Mean field games and interacting particle systems

Daniel Lacker

Spring 2018

Contents

Preface

These are the lecture notes from the Spring 2018 PhD course offered at Columbia University IEOR on mean field games and interacting particle systems. Be warned that the notes are not very polished, nor are they mathematically completely rigorous. The goal was to provide a crash course on stochastic differential mean field games and interacting SDE systems of McKean-Vlasov type, along with many of the pre-requisites a first-year PhD student may not yet have encountered. In particular, the notes include brief treatments of weak convergence, Wasserstein metrics, stochastic optimal control theory, and stochastic differential games.

The course is oriented more toward breadth than depth. While many fairly complete proofs are given, many details are left out as well (e.g., checking that a local martingale is a true martingale, or even spelling out complete assumptions on the coefficients). Notably, our coverage of stochastic control and games is limited to the analytic approach; that is, we emphasize how to derive and apply verification theorems, and we completely omit the popular "probabilistic approach" based on the maximum principle and BSDEs. The course may not prepare theoretically-oriented students to prove new theorems about mean field games, but hopefully they will take away a strong big-picture understanding of the flexibility of the modeling framework and how to build and analyze various models.

Many thanks to the students who transcribed and typed up the majority of these notes: Yaarit Even, Miguel Garrido Garci, Camilo Hernandez, Kleanthis Karakolios, Nikhil Kotecha, Enrique Lelo de Larrea Andrade, Agathe Soret, and Florian Stebegg.

1 Introduction

This course is about the analysis of certain kinds of interacting particle systems, though the scope of application of the models we study has reached far beyond its origin in statistical mechanics. So, while we often stick with the term "particle," it should not be interpreted too literally.

More specifically, this course is primarily about *mean field* models in which each particle is represented by a stochastic differential equation. The term *mean field* refers to the highly symmetric form of interaction between the particles. The first half of the course focuses on zero-intelligence models, in which particles follow prescribed laws of motion. In the second half of the course, we study *mean field games*, in which the particles follow *controlled* dynamics and each tries to optimize some objective. In both cases, the emphasis will be on formulating and analyzing a suitable infinite-particle (or continuum) limit.

The general theme of the analysis is to see how to efficiently move from a microscopic description of a model (i.e., individual-level laws of motion or incentives) to a macroscopic one.

1.1 Part 1: Interacting diffusion models

The main family of models studied in the first half of the course take the following form: Particles $X^1, \ldots, X^n$ are Itô processes, evolving according to the stochastic differential equation (SDE) system

$$dX_t^i = b(X_t^i, \widehat{\mu}_t)dt + \sigma(X_t^i, \widehat{\mu}_t)dW_t^i, \tag{1.1}$$

where

$$\widehat{\mu}_t = \frac{1}{n}\sum_{k=1}^n \delta_{X_t^k}$$

is the *empirical measure* of the particles. The coefficients b and σ are the same for each particle, and the driving Brownian motions W^i are independent. Typically, the initial positions X_0^i are i.i.d. The particles would be i.i.d. if not for the interaction coming through the empirical measure, and this is the only source of interaction between the particles.

For a concrete example, consider the toy model

$$dX_t^i = a(\overline{X}_t - X_t^i)dt + \sigma dW_t^i,$$

where $a, \sigma > 0$ and $\overline{X}_t = \frac{1}{n}\sum_{k=1}^n X_t^k$ is the empirical average. Each particle faces an independent noise, and the drift term pushes each particle toward the empirical average. Notice that the entire system is symmetric, in the sense that if we permute the "names" $i = 1, \ldots, n$, we end up with the same particle system. In other words, $(X^1, \ldots, X^n)$ is *exchangeable*. To get a first vague sense of how a mean field limit works in a model of this form, simply average the n particles to find the dynamics for the empirical mean:

$$d\overline{X}_t = \frac{\sigma}{n}\sum_{k=1}^n dW_t^k.$$

In integrated form, we have

$$\overline{X}_t = \frac{1}{n}\sum_{k=1}^n X_0^k + \frac{\sigma}{n}\sum_{k=1}^n W_t^k.$$

If (X_0^k) are i.i.d. with mean m, then the law of large numbers tells us that $\overline{X}_t \to m$ almost surely as $n \to \infty$, since of course Brownian motion has mean zero. If we focus now on a fixed particle i in the n-particle system we find that as $n \to \infty$ the behavior of particle i should look like

$$dX_t^i = a(m - X_t^i)dt + \sigma dW_t^i.$$

Since m is constant, this "limiting" evolution consists of i.i.d. particles. In summary, as $n \to \infty$, the particles become *asymptotically i.i.d.*, and the behavior of each one is described by an Ornstein-Uhlenbeck process.

The $n \to \infty$ limit in this toy model can be studied quite easily by taking advantage of the special form of the model. The goal of the first part of the course is to see how to identify and analyze an $n \to \infty$ limit for the more general kind of interacting diffusion model like (1.1), when explicit calculations are unavailable.

1.2 Part 2: (Stochastic differential) Mean field games

One we understand the interacting diffusion models of the previous section, we will study the controlled version: Particles $X^1, \ldots, X^n$ are Itô processes, evolving according to the stochastic differential equation (SDE) system

$$dX_t^i = b(X_t^i, \widehat{\mu}_t, \alpha_t^i)dt + \sigma(X_t^i, \widehat{\mu}_t, \alpha_t^i)dW_t^i,$$

where again

$$\widehat{\mu}_t = \frac{1}{n}\sum_{k=1}^{n} \delta_{X_t^k}$$

is the *empirical measure* of the particles. The coefficients b and σ are the same for each particle, and the driving Brownian motions W^i are independent. But now each particle i gets to choose a control process α^i. For instance, X^i could the the velocity and α^i the acceleration, or X^i could be the wealth of an investor with α^i describing the allocation of wealth between various assets or investment vehicles. We typically call them *agents* instead of *particles* with the processes are controlled.

Each agent i endeavors to maximize some objective criterion, typically of the form

$$J_i(\alpha_1, \ldots, \alpha_n) = \mathbb{E}\left[\int_0^T f(X_t^i, \widehat{\mu}_t, \alpha_t^i)dt + g(X_T^i, \widehat{\mu}_T)\right],$$

where $T > 0$ is a finite time horizon, though there are many other natural forms of objective criterion (infinite time horizon, etc.). Again, the cost functions f and g are the same for each agent.

Unlike the particle systems of the previous section, these controlled systems do not yet specify the dynamics, as we must first resolve the optimization problems faced by each particle. These optimization problems are of course interdependent, owing to the dependence on the empirical measure. To do this, we use the canonical notion of *Nash equilibrium* from game theory, which means that $(\alpha^1, \ldots, \alpha^n)$ are an equilibrium if they satisfy

$$J_i(\alpha_1, \ldots, \alpha_n) \geq J_i(\alpha_1, \ldots, \alpha^{i-1}, \beta, \alpha^{i+1}, \ldots, \alpha_n),$$

for every alternative control β and every $i = 1, \ldots, n$. That is, each agent is acting optimally, given the behavior of the other agents–no single agent has any incentive to switch strategies, given that the other agents strategies are fixed. We are for now ignoring some important modeling decisions pertaining to what an "admissible control" really is, and what information is available to each agent.

To analyze the in-equilibrium behavior of this system for large n is typically quite difficult, and for this reason we again study a large n limit. We will see how to identify and analyze these continuum limits, leading us to a fascinating extension of the more classical models of the previous section.

1.3 Organization of the notes

We begin in Section 2 with a review of the basics of weak convergence of probability measures, along with some discussion of Wasserstein metrics and convergence of empirical measures of i.i.d. samples. Section 3 covers interacting diffusions and the McKean-Vlasov limit, as described in Section 1.1 above. Before turning to (dynamic) mean field games, we devote Section 4 to the analysis of a very simple yet instructive model of a large static game, in which each of n agents chooses a single action deterministically. Finally Section 8 is devoted to mean field games, as introduced in Section 1.2 above.

2 Weak convergence and Wasserstein metrics

Many of the main results of this course will be stated in terms of convergence in distribution, of random variables, vectors, processes, and measures. This section covers the very basics of the theory weak convergence of probability measures on metric spaces. In order to get to the meat of the course, we will cover this material far too quickly. For more details, refer to the classic textbook of Billingsley [11], and a more quick and concise treatments can be found in Kallenberg's tome [75, Chapter 14].

Throughout the section, let $(\mathcal{X}, d)$ denote a metric space. We always equip $\mathcal{X}$ with the Borel σ-field, meaning the σ-field generated by the open sets of $\mathcal{X}$. We will write $\mathcal{P}(\mathcal{X})$ for the set of (Borel) probability measures on X. Let $C_b(\mathcal{X})$ denote the set of bounded continuous real-valued functions on X. The fundamental definition is the following, which we state in two equivalent forms, one measure-theoretic and one probabilistic:

Definition 2.1. Given a probability measure $\mu \in \mathcal{P}(\mathcal{X})$ and a sequence $(\mu_n) \subset \mathcal{P}(\mathcal{X})$, we say that μ_n *converges weakly to* μ, or $\mu_n \to \mu$, if

$$\lim_{n\to\infty} \int_{\mathcal{X}} f \, d\mu_n = \int_{\mathcal{X}} f \, d\mu, \text{ for every } f \in C_b(\mathcal{X}).$$

Definition 2.2. Given a sequence of $\mathcal{X}$-valued random variables (X_n), we say that X_n *converges weakly (or in distribution)* to another $\mathcal{X}$-valued random variable X (often denoted $X_n \Rightarrow X$) if

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)], \text{ for every } f \in C_b(\mathcal{X}).$$

Throughout the course, when we say X *is a* $\mathcal{X}$*-valued random variable*, we mean the following: Behind the scenes, there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a function $X : \Omega \to \mathcal{X}$, measurable with respect to the Borel σ-field on $\mathcal{X}$. We will rarely need to be explicit about the choice of probability space. Unlike other modes of convergence of random variables, such as convergence in probability or almost-sure convergence, weak convergence does not require the random variable X_n to be defined on the same probability space! That is, we could have X_n defined on its own probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ for each n, and X defined on its own $(\Omega, \mathcal{F}, \mathbb{P})$, and the Definition 2.2 still makes sense. Note then that $X_n \Rightarrow X$ if and only if $\mathbb{P}_n \circ X_n^{-1} \to \mathbb{P} \circ X^{-1}$ weakly.[1]

As a first trivial case, suppose we are given a deterministic sequence $(x_n) \subset \mathcal{X}$ converging to a point $x \in \mathcal{X}$. Write $\delta_x \in \mathcal{P}(\mathcal{X})$ for the Dirac probability measure at x, meaning $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise, for a set $A \subset \mathcal{X}$. Then $\delta_{x_n} \to \delta_x$ weakly, simply because $f(x_n) \to f(x)$ for every $f \in C_b(\mathcal{X})$.

Another fact to notice is that if X_n converges in probability to X in the sense that

$$\lim_{n\to\infty} \mathbb{P}(d(X_n, X) > \epsilon) = 0, \quad \text{for all } \epsilon > 0,$$

then $X_n \Rightarrow X$. This follows from the dominated convergence theorem. Of course, almost sure convergence implies convergence in probability, and so almost sure convergence implies convergence in distribution.

One's first genuine encounter with weak convergence is the central limit theorem: If (X_n) is a sequence of real-valued random variables with mean μ and variance $\sigma^2 > 0$ then

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_n - \mu}{\sigma} = Z,$$

where Z is a standard Gaussian. This is often stated in a different but equivalent form,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_n - \mu}{\sigma} \le x \right) = \mathbb{P}(Z \le x), \quad \text{for all } x \in \mathbb{R}.$$

[1] We use the standard notation for image measures: $\mathbb{P} \circ X^{-1}(A) := \mathbb{P}(X \in A)$.

The equivalence between these two forms will follow from Theorem 2.3 below. One early motivation for studying weak convergence on metric spaces, as opposed to simply Euclidean spaces, was the study of weak convergence of random walks to Brownian motion, leading to Donsker's theorem.

It is important to notice that the weak convergence $\mu_n \to \mu$ does not imply the stronger convergence $\int_{\mathcal{X}} f \, d\mu_n \to \int_{\mathcal{X}} f \, \mu$ for every bounded measurable function f. Stronger still is *convergence in total variation*, which requires that $\int_{\mathcal{X}} f \, d\mu_n \to \int_{\mathcal{X}} f \, d\mu$ *uniformly* over measurable functions f with $|f| \leq 1$. It is worth noting, however, that if the metric space $\mathcal{X}$ is finite, then $\mathcal{P}(\mathcal{X})$ can be identified with a compact subset of $\mathbb{R}^n$, where $n = |\mathcal{X}|$. In this case, the three aforementioned modes of convergence all coincide. Nonetheless, the following famous theorem clarifies what weak convergence does tell us about setwise convergence:

Theorem 2.3 (Portmaneau theorem). *Let $\mu, \mu_n \in \mathcal{P}(\mathcal{X})$. The following are equivalent:*

(i) $\mu_n \to \mu$.

(ii) $\liminf_{n\to\infty} \mu_n(U) \geq \mu(U)$ *for every open set* $U \subset \mathcal{X}$.

(iii) $\limsup_{n\to\infty} \mu_n(C) \geq \mu(C)$ *for every closed set* $C \subset \mathcal{X}$.

(iv) $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ *for every Borel set* $A \subset \mathcal{X}$ *with* $\mu(A^\circ) = \mu(\overline{A})$.[2]

(v) $\int f \, d\mu_n \to \int f \, d\mu$ *for every bounded uniformly continuous function f on $\mathcal{X}$.*

We omit the proof of Theorem 2.3, as it is entirely classical (see [11, Theorem 2.1] or [75, Theorem 3.25]. We do prove an easy but important theorem we will make good use:

Theorem 2.4 (Continuous mapping theorem). *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces, and (X_n) is a sequence of $\mathcal{X}$-valued random variables converging in distribution to another $\mathcal{X}$-valued random variable X. Suppose $g : \mathcal{X} \to \mathcal{Y}$ is a continuous function. Then $g(X_n) \Rightarrow g(X)$.*

Proof. For any $f \in C_b(\mathcal{Y})$, the function $f \circ g$ belongs to $C_b(\mathcal{X})$. Hence, since $X_n \Rightarrow X$,

$$\lim_{n\to\infty} \mathbb{E}[f(g(X_n))] = \mathbb{E}[f(g(X))].$$

□

[2]Here A° denotes the interior of the set A and $\overline{A}$ the closure.

We will not make too much use of this, but an extremely important theorem of Prokhorov characterizes pre-compact sets in $\mathcal{P}(\mathcal{X})$. Given a set $S \subset \mathcal{P}(\mathcal{X})$, we say that the family S of probability measures is *tight* if for all ϵ there exists a compact set $K \subset X$ such that

$$\sup_{\mu \in K} \mu(K^c) \leq \epsilon.$$

The importance of this definition lies in the following theorem, the proof of which can be found in [11, Theorem 6.1, 6.2] [75, Theorem 14.3]

Theorem 2.5 (Prokhorov's theorem). *Suppose $(\mu_n) \subset \mathcal{P}(\mathcal{X})$. If (μ_n) is tight, then it is pre-compact in the sense that every subsequence admits a further subsequence which converges weakly to some $\mu \in \mathcal{P}(\mathcal{X})$. Conversely, if (μ_n) is pre-compact, and if the metric space $(\mathcal{X}, d)$ is separable and complete, then (μ_n) is tight.*

We state one more theorem without giving the proof, which can be found in [75, Theorem 3.30]:

Theorem 2.6 (Skorokhod's representation theorem). *Suppose $(\mathcal{X}, d)$ is separable. Suppose μ_n converges weakly to μ in $\mathcal{P}(\mathcal{X})$. Then there exists a probability space $(\Omega, F, \mathbb{P})$ supporting $\mathcal{X}$-valued random variables X_n and X, with $X_n \sim \mu_n$ and $X \sim \mu$, such that $X_n \to X$ almost surely.*

Theorem 2.6 is quite useful in that it lets us "cheat" by proving things about weak convergence using what we already know about almost sure convergence. For example, we will give a simple proof of the following important theorem, which tells us when we can extend the convergence of Definition 2.1 to cover unbounded functions:

Theorem 2.7. *Suppose μ_n converges weakly to μ in $\mathcal{P}(\mathcal{X})$. If $f : \mathcal{X} \to \mathbb{R}$ is continuous and* uniformly integrable *in the sense that*

$$\lim_{r \to \infty} \sup_n \int_{\{|f| \geq r\}} |f| \, d\mu_n = 0,$$

then we have $\lim_{n \to \infty} \int_{\mathcal{X}} f \, d\mu_n = \int_{\mathcal{X}} f \, d\mu$.

Proof. By the continuous mapping Theorem 2.4, we know that $\mu_n \circ f^{-1} \to \mu \circ f^{-1}$ weakly in $\mathcal{P}(\mathbb{R})$. By Theorem 2.6, we may find a probability space $(\Omega, F, \mathbb{P})$ supporting $\mathcal{X}$-valued random variables X_n and X, with $X_n \sim \mu_n \circ$

f^{-1} and $X \sim \mu \circ f^{-1}$, such that $X_n \to X$ almost surely. Changing variables, our assumption reads

$$\lim_{r\to\infty} \sup_n \mathbb{E}[|X_n| 1_{\{|X_n|\geq r\}}] = 0,$$

where expectation is now on the space $(\Omega, \mathcal{F}, \mathbb{P})$. This means that $|X_n|$ are uniformly integrable. Since $X_n \to X$ almost surely, we conclude from the dominated convergence theorem that

$$\begin{aligned}\lim_{n\to\infty} \int_{\mathcal{X}} f \, d\mu_n &= \lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X] \\ &= \int_{\mathcal{X}} f \, d\mu.\end{aligned}$$

□

Remark 2.8. One should be careful in applying Skorokhod's Representation Theorem 2.6, which is a bit prone to misapplication. We can say nothing at all about the *joint* distribution of the resulting X_n's. If we are given a sequence of random variables (X_n) defined on a common probably space, then $X_n \Rightarrow X$ certainly does not imply $X_n \to X$ a.s.! Skorokhod's theorem simply says we can find, on some other probability space, random variables $Y_n \sim X_n$ and $Y \sim X$ such that $Y_n \to Y$ a.s. While Skorokhod's theorem is a convenient shortcut in some of the proofs we will see, it is good practice to find alternative proofs, from first principles.

2.1 Weak convergence of empirical measures

As a first exercise on weak convergence, we study the convergence of empirical measures. Suppose (X_i) are i.i.d. $\mathcal{X}$-valued random variables. Define the $\mathcal{P}(\mathcal{X})$-valued random variable

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}.$$

This random probability measure is called an *empirical measure*, and it is worth emphasizing that the integral of a test function takes the following form:

$$\int_{\mathcal{X}} f \, d\mu_n = \frac{1}{n} \sum_{i=1}^n f(X_i).$$

The law of large numbers implies that

$$\mathbb{P}\left(\lim_{n\to\infty}\int_{\mathcal{X}} f\,d\mu_n = \mathbb{E}[f(X_1)]\right) = 1, \quad \text{for every } f \in C_b(\mathcal{X}). \tag{2.1}$$

In fact, we can prove the

$$\mathbb{P}\left(\lim_{n\to\infty}\int_{\mathcal{X}} f\,d\mu_n = \mathbb{E}[f(X_1)] \quad \text{for every } f \in C_b(\mathcal{X})\right) = 1, \tag{2.2}$$

which is equivalent to saying that $\mu_n \to \mu$ weakly with probability 1. Of course, the difference between (2.1) and (2.2) is that we have exchanged the order of the quantifiers "for every f" and "with probability 1." To be completely clear, if we define an $N_f \subset \Omega$ by

$$N_f = \left\{\lim_{n\to\infty}\int_{\mathcal{X}} f\,d\mu_n = \mathbb{E}[f(X_1)]\right\},$$

for $f \in C_b(\mathcal{X})$, then (2.1) says that $\mathbb{P}(N_f) = 1$ for all $f \in C_b(\mathcal{X})$, whereas (2.2) says that $\mathbb{P}(\cup_{f\in C_b(\mathcal{X})} N_f) = 1$. It is not obvious that these are equivalent statements because the set $C_b(\mathcal{X})$ is uncountable, and, in general, uncountable union of probability-one events need not have probability one.

The complete proof of the following theorem would use somewhat heavy machinery from metric space theory, a bit out of character with the rest of the course. Instead, note that, in light of the above discussion, the proof is immediate if we take for granted one essential fact: For any separable metric space $(\mathcal{X}, d)$, there exists a *countable* family $(f_n) \subset C_b(\mathcal{X})$ such that $\mu_n \to \mu$ weakly if and only if $\int f_k\,d\mu_n \to \int f_k\,d\mu$ for every k. This is clear when $\mathcal{X}$ is compact, because then $C_b(\mathcal{X}) = C(\mathcal{X})$ is a separable Banach space when equipped with the supremum norm, but the general case takes some work; see [103, Theorem 6.6] for a proof. In any case, once we know this, the exchange of quantifiers above is straightforward.

Theorem 2.9. *If $(\mathcal{X}, d)$ is separable, then it holds with probability 1 that $\mu_n \to \mu$ weakly.*

2.2 Wasserstein metrics

We next want to study how to place a metric on $\mathcal{P}(\mathcal{X})$ which is compatible with weak convergence, so that we can view $\mathcal{P}(\mathcal{X})$ itself as a metric space. There are many choices, the most common of which are known as the Lévy-Prokhorov metric, the bounded-Lipschitz metric, or the Wasserstein metrics. We will work mostly with the latter.

For technical reasons, we will henceforth assume the metric space $(\mathcal{X}, d)$ is separable. The main reason for this is as follows. Unless $\mathcal{X}$ is separable, the Borel σ-field of the product space $\mathcal{X} \times \mathcal{X}$ (equipped with the usual product topology) is different from the product σ-field generated by $\mathcal{X}$. More precisely, if $\mathcal{B}_{\mathcal{Y}}$ denotes the Borel σ-field of a metric space $\mathcal{Y}$, then it is well known that $\mathcal{B}_{\mathcal{X}\times\mathcal{X}} = \mathcal{B}_{\mathcal{X}} \otimes \mathcal{B}_{\mathcal{X}}$ holds if $\mathcal{X}$ is separable, but this may fail otherwise. It particular, the metric d (viewed as a function from $\mathcal{X} \times \mathcal{X}$ to $\mathbb{R}$) may not be measurable with respect to the product σ-field $\mathcal{B}_{\mathcal{X}} \otimes \mathcal{B}_{\mathcal{X}}$ if $\mathcal{X}$ is not separable! As all of the spaces we encounter in this course will be separable, we impose this assumption throughout.

The definition of the Wasserstein metric is based on the idea of a *coupling*. For $\mu, \nu \in \mathcal{P}(\mathcal{X})$, we write $\Pi(\mu, \nu)$ to denote the set of Borel probability measures π on $\mathcal{X} \times \mathcal{X}$ with first marginal μ and second marginal ν. Precisely, $\pi(A \times \mathcal{X}) = \mu(A)$ and $\pi(\mathcal{X} \times A) = \nu(A)$ for every Borel set $A \subset \mathcal{X}$.

