol{v}) - u^{N,i}(t,\boldsymbol{x}) &= U(t,x_i,m_{\boldsymbol{x}+\boldsymbol{v}}^{N,i}) - U(t,x_i,m_{\boldsymbol{x}}^{N,i}) \\ &= U(t,x_i,(\mathrm{id}+\phi)\#m_{\boldsymbol{x}}^{N,i})) - U(t,x_i,m_{\boldsymbol{x}}^{N,i}). \end{aligned}$$

By the regularity of the solution of the Master Equation as given by Theorem 4.4, and the estimate from Corollary 4.14, we satisfy the requirements for Proposition 5.3

$$\left\| U(\cdot,(\mathrm{id}+\phi)\#m_{\boldsymbol{x}}^{N,i}) - U(\cdot,m_{\boldsymbol{x}}^{N,i}) - \int_{\mathbb{R}^d} D_m U(\cdot,m_{\boldsymbol{x}}^{N,i},y)\cdot\phi(y)m_{\boldsymbol{x}}^{N,i}(dy)\right\|_{2+\alpha} \leq C'\|\phi\|^2_{L^2(m_{\boldsymbol{x}}^{N,i})},$$

where

$$\|\phi\|^2_{L^2(m_{\boldsymbol{x}}^{N,i})} = \frac{1}{N-1}\sum_{j\neq i}^{N}|\phi(x_j)|^2 = \frac{1}{N-1}\sum_{j\neq i}^{N}|v_j|^2 = \frac{1}{N-1}|\boldsymbol{v}|^2.$$

By this bound, we get that

$$\begin{aligned} u^{N,i}(t,\boldsymbol{x}+\boldsymbol{v}) - u^{N,i}(t,\boldsymbol{x}) &= \int_{\mathbb{R}^d} D_m U(x_i,m_{\boldsymbol{x}}^{N,i},y)\cdot\phi(y)m_{\boldsymbol{x}}^{N,i}(dy) + \frac{1}{N-1}O(|\boldsymbol{v}|^2) \\ &= \frac{1}{N-1}\sum_{j\neq i}^{N} D_m U(x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot\phi(x_j) + \frac{1}{N-1}O(|\boldsymbol{v}|^2) \\ &= \frac{1}{N-1}\sum_{j\neq i}^{N} D_m U(x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot v_j + \frac{1}{N-1}O(|\boldsymbol{v}|^2). \end{aligned}$$

Now, denote by $x_{j,k}$ the k-th component of x_j, and define the standard basis vector $e_{j,k} = (0,...,0,1,0,...,0) \in (\mathbb{R}^d)^N$ as the vector with 1 in the $x_{j,k}$-slot, and zero otherwise. Likewise, denote by $e_k = (0,...,0,1,0,...,0) \in \mathbb{R}^d$ the vector with 1 in the k-th slot. Choosing $\boldsymbol{v} = he_{j,k}$ and applying the definition of the derivative in $\mathbb{R}^d$ yields

$$\begin{aligned} \partial_{x_{j,k}} u^{N,i}(t,\boldsymbol{x}) &= \lim_{h\to 0^+}\frac{u^{N,i}(t,\boldsymbol{x}+he_{j,k}) - u^{N,i}(t,\boldsymbol{x})}{h} \\ &= \lim_{h\to 0^+}\frac{1}{h}\frac{1}{N-1}\sum_{j\neq i}^{N} D_m U(x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot(he_{j,k})_j + \frac{1}{h}\frac{1}{N-1}O(h^2) \\ &= \frac{1}{N-1}D_m U(\cdot,m_{\boldsymbol{x}}^{N,i},x_j)\cdot e_k, \end{aligned}$$

which implies that

$$D_{x_j} u^{N,i}(t,\boldsymbol{x}) = \frac{1}{N-1} D_m U(t, x_i, m_{\boldsymbol{x}}^{N,i}, x_j).$$

For the second equality in this lemma, we observe that D_m only depends on x_i through the x-variable, and hence the derivative follows directly from the regularity of the Master Equation.