The Wasserstein metric also requires integrating unbounded functions, so we must restrict the space on which it is defined. For $p \geq 1$, define $\mathcal{P}^p(\mathcal{X})$ to be the set of probability measures $\mu \in \mathcal{P}(\mathcal{X})$ satisfying

$$\int_{\mathcal{X}} d(x, x_0)^p \mu(dx) < \infty,$$

where $x_0 \in \mathcal{X}$ is an arbitrary reference point. (By the triangle inequality, the choice of x_0 is inconsequential.)

Definition 2.10. The p-Wasserstein metric on $\mathcal{P}^p(\mathcal{X})$ is defined by

$$\mathcal{W}_{\mathcal{X},p}(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} d(x, y)^p \, \pi(dx, dy) \right)^{1/p}.$$

If the space $\mathcal{X}$ is understood, we write simply $\mathcal{W}_p$ instead of $\mathcal{W}_{\mathcal{X},p}$.

An equivalent and more probabilistic definition reads

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{X \sim \mu,\, Y \sim \nu} \mathbb{E}[d(X, Y)^p] \right)^{1/p},$$

where the infimum is over all pairs of $\mathcal{X}$-valued random variables X and Y with given marginals μ and ν. The Wasserstein metric is very convenient in that it involves an infimum, which makes it quite easy to bound. That is, for any choice of coupling, we have an upper bound on the Wasserstein metric. We will see this much more clearly when we begin to study particle systems.

We will not prove that $\mathcal{W}_p$ is a metric; refer to Villani [115, Theorem 7.3] and Bolley [14] for proof of the following:

Theorem 2.11. *If $(\mathcal{X}, d)$ is complete and separable, then $\mathcal{W}_p$ defines a metric on $\mathcal{P}^p(\mathcal{X})$. Moreover, $(\mathcal{P}^p(\mathcal{X}), \mathcal{W}_p)$ is itself a complete and separable metric space.*

It is worth noting that the minimization problem appearing in the definition of the Wasserstein distance is an example of an *optimal transport problem*. The theory of optimal transport is rich and beyond the scope of this course, but the interested reader is referred to the excellent book of Villani for [115] for a careful introduction.

Remark 2.12. It is important to note that Jensen's inequality implies $\mathcal{W}_p \le \mathcal{W}_q$ whenever $q \ge p$, and of course $\mathcal{P}^p(\mathcal{X}) \subset \mathcal{P}^q(\mathcal{X})$ as well. This means that the metric $\mathcal{W}_q$ generates a *finer* topology. If $\mathcal{W}_q(\mu_n, \mu) \to 0$, then $\mathcal{W}_p(\mu_n, \mu) \to 0$, but the converse is not necessarily true. (Exercise: Find an example.) On the other hand, if a function $F : \mathcal{P}^q(\mathcal{X}) \to \mathbb{R}$ is continuous with respect to the metric $\mathcal{W}_p$, then it is also continuous with respect to $\mathcal{W}_q$.

We summarize one important fact about Wasserstein convergence in the following theorem. Essentially, it shows that convergence in p-Wasserstein distance is the same as weak convergence plus convergence of p^{th} order moments. Alternatively, we can characterize p-Wasserstein convergence in terms of the convergence of integrals $\int f \, d\mu_n \to \int f \, d\mu$, but for a larger class of test functions than simply $C_b(\mathcal{X})$.

Theorem 2.13. *Let $\mu, \mu_n \in \mathcal{P}^p(\mathcal{X})$ for some $p \ge 1$. The following are equivalent:*

(i) $\mathcal{W}_p(\mu_n, \mu) \to 0$.

(ii) For every continuous function $f : \mathcal{X} \to \mathbb{R}$ with the property that there exist $x_0 \in \mathcal{X}$ and $c > 0$ such that $|f(x)| \le c(1 + d(x, x_0)^p)$ for all $x \in \mathcal{X}$, we have

$$\int_{\mathcal{X}} f \, d\mu_n \to \int_{\mathcal{X}} f \, d\mu.$$

(iii) $\mu_n \to \mu$ weakly and $\int_{\mathcal{X}} d(x, x_0)^p \, \mu_n(dx) \to \int_{\mathcal{X}} d(x, x_0)^p \mu(dx)$ for some $x_0 \in \mathcal{X}$.

(iv) $\mu_n \to \mu$ weakly and

$$\lim_{r \to \infty} \sup_n \int_{\{d(\cdot, x_0) \ge r\}} d(x, x_0)^p \, \mu_n(dx) = 0.$$

See [115, Theorem 7.12] for a proof. These various characterizations are often more convenient to work with than Wasserstein-convergence itself. For instance, we can quickly prove the following general fact, which we will use in our study of particle systems:

Corollary 2.14. *Suppose $\mathcal{X}$ is separable. Suppose (X_i) are i.i.d. $\mathcal{X}$-valued random variables with law μ, and let $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ denote the empirical measure. Let $p \geq 1$. If $\mu \in \mathcal{P}^p(\mathcal{X})$, then $\mathcal{W}_p(\mu_n, \mu) \to 0$ almost surely, and also*

$$\mathbb{E}[\mathcal{W}_p^p(\mu_n, \mu)] \to 0.$$

Proof. We know from Theorem 2.9 that $\mu_n \to \mu$ weakly, with probability 1. On the other hand, since $\mu \in \mathcal{P}^p(\mathcal{X})$, we have $\mathbb{E}[d(X_1, x_0)^p] < \infty$ for any $x_0 \in \mathcal{X}$, and the law of large numbers implies that

$$\begin{aligned}
\lim_{n\to\infty} \int_{\mathcal{X}} d(x, x_0)^p \, \mu_n(dx) &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n d(X_i, x_0)^p \\
&= \mathbb{E}[d(X_1, x_0)^p] \\
&= \int_{\mathcal{X}} d(x, x_0)^p \, \mu(dx), \quad a.s.
\end{aligned}$$

The claimed almost sure convergence now follows from the implication $(iii) \Rightarrow (i)$ of Theorem 2.13.

To prove the second claim, we will apply dominated convergence. To do this, we fix $x_0 \in \mathcal{X}$ and use the triangle inequality and the elementary inequality $(a+b)^p \leq 2^{p-1}(a^p + b^p)$ to estimate

$$\begin{aligned}
\mathcal{W}_p^p(\mu_n, \mu) &\leq 2^{p-1}\mathcal{W}_p^p(\mu_n, \delta_{x_0}) + 2^{p-1}\mathcal{W}_p^p(\delta_{x_0}, \mu) \\
&= \frac{2^{p-1}}{n}\sum_{i=1}^n d(X_i, x_0)^p + 2^{p-1}\int_{\mathcal{X}} d(x, x_0)^p \, \mu(dx).
\end{aligned}$$

The second term is finite by assumption, so we need only to show the first term is uniformly integrable. This follows immediately from the general fact that if Z_i are i.i.d. nonnegative real-valued random variables with $\mathbb{E}[Z_1] < \infty$ then the partial averages $S_n = \frac{1}{n}\sum_{i=1}^n Z_i$ are uniformly integrable. This is perhaps most efficiently argued using the criterion of de la Vallée Poussin: Integrability of Z_1 implies that there exists a convex increasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{t\to\infty} \psi(t)/t = \infty$ such that $\mathbb{E}[\psi(Z_1)] < \infty$. Then, convexity implies

$$\sup_n \mathbb{E}[\psi(S_n)] \leq \sup_n \frac{1}{n}\sum_{i=1}^n \mathbb{E}[\psi(Z_i)] = \mathbb{E}[\psi(Z_1)] < \infty,$$

which yields the uniform integrability of S_n again by the criterion of de la Vallée Poussin. $\square$

While Wasserstein metrics do not precisely metrize weak convergence, it is occasionally useful to note that they can be forced to by replacing the metric on $\mathcal{X}$ with an equivalent bounded metric. That is, if $(\mathcal{X}, d)$ is a separable metric space, then we can define a new metric

$$\overline{d}(x,y) := 1 \wedge d(x,y) = \min\{1, d(x,y)\}, \quad \text{for } x, y \in \mathcal{X}.$$

Then $\overline{d}$ is also a metric on $\mathcal{X}$ generating the same topology. Notice that weak convergence of probability measures on $\mathcal{X}$ depends on the topology of $\mathcal{X}$ but not the particular metric! On the other hand, the Wasserstein metric explicitly involves the choice of metric on $\mathcal{X}$. Regardless, we may define the p-Wasserstein metric on $\mathcal{P}(\mathcal{X})$ relative to $\overline{d}$ instead of d:

$$\overline{\mathcal{W}}_p(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{X}} [1 \wedge d(x,y)]^p \, \pi(dx, dy) \right)^{1/p}.$$

Now notice that $x \mapsto \overline{d}(x, x_0)$ is a bounded continuous function on $\mathcal{X}$. Hence, using the equivalence $(i) \Leftrightarrow (iii)$, we conclude that $\overline{\mathcal{W}}_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \to \mu$ weakly. In summary, if we work with a bounded metric on $\mathcal{X}$, then the Wasserstein distance (of any order p) provides a metric on $\mathcal{P}(\mathcal{X})$ which is compatible with weak convergence.

2.3 Kantorovich duality

One cannot in good faith discuss Wasserstein metrics without discussing the fundamental duality theorem, originally due to Kantorovich. We will not make much use for this in the course, but nonetheless is very often useful when working with Wasserstein metrics or optimal transport problems more generally. It takes its simplest form for the 1-Wasserstein metric, worth singling out. The proof ultimately boils down to the Fenchel-Rockafellar theorem, but we will not go into this; see [115, Theorems 1.3 and 1.14].

Theorem 2.15. *Suppose $(\mathcal{X}, d)$ is a complete and separable metric space, and define $\mathcal{W}_p$ for $p \geq 1$ as in Definition 2.10. Then, for any $\mu, \nu \in \mathcal{P}^p(\mathcal{X})$,*

$$\mathcal{W}_p^p(\mu,\nu) = \sup\left\{ \int f \, d\mu + \int g \, d\nu : f, g \in C_b(E),\ f(x) + g(y) \leq d(x,y)^p \ \forall x, y \in \mathcal{X} \right\}.$$

Moreover, for $p = 1$ and $\mu, \nu \in \mathcal{P}^1(\mathcal{X})$, we have

$$\mathcal{W}_1(\mu, \nu) = \sup_f \left(\int f \, d\mu - \int f \, d\nu \right), \tag{2.3}$$

where the supremum is over all functions $f : \mathcal{X} \to \mathbb{R}$ which are 1-Lipschitz in the sense that $|f(x) - f(y)| \le d(x, y)$ for all $x, y \in \mathcal{X}$.

As is typical in duality theory, one inequality (known as weak duality) is easy to prove. For instance, in the $p = 1$ case, if $\pi \in \Pi(\mu, \nu)$ is any coupling and $f : \mathcal{X} \to \mathbb{R}$ is any 1-Lipschitz function, then

$$\begin{aligned} \int_{\mathcal{X}} f \, d\mu - \int_{\mathcal{X}} f \, d\nu &= \int_{\mathcal{X} \times \mathcal{X}} (f(x) - f(y)) \, \pi(dx, dy) \\ &\le \int_{\mathcal{X} \times \mathcal{X}} d(x, y) \, \pi(dx, dy). \end{aligned}$$

Take the supremum over f on the left-hand side and the infimum over $\pi \in \Pi(\mu, \nu)$ on the right-hand side to get the inequality ($\ge$) in (2.3).

2.4 Interaction functions

When we turn to our study of mean field games and interacting particle systems, our models will involve functions defined on $\mathcal{X} \times \mathcal{P}^p(\mathcal{X})$. We will think of such a function $F = F(x, \mu)$ as determining an *interaction* of a particle x with a distribution of particles μ. It will be important to understand examples and continuity properties of such functions. This section catalogs some examples. Throughout, we assume $(\mathcal{X}, d)$ is a complete and separable metric space.

Example 2.16. Suppose $f : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is jointly continuous. Suppose there exist $c > 0$, $x_0 \in \mathcal{X}$, $y_0 \in \mathcal{X}$, and a continuous function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$|f(x, y)| \le c(1 + \omega(d(x, x_0)) + d(y, y_0)^p)$$

Define $F : \mathcal{X} \times \mathcal{P}^p(\mathcal{X}) \to \mathbb{R}$ by

$$F(x, \mu) = \int_{\mathcal{X}} f(x, y) \, \mu(dy).$$

Then F is well-defined and jointly continuous. To see this, first note assumption (2.16) ensures that

$$\int_{\mathcal{X}} |f(x, y)| \mu(dy) \le c \left(1 + \omega(d(x, x_0)) + \int_{\mathcal{X}} d(y, y_0)^p \, \mu(dy) \right) < \infty,$$

for every $\mu \in \mathcal{P}^p(\mathcal{X})$, which shows that F is indeed well-defined for $\mu \in \mathcal{P}^p(\mathcal{X})$. To prove continuity, suppose $\mu_n \to \mu$ in $\mathcal{P}^p(\mathcal{X})$. Since $\mu_n \to \mu$ weakly, the Skorokhod representation Theorem 2.6 implies that there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting $\mathcal{X}$-valued random variables $Y_n \sim \mu_n$ and $Y \sim \mu$ such that $Y_n \to Y$ a.s. Now, for each n, we have

$$f(x_n, Y_n) \leq c \left(1 + \sup_k \omega(d(x_k, x_0)) + d(Y_n, y_0)^p \right) =: Z_n.$$

Thanks to Theorem 2.13(iv), we know that

$$\lim_{r\to\infty} \sup_n \mathbb{E}[1_{\{d(Y_n,y_0)\geq r\}} d(Y_n, y_0)^p] = 0,$$

and thus

$$\lim_{r\to\infty} \sup_n \mathbb{E}[1_{\{Z_n \geq r\}} Z_n] = 0.$$

In other words, (Z_n) are uniformly integrable, and thus so are $f(x_n, Y_n)$. Since $f(x_n, Y_n)$ converges a.s. to $f(x, Y)$, we conclude from dominated convergence that

$$\begin{aligned}\lim_{n\to\infty} \int_{\mathcal{X}} f(x_n, y)\, \mu_n(dy) &= \lim_{n\to\infty} \mathbb{E}[f(x_n, Y_n)] = \mathbb{E}[f(x, Y)] \\ &= \int_{\mathcal{X}} f(x, y)\, \mu(dy).\end{aligned}$$

Example 2.17 (Convolution)**.** A common special case of Example 2.16 is when the function f is of the form $f(x, y) = h(d(x, y))$, for some continuous function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $|h(t)| \leq c(1 + t^p)$ for all $t \geq 0$, for some $c > 0$. Even more specifically, if $\mathcal{X} = \mathbb{R}^d$ is a Euclidean space we often take $f(x, y) = h(x - y)$. In this case,

$$F(x, \mu) = \int_{\mathcal{X}} h(x - y)\mu(dy)$$

is precisely the convolution of the function h with the probability measure μ, often written as $F(x, \mu) = h * \mu(x)$.

Example 2.18 (Rank-based interaction)**.** Suppose $\mathcal{X} = \mathbb{R}$, and define F on $\mathbb{R} \times \mathcal{P}(\mathbb{R})$ by

$$F(x, \mu) = G(\mu(-\infty, x]),$$

for some function $G : [0,1] \to \mathbb{R}$. In other words, $F(x,\mu) = G(\mathbb{P}(X \le x))$ if X is a random variable with law μ. This kind of function models a *rank-based interaction*, in the following sense: The value $\mu(-\infty, x]$ gives the mass assigned below x by the distribution μ, which can be seen as the "rank" of x in the distribution μ. This is most clear when evaluated at empirical measures. If $x_1, \ldots, x_n \in \mathbb{R}$ then, for $k = 1, \ldots, n$, we have

$$\begin{aligned} F\left(x_k, \frac{1}{n}\sum_{i=1}^n \delta_{x_i}\right) &= G\left(\frac{1}{n}\sum_{i=1}^n 1_{\{x_i \le x_k\}}\right) \\ &= G\left(\frac{1}{n}\#\left\{i \in \{1,\ldots,n\} : x_i \le x_k\right\}\right), \end{aligned}$$

where $\#A$ denotes the cardinality of a set A. In other words, inside of G we see $1/n$ times the number of x_i's below x_k, which is naturally interpreted as the "rank" of x_k among the vector $(x_1, \ldots, x_n)$. Unfortunately, rank-based interactions of this form are not continuous on all of $\mathbb{R} \times \mathcal{P}(\mathbb{R})$, as we have already seen that typically functions of the form $\mu \mapsto \mu(A)$ are not weakly continuous for sets A (see the Portmanteau theorem 2.3).

Example 2.19 (Local interactions)**.** Suppose $\mathcal{X} = \mathbb{R}^d$, and let $\mathcal{A}^d \subset \mathcal{P}(\mathbb{R}^d)$ denote the set of probability measures on $\mathbb{R}^d$ which are absolutely continuous with respect to Lebesgue measure. In particular, for $\mu \in \mathcal{A}^d$, we may write $\mu(dx) = f_\mu(x)dx$ for some measurable nonnegative function f_μ on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} f_\mu(x)dx = 1$. A *local interaction* is a function $F : \mathbb{R}^d \times \mathcal{A}^d \to \mathbb{R}$ of the form

$$F(x,\mu) = G(f_\mu(x)),$$

for some function G on $\mathbb{R}$. The term "local" comes from the fact that the value $F(x,\mu)$ depends on the measure μ only through its infinitesimal mass around the point x, as captured by the density. These kinds of functions fail miserably to be continuous, and one needs to be careful even in the definition, since the density $f_\mu(x)$ is only defined uniquely up to almost-everywhere equality.

Example 2.20 (Geometric mean)**.** The following example appears in a model of competitive optimal investment from [85], which we will study later in the course as time permits. Given a family of positive real numbers $x_1, \ldots, x_n$, the geometric mean is defined as

$$\left(\prod_{i=1}^n x_i\right)^{1/n} = \exp\left(\frac{1}{n}\log\sum_{i=1}^n x_i\right).$$

This extends naturally to a general probability measure $\mu \in \mathcal{P}((0,\infty))$ by setting

$$G(\mu) = \exp\left(\int_{(0,\infty)} \log x \, \mu(dx)\right),$$

provided the integral is well-defined in the sense that $\int |\log x|\mu(dx) < \infty$. This function G is not continuous with respect to any Wasserstein metric, unless we restrict to a subset of measures for which $|\log x|$ is uniformly integrable. This kind of example motivates the study of more general kinds of topologies on subsets of $\mathcal{P}(\mathcal{X})$, induced by *gauge functions*. To wit, if $\psi : \mathcal{X} \to \mathbb{R}_+$ is a continuous function, we may let $\mathcal{P}^\psi(\mathcal{X})$ denote the set of $\mu \in \mathcal{P}(\mathcal{X})$ for which $\int \psi \, d\mu < \infty$, and the relevant topology on $\mathcal{P}^\psi(\mathcal{X})$ is the coarsest one for which the map $\mu \mapsto \int f \, d\mu$ is continuous for every continuous function $f : \mathcal{X} \to \mathbb{R}$ satisfying $|f(x)| \le 1 + \psi(x)$ for all $x \in \mathcal{X}$. In light of Theorem 2.13(ii), we know that if $\psi(x) = d(x, x_0)^p$ for some $x_0 \in \mathcal{X}$ and some $p \ge 1$, then $\mathcal{P}^\psi(\mathcal{X}) = \mathcal{P}^p(\mathcal{X})$, and this topology is precisely the one induced by the p-Wasserstein metric. We will not make use of these ideas, but see [52, Appendix A.6] for a systematic study.

Example 2.21 (Quantile interactions)**.** Define the quantile function of a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ at a point $x \in \mathbb{R}^d$ by

$$R(\mu, x, u) = \inf\{r > 0 : \mu(B(x,r)) \ge u\},$$

where $B(x,r)$ is the closed ball of radius r centered at x. Think of x as the location of a bird in a flock and μ as a distribution of bird locations. The value $R(\mu, x, u)$ represents the minimal radius r for which a fraction of at least u of the birds (distributed according to μ) lies within r of the point x. Fix a number $\alpha \in (0,1)$, and define a function F on $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ by

$$F(x,\mu) = \frac{1}{\mu(B(x, R(\mu, x, \alpha)))} \int_{B(x, R(\mu,x,\alpha))} f(x,y) \, \mu(dy),$$

for some bounded continuous function f on $\mathbb{R}^d \times \mathbb{R}^d$. This is a very similar interaction to what we saw in Example 2.16, but now the bird at position x only interacts with the α percent of birds which are closest to it. In dimension $d = 1$, note that the function R is closely related to the quantile function. In turns out that R (and thus, with some work, F) is continuous in μ when restricted to the subset of probability measures $\mathcal{A}^d$ defined in Example 2.19. This can be proven (exercise) using a remarkable result of R. Rao [105, Theorem 4.2]: If $\mu_n \to \mu$ weakly, and if μ is absolutely continuous with respect to Lebesgue measure, then $\mu_n(B) \to \mu(B)$ for every convex Borel set $B \subset \mathbb{R}^d$, *and the convergence is uniform over all such B.*

3 Interacting diffusions and McKean-Vlasov equations

Now that we understand the fundamentals of weak convergence and Wasserstein metrics, we begin our study of interacting particle systems. For a review of stochastic calculus, the reader is referred to the classic book of Karatzas and Shreve [76], or even the first chapter of Pham's book [104] which we will refer to later when we discuss stochastic control theory. Behind the scenes throughout this section is a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{t \ge 0}$ satisfying the usual conditions. This space should support (at the very least) an i.i.d. sequence (ξ^i) of $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random variables as well as a sequence (W^i) of independent $\mathbb{F}$-Brownian motions.

The main object of study will be a system of n interacting particles $(X_t^{n,1}, \ldots, X_t^{n,n})$, driven by stochastic differential equations (SDEs) of the form

$$\begin{aligned} dX_t^{n,i} &= b(X_t^{n,i}, \mu_t^n)dt + \sigma(X_t^{n,i}, \mu_t^n)dW_t^i, \quad X_0^{n,i} = \xi^i, \\ \mu_t^n &= \frac{1}{n}\sum_{k=1}^n \delta_{X_t^{n,k}}. \end{aligned} \tag{3.1}$$

Driving this SDE system are n independent Brownian motions, $W^1, \ldots, W^n$, and we typically assume the initial states $\xi^1, \ldots, \xi^n$ are i.i.d. We think of $X_t^{n,i}$ as the position of particle i at time t, in a Euclidean space $\mathbb{R}^d$.