For the inequality we perform the same technique, now with a second order difference quotient with $k,l \in \{1,...,d\}$

$$\begin{aligned}
&\partial_{x_{j,l}} \partial_{x_{j,k}} u^{N,i}(t,\boldsymbol{x}) \\
&= \lim_{h\to 0^+} \frac{u^{N,i}(t,\boldsymbol{x}+he_{j,k}+he_{j,l}) - u^{N,i}(t,\boldsymbol{x}+he_{j,l}) - u^{N,i}(t,\boldsymbol{x}+he_{j,k}) + u^{N,i}(t,\boldsymbol{x})}{h^2} \\
&= \lim_{h\to 0^+} \frac{1}{h^2}\Big(\sum_{j\neq i}^{N} D_m U(x_i, m^{N,i}_{\boldsymbol{x}+he_{j,l}}, x_j + he_l)\cdot (he_{j,k})_j - D_m U(x_i, m_{\boldsymbol{x}}^{N,i}, x_j)\cdot (he_{j,k})_j \\
&+ \frac{1}{N-1} O(h^2)\Big) = \frac{1}{N-1} \partial_{y_l} D_m U(x_i, m_{\boldsymbol{x}}^{N,i}, x_j)\cdot e_k + O(\frac{1}{N}),
\end{aligned}$$

where we have used that due to Proposition 4.16

$$\big|D_m U(x_i, m^{N,i}_{\boldsymbol{x}+he_{j,l}}, x_j) - D_m U(x_i, m_{\boldsymbol{x}}^{N,i}, x_j)\big| \le C d_1(m^{N,i}_{\boldsymbol{x}+he_{j,l}}, m_{\boldsymbol{x}}^{N,i}) \le \frac{C}{N-1}|he_l| = C\frac{h}{N-1}.$$

Since this holds for all choices of $k,l \in \{1,...,d\}$, the inequality in the lemma statement holds. □

With these estimates in hand, we can show that the projection $\left(u^{N,i}\right)_{i\in\{1,\dots,N\}}$ is an approximate solution to the Nash system (45).

Proposition 5.6. *Assume the coupling term assumptions* (F),(G) *hold. Then one has, for any* $i \in \{1,...,N\}$,

$$\begin{cases}
-\partial_t u^{N,i}(t,\boldsymbol{x}) - \varepsilon \sum_{j=1}^{N} \Delta_{x_j} u^{N,i}(t,\boldsymbol{x}) + \frac{1}{2}|D_{x_i} u^{N,i}(t,\boldsymbol{x})|^2 \\
\quad + \sum_{j\neq i} D_{x_j} u^{N,j}(t,\boldsymbol{x}) \cdot D_{x_j} u^{N,i}(t,\boldsymbol{x}) \\
= F(x_i, m_{\boldsymbol{x}}^{N,i}) + r^{N,i}(t,\boldsymbol{x}) & \text{in } [0,T]\times \left(\mathbb{R}^d\right)^N \\
u^{N,i}(T,\boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}) & \text{in } \left(\mathbb{R}^d\right)^N
\end{cases}$$

where $r^{N,i} \in L^\infty\Big([0,T]\times \left(\mathbb{R}^d\right)^N\Big)$ *with*

$$\|r^{N,i}\|_\infty \le C\left(\frac{1}{N} + \frac{1}{N^2}\sum_{j\neq i}^{N} |x_j - x_i|\right)$$

Proof. By the definition of the projection (46), $u^{N,i}(t,\boldsymbol{x})$ corresponds to the solution of the Master Equation evaluated at a point $(t, x_i, m_{\boldsymbol{x}}^{N,i})$. We insert the point into the Master Equation (24), and get