We think of the number n of particles as very large, and ultimately we will send it to infinity. There is a key structural feature that makes this system amenable to *mean field* analysis: The coefficients b and σ are the same for each particle, and the only dependence of particle i on the rest of the particles $k \neq i$ is through the empirical measure μ_t^n. Let us build some intuition with a simple example:

3.1 A first example

In this section we study in more detail a warm-up model mentioned in the introduction. Consider the SDE system (3.1), with $d = 1$-dimensional particles, with the coefficients

$$b(x, \mu) = a\left(x - \int_{\mathbb{R}} y\, \mu(dy)\right), \qquad \sigma \equiv 1.$$

Here $a > 0$, and we can write more explicitly

$$\begin{aligned} dX_t^{n,i} &= a\left(\overline{X}_t^n - X_t^{n,i}\right) dt + dW_t^i, \quad i = 1, 2, \ldots, n, \\ \overline{X}_t^n &= \frac{1}{n}\sum_{k=1}^n X_t^{n,k}. \end{aligned} \tag{3.2}$$

The drift pushes each particle toward the empirical average $\overline{X}_t^n$. This is like an Ornstein-Uhlenbeck equation, but the target of mean-reversion is dynamic. To understand how the system behaves for large n, a good way to start is by noticing that if we average the n particles we get very simple dynamics for $\overline{X}_t^n$:

$$d\overline{X}_t^n = \frac{1}{n}\sum_{i=1}^n dW_t^i.$$

In particular,

$$\overline{X}_t^n = \overline{X}_0^n + \frac{1}{n}\sum_{i=1}^n W_t^i.$$

Sending $n \to \infty$, the average of the n Brownian motions vanishes thanks to the law of large numbers. Moreover, if the initial states $X_0^{n,i} = \xi^i$ are i.i.d., the empirical average $\overline{X}_0^n$ converges to the true mean $\mathbb{E}[\xi^1]$. Hence, when $n \to \infty$, the empirical average becomes $\lim_n \overline{X}_t^n = \mathbb{E}[\xi^1]$ for all t, almost surely. Plugging this back in to the original equation (3.2), we find that $\lim_n X^{n,i} = Y^i$ a.s., where

$$dY_t^i = a\left(\mathbb{E}[\xi^i] - Y_t^i\right) dt + dW_t^i, \quad Y_0^i = \xi^i. \tag{3.3}$$

We can solve this equation pretty easily. First, writing it in integral form, we have

$$Y_t^i = \xi^i + a\int_0^t \left(\mathbb{E}[\xi^i] - Y_s^i\right) ds + W_t^i.$$

Take expectations to get

$$\mathbb{E}[Y_t^i] = \mathbb{E}[\xi^i] + a\int_0^t \left(\mathbb{E}[\xi^i] - \mathbb{E}[Y_s^i]\right) ds.$$

Differentiate in t to get

$$\frac{d}{dt}\mathbb{E}[Y_t^i] = a\left(\mathbb{E}[\xi^i] - \mathbb{E}[Y_t^i]\right), \qquad \mathbb{E}[Y_0^i] = \mathbb{E}[\xi^i].$$

This shows that the function $t \mapsto \mathbb{E}[Y_t^i]$ solves a very simple ordinary differential equation, the solution of which is constant, $\mathbb{E}[Y_t^i] = \mathbb{E}[\xi^i]$. Hence, we may rewrite (3.3) as

$$dY_t^i = a\left(\mathbb{E}[Y_t^i] - Y_t^i\right) dt + dW_t^i, \quad Y_0^i = \xi^i. \tag{3.4}$$

This is our first example of a *McKean-Vlasov* equation, an SDE in which the coefficients depend on the law of the solution.

It important to observe that the resulting limiting processes $(Y^i)_{i\in\mathbb{N}}$ are i.i.d. They solve the same SDEs, driven by i.i.d. Brownian motions and i.i.d. initial states. This is a general phenomenon with McKean-Vlasov limits; the particles become asymptotically i.i.d. as $n \to \infty$, in a sense we will later make precise.

3.2 Deriving the McKean-Vlasov limit

While the example of Section 3.1 was simple enough to allow explicit computation, we now set to work on a general understanding of the $n \to \infty$ behavior of systems of the form (3.1). First, we should specify assumptions on the coefficients to let us ensure at the very least that the n-particle SDE system (3.1) is well-posed. In this section, we work with the following set of nice assumptions, recalling from Definition 2.10 the notation for Wasserstein metrics:

Assumption 3.1. Assume the initial states $(\xi^i)_{i\in\mathbb{N}}$ are i.i.d. with $\mathbb{E}[|\xi^1|^2] < \infty$. The Brownian motions $(W^i)_{i\in\mathbb{N}}$ are independent and m-dimensional. Assume $b : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}^{d\times m}$ are Lipschitz, in the sense that there exists a constant $L > 0$ such that

$$|b(x, m) - b(x', m')| + |\sigma(x, m) - \sigma(x', m')| \le L(|x - x'| + \mathcal{W}_2(m, m')). \tag{3.5}$$

Note that we always write $|\cdot|$ for the Euclidean norm on $\mathbb{R}^d$ and $|\cdot|$ for the Frobenius norm on $\mathbb{R}^{d\times m}$.

This assumption immediately lets us check that the SDE system (3.1) is well-posed. We make heavy use of the following well-known well-posedness result:

Lemma 3.2. *Under Assumption 3.1, the n-particle SDE system* (3.1) *admits a unique strong solution, for each n.*

Proof. We will fall back on Itô's classical existence and uniqueness result for Lipschitz SDEs. Define the $\mathbb{R}^{nd}$-valued process $\boldsymbol{X}_t = (X_t^{n,1}, \ldots, X_t^{n,n})$, and similar define the nm-dimensional Brownian motion $\boldsymbol{W}_t = (W_t^1, \ldots, W_t^n)$. We may write

$$d\boldsymbol{X}_t = B(\boldsymbol{X}_t)dt + \Sigma(\boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{3.6}$$

if we make the following definitions: $L_n : (\mathbb{R}^d)^n \to \mathcal{P}(\mathbb{R}^d)$ denotes the empirical measure map,

$$L_n(x_1, \ldots, x_n) = \frac{1}{n}\sum_{k=1}^n \delta_{x_k},$$

where we notice that the range of L_n actually lies in $\mathcal{P}^p(\mathbb{R}^d)$ for any exponent p. Define also $B : \mathbb{R}^{nd} \to \mathbb{R}^{nd}$ and $\Sigma : \mathbb{R}^{nd} \to \mathbb{R}^{nd} \times \mathbb{R}^{nm}$, for $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^{nd}$, by

$$B(\boldsymbol{x}) = \begin{pmatrix} b(x_1, L_n(\boldsymbol{x})) \\ b(x_2, L_n(\boldsymbol{x})) \\ \vdots \\ b(x_n, L_n(\boldsymbol{x})) \end{pmatrix},$$

$$\Sigma(\boldsymbol{x}) = \begin{pmatrix} \sigma(x_1, L_n(\boldsymbol{x})) & & & \\ & \sigma(x_2, L_n(\boldsymbol{x})) & & \\ & & \ddots & \\ & & & \sigma(x_n, L_n(\boldsymbol{x})) \end{pmatrix},$$

where Σ contains all zeros except for $d \times m$ blocks on the main diagonal. Then, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{nd}$, we have

$$\begin{aligned} |B(\boldsymbol{x}) - B(\boldsymbol{y})|^2 &= \sum_{i=1}^n |b(x_i, L_n(\boldsymbol{x})) - b(y_i, L_n(\boldsymbol{y}))|^2 \\ &\le 2L^2 \sum_{i=1}^n \left(|x_i - y_i|^2 + \mathcal{W}_2^2(L_n(\boldsymbol{x}), L_n(\boldsymbol{y}))\right) \\ &= 2L^2|\boldsymbol{x} - \boldsymbol{y}|^2 + 2L^2 n \mathcal{W}_2^2(L_n(\boldsymbol{x}), L_n(\boldsymbol{y})), \end{aligned}$$

where the second line used the elementary inequality $(a+b)^2 \le 2a^2 + 2b^2$. Next, to bound the Wasserstein distance between empirical measures, we use the most natural coupling; namely, the empirical measure

$$\pi = \frac{1}{n}\sum_{i=1}^n \delta_{(x_i, y_i)}$$

is a coupling of $L_n(\boldsymbol{x})$ and $L_n(\boldsymbol{y})$. Hence,

$$\begin{aligned}\mathcal{W}_2^2(L_n(\boldsymbol{x}), L_n(\boldsymbol{y})) &\le \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \pi(dx,dy) \\ &= \frac{1}{n}\sum_{i=1}^n |x_i - y_i|^2 \\ &= \frac{1}{n}|\boldsymbol{x}-\boldsymbol{y}|^2.\end{aligned}$$

We conclude that $|B(\boldsymbol{x}) - B(\boldsymbol{y})|^2 \le 4L^2|\boldsymbol{x}-\boldsymbol{y}|^2$, which means that B is $2L$-Lipschitz. A similar estimate is available for Σ. Conclude that the SDE (3.6), which was simply a rewriting of our original SDE system, has a unique strong solution. $\square$

Now, to send $n \to \infty$ in our particle system, we start by "guessing" what the limit should look like. As $n \to \infty$, the interaction becomes weaker and weaker in the sense that the contribution of a given particle i is of order $1/n$, intuitively. Hence, as $n \to \infty$, we expect the interaction to vanish, in the sense that a given particle i does not appear in the measure term anymore. If particles do not affect the measure flow, then the particles should be i.i.d., as they have the same coefficients b and σ and are driven by independent Brownian motions. For a better, more mathematical derivation, see Section 3.4.

For now, the above guess leads us to expect that if $\mu_t^n \to \mu_t$ in some sense, where μ_t is a *deterministic* measure flow (i.e., a function $\mathbb{R}_+ \ni t \mapsto \mu_t \in \mathcal{P}^2(\mathbb{R}^d)$), then the dynamics of any particle particle will look like

$$dY_t^i = b(Y_t^i, \mu_t)dt + \sigma(Y_t^i, \mu_t)dW_t^i.$$

These particles are i.i.d., but μ_t should still somehow represent their distribution. We know that the empirical measure of i.i.d. samples converges to the true distribution (see Theorem 2.9), so we should expect that μ_t is actually the law of Y_t^i, for any i. In other words, the law of the solution shows up in the coefficients of the SDE! We call this a *McKean-Vlasov equation*, after the seminal work of McKean [94, 95].

To formulate the McKean-Vlasov equation more precisely, it is convenient and often more general to lift the discussion to the path space, in the following sense. For convenience, we will fix a time horizon $T > 0$, but it is straightforward to extend much of the discussion to follow to the infinite time interval. Let $\mathcal{C}^d = C([0,T];\mathbb{R}^d)$ denote the set of continuous $\mathbb{R}^d$-valued functions of time, equipped with the supremum norm

$$\|x\| = \sup_{t \in [0,T]} |x_t|$$

and the corresponding Borel σ-field. Rather than work with measure flows, as elements of $C([0,T];\mathcal{P}^2(\mathbb{R}^d))$, we will work with probability measures on $\mathcal{C}^d$. There is a natural surjection

$$\mathcal{P}^2(\mathcal{C}^d) \ni \mu \to (\mu_t)_{t \in [0,T]} \in C([0,T];\mathcal{P}^2(\mathbb{R}^d)),$$

where μ_t is defined as the image of the measure μ through the map $\mathcal{C}^d \ni x \mapsto x_t \in \mathbb{R}^d$. In other words, if $\mu = \mathcal{L}(X)$ is the law of a $\mathcal{C}^d$-valued random variable, then $\mu_t = \mathcal{L}(X_t)$ is the time-t marginal. Note that this surjection is continuous and, in fact, 1-Lipschitz, if $C([0,T];\mathcal{P}^2(\mathbb{R}^d))$ is endowed with the supremum distance induced by the metric $\mathcal{W}_2$ on $\mathbb{R}^d$; explicitly, we have

$$\sup_{t \in [0,T]} \mathcal{W}_{\mathbb{R}^d,2}(\mu_t, \nu_t) \le \mathcal{W}_{\mathcal{C}^d,2}(\mu, \nu) \tag{3.7}$$

for every $\mu, \nu \in \mathcal{P}^2(\mathcal{C}^d)$.

The McKean-Vlasov equation is defined precisely as follows:

$$\begin{aligned} dY_t &= b(Y_t, \mu_t)dt + \sigma(Y_t, \mu_t)dW_t, \quad t \in [0,T], \quad Y_0 = \xi, \\ \mu &= \mathcal{L}(Y), \quad \forall t \ge 0. \end{aligned} \tag{3.8}$$

Here we write $\mathcal{L}(Z)$ for the law or distribution of a random variable Z. Here W is a Brownian motion, ξ is an $\mathbb{R}^d$-valued random variable with the same law as ξ^i, and both are (say) supported on the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Of course, ξ should be $\mathcal{F}_0$-measurable, and W should be an $\mathbb{F}$-Brownian motion. A strong solution of the McKean-Vlasov equation is a pair (Y, μ), where Y is a continuous $\mathbb{F}$-adapted $\mathbb{R}^d$-valued process (alternatively, a $C([0,T];\mathbb{R}^d)$-valued random variable), and μ is a probability measure on $C([0,T];\mathbb{R}^d)$, such that both equations (3.8) hold simultaneously. More compactly, one could simply write

$$dY_t = b(Y_t, \mathcal{L}(Y_t))dt + \sigma(Y_t, \mathcal{L}(Y_t))dW_t, \quad t \in [0,T], \quad Y_0 = \xi.$$

Under the Lipschitz assumption, we can always uniquely solve for Y if we know μ, and so we sometimes refer to μ itself (instead of the pair (Y, μ)) as the solution of the McKean-Vlasov equation.

We formalize all of this in the following theorem. We will make use of the 2-Wasserstein distance $\mathcal{W}_2 = \mathcal{W}_{\mathcal{C}^d,2}$, which we recall is defined by

$$\mathcal{W}_2^2(m, m') = \inf_{\pi \in \Pi(m,m')} \int_{\mathcal{C}^d \times \mathcal{C}^d} \|x - y\|^2 \, \pi(dx, dy).$$

In the n-particle system (3.1), let

$$\mu^n = \frac{1}{n} \sum_{i=1}^n \delta_{X^{n,i}}$$

denote the lifted empirical measure, i.e., the random probability measure on $\mathcal{C}^d$ obtained as the empirical measure of the n particle trajectories.

Theorem 3.3. *Suppose Assumption 3.1 holds. There exists a unique strong solution of the McKean-Vlasov equation* (3.8)*. Moreover, the n-particle system converges in the following two senses. First,*

$$\lim_{n\to\infty} \mathbb{E}\left[\mathcal{W}_{\mathcal{C}^d,2}^2(\mu^n, \mu)\right] = 0. \tag{3.9}$$

Second, for a fixed $k \in \mathbb{N}$, we have the weak convergence

$$(X^{n,1}, \ldots, X^{n,k}) \Rightarrow (Y^1, \ldots, Y^k), \tag{3.10}$$

where $Y^1, \ldots, Y^k$ are independent copies of the solution of the McKean-Vlasov equation.

We interpret the second form of the limit (3.10) as saying that the particles $X^{n,i}$ become asymptotically i.i.d. as $n \to \infty$. Indeed, for any fixed k, the first k particles converge in distribution to i.i.d. random variables. The choice of the "first k" particles here is inconsequential, in light of the fact that the n-particle system (3.1) is *exchangeable*, in the sense that

$$(X^{n,\pi(1)}, \ldots, X^{n,\pi(n)}) \stackrel{d}{=} (X^{n,1}, \ldots, X^{n,n})$$

for any permutation π of $\{1, \ldots, n\}$. (Here $\stackrel{d}{=}$ means equal in distribution.) It is left as an exercise for the reader to prove this claimed exchangeability, using the uniqueness of the solution of the SDE system (3.1).

Remark 3.4. The two kinds of limits described in Theorem 3.3 are sometimes referred to as *propagation of chaos*, though this terminology has somewhat lost its original meaning. One would say that propagation of chaos holds for the interacting particle system (3.1) if the following holds: For any $m_0 \in \mathcal{P}(\mathbb{R}^d)$ and any choice of deterministic initial states $(X_0^{n,i})$ satisfying

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{X_0^{n,i}} \to m_0, \tag{3.11}$$

we have the weak limit $\frac{1}{n}\sum_{i=1}^{n} \delta_{X_t^{n,i}} \to \mu_t$ in probability in $\mathcal{P}(\mathbb{R}^d)$, where (Y, μ) is the solution of the McKean-Vlasov equation (3.8) with initial state $\xi \sim m_0$. Initial states which converge weakly as in (3.11) are called "m_0-chaotic," and the term *propagation of chaos* means that the "chaoticity" of the initial distributions "propagates" to later times $t > 0$.

We break up the proof of Theorem 3.3 into a two major steps. First, we show existence and uniqueness:

Proof of existence and uniqueness

Define the truncated supremum norm

$$\|x\|_t := \sup_{0 \le s \le t} |x_s|$$

for $x \in \mathcal{C}^d$ and $t \in [0, T]$. Using this, define the truncated Wasserstein distance on $\mathcal{P}^2(\mathcal{C}^d)$ by

$$d_t^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{C}^d \times \mathcal{C}^d} \|x - y\|_t^2 \, \pi(dx, dy). \tag{3.12}$$

Notice that for any fixed $\mu \in \mathcal{P}^2(\mathcal{C}^d)$ we have the Lipschitz condition

$$|b(x, \mu_t) - b(y, \mu_t)| + |\sigma(x, \mu_t) - \sigma(y, \mu_t)| \le L|x - y|,$$

and thus classical theory ensures that there exists a unique (square-integrable) solution Y^μ of the SDE

$$dY_t^\mu = b(Y_t^\mu, \mu_t)dt + \sigma(Y_t^\mu, \mu_t)dW_t, \quad Y_0^\mu = \xi.$$

Define a map $\Phi : \mathcal{P}^2(\mathcal{C}^d) \to \mathcal{P}^2(\mathcal{C}^d)$ by setting $\Phi(\mu) = \mathcal{L}(Y^\mu)$. That is, for a fixed μ, we solve the above SDE, and set $\Phi(\mu)$ to be the law of the solution. Note that (Y, μ) is a solution of the McKean-Vlasov equation if and only if $Y = Y^\mu$ and $\mu = \Phi(\mu)$. That is, fixed points of Φ are precisely solutions of the McKean-Vlasov equation.

Let $\mu, \nu \in \mathcal{P}^2(\mathcal{C}^d)$, and use Jensen's inequality to get, for $t \in [0,T]$,

$$\begin{aligned} |Y_t^\mu - Y_t^\nu|^2 \leq 2t \int_0^t |b(Y_r^\mu, \mu_r) - b(Y_r^\nu, \nu_r)|^2 \, dr \\ + 2 \left| \int_0^t (\sigma(Y_r^\mu, \mu_r) - \sigma(Y_r^\nu, \nu_r)) \, dW_r \right|^2 . \end{aligned}$$

Take the supremum and use Doob's inequality followed by the Lipschitz assumption to find

$$\begin{aligned} \mathbb{E}\left[\|Y^\mu - Y^\nu\|_t^2\right] &\leq 2t\mathbb{E}\left[\int_0^t |b(Y_r^\mu, \mu_r) - b(Y_r^\nu, \nu_r)|^2 \, dr\right] \\ &\quad + 2\mathbb{E}\left[\sup_{0\leq s\leq t} \left| \int_0^s (\sigma(Y_r^\mu, \mu_r) - \sigma(Y_r^\nu, \nu_r)) \, dW_r \right|^2\right] \\ &\leq 2t\mathbb{E}\left[\int_0^t |b(Y_r^\mu, \mu_r) - b(Y_r^\nu, \nu_r)|^2 \, dr\right] \\ &\quad + 8\mathbb{E}\left[\int_0^t |\sigma(Y_r^\mu, \mu_r) - \sigma(Y_r^\nu, \nu_r)|^2 \, dr\right] \\ &\leq 2(8+2t)L^2\mathbb{E}\left[\int_0^t \left(|Y_r^\mu - Y_r^\nu|^2 + \mathcal{W}_2^2(\mu_r, \nu_r)\right) dr\right] \\ &\leq 2(8+2t)L^2\mathbb{E}\left[\int_0^t \left(\|Y^\mu - Y^\nu\|_r^2 + \mathcal{W}_2^2(\mu_r, \nu_r)\right) dr\right]. \end{aligned}$$

Use Fubini's theorem and Gronwall's inequality[3] to conclude that

$$\mathbb{E}\left[\|Y^\mu - Y^\nu\|_t^2\right] \leq C \int_0^t \mathcal{W}_2^2(\mu_r, \nu_r) dr,$$

for $t \in [0,T]$, where $C = 2(8+2T)L^2 \exp(2(8+2T)L^2)$. Use the inequality (3.7) to get

$$\mathbb{E}\left[\|Y^\mu - Y^\nu\|_t^2\right] \leq C \int_0^t d_r^2(\mu, \nu) dr.$$

[3]The form of Gronwall's inequality we use is as follows: If we are given a constant $c > 0$ and continuous nonnegative functions f and g such that

$$f(t) \leq g(t) + c\int_0^t f(r)dr, \quad \forall t \in [0,T],$$

then $f(t) \leq e^{ct} g(t)$ for all $t \in [0,T]$.

Finally, recall the definition of d_t^2, and notice that the joint distribution of (Y^μ, Y^ν) is a coupling of $(\Phi(\mu), \Phi(\nu))$. Hence,

$$d_t^2(\Phi(\mu), \Phi(\nu)) \leq \mathbb{E}\left[\|Y^\mu - Y^\nu\|_t^2\right] \leq C \int_0^t d_r^2(\mu, \nu) dr.$$

The proof of existence and uniqueness now follows from the usual Picard iteration argument. In particular, uniqueness follows from the above inequality and another application of Gronwall's inequality, whereas existence is derived by choosing arbitrarily $\mu^0 \in \mathcal{P}^2(\mathcal{C}^d)$, setting $\mu^{k+1} = \Phi(\mu^k)$ for $k \geq 0$, and showing using the above inequality that (μ^k) forms a Cauchy sequence whose limit must be a fixed point of Φ. $\square$

Proof of McKean-Vlasov limit

We now move to the second part of Theorem 3.3, proving the claimed limit theorem. The idea is what is by a coupling argument, where we construct i.i.d. copies of the unique solution of the McKean-Vlasov equation which has desirable joint-distributional properties with the original particle system. To do this, let μ denote the unique solution of the McKean-Vlasov equation (3.8). Using the same Brownian motions (W^i) and initial states (ξ^i) as our original particle system (3.1), define Y^i as the solution of the SDE

$$dY_t^i = b(Y_t^i, \mu_t)dt + \sigma(Y_t^i, \mu_t)dW_t^i, \quad Y_0^i = \xi^i.$$

Because the initial states and Brownian motions are i.i.d., so are (Y^i). We now want to estimate the difference $|X_t^{n,i} - Y_t^i|$. To do this, we proceed as in the previous step, starting with a fixed i and n:

$$\begin{aligned}|X_t^{n,i} - Y_t^i|^2 \leq 2t &\int_0^t \left|b(X_r^{n,i}, \mu_r^n) - b(Y_r^i, \mu_r)\right|^2 dr \\ &+ 2\left|\int_0^t \left(\sigma(X_r^{n,i}, \mu_r^n) - \sigma(Y_r^i, \mu_r)\right) dW_r^i\right|^2.\end{aligned}$$

Take the supremum and use Doob's inequality followed by the Lipschitz assumption to find

$$\begin{aligned}
\mathbb{E}\left[\|X^{n,i}-Y^i\|_t^2\right] &\le 2t\mathbb{E}\left[\int_0^t \left|b(X_r^{n,i},\mu_r^n)-b(Y_r^i,\mu_r)\right|^2 dr\right] \\
&\quad + 2\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \left(\sigma(X_r^{n,i},\mu_r^n)-\sigma(Y_r^i,\mu_r)\right) dW_r^i\right|^2\right] \\
&\le 2t\mathbb{E}\left[\int_0^t \left|b(X_r^{n,i},\mu_r^n)-b(Y_r^i,\mu_r)\right|^2 dr\right] \\
&\quad + 8\mathbb{E}\left[\int_0^t \left|\sigma(X_r^{n,i},\mu_r^n)-\sigma(Y_r^i,\mu_r)\right|^2 dr\right] \\
&\le 2(8+2t)L^2\mathbb{E}\left[\int_0^t \left(\|X^{n,i}-Y^i\|_r^2+\mathcal{W}_2^2(\mu_r^n,\mu_r)\right) dr\right].
\end{aligned}$$

Use Fubini's theorem, Gronwall's inequality, and the inequality (3.7) to get

$$\begin{aligned}
\mathbb{E}\left[\|X^{n,i}-Y^i\|_t^2\right] &\le C\mathbb{E}\left[\int_0^t \mathcal{W}_2^2(\mu_r^n,\mu_r)dr\right] \\
&\le C\mathbb{E}\left[\int_0^t d_r^2(\mu^n,\mu)dr\right],
\end{aligned} \tag{3.13}$$

for $t\in[0,T]$, where $C=2(8+2T)L^2\exp(2(8+2T)L^2)$, and where we recall the definition of the truncated Wasserstein distance from (3.12).

Define now the empirical measure of the Y^i's,

$$\nu^n=\frac{1}{n}\sum_{i=1}^n \delta_{Y^i}.$$

The empirical measure $\frac{1}{n}\sum_{i=1}^n \delta_{(X^{n,i},Y^i)}$ is a coupling of the (random) empirical measures μ^n and ν^n, and so

$$d_t^2(\mu^n,\nu^n)\le \frac{1}{n}\sum_{i=1}^n \|X^{n,i}-Y^i\|_t^2, \quad a.s.$$

Combine this with (3.13) to get

$$\mathbb{E}[d_t^2(\mu^n,\nu^n)]\le C\mathbb{E}\left[\int_0^t d_r^2(\mu^n,\mu)dr\right].$$

Use the triangle inequality and the previous inequality to get

$$\begin{aligned}\mathbb{E}[d_t^2(\mu^n,\mu)] &\le 2\mathbb{E}[d_t^2(\mu^n,\nu^n)] + 2\mathbb{E}[d_t^2(\nu^n,\mu)] \\ &\le 2C\mathbb{E}\left[\int_0^t d_r^2(\mu^n,\mu)dr\right] + 2\mathbb{E}[d_t^2(\nu^n,\mu)].\end{aligned}$$

Apply Gronwall's inequality once again to get

$$\mathbb{E}[d_t^2(\mu^n,\mu)] \le 2e^{2CT}\mathbb{E}\left[d_t^2(\nu^n,\mu)\right].$$

In particular, setting $t = T$, we have

$$\mathbb{E}[\mathcal{W}_2^2(\mu^n,\mu)] \le 2e^{2CT}\mathbb{E}[\mathcal{W}_2^2(\nu^n,\mu)].$$

But ν^n are the empirical measures of i.i.d. samples from the law μ, and the claimed limit (3.9) follows from the law of large numbers in the form of Corollary 2.14.

Finally, to prove the second claimed limit (3.10), we use (3.13) to find

$$\begin{aligned}\mathbb{E}\left[\max_{i=1,\dots,k}\|X^{n,i}-Y^i\|_t^2\right] &\le \sum_{i=1}^k \mathbb{E}\left[\|X^{n,i}-Y^i\|_t^2\right] \\ &\le Ck\mathbb{E}\left[\int_0^t d_r^2(\mu^n,\mu)dr\right] \\ &\le CkT\mathbb{E}[\mathcal{W}_2^2(\mu^n,\mu)].\end{aligned}$$

We just showed that this converges to zero, and the claimed convergence in distribution follows.

3.3 The PDE form of the McKean-Vlasov equation

In a first course on stochastic calculus, one encounters the Kolmogorov forward and backward equations, which describe the behavior of the distribution of the solution of an SDE. Both of these PDEs are linear in the unknown. A McKean-Vlasov equation like (3.8) admits a similar Kolmogorov forward-equation, but it is markedly different in that it is both nonlinear and nonlocal, in a sense we will soon make clear.

Suppose (Y,μ) solves the McKean-Vlasov equation (3.8). Apply Itô's formula to $\varphi(Y_t)$, where φ is a smooth function with compact support, to get

$$\begin{aligned}d\varphi(Y_t) = &\left(b(Y_t,\mu_t)\cdot\nabla\varphi(Y_t) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(Y_t,\mu_t)\nabla^2(Y_t)]\right)dt \\ &+ \nabla\varphi(Y_t)\cdot\sigma(Y_t,\mu_t)dW_t,\end{aligned}$$

where ∇ and ∇^2 denote the gradient and Hessian operators, respectively, and $\sigma\sigma^\top(x,m) = \sigma(x,m)\sigma(x,m)^\top$, where $^\top$ denotes the transpose of a matrix. Integrating this equation, taking expectations to kill the martingale term, and then differentiating, we find

$$\frac{d}{dt}\mathbb{E}[\varphi(Y_t)] = \mathbb{E}\left[b(Y_t,\mu_t)\cdot\nabla\varphi(Y_t) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(Y_t,\mu_t)\nabla^2(Y_t)]\right].$$

Now, we know that $Y_t \sim \mu_t$. Suppose in addition that μ_t has a density (with respect to Lebesgue measure), which we write as $\mu(t,x)$. Assume in addition that it admits one continuous derivative in t and two in x. We may then rewrite the above equation as

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^d}\varphi(x)\mu(t,x)dx \\
&\quad = \int_{\mathbb{R}^d}\left(b(x,\mu_t)\cdot\nabla\varphi(x) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x,\mu_t)\nabla^2\varphi(x)]\right)\mu(t,x)dx \qquad (3.14)\\
&\quad = \int_{\mathbb{R}^d}\left(-\mathrm{div}_x(b(x,\mu_t)\mu(t,x)) + \frac{1}{2}\mathrm{Tr}\Big[\nabla^2\big(\sigma\sigma^\top(x,\mu_t)\mu(t,x)\big)\Big]\right)\varphi(x)dx.
\end{aligned}$$

Because this must hold for every test function φ, we conclude (formally) that $\mu(t,x)$ solves the PDE

$$\partial_t\mu(t,x) = -\mathrm{div}_x(b(x,\mu_t)\mu(t,x)) + \frac{1}{2}\mathrm{Tr}\Big[\nabla^2\big(\sigma\sigma^\top(x,\mu_t)\mu(t,x)\big)\Big]. \qquad (3.15)$$

Notice that if b and σ did not depend on μ, this would be a usual (linear) Kolmogorov forward equation, also known as a Fokker-Planck equation. Note in addition that the nonlinear dependence on μ_t is also typically *nonlocal*, in the sense that $b = b(x,\mu_t)$ is a function not of the value $\mu(t,x)$ but of the entire distribution $(\mu(t,x))_{x\in\mathbb{R}^d}$; for instance, if $b(x,\mu_t)$ involves the mean of the measure μ_t, then the interaction is nonlocal. In fact, assuming that b is continuous with respect to weak convergence or a Wasserstein metric already prohibits local interactions; we briefly discussed *local* interactions in Example 2.19, and these can (and often are) incorporated into McKean-Vlasov equations, though it takes more care to represent a local McKean-Vlasov equation as a meaningful n-particle system which converges to it.

Even if the density of μ_t does not exist or is not sufficiently smooth, is it still clear that (μ_t) is a weak solution of the PDE (3.15), in the sense that it holds integrated against test functions. That is, the first equation of (3.14) is always valid, for every smooth test function φ of compact support. What we have shown with the above argument is that if μ is a solution of the McKean-Vlasov equation (3.8), then it is also a weak solution of the PDE (3.15) in a suitable sense. The converse can often be shown as well. In particular, it suffices to show that the PDE (3.15) has a unique weak solution. This can be done in many cases, and for this reason the PDE (3.15) is often itself the main object of study, instead of the McKean-Vlasov SDE (3.8).

3.4 An alternative derivation of the McKean-Vlasov limit

The PDE (3.15) can be used as the basis for studying the $n \to \infty$ limit of the n-particle system. This approach requires some more machinery from weak convergence theory for stochastic processes, which we will not develop. Instead, the argument is merely sketched, with some warnings when the arguments become hand-wavy.

A good place to start is to study this behavior of the $\mathcal{P}(\mathbb{R}^d)$-valued stochastic process $(\mu_t^n)_{t\ge 0}$ through the integrals of test functions. Let us make use of the shorthand notation

$$\langle \nu, \varphi \rangle := \int_{\mathbb{R}^d} \varphi \, d\nu,$$

for a measure ν on $\mathbb{R}^d$ and a ν-integrable function φ.

Fix a smooth function φ on $\mathbb{R}^d$ with compact support. To identify the behavior of

$$\langle \mu_t^n, \varphi \rangle = \frac{1}{n} \sum_{i=1}^n \varphi(X_t^{n,i}),$$

we first use Itô's formula to write, for each $i = 1, \ldots, n$,

$$\begin{aligned} d\varphi(X_t^{n,i}) = &\left(b(X_t^{n,i}, \mu_t^n) \cdot \nabla \varphi(X_t^{n,i}) + \frac{1}{2} \mathrm{Tr}[\sigma\sigma^\top (X_t^{n,i}, \mu_t^n) \nabla^2 \varphi(X_t^{n,i}] \right) dt \\ &+ \nabla \varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i. \end{aligned}$$

It is convenient now to define the infinitesimal generator by setting

$$L_m \varphi(x) = b(x,m) \cdot \nabla \varphi(x) + \frac{1}{2} \mathrm{Tr}[\sigma\sigma^\top (x,m) \nabla^2 \varphi(x)],$$

for each $m \in \mathcal{P}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. We may then write

$$d\varphi(X_t^{n,i}) = L_{\mu_t^n} \varphi(X_t^{n,i}) dt + \nabla \varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i.$$

Average over $i = 1, \ldots, n$ to get

$$\begin{aligned}
d\langle \mu_t^n, \varphi \rangle &= \frac{1}{n} \sum_{i=1}^n d\varphi(X_t^{n,i}) \\
&= \left\langle \mu_t^n, L_{\mu_t^n} \varphi \right\rangle dt + \frac{1}{n} \sum_{i=1}^n \nabla \varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i \\
&=: \left\langle \mu_t^n, L_{\mu_t^n} \varphi \right\rangle dt + dM_t^n,
\end{aligned}$$

where the last line defines the local martingale M^n. In the simplest case, the function σ is uniformly bounded, and so there exists a constant C such that $|\sigma^\top \nabla \varphi| \leq C$ uniformly. Then, M^n is a martingale with quadratic variation

$$[M^n]_t = \frac{1}{n^2} \sum_{i=1}^n \int_0^t |\sigma^\top(X_s^{n,i}, \mu_s^n) \nabla \varphi(X_s^{n,i})|^2 ds \leq \frac{tC^2}{n}.$$

In particular, this implies $\mathbb{E}[(M_t^n)^2] \leq tC^2/n$.

At this point we begin skipping some crucial steps. One should show next that $(\mu_t^n)_{t \geq 0}$ is tight as a sequence of random elements of $C(\mathbb{R}_+; \mathcal{P}^2(\mathbb{R}^d))$, endowed with an appropriate topology. Then, by Prokhorov's theorem, it admits a subsequential limit point, and we would like to describe all such limit points. So suppose the process $(\mu_t^n)_{t \geq 0}$ converges (along a subsequence is enough) to another (measure-valued) process $(\mu_t)_{t \geq 0}$. Then, since the martingale term M^n vanishes as $n \to \infty$, we should have

$$d\langle \mu_t, \varphi \rangle = \langle \mu_t, L_{\mu_t} \varphi \rangle \, dt. \tag{3.16}$$

The identity (3.16) holds a.s. for all smooth φ of compact support, and with some work one can switch the order of quantifiers to show that in fact, with probability 1, the equation holds (3.16) for all φ and all $t \geq 0$. (That is, the null set does not depend on t or φ.) This shows that the limit point μ is almost surely a weak solution of the PDE (3.15).

If we knew, via different arguments, that the weak solution of the PDE (3.15) is unique, then we would be in good shape. Let us write ν for the unique solution. Every subsequential limit in distribution μ of μ^n is almost surely a weak solution, and thus $\mu = \nu$ a.s. This implies that the full sequence μ^n converges in distribution to the deterministic ν.

In fact, this line of argument reveals, interestingly, that even if the PDE (3.15) is not unique, we can still say that every subsequential limit point of the μ^n is supported on the set of solutions of PDE (3.15).

Remark 3.5. A variation of this same argument uses the theory of martingale problems [112] to show essentially the same thing, but replacing the PDE (3.15) with the martingale problem form of the McKean-Vlasov equation. Alternatively, a similar argument can work directly with the SDEs, without resorting to PDEs or martingale problems, using the result of Kurtz and Protter [77] on weak convergence of stochastic integrals. These kinds of arguments are more advanced but tend to be much more versatile than those presented in the previous sections (in Theorem 3.3 and its proof), which required the coefficients to be Lipschitz.

3.5 Loss of the Markov property

This short section highlights how the Markov properties ties in with McKean-Vlasov SDEs, which in fact are sometimes called *nonlinear Markov processes*, essentially because (as we saw in Section 3.3) the corresponding Fokker-Planck equation is nonlinear. Let us start by briefly reviewing the situation for standard SDEs.

Suppose we are given coefficients b and σ, and suppose the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x,$$

has a unique solution for any initial state x. (Existence of a weak solution and uniqueness in law is enough here.) Then the resulting process $(X_t)_{t\geq 0}$ is a Markov process, meaning that $\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s]$ a.s. for every $s < t$, where $\mathcal{F}_s = \sigma(X_r : r \leq s)$ is the natural filtration of X.

In fact, the SDE above has the even stronger property of defining what is often called a *Markov family.* letting $\mathcal{C}^d = C([0,\infty); \mathbb{R}^d)$ denote the path space, suppose we define $P_x \in \mathcal{P}(\mathcal{C}^d)$ as the law of the solution starting from $x \in \mathbb{R}^d$. Now suppose we randomize the initial state. Consider an initial distribution $m \in \mathcal{P}(\mathbb{R}^d)$. Then the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 \sim m,$$

has a unique (in law) solution, and we may write $P_m \in \mathcal{P}(\mathcal{C}^d)$ for the law of the solution. An instance of the "Markov family" property is the statement that for any Borel set $A \subset \mathcal{C}^d$ we have

$$P_m(A) = \int_{\mathbb{R}^d} P_x(A)\, m(dx). \tag{3.17}$$

In other words, if we solve the SDE from every deterministic initial state, this is enough to determine the behavior of the SDE started from a random initial state, simply by integrating over the distribution of the initial state. Another way to say this: under P_m, the conditional law of the process X given $X_0 = x$ is equal to P_x.

With McKean-Vlasov equations, the situation is different. Consider the McKean-Vlasov SDE

$$dX_t = b(X_t, \mathcal{L}(X_t))dt + \sigma(X_t, \mathcal{L}(X_t))dB_t, \quad X_0 = x. \tag{3.18}$$

Suppose b and σ satisfy the Lipschitz condition of 3.1. First notice that if we let $\mu_t = \mathcal{L}(X_t)$, and we let Y denote the unique (strong) solution of the SDE

$$dY_t = b(Y_t, \mu_t)dt + \sigma(Y_t, \mu_t)dB_t, \quad Y_0 = x,$$

then we must have $Y \equiv X$. In the latter equation, we simply treat (μ_t) as a time-dependence of the coefficients, and well-posedness follows from standard SDE theory. We also conclude that Y (and thus X) is a Markov process, though this time-inhomogeneous, because the coefficients of the latter equation only depend on (t, Y_t).

However, the "Markov family" property described above is lost. Suppose $P_x \in \mathcal{P}(\mathcal{C}^d)$ denotes the law of the solution of (3.18) started from $x \in \mathbb{R}^d$. For $m \in \mathcal{P}(\mathbb{R}^d)$ let $P_m \in \mathcal{P}(\mathcal{C}^d)$ denote the law of the unique solution of the same SDE but started with initial distribution $X_0 \sim m$. Then, in general, the equation (3.17) does not hold! If we solve the McKean-Vlasov SDE with a random initial state, and then condition on $X_0 = x$ for some $x \in \mathbb{R}^d$, the resulting distribution is NOT the same as the law P_x obtained by solving the McKean-Vlasov SDE with deterministic initial state $X_0 = x$.

3.6 Common noise

In this section we discuss briefly, without proof, an important extension of the main model to allow some correlations between the driving Brownian motions. The way this is typically done is as follows: In the n-particle system, there are now $n+1$ independent Brownian motions, B and $W^1, \ldots, W^n$. As before, W^i is of dimension m, while now B is of dimension m_0. We are given a volaility coefficient $\sigma_0 : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}^{d \times m_0}$ for the common noise term. The particles $(X_t^{n,1}, \ldots, X_t^{n,n})$ evolve according to the SDE system

$$dX_t^{n,i} = b(X_t^{n,i}, \mu_t^n)dt + \sigma(X_t^{n,i}, \mu_t^n)dW_t^i + \sigma_0(X_t^{n,i}, \mu_t^n)dB_t, \quad X_0^{n,i} = \xi^i, \tag{3.19}$$

where as usual we write

$$\mu^n = \frac{1}{n}\sum_{k=1}^n \delta_{X^{n,k}}$$

for the empirical measure of the particles, viewed as a random element of $\mathcal{P}(\mathcal{C}^d)$, and $\mu^n_t = \frac{1}{n}\sum_{k=1}^n \delta_{X^{n,k}_t}$ denotes the time-t marginal.

The key difference between (3.1) and (3.19) is that in the latter case the particles are not driven by their own independent Brownian motions. The "common noise" B can model what economists might call "aggregate shocks," which effect the system as a whole as opposed to a single particle.

We saw in Theorem 3.3 that in models without common noise the particles become asymptotically i.i.d., and this is no longer the case when there is common noise. The correlations induced by the common noise persist in the limit, and the limiting μ is stochastic.

3.6.1 The limit theorem

The *conditional* (or *stochastic*) McKean-Vlasov equation is defined precisely as follows:

$$\begin{aligned} dY_t &= b(Y_t,\mu_t)dt + \sigma(Y_t,\mu_t)dW_t + \sigma_0(Y_t,\mu_t)dB_t, \quad t\in[0,T], \quad Y_0=\xi, \\ \mu &= \mathcal{L}(Y|B), \quad \forall t \ge 0. \end{aligned} \tag{3.20}$$

Here W and B are independent Brownian motions, ξ is an $\mathbb{R}^d$-valued random variable with the same law as ξ^i, and both are (say) supported on the same filtered probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$. Again, ξ should be $\mathcal{F}_0$-measurable, and W and B are $\mathbb{F}$-Brownian motions. We write also $\mathcal{L}(Z|\widetilde{Z})$ for the law or distribution of a random variable Z given another random variable $\widetilde{Z}$, and we stress that $\mathcal{L}(Z|B)$ means we are conditioning on the entire $\mathcal{C}^{m_0}$-valued random variable B, or equivalently the entire trajectory $(B_t)_{t\in[0,T]}$.[4]

The last equality $\mu = \mathcal{L}(Y|B)$ in (3.20) is meant to be understood in the almost sure sense. That is, μ is a $\mathcal{P}^2(\mathcal{C}^d)$-valued random variable, and μ is a version of the regular conditional law of Y given B. Equivalently,

$$\int \varphi \, d\mu = \mathbb{E}[\varphi(Y)\,|\,B], \quad a.s.,$$

for every bounded measurable function φ on $\mathcal{C}^d$. It is important to note here that if Y is $\mathbb{F}$-adapted and B is an $\mathbb{F}$-Brownian motion, and if $\mu = \mathcal{L}(Y|B)$, then it holds automatically that

$$\mathcal{L}(Y_t|B) = \mathcal{L}(Y_t|\mathcal{F}^B_t), \quad a.s., \ \forall t\in[0,T],$$

[4] We may say "equivalently" here because it is well known (and a good exercise to show) that the Borel σ-field on $\mathcal{C}^k = C([0,T];\mathbb{R}^k)$ is precisely the σ-field generated by the family of coordinate maps $\mathcal{C}^k \ni x \mapsto x_t \in \mathbb{R}^k$, for $t\in[0,T]$.

where $\mathcal{F}_t^B = \sigma(B_s : s \le t)$ is the filtration generated by B. To see this, fix any bounded measurable function φ on $\mathbb{R}^d$. Because Y_t is $\mathcal{F}_t$-measurable and the future increments $(B_s - B_t)_{s\in[t,T]}$ are independent of $\mathcal{F}_t$, we have

$$\begin{aligned}\mathbb{E}[f(Y_t)\,|\,B] &= \mathbb{E}[f(Y_t)\,|\,\mathcal{F}_t^B \vee \sigma(B_s - B_t : s\in[t,T])] \\ &= \mathbb{E}[f(Y_t)\,|\,\mathcal{F}_t^B], \quad a.s.,\end{aligned}$$

where $\mathcal{G}\vee\mathcal{H} := \sigma(\mathcal{G}\cup\mathcal{H})$ denotes the "join" of two σ-fields $\mathcal{G}$ and $\mathcal{H}$.

The discussion of the previous paragraph tells us that, if Y is $\mathbb{F}$-adapted and B is an $\mathbb{F}$-Brownian motion, and if $\mu = \mathcal{L}(Y|B)$, then μ_t is $\mathcal{F}_t^B$-measurable, at least *almost surely*. That is, μ_t agrees a.s. with an $\mathcal{F}_t^B$-measurable random variable. Hence, if we are careful to assume our filtration $\mathbb{F}$ is complete, we may safely conclude that the $\mathcal{P}^2(\mathbb{R}^d)$-valued process $(\mu_t)_{t\in[0,T]}$ is adapted to the filtration $\mathbb{F}^B = (\mathcal{F}_t^B)_{t\in[0,T]}$ generated by the common noise.

A strong solution of the conditional McKean-Vlasov equation is a pair (Y,μ), where Y is a continuous $\mathbb{F}$-adapted $\mathbb{R}^d$-valued process (alternatively, a $C([0,T];\mathbb{R}^d)$-valued random variable), and μ is a probability measure on $C([0,T];\mathbb{R}^d)$, such that both equations (3.8) hold.

Note that for a fixed $\mathbb{F}^B$-adapted process $(\mu_t)_{t\in[0,T]}$, the SDE for Y in the first line of (3.20) is just an SDE with random coefficients. If b and σ are uniformly Lipschitz in the Y_t variable, then there is no problem with existence and uniqueness.