$$\begin{aligned}
-\partial_t U(t,x_i,m_{\boldsymbol{x}}^{N,i}) - \varepsilon\Delta_x U(t,x_i,m_{\boldsymbol{x}}^{N,i}) &+ \frac{1}{2}|D_x U(x_i,t,m_{\boldsymbol{x}}^{N,i})|^2 \\
&- \varepsilon\int_{\mathbb{R}^d} \operatorname{div}_y[D_m U](t,x_i,m_{\boldsymbol{x}}^{N,i},y) m_{\boldsymbol{x}}^{N,i}(dy) \\
&+ \int_{\mathbb{R}^d} D_m U(t,x_i,m_{\boldsymbol{x}}^{N,i},y)\cdot D_x U(t,y,m_{\boldsymbol{x}}^{N,i}) m_{\boldsymbol{x}}^{N,i}(dy), \\
&= F(x_i, m_{\boldsymbol{x}}^{N,i})
\end{aligned}$$

which after applying the definition for integration with respect to $m_{\boldsymbol{x}}^{N,i}$ yields

$$
\begin{aligned}
-\partial_t u^{N,i}(t,\boldsymbol{x}) - \varepsilon\Delta_{x_i} u^{N,i}(t,\boldsymbol{x}) + \frac{1}{2}|D_{x_i}u^{N,i}(t,\boldsymbol{x})|^2 \\
- \varepsilon\frac{1}{N-1}\sum_{j\neq i}^{N} \mathrm{div}_y[D_mU](t,x_i,m_{\boldsymbol{x}}^{N,i},x_j) \\
+ \frac{1}{N-1}\sum_{j\neq i}^{N} D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot D_xU(t,x_j,m_{\boldsymbol{x}}^{N,i}) \\
= F(x_i,m_{\boldsymbol{x}}^{N,i}).
\end{aligned}
\tag{47}
$$

We first consider the term containing the sum of the dot products. From (27) in Theorem 4.4

$$
\begin{aligned}
\left|D_xU(t,x_i,m_{\boldsymbol{x}}^{N,i}) - D_xU(t,x_i,m_{\boldsymbol{x}}^{N,j}))\right| &\le Cd_1(m_{\boldsymbol{x}}^{N,i},m_{\boldsymbol{x}}^{N,j}) = C\sup_{\psi\in 1\text{-Lip}}\int_{\mathbb{R}^d}\psi(y)(m_{\boldsymbol{x}}^{N,i}-m_{\boldsymbol{x}}^{N,j})(dy) \\
&= \frac{C}{N-1}\sup_{\psi\in 1\text{-Lip}}(\psi(x_j)-\psi(x_k)) \le \frac{C}{N-1}|x_j-x_i|,
\end{aligned}
$$

and thus

$$
\left|D_xU(t,x_j,m_{\boldsymbol{x}}^{N,i}) - D_{x_j}u^{N,j}(t,\boldsymbol{x})\right| = \left|D_xU(t,x_j,m_{\boldsymbol{x}}^{N,i}) - D_xU(t,x_j,m_{\boldsymbol{x}}^{N,j})\right| \le \frac{C}{N-1}|x_j-x_i|.
$$

In combination with the first equality from Lemma 5.4, this yields

$$
\begin{aligned}
&\frac{1}{N-1}\sum_{j\neq i}^{N} D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot D_xU(t,x_j,m_{\boldsymbol{x}}^{N,i}) \\
&= \frac{1}{N-1}\sum_{j\neq i}^{N} D_mU(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot\left(D_{x_j}u^{N,j}(t,\boldsymbol{x}) + O\left(\frac{|x_j-x_i|}{N-1}\right)\right) \\
&= \sum_{j\neq i}^{N} D_{x_j}u^{N,i}(t,\boldsymbol{x})\cdot D_{x_j}u^{N,j}(t,\boldsymbol{x}) + \frac{1}{(N-1)^2}\sum_{j\neq i}^{N} O(|x_i-x_j|)
\end{aligned}
$$

For the next term in the system, we compute using the inequality from Lemma 5.4

$$
\begin{aligned}
\sum_{j=1}^{N}\Delta_{x_j}u^{N,i}(t,\boldsymbol{x}) &= \Delta_{x_i}u^{N,i}(t,\boldsymbol{x}) + \sum_{j\neq i}^{N}\Delta_{x_j}u^{N,i}(t,\boldsymbol{x}) \\
&= \Delta_{x_i}u^{N,i}(t,\boldsymbol{x}) + \frac{1}{N-1}\sum_{j\neq i}^{N}\mathrm{div}_y[D_mU](t,x_i,m_{\boldsymbol{x}}^{N,i},x_j) + O\big(\frac{1}{N}\big).
\end{aligned}
$$

Inserting these identities into (47) leaves us with

$$
\begin{aligned}
-\partial_t u^{N,i}(t,\boldsymbol{x}) - \varepsilon\sum_{j=1}^{N}\Delta_{x_j}u^{N,i}(t,\boldsymbol{x}) + \frac{1}{2}|D_{x_i}u^{N,i}(t,\boldsymbol{x})|^2 + \sum_{j\neq i}^{N} D_{x_j}u^{N,i}(t,\boldsymbol{x})\cdot D_{x_j}u^{N,j}(t,\boldsymbol{x}) \\
= F(x_i,m_{\boldsymbol{x}}^{N,i}) + O\big(\frac{1}{N} + \frac{1}{N^2}\sum_{j\neq i}^{N}|x_j-x_i|\big),
\end{aligned}
$$

and we collect the $O(\frac{1}{N} + \frac{1}{(N-1)^2}\sum_{j\neq i}^{N}|x_j-x_i|$ as the term $r^{N,i}\in L^\infty\Big([0,T]\times\big(\mathbb{R}^d\big)^N\Big)$, finalising the proof. □

Remark 5.7. The bound upon the error term

$$\|r^{N,i}\|_\infty \leq C\left(\frac{1}{N}+\frac{1}{N^2}\sum_{j\neq i}^{N}|x_j-x_i|\right),$$

might seem a bit pathological at first glance. Indeed its form is due to Proposition 5.6 mirroring a proof performed on the torus in [7]. For the case on the torus, we have $\boldsymbol{x}\in\left(\mathbb{T}^d\right)^N$ and hence $\frac{1}{N^2}\sum_{j\neq i}^{N}|x_j-x_i|\leq\frac{1}{N}$. Furthermore we observe that if the condition

$$\sup_{N\in\mathbb{N}}\left(\frac{1}{N}\sum_{i=1}^{N}|x_i|\right)<\infty,$$

is satisfied, then we have that $\|r^{N,i}\|_\infty\leq\frac{C}{N}\to 0$ as $N\to\infty$. As was noted in Remark 5.2, this is achieved if $\{x_1,...,x_N\}$ are the i.i.d. samples of a distribution in $\mathcal{P}_1(\mathbb{R}^d)$.

5.3 Convergence and Further Work

In this final section, we will cover the convergence of the N-dimensional Nash system to the Master Equation. The suitable form of convergence in this case will be the following.

$$\sup_{i\in\{1,\dots,N\}}\left|v^{N,i}(t_0,\boldsymbol{x})-U(t_0,x_i,m_{\boldsymbol{x}}^{N,i})\right|\leq\frac{C}{N} \tag{48}$$

This "convergence" is a bit strange and subtle, which according to [7], is due to the qualitative difference between the agents in the Nash system having to observe each other, while in the limit system given by the Master Equation, the agents only need to observe a distribution of the populations, and hence do not need to "react" to players specific behaviour.

We will not perform any proofs for the convergence, as the stochastic analysis for the system on $\mathbb{R}^d$ is outside the scope of this thesis, and will instead perform a brief revue of the method [7] employs in order to prove convergence for the system on the torus $\mathbb{T}^d$.

The proof of convergence starts off by of comparing the "optimal trajectories" generated by the projected Master Equation and the Nash system, a pair of two similar SDEs

$$\begin{cases} dX_{i,t}=-D_{x_i}u^{N,i}(t,\boldsymbol{X}_t)dt+\sqrt{2\varepsilon}dB_t^i, & t\in(t_0,T],\\ X_{i,t_0}=Z_i, \end{cases}$$

and

$$\begin{cases} dY_{i,t}=-D_{x_i}v^{N,i}(t,\boldsymbol{Y}_t)dt+\sqrt{2\varepsilon}dB_t^i, & t\in(t_0,T],\\ Y_{i,t_0}=Z_i, \end{cases}$$