Theorem 3.6. *Suppose Assumption 3.1 holds, with the new coefficient σ_0 satisfying the same Lipschitz assumption. There exists a unique strong solution of the conditional McKean-Vlasov equation* (3.8)*. Moreover, the n-particle system converges in the following two senses. First,*

$$\lim_{n\to\infty} \mathbb{E}\left[\mathcal{W}^2_{\mathcal{C}^d,2}(\mu^n,\mu)\right] = 0. \tag{3.21}$$

Second, for a fixed $k\in\mathbb{N}$, we have the weak convergence

$$\mathcal{L}\Big((X^{n,1},\ldots,X^{n,k})\,|\,B\Big) \to \mu^{\otimes k}, \quad \textit{weakly in probability}, \tag{3.22}$$

where $\mu^{\otimes k}$ denotes the k-fold product measure of μ with itself.

We will not write out the proof of Theorem 3.6, which follows very closely the proof of Theorem 3.3. The key idea of the limit theorem is now to construct not i.i.d. but rather *conditionally i.i.d.* copies of the solution of the McKean-Vlasov equation, driven by the same Brownian motions as the n-particle systems. That is, let

$$dY_t^i = b(Y_t^i, \mu_t)dt + \sigma(Y_t^i, \mu_t)dW_t^i + \sigma_0(Y_t^i, \mu_t)dB_t, \quad t \in [0,T], \quad Y_0 = \xi^i,$$

where μ solves the McKean-Vlasov equation. Then $Y^i = \mathcal{L}(Y^i|B)$ a.s., for each i, and in fact (Y^i) are conditionally i.i.d. with common conditional law μ. The strategy is again to estimate $\|X^{n,i} - Y^i\|$.

The two most pertinent details to check are as follows. First, to estimate the distance between two conditional measure of the form $\mathcal{L}(Y^1|B)$ and $\mathcal{L}(Y^2|B)$, where Y^1 and Y^2 are $\mathcal{C}^d$-valued random variables, simply note that $\mathcal{L}((Y^1, Y^2)|B)$ defines a coupling of the two, almost surely. Hence,

$$\mathcal{W}_2^2(\mathcal{L}(Y^1|B),\, \mathcal{L}(Y^2|B)) \le \mathbb{E}\left[\|Y^1 - Y^2\|^2 \,\middle|\, B\right], \quad a.s.$$

The second point to be careful about is that we invoked the law of large numbers for an i.i.d. sequence, and now we need to invoke a *conditional* law of large numbers. But this is a straightforward extension of Theorem 2.9, so we will not go into the details.

3.6.2 An alternative derivation of the common noise limit

Returning now to the n-particle system, let us next sketch a direct derivation of the McKean-Vlasov limit in the common noise setting, analogous to Section 3.4. Let us assume that we have shown the sequence of processes $(\mu_t^n)_{t\in[0,T]}$ is tight in $C([0,T];\mathcal{P}(\mathbb{R}^d))$, with $\mathcal{P}(\mathbb{R}^d)$ equipped with some metric of weak convergence. Fix a smooth function φ on $\mathbb{R}^d$ with compact support. As before, define the infinitesimal generator by setting

$$L_m\varphi(x) = b(x,m)\cdot\nabla\varphi(x) + \frac{1}{2}\mathrm{Tr}[(\sigma\sigma^\top + \sigma_0\sigma_0^\top)(x,m)\nabla^2\varphi(x)],$$

for each $m \in \mathcal{P}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Use Itô's formula to write, for each $i = 1, \ldots, n$,

$$\begin{aligned} d\varphi(X_t^{n,i}) = L_{\mu_t^n}\varphi(X_t^{n,i})dt + \nabla\varphi(X_t^{n,i})\cdot\sigma(X_t^{n,i}, \mu_t^n)dW_t^i \\ + \nabla\varphi(X_t^{n,i})\cdot\sigma_0(X_t^{n,i}, \mu_t^n)dB_t. \end{aligned}$$

Average over $i = 1, \dots, n$ to get

$$\begin{aligned} d\langle \mu_t^n, \varphi \rangle &= \frac{1}{n} \sum_{i=1}^n d\varphi(X_t^{n,i}) \\ &= \langle \mu_t^n, L_{\mu_t^n} \varphi \rangle \, dt + \frac{1}{n} \sum_{i=1}^n \nabla \varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i \\ &\quad + \langle \mu_t^n, \nabla \varphi(\cdot)^\top \sigma(\cdot, \mu_t^n) \rangle dB_t. \end{aligned}$$

Consider the martingale

$$M_s^n = \int_0^s \frac{1}{n} \sum_{i=1}^n \nabla \varphi(X_t^{n,i}) \cdot \sigma(X_t^{n,i}, \mu_t^n) dW_t^i.$$

If σ is uniformly bounded, then there exists a constant C such that $|\sigma^\top \nabla \varphi| \leq C$ uniformly. Then, M^n is a martingale with quadratic variation

$$[M^n]_t \leq \frac{tC^2}{n}.$$

Now, if $(\mu_t)_{t \in [0,T]}$ is the weak limit of a convergent subsequence of $(\mu_t^n)_{t \in [0,T]}$ in $C([0,T]; \mathcal{P}(\mathbb{R}^d))$, then we expect to be able to pass to the limit in the above equation to get

$$d\langle \mu_t, \varphi \rangle = \langle \mu_t, L_{\mu_t} \varphi \rangle \, dt + \langle \mu_t, \nabla \varphi(\cdot)^\top \sigma(\cdot, \mu_t) \rangle dB_t. \tag{3.23}$$

Technically, one should consider joint weak limits of (μ^n, B). If $L_m \varphi(x)$ and $\sigma(x, m)$ are continuous and uniformly bounded as functions of $(x.m)$, then this passage to the limit can be justified by writing the equation in integral form and taking advantage of weak convergence results for stochastic integrals of Kurtz-Protter [78].

Details aside, we now expect that any limit of μ^n solves the above *stochastic PDE in weak form*. Indeed, suppose for the moment that (3.23) holds for every nice test function φ with probability 1, and suppose that $\mu_t(dx) = \mu(t,x)dx$ for some smooth (random!) function $\mu = \mu(t,x)$. Then just as we derived (3.15) we may deduce from (3.23) the stochastic PDE

$$\begin{aligned} \partial_t \mu(t,x) = &- \operatorname{div}_x(b(x, \mu_t)\mu(t,x)) + \frac{1}{2} \operatorname{Tr}\Big[\nabla^2 \big((\sigma\sigma^\top + \sigma_0\sigma_0^\top)(x, \mu_t)\mu(t,x)\big)\Big] \\ &- \sum_{j=1}^{m_0} \operatorname{div}_x(\mu(t,x)\sigma_0^j(t,x) \cdot dB_t^j, \end{aligned} \tag{3.24}$$

where σ_0^j denotes the j^{th} row of σ_0 and B^j denotes the j^{th} component of the m_0-dimensional Brownian motion B.

3.7 Long-time behavior

We now understand the $n \to \infty$ limit theory for interacting diffusion models, and the goal of this section is to give an informal discussion of some of the interesting features of the $t \to \infty$ limit theory. The analysis of this section will be quite far from self-contained or complete but is meant to give a brief view of an important topic, interesting both theoretically and in applications of McKean-Vlasov equations.

To keep things concrete, we focus on the following specific model of n one-dimensional ($d = 1$) particles:

$$\begin{aligned} dX_t^{n,i} &= \left[a(\overline{X}_t^n - X_t^{n,i}) + (X_t^{n,i} - |X_t^{n,i}|^3)\right] dt + \sigma dW_t^i, \qquad (3.25) \\ \overline{X}_t^n &= \frac{1}{n}\sum_{k=1}^n X_t^{n,k}. \end{aligned}$$

This model has been studied in a wide range of applications, perhaps most recently in a model of systemic risk [57], and perhaps originally as a model of dynamic phase transitions in statistical physics; see [41] for references and the details of the analysis that we will omit.

3.7.1 A brief discussion of ergodicity for SDEs

Consider a smooth function V on $\mathbb{R}^k$, and consider the k-dimensional SDE

$$dX_t = \nabla V(X_t)dt + \sigma dW_t,$$

where $\sigma > 0$ is scalar and W a k-dimensional Brownian motion. Under suitable conditions on V, it is well known that X ergodic with invariant distribution

$$\rho(dx) = \frac{1}{Z}e^{\frac{2}{\sigma^2}V(x)}dx,$$

where $Z > 0$ is a normalizing constant rendering ρ a probability measure. That is, no matter the distribution of X_0, the law of X_t converges weakly to ρ.

Let us argue at least that ρ is stationary. Using Itô's formula, one quickly derives the Fokker-Planck equation associated to the SDE, which says that for any smooth function φ of compact support we have

$$\frac{d}{dt}\int_{\mathbb{R}^k} \varphi\, d\mu_t = \int_{\mathbb{R}^k} \left(\nabla V(x) \cdot \nabla\varphi(x) + \frac{\sigma^2}{2}\Delta\varphi(x)\right) \mu_t(dx),$$

where μ_t is the law of X_t. Assuming V decays quickly enough at infinity, we may show that $\mu_t = \rho$ provides a constant solution of the Fokker-Planck equation. Indeed, plugging in $\mu_t = \rho$, the right-hand side of this equation becomes

$$\begin{aligned}
&\frac{1}{Z}\int_{\mathbb{R}^k}\left(\nabla V(x)\cdot\nabla\varphi(x) + \frac{\sigma^2}{2}\Delta\varphi(x)\right)e^{\frac{2}{\sigma^2}V(x)}dx \\
&\qquad = \frac{1}{Z}\int_{\mathbb{R}^k}\left(\nabla V(x)\cdot\nabla\varphi(x) - \frac{\sigma^2}{2}\nabla\varphi(x)\cdot\frac{2}{\sigma^2}\nabla V(x)\right)e^{\frac{2}{\sigma^2}V(x)}dx \\
&\qquad = 0,
\end{aligned}$$

where the first step is simply integration by parts, justified if V decays quickly enough at infinity. Hence $\frac{d}{dt}\mathbb{E}[\varphi(X_t)] = 0$, which means that the distribution ρ is stationary. There is more work in proving ergodicity, but we will not go into the details; let us take for granted that for reasonable V it holds not only that ρ is stationary but that no matter what the initial distribution of X_0 is the law of the solution X_t converges weakly to ρ.

An important intuition falls out of this result regarding the effect of the noise parameter σ. As it decreases toward zero, the exponent blows up, but of course the normalizing constant Z adapts to keep ρ a probability measure. The result is that, for small σ (the *low-temperature* regime), the measure ρ puts much more mass in the regions where V is the smallest. In other words, as σ decreases, we expect ρ to concentrate around the minimizers of V. For example, if V is a double-well potential like $V(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2$ with multiple minimizers (namely, ± 1), then for small σ the particle overwhelmingly likely to be near one of these two minimizers. Notice that the function $V'(x) = x^3 - x$ appears in the drift of (3.25).

3.7.2 Long-time behavior of (3.25)

Applying standard ergodicity results as in the previous paragraph to our particle system (3.25), it can be shown for a fixed n that $(X^{n,1}, \ldots, X^{n,n})$ is ergodic, with invariant measure

$$\rho_n(dx_1, \ldots, dx_n) = \frac{1}{Z_n}\exp\left(\frac{a}{\sigma^2 n}\sum_{i,j=1}^n x_i x_j + \frac{1}{\sigma^2}\sum_{i=1}^n\left((1-a)x_i^2 - \frac{1}{2}x_i^4\right)\right)dx_1 \cdots dx_n.$$

For large n, these measures are quite hard to analyze, as the normalizing constant Z_n is quite hard to estimate. However, we can learn quite a bit about ρ_n by studying the McKean-Vlasov limit!

Now, the model (3.25) does not fit the assumptions of Theorem 3.3, but it can be shown nonetheless (see [41, 58]) that the law of large numbers holds here. That is, there is a unique solution of the McKean-Vlasov equation,

$$dX_t = \left[a\left(\mathbb{E}[X_t] - X_t\right) + (X_t - X_t^3)\right] dt + \sigma dW_t.$$

Moreover, letting $\mu = \mathcal{L}(X)$ denote the law of the solution, the empirical $\mu^n = \frac{1}{n}\sum_{k=1}^n \delta_{X^{n,k}}$ converges weakly in probability to μ. In other words, $\mathcal{L}(\mu^n)$ converges weakly to δ_μ. However, this does NOT imply anything about convergence of stationary distributions! To be clear, notice that ergodicity of the n-particle system lets us can identify the weak limit

$$\lim_{t\to\infty} \mathcal{L}(\mu_t^n) = \rho_n \circ L_n^{-1},$$

where $L_n : \mathbb{R}^n \to \mathcal{P}(\mathbb{R})$ is the empirical measure map, defined by $L_n(x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. One is tempted to conclude that

$$\lim_{t\to\infty} \delta_{\mu_t} = \lim_{t\to\infty}\lim_{n\to\infty} \mathcal{L}(\mu_t^n) = \lim_{n\to\infty}\lim_{t\to\infty} \mathcal{L}(\mu_t^n) = \lim_{t\to\infty} \rho_n \circ L_n^{-1},$$

but in general these two limits cannot be interchanged.

For some models, this is possible; the way to prove it is by showing that the limit $\lim_{n\to\infty} \mu_t^n = \mu_t$ holds uniformly in time in a suitable sense. But in the model (3.25), this only works if σ is sufficiently large. For small σ, there are multiple stationary distributions for the McKean-Vlasov equation, and the two limits do not commute.

3.7.3 Stationary measures for the McKean-Vlasov equation

To dig into this, let us try to find stationary solutions of the McKean-Vlasov equation. In PDE form, μ_t satisfies

$$\frac{d}{dt}\langle \mu_t, \varphi\rangle = \int_{\mathbb{R}} \left[a(m_t - x)\varphi'(x) + (x - x^3)\varphi'(x) + \frac{\sigma^2}{2}\varphi''(x)\right] \mu_t(dx),$$
$$\text{with } \ m_t = \int_{\mathbb{R}} y\, \mu_t(dy).$$

As before, if we set

$$\mu_t(dx) = \frac{1}{Z}\exp\left(\frac{2}{\sigma^2}\left(am_t x - \frac{a}{2}x^2 + \frac{1}{2}x^2 - \frac{1}{4}x^4\right)\right),$$

then integration by parts shows that μ_t is a solution. But for this to truly be a stationary distribution, the mean path m_t should be constant. Define $\rho_m \in \mathcal{P}(\mathbb{R})$ for $m \in \mathbb{R}$ by

$$\rho_m(dx) = \frac{1}{Z_m} \exp\left(\frac{2}{\sigma^2}\left(amx - \frac{a}{2}x^2 + \frac{1}{2}x^2 - \frac{1}{4}x^4\right)\right),$$

where as usual $Z_m > 0$ is a normalizing constant. The goal then is to find $m \in \mathbb{R}$ such that $m = \int_{\mathbb{R}} y\,\rho(dy)$, i.e., m is the mean of ρ_m. Notice that if $m = 0$ then ρ_0 is symmetric about the origin and thus has mean zero. So $m = 0$ is always a solution. It is shown in [41] that there exists a critical value $\sigma_c > 0$ such that, if $\sigma \geq \sigma_c$, then the only solution is $m = 0$, whereas if $\sigma < \sigma_c$ then there exists $m_1 > 0$ such that $\pm m_1$ are solutions. In other words, for small σ, there are three stationary distributions for the McKean-Vlasov equation.

An intuitive explanation for this phase transition is as follows, and it might help at this point to review the last paragraph of Section 3.7.1. We might expect to see most of the particles near the minimizers ± 1 of the potential function $V(x) = x^4/4 - x^2/2$. When σ is large, the noise σdW_t^i dominates the attractive drift term $a(\overline{X}_t^n - X_t^{n,i})dt$, and we see about an even split of particles at these two energy-minimizing locations, for a mean location of 0. When σ is small, the interaction through the drift is much stronger, and the particles tend to cluster together at one of these two locations, and the mean-zero equilibrium is no longer stable. This is an example of *metastability*.

These ideas lead to a mathematical model of a phenomenon sometimes referred to as (quantum) *tunneling*. The idea is that, in the n-particle system for $\sigma < \sigma_c$, we expect to see the particles moving around one of the points, say -1. If one of the n noises takes an extreme value and kicks the corresponding particle to the other stable state $+1$, then it should not take long for the dominant attractive force in the drift to pull this particle back to the average particle location, near -1. However, there is a small probability (decaying to zero in n) that the noises conspire to kick many of the particles to $+1$, after which the attractive force in the drift will pull the rest of the particles over to $+1$. The probabilities of rare transitions of this form can be studied quantitatively by means of a large deviation analysis for the limit $\mu^n \to \mu$.

This notion of tunneling has a natural physical interpretation, but it was also applied in a model of systemic risk in [57]. There, the process $X_t^{n,i}$ represents the "state of risk" of agent i, and the two stable states ± 1

represent a good state and a bad or failed state. The system as a whole is in trouble if "most" of the agents are around the failed state.

3.8 Bibliographic notes and related models

There is by now a rich literature on McKean-Vlasov systems of the kind studied in this section, and to truly do the subject justice would require an entire course of its own. This class of models was originally introduced by McKean [95], building on ideas of Kac [73],in an effort to rigorously derive certain reduced equations (e.g., Burger's or Boltzmann's) from finite-particle systems. The particularly nice Lipschitz models we studied in detail originate in Sznitman's monograph [113], but this is really just the tip of the iceberg. The toy model of Section 3.1 was borrowed from the systemic risk model of [32].

The more powerful martingale method for studying the $n \to \infty$ limit was developed in [101, 58, 96] for models with continuous coefficients (b, σ). Models based on jump-diffusions have been studied a good deal as well, for instance in [61]. In the context of stochastic portfolio theory, Shkolnikov [109] analyzed an interesting family of rank-based interactions, pushing through the martingale method in spite of the inherent discontinuities (recall the discussion of Example 2.18).

An important direction of generalization of the models studied in this section would allow correlations between the driving Brownian motions. This is most easily accomplished by introducing a *common noise*. Given a Brownian motion B, independent of (W^i), consider the n-particle system of the form

$$dX_t^{n,i} = b(X_t^{n,i}, \mu_t^n)dt + \sigma(X_t^{n,i}, \mu_t^n)dW_t^i + \sigma_0(X_t^{n,i}, \mu_t^n)dB_t.$$

The best way to think about these systems intuitively (though bearing in mind that this is horribly non-rigorous) is to imagine "conditioning on" or "freezing" a realization of the common noise B. With this trajectory of B fixed, think of the dB_t term as part of the drift, and pass to the limit $n \to \infty$ for the resulting interacting diffusion model *without common noise*. Once in the limit, recall that everything is happening inside of this conditioning on B. Hence, the McKean-Vlasov equation (3.8) should now read

$$\begin{aligned} dY_t &= b(Y_t, \mu_t)dt + \sigma(Y_t, \mu_t)dW_t + \sigma_0(Y_t, \mu_t)dB_t, \\ \mu &= \mathcal{L}(Y \mid B), \end{aligned}$$

where $\mu = \mathcal{L}(Y \mid B)$ means that μ is now a random probability measure, and it represents a version of the conditional law of Y given B. The PDE form of the McKean-Vlasov equation (3.15) now becomes a *stochastic* PDE (or SPDE), driven by the Brownian motion B. See [40, 79, 80]. The analysis of common noise models tends to be much more delicate, but these models are increasingly relevant in many areas of application, for instance in capturing *aggregate shocks* in economics.

Another natural modification of our setup would permit *local interactions*, as discussed in Example 2.19, in which, for instance, the drift b may take the form $b(x, m) = \tilde{b}(\frac{dm}{dx}(x))$, defined only for those measures m for which the (Lebesgue) density $\frac{dm}{dx}$ exists. Given that an empirical measure of n points can never be absolutely continuous with respect to Lebesgue measure, the natural n-particle model (say, with constant volatility σ and particles taking values in $\mathbb{R}^d$) would look like

$$dX_t^{n,i} = \tilde{b}\left(\rho_n * \mu_t^n(X_t^{n,i})\right) dt + \sigma dW_t^i,$$

where

$$\rho_n * \mu_t^n(x) = \int \rho_n(x-y)\, \mu_t^n(dy) = \frac{1}{n}\sum_{i=1}^n \rho_n(x - X_t^{n,i})$$

is the convolution of ρ_n with the empirical measure, and ρ_n is a smooth approximation of the identity. For instance, take $\rho_n(x) = n^d\rho(nx)$ for some nonnegative continuous function ρ on $\mathbb{R}^d$ with compact support and $\int_{\mathbb{R}^d} \rho(x)dx = 1$. This is sometimes called a particle system with *moderate interactions*, and the $n \to \infty$ was studied in [102, 97].

The $n \to \infty$ limit for the empirical measure should be seen as a *law of large numbers*, and it can be complemented with different limit probabilistic theorems. Several authors have studied central limit theorems, showing that the recentered measure flow $\sqrt{n}(\mu_t^n - \mu_t)$ converges in a sense to the solution of a linear stochastic partial differential equation [68, 96]. For a large deviations principle, refer to the seminal paper of Dawson-Gärtner [42] where the study of process-level large deviations led to their discovery of the important *projective limit* method; see also the more recent [17] for a powerful and completely different approach. Lastly, concentration of measure techniques have been used to derive non-asymptotic estimates on the distance between μ^n and μ; see [15].

The long-time behavior and phase transitions discussed in Section 3.7 has been studied by now by a number of authors [41, 43, 33, 67, 114], though this is not a direction that we will develop (or has really been explored) in the controlled (mean field game(models we study later.

Beyond our brief discussion of models from statistical physics and systemic risk mentioned above, there are many more areas of application, and we mention only two particular areas which have seen recent activity, namely models of neuron networks [19, 44, 45] and animal flocking [39, 63].

4 Static mean field games

In this section we warm up by studying *static* games, in which there is no time component. This serves in large part as a warm up, but in the last two Sections 4.4 and 4.5 we will discuss some more modern topics. We work with a somewhat abstract setup, yet we will still make strong enough assumptions to render the analysis fairly straightforward. The discussion of this section is borrowed in part from the lecture notes of Cardaliaguet [20, Section 2].

There are n agents, and each agent chooses an action at the same time. An action is an element of a given set A, called the *action space*. We will assume throughout that A is a compact metric space (in fact, it is often finite in applications). A *strategy profile* is a vector $(a_1, \dots, a_n) \in A^n$. In the n-player game, player i has an *objective function* $J_i^n : A^n \to \mathbb{R}$, which assigns a "reward" to every possible strategy profile.

The goal of player i is to choose $a_i \in A$ to maximize the reward J_i^n. But J_i^n depends on all of the other agents' actions; the optimal choice of player i depends on the actions of the other players, and vice versa. To resolve inter-dependent optimization problems, we use the concept of Nash equilibrium.

Definition 4.1. A *Nash equilibrium* (for the n-player game) is a strategy profile $(a_1, \dots, a_n) \in A^n$ such that, for every $i = 1, \dots, n$ and every $\tilde{a} \in A$, we have

$$J_i^n(a_1, \dots, a_n) \geq J_i^n(a_1, \dots, a_{i-1}, \tilde{a}, a_{i+1}, \dots, a_n).$$

Similarly, given $\epsilon \geq 0$, an ϵ-*Nash equilibrium* is a strategy profile $(a_1, \dots, a_n) \in A^n$ such that, for every $i = 1, \dots, n$ and every $\tilde{a} \in A$, we have

$$J_i^n(a_1, \dots, a_n) \geq J_i^n(a_1, \dots, a_{i-1}, \tilde{a}, a_{i+1}, \dots, a_n) - \epsilon.$$

Note that a 0-Nash equilibrium and a Nash equilibrium are the same thing.

Intuitively, in Nash equilibrium, each player i is choosing a_i optimally, *given the other agents' choices.* In an ϵ-Nash equilibrium, each player could improve their reward, but by no more than ϵ.

Nash equilibria can be difficult to compute when n is large, and there is in fact a rich literature (which we will not discuss) on the computational complexity of Nash equilibria. It is often simpler to work with the $n \to \infty$ limit and study a game with a continuum of players. Such a limiting analysis is possible for a certain class of games, namely those in which the objective functions are suitably symmetric.

We assume henceforth that there is a single common payoff function of the form $F : A \times \mathcal{P}(A) \to \mathbb{R}$, and in the n-player game the objective function for player i is given by

$$J_i^n(a_1, \ldots, a_n) = F\left(a_i, \frac{1}{n}\sum_{k=1}^n \delta_{a_k}\right). \tag{4.1}$$

Intuitively, $F(a, m)$ represents the reward to a player choosing the action a when the distribution of actions chosen by other players is m. This cost structure renders the game symmetric, in the sense that for any $(a_1, \ldots, a_n)$ and any permutation π of $\{1, \ldots, n\}$, we have

$$J_i^n(a_1, \ldots, a_n) = J_{\pi(i)}^n(a_{\pi(1)}, \ldots, a_{\pi(n)}).$$

The objective of player i depends on the actions of the other agents only through their empirical distribution. In particular, the "names" or "labels" assigned to the players are irrelevant: player 1 does not care what player 2 is doing any more than what player 3 is doing. All that matters is the distribution of actions. For this reason, a game of this form is sometimes called *anonymous*.

Remark 4.2. In the specification (4.1) of J_i^n, it is arguably more natural to use the empirical measure $\frac{1}{n-1}\sum_{k \neq i} \delta_{a_k}$ of the other agents, not including agent i. This leads to the same $n \to \infty$ limit, the sense that Theorem 4.3 holds with no change (exercise).

The intuition behind the $n \to \infty$ limit is as follows: When n is very large, agent does not contribute much to the empirical measure, since δ_{a_i} is multiplied by a factor of $1/n$. Hence, when $n \to \infty$, we expect the optimization problems to decouple in some sense. The Nash equilibrium property will be reflected in a consistency between the limiting distribution obtained from the empirical measure and the action of a typical player. Throughout this section we will assume the following:

Standing assumption: A is a compact metric space, and $F : A \times \mathcal{P}(A) \to \mathbb{R}$ is jointly continuous, using the weak convergence topology on $\mathcal{P}(A)$.

Theorem 4.3. *Assume F is jointly continuous, using the weak convergence topology on $\mathcal{P}(A)$. Suppose that for each n we are given $\epsilon_n \geq 0$ and an ϵ_n-Nash equilibrium $(a_1^n, \ldots, a_n^n)$. Suppose $\lim_n \epsilon_n = 0$, and let*

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{a_i^n}.$$

Then $(\mu_n) \subset \mathcal{P}(A)$ is tight, and, for every weak limit point $\mu \in \mathcal{P}(A)$, we have that μ is supported on the set $\{a \in A : F(a, \mu) = \sup_{b \in A} F(b, \mu)\}$.[5]

Proof. Because A is compact, tightness of (μ_n) is automatic. Prokhorov's theorem 2.5 ensures that (μ_n) therefore admits a weak limit, along a subsequence. That is, there exist (n_k) and $\mu \in \mathcal{P}(A)$ such that $\lim_k \mu_{n_k} = \mu$ weakly. Fix any alternative action $b \in A$. Because $(a_1^n, \ldots, a_n^n)$ is ϵ_n-Nash, we have for each $i = 1, \ldots, n$

$$F(a_i^n, \mu_n) \geq F(b, \mu_n^i[b]) - \epsilon_n,$$

where we define

$$\mu_n^i[b] = \frac{1}{n}\left(\delta_b + \sum_{k \neq i} \delta_{a_k^n}\right) = \mu_n + \frac{1}{n}\left(\delta_b - \delta_{a_i^n}\right).$$

This is the empirical measure with the i^{th} player's action swapped out for b. Averaging over $i = 1, \ldots, n$, we find

$$\int_A F(a, \mu_n)\, \mu_n(da) = \frac{1}{n} \sum_{i=1}^n F(a_i^n, \mu_n) \geq \frac{1}{n} \sum_{i=1}^n F(b, \mu_n^i[b]) - \epsilon_n.$$

The goal now is to take limits on both sides of this inequality. On the left-hand side, we use joint continuity of F to find

$$\lim_k \int_A F(a, \mu_{n_k})\, \mu_{n_k}(da) = \int_A F(a, \mu)\, \mu(da).$$

See Example 2.16 for a justification of this step. To take limits on the right-hand side, we estimate the difference between μ_n and $\mu_n^i[b]$ as follows. First notice that the (A, d) is a compact metric space and thus bounded in the sense that there exists $C > 0$ such that $d(x, y) \leq C$ for all $x, y \in A$. Noting that the joint empirical measure

$$\frac{1}{n} \sum_{k \neq i} \delta_{(a_k^n, a_k^n)} + \frac{1}{n} \delta_{(a_i^n, b)}$$

[5] We say a probability measure m is *supported on* a set B if $m(B) = 1$.

is a coupling of μ_n and $\mu_n^i[b]$. Hence, we may use it to bound the 1-Wasserstein distance:

$$\mathcal{W}_1(\mu_n, \mu_n^i[b]) \leq \frac{1}{n} d(a_i^n, b) \leq \frac{C}{n}.$$

Because A is compact, so is $\mathcal{P}(A)$, and so the continuous function F is in fact uniformly continuous. Thus

$$\lim_{n\to\infty} \sup_{b\in A} \max_{i=1,\ldots,n} |F(b, \mu_n^i[b]) - F(b, \mu_n)| = 0.$$

Since $\epsilon_n \to 0$, we find

$$\lim_k \frac{1}{n_k} \sum_{i=1}^{n_k} F(b, \mu_{n_k}^i[b]) - \epsilon_{n_k} = \lim_k F(b, \mu_{n_k}) = F(b, \mu).$$

We conclude that

$$\int_A F(a, \mu)\, \mu(da) \geq F(b, \mu).$$

This is valid for every choice of $b \in A$, so we may take the supremum to get

$$\int_A F(a, \mu)\, \mu(da) \geq \sup_{b\in A} F(b, \mu).$$

Since trivially $F(a, \mu) \leq \sup_{b\in A} F(b, \mu)$ for all $a \in A$, the above inequality can only happen if

$$\mu \left\{ a \in A : F(a, \mu) = \sup_{b\in A} F(b, \mu) \right\} = 1.$$

□

Remark 4.4. We are taking for granted the existence of (ϵ-) Nash equilibria for the n-player games. This is not necessarily possible, though the famous work of Nash [99] ensures that it is possible if we work with *mixed strategies*. We will not go down this road, but refer to [20, Section 2] for a discussion of what happens to the $n \to \infty$ limit when we used mixed strategies.

Theorem 4.3 leads naturally to the following definition:

Definition 4.5. A probability measure $\mu \in \mathcal{P}(A)$ is called a *mean field equilibrium (MFE)* if

$$\mu\left\{a \in A : F(a,\mu) = \sup_{b\in A} F(b,\mu)\right\} = 1.$$

In other words, μ is an MFE if it is supported on the set of maximizers of the function $F(\cdot,\mu)$.

Intuitively, in the definition of an MFE, a point $a \in A$ in the support of μ represents one agent among a continuum, with μ representing the distribution of the continuum of agents' actions. In the presence of infinitely many agents, no single agent can influence μ by changing actions. Hence, the optimality of a typical agent a is expressed by the inequality $F(a,\mu) \geq F(b,\mu)$ for all $b \in A$.

4.1 Uniqueness

Theorem 4.3 shows that, in the sense of empirical measure convergence, the limit points of n-player Nash equilibria are always MFE. If it happens that we can prove by other means that there is at most one mean field equilibrium μ, then it would follow that (in the notation of Theorem 4.3) every subsequence of (μ_n) has a further subsequence that converges to μ, which means that in fact the full sequence μ_n converges to μ. This is an extremely common line of argument in applications of weak convergence theory: First prove tightness, then characterize (i.e., find a set of properties satisfied by) the subsequential limit points, and then finally (via separate arguments) prove uniqueness of this characterization.

Regarding uniqueness, one should first observe that it is not enough just to know that the function $a \mapsto F(a,m)$ has a unique maximizer for each a. If this were true, we could conclude that every mean field equilibrium must be a delta. That is, if $\hat{a}(m)$ is the unique maximizer of $F(\cdot,m)$ for each $m \in \mathcal{P}(A)$, then any MFE μ must satisfy $\mu = \delta_{\hat{a}(\mu)}$. But this does not mean the MFE is unique. The following theorem gives an example of a checkable assumption for uniqueness, due to Lasry-Lions [89]:

Theorem 4.6. *Suppose the objective function F satisfies the monotonicity condition,*

$$\int_A \left(F(a,m_1) - F(a,m_2)\right)(m_1 - m_2)(da) < 0, \tag{4.2}$$

for all $m_1, m_2 \in \mathcal{P}(A)$ with $m_1 \neq m_2$. Then there is at most one MFE.

Proof. Suppose $m_1, m_2 \in \mathcal{P}(A)$ are both MFE, and suppose they are distinct. Then

$$\int_A F(a, m_1)\, m_1(da) - \int_A F(a, m_1)\, m_2(da) \geq 0,$$
$$\int_A F(a, m_2)\, m_2(da) - \int_A F(a, m_2)\, m_1(da) \geq 0.$$

Add these inequalities to get

$$\int_A (F(a, m_1) - F(a, m_2))\, (m_1 - m_2)(da) \geq 0,$$

which, in light of (4.2), contradicts the assumption that $m_1 \neq m_2$. $\square$

For an example of a function F satisfying (4.2), suppose the space A is finite, with cardinality $|A| = d$. Then $\mathcal{P}(A)$ can be identified with the simplex in $\mathbb{R}^d$, namely, the set Δ^d of $(m_1, \ldots, m_n) \in \mathbb{R}^d$ with $m_i \geq 0$ and $\sum_{i=1}^d m_i = 1$. For $i \in A$, write $F_i(m) = F(i, m)$. We call the game a *potential game* if there exists a function $G : \Delta^d \to \mathbb{R}$, such that $\nabla G = (F_1, \ldots, F_d)$, and we call G the *potential function.* Suppose that we have a *strictly concave* potential G. One of the many characterizations of strict concavity reads as

$$(\nabla G(m) - \nabla G(m')) \cdot (m - m') < 0, \quad \text{for all } m_1, m_2 \in \Delta^d,\ m_1 \neq m_2.$$

This inequality may be written as

$$\sum_{i=1}^d (F_i(m) - F_i(m'))(m_i - m_i') < 0,$$

which is exactly the assumption (4.2).

We can go a bit further with this idea. Suppose our game admits a potential function G, not necessarily concave. Then the directional derivative of G at $m \in \Delta^d$ in the direction of $m' \in \Delta^d$ is given by

$$\begin{aligned} D_{m'}G(m) &:= \frac{d}{d\epsilon} G(m + \epsilon(m' - m)) = \nabla G(m) \cdot (m' - m) \\ &= \sum_{i=1}^d F_i(m)(m_i' - m_i) \\ &= \int_A F(a, m)\, m'(da) - \int_A F(a, m)\, m(da). \end{aligned}$$

Now, $m \in \Delta^d$ is a mean field equilibrium if and only if

$$\int_A F(a,m)\, m'(da) \le \int_A F(a,m)\, m(da),$$

if and only if $D_{m'}G(m) \le 0$ for every $m' \in \Delta^d$. In other words, m is a mean field equilibrium if and only if it locally maximizes the potential function G. If G is assumed concave, we conclude that m is a mean field equilibrium if and only if it maximizes G globally. This shows the true power of potential games: The competitive equilibrium is equivalent to the maximization of a function, and maximizing a function is typically much easier than computing a Nash equilibrium.

The classic reference for potential games is the treatise of Monderer and Shapley [98]. This idea can be generalized to the situation where A is not finite but requires a concept of the derivative of a function of a probability measure; we may return to this later in the course as time permits.

4.2 A converse to the limit theorem

We saw in Theorem 4.3 that limit points of n-player equilibria are always mean field equilibria (granted a continuous objective function). But to convince ourselves we have truly identified the limiting behavior of n-player equilibria, we must answer the natural followup question: Does every MFE arise as the limit of some sequence of n-player Nash equilibria? The following theorem and example provide answers.

Theorem 4.7. *Suppose $\mu \in \mathcal{P}(A)$ is an MFE. Then there exist $\epsilon_n \ge 0$ a sequence of strategy profiles $(a_i)_{i\in\mathbb{N}}$ such that*

$$\lim_n \frac{1}{n}\sum_{i=1}^n \delta_{a_i} = \mu, \qquad \lim_n \epsilon_n = 0.$$

Proof. Let $(X_i)_{i\in\mathbb{N}}$ be i.i.d. A-valued random variables, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define the (random) empirical measure

$$\mu_n = \frac{1}{n}\sum_{k=1}^n \delta_{X_k}.$$

For $a \in A$ and $i = 1, \ldots, n$, define

$$\mu_n^{(i)}[a] = \frac{1}{n}\left(\delta_a + \sum_{k\neq i}\delta_{X_k}\right) = \mu_n + \frac{1}{n}\left(\delta_a - \delta_{X_k}\right).$$

We know from Theorem 2.9 that $\mu_n \to \mu$ weakly a.s. Now, define

$$\epsilon_n = \max_{i=1,\ldots,n} \left(\sup_{a \in A} F(a, \mu_n^{(i)}[a]) - F(X_i, \mu_n) \right).$$

Clearly $\epsilon_n \geq 0$, and by construction $(X_1, \ldots, X_n)$ is an ϵ_n-Nash equilibrium for each n, almost surely (noting that ϵ_n is random). The proof will be complete if we show that $\epsilon_n \to 0$ a.s. Indeed, we can then select any realization ω for which both $\epsilon_n(\omega) \to 0$ and $\mu_n(\omega) \to \mu$ weakly, and set $a_i = X_i(\omega)$.

To show that $\epsilon_n \to 0$ a.s., we proceed as follows, making repeated use of uniform continuity of F and the convergence of μ_n to μ. We first swap out $\mu_n^{(i)}[a]$ for μ_n, noting that

$$\lim_{n \to \infty} |\epsilon_n - \widetilde{\epsilon}_n| = 0, \quad a.s.,$$

where

$$\widetilde{\epsilon}_n = \max_{i=1,\ldots,n} \left(\sup_{a \in A} F(a, \mu_n) - F(X_i, \mu_n) \right).$$

Again use uniform continuity of F to deduce that

$$\lim_{n \to \infty} \max_{i=1,\ldots,n} \left| \left(\sup_{a \in A} F(a, \mu_n) - F(X_i, \mu_n) \right) - \left(\sup_{a \in A} F(a, \mu) - F(X_i, \mu) \right) \right| = 0.$$

The key observation now is that, because μ is an MFE and X_i is a sample from μ, we have

$$\sup_{a \in A} F(a, \mu) = F(X_i, \mu), \quad a.s.$$

Hence, $\lim_n \epsilon_n = \lim_n \widetilde{\epsilon}_n = 0$ a.s., and the proof is complete. $\square$

A simple example shows that Theorem 4.7 is false if we require $\epsilon_n = 0$ for all n. In other words, there can exist MFE μ for which there is no sequence of Nash equilibria converging to μ.

Example 4.8. Suppose $A = [0, 1]$ and

$$F(a, m) = a\overline{m}, \qquad \text{where} \qquad \overline{m} := \int_{[0,1]} x \, m(dx).$$

There are exactly two MFE, namely $m = \delta_0$ and $m = \delta_1$. To see that δ_0 is a MFE, note simply that $F(a, \delta_0) = 0$ for all a, so δ_0 is certainly supported on the set $\{a \in [0, 1] : F(a, \delta_0) \geq F(b, \delta_0) \ \forall b \in [0, 1]\} = 1$. To see that δ_1 is a MFE, note that $F(a, \delta_1) = a$ is maximized when $a = 1$. To see that there are no other MFE, notice that if $0 < \overline{m} < 1$, then m cannot be supported on $\{1\}$, whereas the only maximizer of $F(a,m) = a\overline{m}$ is $a = 1$.

We claim that only the $m = \delta_1$ equilibrium arises as the limit of n-player equilibria. To see this, notice that the *only* strategy profile $(a_1, \ldots, a_n)$ which is a Nash equilibrium for the n-player game is $a_1 = a_2 = \ldots = a_n = 1$. It is clear that if any single agent deviates from this choice, the reward will decrease. Moreover, for any other choice of strategy profile, and agent i with $a_i < 1$ can improve his reward by switching to $a_i = 1$. Notably, $a_1 = \ldots = a_n = 0$ is *not* a Nash equilibrium for the n-player game (though it is a $(1/n)$-Nash equilibrium!), because when a single agent i switches to $a_i = 1$, the mean of the empirical measure increases from 0 to $1/n$, and thus the reward to agent i increases from 0 to $1/n$.

The astute reader might be dissatisfied with this counterexample on the grounds that it is no longer a counterexample if the n-player game reward functions are defined as

$$J_i^n(a_1, \ldots, a_n) = F\left(a_i, \frac{1}{n-1}\sum_{k \neq i} \delta_{a_k}\right),$$

with the i^{th} player excluded from the empirical measure. Indeed, with this alternative definition, which we stated in Remark 4.2 to have little consequence, the n-player strategy profile $a_1 = \ldots = a_n = 0$ is in fact a Nash equilibrium. However, this example can be modified to produce one which is not sensitive to this difference in objective function. It is left as an exercise for the reader to analyze the related example in which

$$F(a,m) = a(\overline{m} - r),$$

for a given $r \in (0,1)$, with the action set constrained now to $A = \{0,1\}$. One should also ask how the answer changes if we allow mixed strategies in our n-player equilibrium, but this is taking us too far off-topic.

4.3 Existence

We have seen in Theorem 4.3 that any convergent subsequence of n-player approximate equilibria converges to a MFE, and we have seen in Theorem 4.7 that, conversely, every MFE is the limit of some sequence of approximate equilibria. In each theorem we take existence for granted, we have said nothing so far about the existence of n-player equilibria or MFE. Existence of n-player is a classical topic which we will not discuss any further than

Remark 4.4. Existence of MFE, however, holds automatically under our standing assumptions (see just before Theorem 4.3):

Theorem 4.9. *There exists a MFE.*

Like Nash's proof of existence of mixed strategy equilibria for n-player games, the proof of Theorem 4.9 makes use of a famous fixed point theorem of Kakutani, which we state without proof:

Theorem 4.10. *Suppose K is a convex compact subset of a locally convex topological vector space. Suppose $\Gamma : K \to 2^K$ is a set-valued function (where 2^K is the set of subsets of K) satisfying the following conditions:*

(i) $\Gamma(x)$ is nonempty and convex for every $x \in K$.

(ii) The graph $\mathrm{Gr}(\Gamma) = \{(x, y) \in K \times K : y \in \Gamma(x)\}$ is closed.

Then there exists a fixed point, i.e., a point $x \in K$ such that $x \in \Gamma(x)$.

Proof of Theorem 4.9. Define a map $\Gamma : \mathcal{P}(A) \to \mathcal{P}(A)$ by letting $\Gamma(\mu)$ denote the set of probability measures which are supported on the set of maximizers of $F(\cdot, \mu)$. That is,

$$\Gamma(\mu) = \left\{ m \in \mathcal{P}(A) : m(\{a \in A : F(a, \mu) = \sup_{b \in A} F(b, \mu)\}) = 1 \right\}.$$

We see from the definition that $\mu \in \mathcal{P}(A)$ is an MFE if and only if it is a fixed point of Γ, i.e., $\mu \in \Gamma(\mu)$. Note that we may also write

$$\Gamma(\mu) = \left\{ m \in \mathcal{P}(A) : \int_A F(a, \mu)\, m(da) \geq F(b, \mu), \ \forall b \in A \right\}.$$

We now check the conditions of Kakutani's theorem. Recall from the Riesz representation that the topological dual of the space $C(A)$ of continuous functions on A is precisely the set $M(A)$ of signed measures on A of bounded variation. The corresponding weak* topology is precisely the weak convergence topology, and $\mathcal{P}(A)$ is a convex compact subset of $M(A)$ with this topology. So we can take $K = \mathcal{P}(A)$ in Kakutani's theorem. Let us check the two required properties of the map Γ:

(i) Fix $\mu \in \mathcal{P}(A)$. Let $S \subset A$ denote the set of maximizers of $F(\cdot, \mu)$, defined by
$$\left\{ a \in A : F(a, \mu) = \sup_{b \in A} F(b, \mu) \right\}.$$

We can then write $\Gamma(\mu) = \{m \in \mathcal{P}(A) : m(S) = 1\}$. Because F is continuous, the set S is nonempty, and we conclude that $\Gamma(\mu)$ is also nonempty. Moreover, $\Gamma(\mu)$ is clearly convex: If $m_1, m_2 \in \Gamma(\mu)$ and $t \in (0,1)$, then setting $m = tm_1 + (1-t)m_2$ we have

$$m(S) = tm_1(S) + (1-t)m_2(S) = t + (1-t) = 1,$$

so $m \in \Gamma(\mu)$.

(ii) The graph of Γ can be written as

$$\mathrm{Gr}(\Gamma) = \bigcap_{b \in A} K_b,$$

where we define, for $b \in A$,

$$K_b = \left\{ (\mu, m) \in \mathcal{P}(A) \times \mathcal{P}(A) : \int_A F(a,\mu)\, m(da) \geq F(b,\mu) \right\}.$$

If we show that K_b is closed for each $b \in A$, then it will follow that $\mathrm{Gr}(\Gamma)$ is closed. To this end, fix $b \in A$, and let $(\mu_n, m_n) \in K_b$ for each n. Suppose $(\mu_n, m_n) \to (\mu, m)$ for some $(\mu, m) \in \mathcal{P}(A) \times \mathcal{P}(A)$. We must prove that $(\mu, m) \in K_b$. To do this, note that because F is jointly continuous we have

$$\int_A F(a,\mu)\, m(da) = \lim_n \int_A F(a,\mu_n)\, m_n(da) \geq \lim_n F(b,\mu_n) = F(b,\mu).$$

Indeed, the first limit holds in light of Example 2.16.