which generate the systems of stochastic processes $\{\boldsymbol{X}_t=\{X_{i,t}\}_{i\in\{1,\dots,N\}}\}_{t\in[t_0,T]}$ and $\{\boldsymbol{Y}_t=\{Y_{i,t}\}_{i\in\{1,\dots,N\}}\}_{t\in[t_0,T]}$ respectively. $\boldsymbol{Z}=\{Z_i\}_{i\in\{1,\dots,N\}}$ is a family i.i.d. random variables of law m_0 independent of the i.i.d. family of Brownian motions $\{\{B_t^i\}_{t\in[0,T]}\}_{i\in\{1,\dots,N\}}$ Cardaliaguet et al. [7] then use standard arguments from Itô calculus as well as Grönwall estimates in order to produce the estimates

$$E\left[\sup_{t\in[t_0,T]}\left|Y_{i,t}-X_{i,t}\right|\right]\leq\frac{C}{N},\qquad\forall t\in[t_0,T],$$

and almost surely,

$$\left|u^{N,i}(t_0,\boldsymbol{Z})-v^{N,i}(t_0,\boldsymbol{Z})\right|\leq\frac{C}{N}, \tag{49}$$

for all $i \in \{1, ..., N\}$ with constant C deterministic and independent of t_0, m_0, and N. They then continue by choosing $\boldsymbol{Z}$ as uniformly distributed random variables, and invoke continuity of U and $\{v^{N,i}\}_{i\in\{1,...,N\}}$ to obtain the desired result (48).

As a final remark upon the convergence of the Nash system, we discuss three points that need to be accounted for in order to translate this proof to the domain $\mathbb{R}^d$. Firstly, as discussed in Remark 5.7 of the error term $r^{N,i}$ of Proposition 5.6 we need to make sure that the estimate

$$\|r^{N,i}\|_\infty \leq C \left(\frac{1}{N} + \frac{1}{N^2} \sum_{j\neq i}^{N} |x_j - x_i| \right) \leq \frac{C}{N},$$

for some N-independent constant C in order for the constant in (49) to be N-independent. This is obtained by the sampling assumption remarked earlier, however, this might be too strong of an assumption to be practical. It might also be possible to supply some sort of boundedness from a tightness estimate, since tightness is central to convergence in measure.

Secondly, on the proof on the torus the compactness of the domain means that one is able to choose $\boldsymbol{Z}$ to have an uniform distribution. This approach is not possible for $\mathbb{R}^d$, as a uniform probability distributions cannot exist on the entire space, since $\mathbb{R}^d$ has infinite Lebesgue measure. A possible way to fix this might be to try to perform the proof with a deterministic initial value $\boldsymbol{Z}$, or attempt to perform some sort of approximation by Gaussian distributions.

Lastly, we have to make sure that the Itô calculus and Grönwall estimates employed are well defined. A lot of Itô theory, for example the treatment in [26], is primarily L^2-based, that is, concerning random variables with second moments. As previously mentioned, probability measures defined on $\mathbb{T}^d$ have every moment, and are problem free in this regard. However, for our analysis on $\mathbb{R}^d$ some care needs to be taken.

Bibliography

[1] Yves Achdou et al. *Mean Field Games*. Springer, 2019.

[2] Robert J. Aumann. 'Markets with a Continuum of Traders'. In: *Econometrica* 32.1/2 (1964), pp. 39–50. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1913732 (visited on 22nd July 2022).

[3] Vladimir Bogachev, N. V. Krylov and Michael Röckner. 'Elliptic and parabolic equations for measures'. In: *Russian Mathematical Surveys* 64 (2009), pp. 973–1078.

[4] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2011. ISBN: 9780387709130.

[5] Pierre Cardaliaguet. *Notes on Mean Field Games (from P.-L. Lions' lectures at College de France)*. Sept. 2013.

[6] Pierre Cardaliaguet, Marco Cirant and Alessio Porretta. *Splitting methods and short time existence for the master equations in mean field games*. 2020. DOI: 10.48550/ARXIV.2001.10406. URL: https://arxiv.org/abs/2001.10406.