In summary, we may apply Kakutani's theorem to Γ to obtain the existence of a fixed point. $\square$

Interestingly, our n-player game may fail to have a Nash equilibrium (in pure strategies), but Theorem 4.9 ensures that there still exists an MFE. Using Theorem 4.7, we conclude that there exist $\epsilon_n \to 0$ such that for each n there exists an ϵ_n-Nash equilibrium for the n-player game! So, even though there are not necessarily Nash equilibria, the mean field structure lets us construct approximate equilibria for large n.

4.4 Multiple types of agents

The model studied in this section is admittedly unrealistic in the sense that the agents are extremely homogeneous. A much more versatile framework is obtained by introducing different *types* of agents, with the essential ideas behind the analysis being the same. This section will only briefly state the setup and an extension of the main limit Theorem 4.3, but analogues of Theorems 4.7 and 4.6 are possible; see [84] for details.

In addition to our action space A, let $\mathcal{T}$ be a complete, separable metric space, which we will call the *type space*. (In practice, both A and $\mathcal{T}$ are often finite.) The payoff function is now $F : \mathcal{T} \times A \times \mathcal{P}(\mathcal{T} \times A) \to \mathbb{R}$, acting on a type, an action, and a type-action distribution. If each agent i in the n-player game is assigned a type t_i, then the reward for agent i is

$$F\left(t_i, a_i, \frac{1}{n}\sum_{k=1}^{n} \delta_{(t_k,a_k)}\right),$$

when agents choose actions $a_1, \ldots, a_n$.

An important additional feature we can incorporate is a dependence of the set of allowable actions on the type parameter. That is, as another input to the model, suppose $C : \mathcal{T} \to 2^A$ is a set-valued map, with the interpretation that $C(t)$ is the set of admissible actions available to an agent of type t. We call C the *constraint map*. In the following, let

$$\mathrm{Gr}(C) = \{(t,a) \in \mathcal{T} \times A : a \in C(t)\}$$

denote the graph of C.

We now define an *ϵ-Nash equilibrium associated with types* $(t_1, \ldots, t_n)$ as a vector $(a_1, \ldots, a_n) \in A$, with $a_i \in C(t_i)$ for each i, such that

$$F\left(t_i, a_i, \frac{1}{n}\sum_{k=1}^{n} \delta_{(t_k,a_k)}\right) \geq F\left(t_i, b, \frac{1}{n}\sum_{k\neq i} \delta_{(t_k,a_k)} + \frac{1}{n}\delta_{(t_i,b)}\right) - \epsilon,$$

for each $b \in C(t_i)$, for each $i = 1, \ldots, n$.

This more general setup retains the essential symmetric features of the previous setup, the idea being that in a large-n limit we can still obtain distributional limits if we know something about the distribution of types. This is captured by the following extension of Theorem 4.3.

Theorem 4.11. *Assume that F and C satisfy the following assumptions:*

- *F is jointly continuous on* $\mathrm{Gr}(C) \times \mathcal{P}(\mathrm{Gr}(C))$*, where* $\mathcal{P}(\mathrm{Gr}(C))$ *is shorthand for the set of* $\mu \in \mathcal{P}(\mathcal{T} \times A)$ *with* $\mu(\mathrm{Gr}(C)) = 1$*.*
- *$C(t)$ is nonempty for each $t \in \mathcal{T}$.*
- *The graph* $\mathrm{Gr}(C)$ *is closed.*
- *C is lower hemicontinuous, which means: If $t_k \to t$ in $\mathcal{T}$ and $a \in C(t)$, then there exist a subsequence (k_j) and some $a_j \in C(t_{k_j})$ such that $a_j \to a$.*

Suppose for each n we are given $(t_1^n, \ldots, t_n^n) \in \mathcal{T}^n$, as well as an ϵ_n-Nash equilibrium $(a_1^n, \ldots, a_n^n) \in A^n$ for the corresponding game, where $\epsilon_n \to 0$. Let

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{(t_k^n, a_k^n)}$$

denote the empirical type-action distribution. Suppose finally that the empirical type distribution converges weakly to some $\lambda \in \mathcal{P}(\mathcal{T})$, i.e.,

$$\frac{1}{n} \sum_{k=1}^{n} \delta_{t_k^n} \to \lambda.$$

Then (μ_n) is tight, and every weak limit point $\mu \in \mathcal{P}(\mathcal{T} \times A)$ is supported on the set

$$\left\{ (t, a) \in \mathcal{T} \times A : F(t, a, \mu) = \sup_{b \in A} F(t, a, \mu) \right\}.$$

The proofs of Theorem 4.11 and of the existence of mean field equilibria are very similar to those of Theorems 4.3 and 4.9. The key missing ingredient is (a special case of) Berge's theorem:

Theorem 4.12. *Suppose C satisfies the assumptions of Theorem 4.11. For each $m \in \mathcal{P}(\mathrm{Gr}(C))$ and $t \in \mathcal{T}$, define $C^*(t, m)$ to be the set of admissible maximizers of $F(t, \cdot, m)$, i.e.,*

$$C^*(t, m) = \left\{ a \in C(t) : F(t, a, m) = \sup_{b \in C(t)} F(t, b, m) \right\}.$$

Define also the maximum value

$$F^*(t, m) = \sup_{b \in C(t)} F(t, b, m).$$

Then F^* is (jointly) continuous, and the graph

$$\mathrm{Gr}(C^*) = \{(t, a, m) \in \mathcal{T} \times A \times \mathcal{P}(\mathrm{Gr}(C)) : a \in C^*(t, m)\}$$

is closed.

The proof of both of these theorems is left as an exercise. A great reference for analysis of set-valued functions, which come up quite a bit in game theory, is Chapter 17 of the textbook of Aliprantis and Border [3]. A general form of Berge's theorem can be found therein as Theorem 17.31.

4.5 Congestion games

Congestion games particularly important class of models fitting into the framework of Section 4.4, in fact with finite action and type spaces. These are standard models of network traffic, which we will think of in terms of their original application in road networks for the sake of intuition, though recent applications focus on different kinds of networks (e.g., the internet).

A finite directed graph $G = (V, E)$ is given, meaning V is some finite set of vertices and E is an arbitrary subset of $V \times V$. The vertices represent locations in the network, and a directed edge (u, v) is a road from u to v. Each edge is assigned an increasing cost function $c_e : [0, 1] \to \mathbb{R}_+$, with $c_e(u)$ representing the speed or efficiency of the road e when the *load* on the road is u, which means that the road is utilized by a fraction u of the population.

The type space $\mathcal{T}$ is a subset of $V \times V$, with an element $(u, v) \in \mathcal{T}$ representing a source-destination pair. An agent of type (u, v) starts from the location u and must get to v.

The action space A is the set of all Hamiltonian paths in the graph G. A Hamiltonian path is simply a subset of E which can be arranged into a path connecting two vertices. We will not bother to spell this out in more mathematical detail, but simply note that A is a subset of 2^E. Finally, the admissible actions to an agent of type $t = (u, v)$ is the set $C(t)$ consisting of all paths connecting u to v.

Finally, the cost function $F : \mathcal{T} \times A \times \mathcal{P}(\mathcal{T} \times A) \to \mathbb{R}$ is defined by setting

$$F(t, a, m) = \sum_{e \in a} c_e(\ell_e(m)), \quad \text{where} \quad \ell_e(m) = m\{(t', a') \in \mathcal{T} \times A : e \in a'\}.$$

Given a distribution of type-action pairs m, the value $\ell_e(m)$ is the fraction of agents who use the road e in their path, and thus it is called the *load* of e. The travel time faced by an agent of type t choosing path a is then calculated by adding, over every road used ($e \in a$), the cost incurred on that road, which is a function (c_e) of the load.

Note that agents are now seeking to *minimize* travel time. We have been maximizers in previous sections, not minimizers, but this can be accounted for of course by taking $-F$ to be our reward function.

One nice feature of congestion games is that they are always potential games. Indeed, the function

$$U(m) := \sum_{e \in E} \int_0^{\ell_e(m)} c_e(s)ds$$

can be shown to be a convex potential function. This means that minimizers of U correspond to mean field equilibria. In fact, U is sometimes even strictly convex.

An important question studied in this context pertains to the effect of network topology on the *efficiency* of Nash equilibria. A common measure of efficiency is the so-called *price of anarchy*, defined as the worst-case ratio of average travel time *in equilibrium* to the minimal possible travel time achievable by a central planner. To be somewhat more: The average travel times over a continuum of agents with type-action distribution m is given by

$$A(m) = \int_{\mathcal{T} \times A} F(t, a, m)\, m(dt, da).$$

If we let $\mathcal{M}$ denote the set of mean field equilibria with a given type distribution, then the price of anarchy is defines as

$$\mathrm{PoA} = \frac{\sup_{m \in \mathcal{M}} A(m)}{\inf_{m \in \mathcal{P}(\mathrm{Gr}(C))} A(m)}.$$

Of course, the price of anarchy is always at least 1 (and we should be careful to avoid dividing by zero). Various upper bounds are known for different kinds of networks and different restrictions on the cost functions c_e.

We will not go any deeper into the analysis of congestion games, but some references are provided in the next section.

4.6 Bibliographic notes

The idea of a game with a continuum of agents essentially originated in the work of Aumann [4] and [108]. The framework of Mas-Colell [93] is close to the one studied here, although extended to cover different *types* of agents, and focused solely on proving existence of equilibrium. See also the recent

Blanchet and Carlier [12, 13] for a reparametrization and some extensions of the framework of Mas-Colell, some discussion of potential games in the setting of an infinite action space, and an interesting connection with the theory of optimal transport.

The limit theory from n players to the continuum is not terribly well-studied in the literature, but some references include [62, 69, 24, 74, 12, 84]. The last of these references [84] also studies large deviation principles associated to this limit.

While dynamic games tend to be more popular in modern applications than the static games studied in this section, the literature on *algorithmic game theory* [100] makes heavy use of the model studied in the previous section. This class of *congestion games*, introduced by Rosenthal [107], have found widespread application in modeling traffic routing, in both physical and communications networks. There are far too many references to begin cover here, but for further references and a discussion in the context of our $n \to \infty$ limits theorems see [84].

5 Stochastic optimal control

This section is a fast-paced and not remotely comprehensive introduction to stochastic optimal control. We focus on the analytic approach, in which one identifies a PDE, known as the Hamilton-Jacobi-Bellman equation, that the value function should solve. For a more thorough and careful study of stochastic optimal control, refer to the classic texts of Fleming-Soner [50] and Yong-Zhou [118], or the more recent books of Pham [104] or Carmona [25]. The first three of these references go into some detail on the theory of viscosity solutions, and the last three include also the so-called *probabilistic approach*, via the (Pontryagin) maximum principle. Another important but less popular techniques include various *weak formulations*, in which the probability space is not specified ahead of time; see [118, 104, 25] for an approach based on backward SDEs (BSDEs). Lastly, it is worth mentioning a powerful abstract method based on compactification [48, 66], which is useful for proving existence of (but not explicitly constructing) optimal Markovian controls.

5.1 Setup

We work on finite time horizon, $T > 0$, and with a given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ that supports a d-dimensional Brownian motion W, and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ where $\mathcal{F}_t = \sigma(W_s : 0 \le s \le t)$. Agents choose actions from a closed set $A \subset \mathbb{R}^k$. We make the following assumptions throughout:

Standing assumptions

1. The drift and volatility functions, $b : \mathbb{R}^d \times A \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times A \to \mathbb{R}^{d\times m}$, are measurable and satisfy a uniform Liptschitz condition in x. That is, there exists $K > 0$ such that, for all $x, y \in \mathbb{R}^d$ and $a \in A$, we have
$$|b(x, a) - b(y, a)| + |\sigma(x, a) - \sigma(y, a)| \leq K|x - y|.$$
Here, $|\cdot|$ denotes absolute value, Euclidean norm, or Frobenius norm, depending on whether it is applied to a scalar, vector, or matrix..

2. The objective functions $f : \mathbb{R}^d \times A \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$ are continuous and bounded from above.

The *state process* X is a d-dimensional stochastic process controlled by the A valued process α and whose dynamics are given by

$$\begin{aligned} dX_t &= b(X_t, \alpha_t)dt + \sigma(X_t, \alpha_t)dW_t \\ X_0 &= x \end{aligned} \tag{5.1}$$

The controller steers the process X by choosing the control α, which is a process taking values in A. The goal of the controller is to choose α to maximize

$$\mathbb{E}\left[\int_0^T f(X_t, \alpha_t)dr + g(X_T)\right].$$

Here, f is called the *running reward/objective function*, and g is the *terminal reward/objective function*. In the rest of the section, we will be more precise about what a control α is exactly, and we will find a recipe for solving such a control problem, in the sense of computing the value and an optimal control.

Remark 5.1. These assumptions are far from necessary and mostly given for convenience, and it is worth stressing that the analysis of this section (namely, the verification theorem) hold in much broader contexts. It is is somewhat of a nuisance, in stochastic optimal control theory, that comprehensive general theorems are hard to prove and tend not to cover many natural models in practice. That said, the theorems we develop under the above restrictive assumptions are, in a sense, typical. Many models encountered in practice obey the same principles, and case-by-case proofs almost always follow essentially the same strategies outlined below.

We will consider the following two natural families of controls:

1. **Open loop:** We denote by $\mathbb{A}$ the set of $\mathbb{F}$-progressively measurable processes $\alpha = (\alpha_t)_{t\geq 0}$ such that

$$\mathbb{E}\left[\int_0^T [|b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2]dt\right] < \infty$$

   Under the above assumptions, classical theory ensures that the state equation (5.1) has a unique strong solution.

2. **Markovian controls:** $\mathbb{A}_M \subset \mathbb{A}$ consist of the set of Markovian controls. That is, $\alpha \in \mathbb{A}_M$ if $\alpha \in \mathbb{A}$ and there exists a measurable function $\hat{\alpha} : [0,T] \times \mathbb{R}^d \to A$ such that $\alpha_t = \hat{\alpha}(t, X_t)$.

One uncommon but natural alternative would be to work with path-dependent controls, of the form $\alpha_t = \hat{\alpha}(t, (X_s)_{s\in[0,t]})$. The terms *closed-loop control* and *feedback control* are variously used by different authors as synonymous with *Markovian control* or even sometimes the latter notion of path-dependent control.

5.2 Dynamic programming

The reader familiar with discrete-time optimal control has likely encountered the notion of *dynamic programming* or *backward induction*. This section develops the continuous-time version of this recursive approach to solving a dynamic optimization problem. The idea is as follows. Suppose that we know how to solve the control problem when it "starts" at a time $t \in (0,T)$ and from any initial state $x \in \mathbb{R}^d$, and we denote the resulting optimal value by $v(t,x)$. Then, to solve the control problem starting from an earlier time $s \in (0,t)$ and starting from another state $y \in \mathbb{R}^d$, we may equivalently solve a control problem on the time horizon (s,t), using the function $x \mapsto v(t,x)$ as our terminal reward function.

To make this precise, we need to explain what it means to solve the control problem starting from $(t,x) \in [0,T] \times \mathbb{R}^d$. To this end, define the state process $X^{t,x} = (X^{t,x}_s)_{s\in[t,T]}$ for any (t,x) by the SDE

$$\begin{aligned} dX^{t,x}_s &= b(X^{t,x,\alpha}_s, \alpha_s)ds + \sigma(X^{t,x}_s, \alpha_s)dW_s, \quad s \in [t,T] \\ X^{t,x}_t &= x \end{aligned} \tag{5.2}$$

To be more clear we should write $X^{t,x,\alpha}$, but it will be clear from context which control we are using.

For $\alpha \in \mathbb{A}$ and $(t,x) \in [0,T] \times \mathbb{R}^d$, define the reward functional

$$J(t,x,\alpha) := \mathbb{E}\left[\int_t^T f(X_r^{t,x}, \alpha_r)dr + g(X_T^{t,x})\right], \tag{5.3}$$

where $X^{t,x}$ is defined as in (5.2), following the control α. Under the standing assumptions, the SDE (5.2) admits a unique strong solution. Moreover, since f and g are bounded from above, the integral and expectation in (5.3) are well-defined, and $J(t,x,\alpha) < \infty$, though it is possible that $J(t,x,\alpha) = -\infty$.

Finally, define the *value function*

$$V(t,x) = \sup_{\alpha \in \mathbb{A}} J(t,x,\alpha). \tag{5.4}$$

This gives the optimal expected reward achievable starting at time t from the state x.

Theorem 5.2 (Dynamic Programming Principle). *Fix $0 \le t < s \le T$ and $x \in \mathbb{R}^d$. Then*

$$V(t,x) = \sup_{\alpha \in \mathbb{A}} \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr + V(s, X_s^{t,x})\right] \tag{5.5}$$

Proof. We first show the inequality ($\le$). Fix $\alpha \in \mathbb{A}$, and define the state process accordingly. The proof exploits the Markov or *flow property*,

$$X_u^{t,x} = X_u^{s,X_s^{t,x}}, \ t \le s \le u.$$

To make technically precises is somewhat annoying, but the intuition is clear: If we start at time t and solve the equation up to time u, this is the same as if we first solve from time t to time s and then, "starting over" at s from whatever position $X_s^{t,x}$ we reached by starting from t, we solve forward from time s to u. Using this, we find

$$\begin{aligned}
J(t,x,\alpha) &= \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr\right] + \mathbb{E}\left[\int_s^T f(X_r^{t,x}, \alpha_r)dr + g(X_T^{t,x})\right] \\
&= \mathbb{E}\left[\int_t^s f(X_r^{t,x,\alpha}, \alpha_r)dr + J(s, X_s^{t,x}, \alpha)\right] \\
&\le \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr + V(s, X_s^{t,x})\right]
\end{aligned}$$

where the second equality follows from the tower property and the flow property. The last inequality follows simply from the definition of V.

To prove the reverse inequality ($\geq$) is a bit trickier. Fix $\alpha \in \mathbb{A}$ and $\epsilon > 0$. From the definition of V we can find, for each $\omega \in \Omega$, a control $\alpha^{\epsilon,\omega} \in \mathbb{A}$ such that

$$V(s, X_s^{t,x}(\omega)) - \epsilon \leq J(s, X_s^{t,x}(\omega), \alpha^{\epsilon,\omega}).$$

That is, for every realization of $X_s^{t,x}(\omega)$, we choose $\alpha^{\epsilon,\omega}$ to be an ϵ-optimal control for the problem starting from time s at the position $X_s^{t,x}(\omega)$. Now, define a new control $\widehat{\alpha}^\epsilon$ by

$$\widehat{\alpha}_r^\epsilon(\omega) = \begin{cases} \alpha_r(\omega) & r \leq s \\ \alpha_r^{\epsilon,\omega}(\omega) & r > s. \end{cases}$$

There is delicate measurability issue here, but it can be shown using a *measurable selection theorem* that the process $\widehat{\alpha}^\epsilon$ is progressively measurable (or, more precisely, a modification is), and so lies in $\mathbb{A}$. Then, use the control $\widehat{\alpha}^\epsilon$ to define the state process $X^{t,x}$, and use the definition of the value function to find

$$\begin{aligned} V(t,x) \geq J(t,x,\widehat{\alpha}^\epsilon) &= \mathbb{E}\left[\int_t^T f(X_r^{t,x}, \widehat{\alpha}_r^\epsilon)dr + g(X_T^{t,x})\right] \\ &= \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr + \int_s^T f(X_r^{t,x}, \widehat{\alpha}_r^\epsilon)dr + g(X_T^{t,x})\right] \\ &= \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr + J(s, X_s^{t,x}, \alpha^\epsilon)\right] \\ &\geq -\epsilon + \mathbb{E}\left[\int_t^s f(X_r^{t,x}, \alpha_r)dr + V(s, X_s^{t,x})\right]. \end{aligned}$$

Indeed, the third equality followed from the flow property, and the last inequality comes from the particular choice of α^ϵ. Because ϵ was arbitrary, this completes the proof. $\square$

The dynamic programming principle (DPP) is an extremely important tool when it comes to the rigorous and general analysis of stochastic control problems. See, for example, [104, Section 3.4] for a heuristic derivation of the HJB equation from the DPP, which is made rigorous using viscosity solution theory in [104, Section 4.3]. We will focus more on verification theorems, which means we will make almost no use of the DPP, but even the most brief of overviews of stochastic control theory would be glaringly incomplete without a discussion of the DPP.

5.3 The verification theorem

Our point of view will be to identify a PDE which, if solvable in the classical sense, must coincide with the value function. We begin by reviewing the uncontrolled analogue, in which the "verification theorem" is nothing but the celebrated Feynman-Kac formula. As this can be found in any text on stochastic differential equations, we will be quite loose about precise assumptions and about dealing carefully with local martingale terms, which we will assume are true martingales.

Theorem 5.3 (Feyman-Kac)**.** *Let b and σ be "nice" coefficients. Let $X^{t,x}$ solve the SDE*

$$\begin{aligned} dX_r^{t,x} =& b(X_r^{t,x})dr + \sigma(X_r^{t,x}) \cdot dW_r \quad r \in (t,T] \\ X_t^{t,x} =& x. \end{aligned} \tag{5.6}$$

Suppose v is a smooth, i.e. $C^{1,2}([0,T],\mathbb{R})$, solution of the PDE

$$\begin{aligned} &\partial_t v(t,x) + b(x) \cdot \nabla v(t,x) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(x)\nabla^2 v(t,x)] + f(t,x) = 0 \\ &v(T,x) = g(x) \end{aligned} \tag{5.7}$$

Then v admits the representation

$$v(t,x) = \mathbb{E}\left[\int_t^T f(r, X_r^{t,x})dr + g(X_T^{t,x})\right]. \tag{5.8}$$

Proof. Apply Itô's formula to $v(r, X_r^{t,x})$ and then use the PDE to get

$$\begin{aligned} g(X_T^{t,x}) =& v(T, X_T^{t,x}) \\ &= v(t, X_t^{t,x}) + \int_t^T \nabla v(r, X_r^{t,x,\hat{\nu}}) \cdot \sigma(X_r^{t,x})dW_r \\ &\quad + \int_t^T \left(v_t(r, X_r^{t,x}) + b(t, X_r^{t,x})\nabla v(r, X_r^{t,x}) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\intercal(X_r^{t,x})\nabla^2 v(r, X_r^{t,x})]\right)dr \\ &= v(t,x) - \int_t^T f(r, X_r^{t,x})dr + \int_t^T \sigma(X_r^{t,x})\nabla v(r, X_r^{t,x,\hat{\nu}})dW_r \end{aligned}$$

Where we made use v fulfills (5.7) in the last step. Taking expectations, the last term vanishes assuming the coefficients are nice enough that the last term in the above equation is a genuine martingale. $\square$

With the Feynman-Kac formula in hand, we now see how to adapt it to the controlled setting. Essentially, the PDE in (5.7) becomes an inequality, with equality holding only along the optimal control. Or, in other words, the PDE involves a pointwise optimization over the control variable. This leads to what is known as the Hamilton-Jacobi-Bellman (HJB) equation. To ease notation, we introduce the *infinitesimal generator* of the controlled process.