[7] Pierre Cardaliaguet et al. *The master equation and the convergence problem in mean field games*. Princeton University Press, 2019.

[8] Rene Carmona and Francois Delarue. *Probabilistic theory of mean field games with applications. I*. Springer, 2018.

[9] Indranil Chowdhury, Espen R. Jakobsen and Miłosz Krupski. *On fully nonlinear parabolic mean field games with examples of nonlocal and local diffusions*. 2021. DOI: 10.48550/ARXIV.2104.06985. URL: https://arxiv.org/abs/2104.06985.

[10] Nelson Dunford and Jacob T. Schwartz. *Linear operators. 1: General theory*. Wiley Interscience Publ, 1988.

[11] Olav Ersland and Espen R. Jakobsen. 'On fractional and nonlocal parabolic mean field games in the whole space'. In: *Journal of Differential Equations* 301 (2021), pp. 428–470. ISSN: 0022-0396. DOI: https://doi.org/10.1016/j.jde.2021.08.026. URL: https://www.sciencedirect.com/science/article/pii/S0022039621005374.

[12] Lawrence C. Evans. *Partial Differential Equations, second edition*. American Mathematical Society, 2010.

[13] C. Gasquet, R. Ryan and P. Witomski. *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets*. Texts in Applied Mathematics. Springer New York, 1998. ISBN: 9780387984858. URL: https://books.google.no/books?id=PXBUBBBmN9oC.

[14] Anders Hoel. 'PDE Analysis of Mean Field Games'. In: (2022).

[15] J. Jacod and P. Protter. *Probability Essentials*. Universitext. Springer Berlin Heidelberg, 2004. ISBN: 9783540438717. URL: https://books.google.no/books?id=3xm1vwEACAAJ.

[16] John L. Kelley. *General Topology*. D. Van Nostrand Company, 1955.

[17] N.V. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. American Mathematical Society, 1996.

[18] O. A. Ladyzhenskaia, V. A. Solonnikov and N. Uraltseva. *Linear and quasi-linear equations of parabolic type*. American Mathematical Society, 1998.

[19] Hans Langtangen and Geir Pedersen. *Scaling of Differential Equations*. Jan. 2016. ISBN: 978-3-319-32725-9. DOI: 10.1007/978-3-319-32726-6.

[20] Jean-Michel Lasry and Pierre-Louis Lions. 'Mean Field Games'. In: *Japanese Journal of Mathematics* 2 (Mar. 2007), pp. 229–260. DOI: 10.1007/s11537-007-0657-8.

[21] Gary M. Lieberman. *Second order parabolic differential equations*. World Scientific, 1996. ISBN: 9789810228835.

[22] Pierre-Louis Lions. *Courses in the mean field games at Collège de France*. 2006–2012.

[23] Gonzalo García Macias. 'A quantitative version of the Arzelà-Ascoli theorem based on the degree of nondensifiability and applications'. In: *Applied General Topology* 20.1 (2019), pp. 265–279. ISSN: 1989-4147. DOI: 10.4995/agt.2019.10930. URL: https://polipapers.upv.es/index.php/AGT/article/view/10930.

[24] John Nash. 'Non-Cooperative Games'. In: *Annals of Mathematics* 54.2 (1951), pp. 286–295. ISSN: 0003486X. URL: http://www.jstor.org/stable/1969529 (visited on 22nd July 2022).

[25] John von Neumann, Oskar Morgenstern and Ariel Rubinstein. *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Princeton University Press, 1944. ISBN: 9780691130613. URL: http://www.jstor.org/stable/j.ctt1r2gkx (visited on 25th July 2022).

[26] Bernt Øksendal. *Stochastic Differential Equations*. Springer, 2003.

[27] Yu. V. Prokhorov. 'Convergence of random processes and limit theorems in probability theory'. In: *Theory Probab. Appl.* 1 (1956), pp. 157–214. DOI: https://doi.org/10.1137/1101016.

[28] Fredi Troltzsch. *Optimal control of partial differential equations: Theory, methods, and applications*. American Mathematical Society, 2010.

NTNU
Norwegian University of
Science and Technology