Definition 5.4 (Infinitesimal generator)**.** For a smooth function $\psi : \mathbb{R}^d \to \mathbb{R}$, and for $(x, a) \in \mathbb{R}^d \times A$, define

$$L^a \psi(x) = b(x, a) \cdot \nabla_x \psi(x) + \frac{1}{2} \mathrm{Tr}[\sigma\sigma^t(x, a) \nabla^2 \psi(x)].$$

Recall in the following that the value function $V = V(t, x)$ was defined in (5.4). The following theorem is absolutely crucial for the developments to come, as it provides a powerful recipe for finding an optimal control.

Theorem 5.5 (Verification theorem)**.** *Suppose $v = v(t, x)$ is $C^{1,2}([0, T] \times \mathbb{R}^d)$ and satisfies $v(T, x) = g(x)$ along with the Hamilton-Jacobi-Bellman equation*

$$\partial_t v(t, x) + \sup_{a \in A} \{L^a v(t, x) + f(x, a)\} = 0, \tag{5.9}$$

where the operator L^a acts on the x variable of v. Assume also that there exists a measurable function $\widehat{\alpha} : [0, T] \times \mathbb{R}^d \to A$ such that

$$\widehat{\alpha}(t, x) \in \arg\max_{a \in A} [L^a v(t, x) + f(x, a)]$$

and the SDE

$$dX_t = b(X_t, \widehat{\alpha}(t, X_t))dt + \sigma(X_t, \widehat{\alpha}(t, X_t))dW_t$$

has a solution from any starting time and state. Then $v(t, x) = V(t, x)$ for all (t, x), and $\widehat{\alpha}(t, X_t)$ is an optimal control.

Proof. Fix $\alpha \in \mathbb{A}$. Apply Itô to $v(r, X_r^{t,x})$ we get

$$\begin{aligned} g(X_T^{t,x}) =& v(T, X_T^{t,x}) \\ &= v(t, X_t^{t,x}) + \int_t^T \nabla v(r, X_r^{t,x}) \cdot \sigma(X_r^{t,x}) dW_r \\ &\quad + \int_t^T \Big(\partial_t v(r, X_r^{t,x}) + L^{\alpha_r} v(r, X_r^{t,x})\Big) dr. \end{aligned}$$

Take expectations to find

$$\begin{aligned} \mathbb{E}[g(X_T^{t,x})] =& w(t,x) + \mathbb{E}\left[\int_t^T \Big(\partial_t v(r, X_r^{t,x}) + L^{\alpha_s} v(r, X_r^{t,x}) + f(X_r^{t,x}, \alpha_r)\Big) dr\right] \\ & - \mathbb{E}\left[\int_t^T f(X_r^{t,x}, \alpha_r) dr\right] \end{aligned}$$

Using the PDE, we have

$$\mathbb{E}\left[\int_t^T \Big(\partial_t v(r, X_r^{t,x}) + L^{\alpha_s} v(r, X_r^{t,x}) + f(X_r^{t,x}, \alpha_r)\Big) dr\right] \leq 0, \tag{5.10}$$

which implies

$$\mathbb{E}[g(X_T^{t,x})] \leq v(t,x) - \mathbb{E}\left[\int_t^T f(X_r^{t,x}, \alpha_r) dr\right].$$

As this holds for any $\alpha \in \mathbb{A}$, we conclude that $V \leq v$. On the other hand, if we use the particular control $\alpha_r^* = \widehat{\alpha}(r, X_r^{t,x})$, then (5.10) holds with equality, and thus

$$\mathbb{E}[g(X_T^{t,x})] = v(t,x) - \mathbb{E}\left[\int_t^T f(X_r^{t,x}, \alpha_r^*) dr\right].$$

This shows that $v(t,x) = J(t,x,\alpha^*) \leq V(t,x)$, and we conclude that $v \equiv V$. Moreover, since the chain of inequalities $V(t,x) \leq v(t,x) \leq J(t,x,\alpha^*) \leq V(t,x)$ collapses to equalities, we conclude that α^* is an optimal control. $\square$

The main technical gap in the above proof is in checking that the stochastic integral is a true martingale, which in fact it may not be even if the coefficients b and σ are nice. Indeed, we only assume that v is $C^{1,2}$, so we know at best that $\nabla v(t,x)$ is Lipschitz in x, uniformly in t. A proper proof is possible using a localization argument, and the curious reader is referred to [104] for details.

Remark 5.6. There is a vast literature on both the theory and implementation of nonlinear PDEs of the form appearing in Theorem 5.5. An important feature of these PDEs is that classical solutions are not to be expected, and a major breakthrough in the '80s was the discovery of a natural solution concept known as a *viscosity solution*, which gives meaning to how a non-differentiable function can be considered a "solution" to a PDE like (5.9). The viscosity solution concept is "natural" in the sense that it often permits existence and uniqueness results and, more importantly, enjoys strong stability properties which ensure that reasonably structured numerical methods (e.g., finite difference schemes) tend to converge [111]. Often, one can directly show that the value function as defined in (5.4) is the unique viscosity solution of the HJB equation, with no need to even check that V has any derivatives. The classic reference for the theory of viscosity solutions is the "user's guide" [38]. For an introduction more focused on applications in control theory, refer to [50, 104].

5.4 Analyzing the HJB equation

We begin with a definition, ubiquitous in the theory of stochastic control:

Definition 5.7. Let $\mathbb{S}_d$ denote the set of symmetric $d \times d$ matrices. The *Hamiltonian* of the control problem is the function $H : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}_d \to \mathbb{R} \cup \{\infty\}$ defined by

$$H(x, y, z) = \sup_{a \in A} \left(b(x, a) \cdot y + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\intercal(x, a)z] + f(x, a) \right). \tag{5.11}$$

The dummy variables y and z are called the *adjoint variables*.

The Hamiltonian encapsulates the nonlinearity of the HJB equation, which we may now write as

$$\partial_t v(t, x) + H(x, \nabla v(t, x), \nabla^2 v(t, x)) = 0.$$

Suppose we can find a maximizer $\widehat{\alpha}(x, y, z) \in A$ in the supremum in (5.11), for each choice of (x, y, z). Then, once we solve the HJB equation, our optimal control (assuming solvability of the SDE) is

$$\alpha(t, x) = \widehat{\alpha}(x, \nabla v(t, x), \nabla^2 v(t, x)).$$

The HJB equation is a *parabolic* PDE. While we will not make use of this, a crucial structural feature in the PDE theory, particularly the theory of viscosity solutions, is that $H(x, y, z)$ is monotone in z in the sense that $H(x, y, z) \le H(x, y, \tilde{z})$ whenever the matrix $\tilde{z} - z$ is positive semidefinite. Notice also that H is always convex in (y, z), as it is a supremum of affine functions.

The astute reader should notice that the Hamiltonian may take the value $+\infty$ in general, at least if the control space A is unbounded. There are two ways this is typically dealt with. First, suppose $\mathrm{dom}(H)$ denotes the (convex) set of (x, y, z) for which $H(x, y, z) < \infty$. Suppose we can find a (smooth) solution v of the HJB equation such that $(x, \nabla v(t, x), \nabla^2 v(t, x)) \in \mathrm{dom}(H)$ for every (t, x). That is, we find a solution that "avoids" the bad points at which $H = \infty$. Then the verification theorem 5.5 can still be applied. A second and more robust approach to this issue is to introduce an auxiliary real-valued continuous function G such that $H < \infty \iff G \ge 0$, and then turn the HJB equation into a variational inequality; see [104, pp. 45-46].

Let us now focus on a few particular cases in which the HJB equation simplifies.

5.4.1 Uncontrolled volatility

Suppose that $\sigma = \sigma(x)$ is not controlled. Then one often omits the z argument and its corresponding term from the definition (5.11) of the Hamiltonian, instead defining

$$H(x, y) = \sup_{a \in A} \Big(b(x, a) \cdot y + f(x, a) \Big),$$

which is sometimes called the *reduced Hamiltonian*. This is justified by the fact that the omitted term $\frac{1}{2}\sigma\sigma^\intercal(x)z$ would not influence the optimization. Of course, this removed term must be added back to get the correct HJB equation:

$$\partial_t v(t, x) + H(x, \nabla v(t, x)) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^t(x)\nabla^2 v(t, x)] = 0.$$

In this case, the PDE is linear in its highest-order derivative $\nabla^2 v$, and the PDE is said to be *semi-linear*. This is in contrast with the general case of controlled volatility problems, for which the HJB equation is *fully nonlinear*.

5.4.2 Linear-convex control

Suppose $b(x, \alpha) = \alpha$ and $\sigma \equiv I$, where I denotes the $d \times d$ identity matrix; here $d = m$, meaning the state process and Brownian motion have the same dimension. The second order derivative term in the HJB equation now becomes simply

$$\frac{1}{2}\mathrm{Tr}[\sigma\sigma^t(x)\nabla^2 v(t,x)] = \frac{1}{2}\Delta v(t,x),$$

where $\Delta = \sum_{i=1}^d \partial_{x_i x_i}$ denotes the Laplacian operator.

Suppose also that the action space $A = \mathbb{R}^d$ is the whole space, and let us denote $f = -L$ for convenience. Then the (reduced) Hamiltonian becomes

$$H(x,y) = \sup_{a\in\mathbb{R}^d}\Big(a\cdot y - L(x,a)\Big).$$

In other words, for fixed x, the function $y \mapsto H(x,y)$ is the *Legendre transform* or *convex conjugate* of the function $y \mapsto L(x,y)$. If L is differentiable in a, then the optimal $a \in \mathbb{R}^d$ satisfies $y = \nabla_a L(x,a)$. We can say even more if $a \mapsto L(x,a)$ is assumed to be strictly convex and continuously differentiable. Indeed, then the function $a \mapsto \nabla_a L(x,a)$ is one-to-one, and its inverse function is precisely $y \mapsto \nabla_y H(x,y)$ (see Corollaries 23.5.1 and 26.3.1 of [106]). Hence, the optimizer in the Hamiltonian is $\widehat{\alpha}(x,y) = \nabla_y H(x,y)$.

For an even more special case, assume now that

$$f(x,a) = -\frac{1}{2}|a|^2 - F(x)$$

for some function F. Then

$$\begin{aligned} H(x,y) &= \sup_{a\in\mathbb{R}^d}\Big(a\cdot y - \frac{1}{2}|a|^2\Big) - F(x) \\ &= \frac{1}{2}|y|^2 - F(x), \end{aligned}$$

and the optimizer is $\widehat{\alpha}(x,y) = y$. The HJB equation becomes

$$\begin{aligned} \partial_t v(t,x) + \frac{1}{2}|\nabla v(t,x)|^2 + \frac{1}{2}\Delta^2 v(t,x) &= F(x) \\ v(t,x) &= g(x). \end{aligned}$$

5.5 Solving a linear quadratic control problem

This section and the next work through explicit solutions of some control problems, starting with a linear-quadratic model. In a linear-quadratic model, the state and control variables appear in a linear fashion in the coefficients b and σ for the state process and in a quadratic fashion in the objective functions f and g. We will not address the most general linear-quadratic model, but a simple special case to illustrate some of the machinery developed above.

Our controlled state process is

$$dX_t = \alpha_t dt + dW_t,$$

where α takes values in the action space $A = \mathbb{R}^d$. The objective function is

$$\mathbb{E}\left[\int_0^T -\frac{1}{2}|\alpha_t|^2 dt - \frac{\lambda}{2}|X_T - z|^2\right],$$

where $\lambda > 0$ is a cost parameter and $z \in \mathbb{R}^d$ is a (deterministic) target. The goal of the controller is to steer X to be as close as possible to z at the final time but also to do so using as little "energy" or "fuel" as possible. The parameter λ controls the relative importance of hitting the target versus conserving energy.

Here, our coefficients are of course

$$b(x, a) = a, \quad f(x, a) = -\frac{1}{2}|a|^2, \quad g(x) = -\frac{\lambda}{2}|x - z|^2.$$

To solve this control problem, we first write the (reduced) Hamiltonian,

$$H(x, y) = \sup_{a \in \mathbb{R}^d}\left(a \cdot y - \frac{1}{2}|a|^2\right) = \frac{1}{2}|y|^2.$$

The maximizer, as we saw in Section 5.4.2, is $\widehat{\alpha}(x, y) = y$. The HJB equation becomes

$$\begin{aligned}
&\partial_t v(t, x) + \frac{1}{2}|\nabla v(t, x)|^2 + \frac{1}{2}\Delta v(t, x) = 0 \\
&v(T, x) = -\frac{\lambda}{2}|x - z|^2.
\end{aligned}$$

Once we solve this equation, the optimal control is nothing but $\widehat{\alpha}(x, \nabla v(t, x)) = \nabla v(t, x)$.

We look for a separable solution, of the form $v(t, x) = f(t)\psi(x) + g(t)$, for some functions f, g, and ψ to be determined. Of course, the terminal condition then implies $\varphi(T)\psi(x) + \tilde{\varphi}(T) = -\frac{\lambda}{2}|x - z|^2$ for all x, so we might as well take

$$v(t, x) = \frac{1}{2}f(t)|x - z|^2 + g(t).$$

To make use of this *ansatz*, we differentiate v and plug it into the HJB equation. The relevant derivatives are

$$\begin{aligned} v_t(t,x) &= \frac{1}{2}f'(t)|x-z|^2 + g'(t) \\ \nabla v(t,x) &= f(t)(x-z) \\ \Delta v(t,x) &= df(t), \end{aligned}$$

where we recall that d is the dimension of the state process. We also impose the boundary conditions $F(T) = -\lambda$ and $g(T) = 0$, to match the HJB boundary condition. Plugging in these derivatives to the PDE, we find

$$\frac{1}{2}f'(t)|x-z|^2 + g'(t) + f(t)^2|x-z|^2 + \frac{1}{2}df(t) = 0.$$

Combining the coefficients multiplied by $|x-z|^2$ and, separately, those with no factor of $|x-z|^2$, we find that f and g must satisfy the following ODEs:

$$\begin{aligned} f'(t) + f^2(t) = 0, &\quad f(T) = -\lambda \\ g'(t) + \frac{d}{2}f(t) = 0, &\quad g(T) = 0 \end{aligned} \tag{5.12}$$

The first one is straightforward to solve:

$$T - t = \int_t^T 1ds = \int_t^T -\frac{f'(t)}{f^2(t)}ds = \frac{1}{f(T)} - \frac{1}{f(t)} = -\frac{1}{\lambda} - \frac{1}{f(t)}.$$

Rearrange to find

$$f(t) = -\frac{1}{\frac{1}{\lambda} + T - t}, \quad t \in [0,T].$$

Plugging this in for g one gets

$$g(t) = g(T) + \frac{d}{2}\int_t^T f(s)ds = \frac{d}{2}\log(1 + \lambda(T-t)).$$

Finally, we conclude that the HJB admits the smooth solution

$$v(t,x) = -\frac{|x-z|^2}{2(\frac{1}{\lambda} + T - t)} + \frac{d}{2}\log(1 + \lambda(T-t)).$$

As remarked above, the optimal control is

$$\nabla v(t,x) = -\frac{|x-z|}{\frac{1}{\lambda} + T - t)},$$

and we may write the optimally controlled state process as

$$dX_t = \frac{z - X_t}{\frac{1}{\lambda} + T - t} dt + dW_t.$$

This drift coefficient is Lipschitz in X_t, uniformly in the other variables, and so this SDE has a unique solution. Finally, the verification theorem 5.5 ensures that we have truly solved the control problem.

The solution makes intuitive sense; the optimal control is mean-reverting, with the control pushing us always toward the target z. As we approach the time horizon T, the speed of mean-reversion becomes stronger.

It is interesting to note what happens in the $\lambda \to \infty$ case. Intuitively, this corresponds to a "hard constraint," in the sense that the controller must hit exactly the target z at the final time. The resulting state process is

$$dX_t = \frac{z - X_t}{T - t} dt + dW_t.$$

There is a singularity as $t \uparrow T$, with an "infinitely strong" push toward z. Indeed, this process is nothing but a Brownian bridge ending at $X_T = z$.

Remark 5.8. Much more general forms of linear-quadratic models are tractable. Indeed, a similar strategy applies to consider control problems of the form

$$\begin{aligned} dX_t &= (b_t^1 + b_t^2 X_T + b_t^3 \alpha_t)dt + (\sigma_t^1 + \sigma_t^2 X_t + \sigma_t^3 \alpha_t)dW_t, \\ \sup_\alpha \mathbb{E}&\left[\int_0^T \left(X_t^\intercal f_t^1 X_t + \alpha_t^\intercal f_t^2 \alpha)_t + X_t^\intercal f_t^3 \alpha_t\right) dt + X_T^\intercal g^1 X_T + g^2 \cdot X_T\right], \end{aligned}$$

for constant or time-dependent coefficients of the appropriate dimension. These models are approached with a similar quadratic ansatz for the value function. The resulting ODEs (5.12) become significantly more complicated, in general becoming a *Riccati equation*. These do not always admit completely explicit solutions, but the solution theory is nonetheless well-understood. Morally speaking, solvability of a linear-quadratic control problem is equivalent to solvability of a Riccati equation.

5.6 The Hopf-Cole transform

We may semi-explicitly solve a more general class of models than what appeared in the previous section, and this is an important class of models for which the HJB is solvable by means of the so-called Hopf-Cole transform.

With the action space $A = \mathbb{R}^d$, consider the problem

$$\begin{cases} \sup_\alpha \mathbb{E}\left[g(X_T) - \frac{1}{2}\int_0^T |\alpha_t|^2 dt\right] \\ dX_t = \alpha_t dt + dW_t. \end{cases}$$

Note that the specification $g(x) = -\frac{\lambda}{2}|x-z|^2$ recovers the linear-quadratic model of the previous section. First, we identify the Hamiltonian again as

$$H(x,y) = \sup_{a \in \mathbb{R}^d}\left(a \cdot y - \frac{1}{2}|a|^2\right) = \frac{1}{2}|y|^2,$$

with maximizer $\hat{\alpha}(x,y) = y$. Then, the HJB equation is

$$\partial_t v(t,x) + \frac{1}{2}|\nabla v(t,x)|^2 + \frac{1}{2}\Delta v(t,x) = 0, \tag{5.13}$$

with boundary condition $v(T,x) = g(x)$. This is a nonlinear pde for $v(t,x)$, but it can be solved with a change of variable. We set $u(t,x) = e^{v(t,x)}$ and identify the partial derivatives

$$\begin{aligned} \partial_t u(t,x) &= e^{v(t,x)}\partial_t v(t,x) \\ \partial_{x_i} u(t,x) &= e^{v(t,x)}\partial_{x_i} v(t,x) \\ \partial_{x_i x_i} u(t,x) &= e^{v(t,x)}\left((\partial_{x_i} v(t,x))^2 + \partial_{x_i x_i} v(t,x)\right), \end{aligned}$$

where the last two can be aggregated into vector form as

$$\begin{aligned} \nabla u(t,x) &= e^{v(t,x)}\nabla v(t,x) \\ \Delta u(t,x) &= e^{v(t,x)}\left(|\nabla v(t,x)|^2 + \Delta v(t,x)\right). \end{aligned}$$

We multiply (5.13) by $e^{v(t,x)}$ to obtain

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) = 0, \tag{5.14}$$

with terminal condition $u(T,x) = e^{g(x)}$. Equation 5.14 is nothing but the *heat equation* which has solution equal to

$$u(t,x) = \mathbb{E}\left[e^{g(W_T)} \mid W_t = x\right] = \int_{\mathbb{R}^d} e^{g(y)} p(T-t, y-x)dy,$$

where we recall that W_t is a d-dimensional Brownian motion, we define its transition density (the heat kernel) by

$$p(s,z) = (2\pi s)^{-d/2} e^{-\frac{|z|^2}{2s}}.$$

In other words, the solution of (5.13) is

$$v(t,x) = \log \int_{\mathbb{R}^d} e^{g(y)} p(T-t, y-x)dy.$$

5.7 The Merton problem

Given a risky asset with price S_t and a risk-free asset with price B_t with dynamics

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$
$$\frac{dB_t}{B_t} = r dt,$$

suppose that we can allocate a fraction α_t of our wealth at time t to the risky asset and the remaining fraction $1 - \alpha_t$ to to the risk-free asset. In this setting, we allow short sales and borrowing, that is, $A = \mathbb{R}$. Assuming $\mathbb{E}\left[\int_0^T |\alpha_t|^2 dt\right] < \infty$, the *wealth* process X_t with initial wealth $X(0) = x_0 > 0$ evolves according to

$$\frac{dX_t}{X_t} = (\alpha_t \mu + (1 - \alpha_t) r)\, dt + \alpha_t \sigma dW_t. \tag{5.15}$$

Given a utility function $U(x)$, the objective of the Merton problem is

$$\sup_{\alpha} \mathbb{E}\left[U(X_T)\right].$$

For exponential, logarithmic and power utilities the Merton problem can be solved explicitly. Here, we focus on the power utility, $U(x) = \frac{1}{p} x^p$, with $p < 1$ and $p \neq 0$.

We identify the Hamiltonian as

$$H(x, y, z) = \sup_{a \in \mathbb{R}} \left[(r + (\mu - r)a) xy + \frac{1}{2}\sigma^2 x^2 z a^2\right].$$

The function we want to maximize is a quadratic function on a whose shape depends on the sign of z. First order conditions quickly imply that

$$H(x, y, z) = \begin{cases} rxy - \frac{(\mu - r)^2 y}{2\sigma^2 z} & \text{if } z < 0 \\ \infty & \text{if } z \geq 0, \end{cases}$$

with maximizer $\widehat{\alpha}(x, y, z) = -\frac{(\mu - r) y}{\sigma^2 x z}$ when $z < 0$. Observe that this is the first example in which we encounter a Hamiltonian which is not always finite. To avoid this inconvenience, we will proceed assuming that $z < 0$ and check in the end that $\partial_{xx} v(t, x) < 0$.

The HJB equation is then

$$\partial_t v(t,x) + rx\partial_x v(t,x) - \frac{(\mu - r)^2(\partial_x v(t,x))^2}{2\sigma^2 \partial_{xx} v(t,x)} = 0. \tag{5.16}$$

To solve this equation we use as ansatz a separable solution $v(t,x) = f(t)h(x)$. In light of the boundary condition $v(T,x) = U(x) = \frac{1}{p}x^p$, we try $h(x) = \frac{1}{p}x^p$. That is, our ansatz is $v(t,x) = \frac{1}{p}f(t)x^p$, for a function $f = f(t)$ to be determined but satisfying the boundary condition $f(T) = 1$. The partial derivatives of v become

$$\begin{aligned}
\partial_t v(t,x) &= \frac{1}{p}f'(t)x^p \\
\partial_x v(t,x) &= f(t)x^{p-1} \\
\partial_{xx} v(t,x) &= (p-1)f(t)x^{p-2}.
\end{aligned}$$

Plugging these into the HJB equation (5.16), we find (miraculously) that every term has a factor of x^p:

$$x^p \left[\frac{1}{p}f'(t) + rf(t) - \frac{(\mu - r)^2 f(t)}{2\sigma^2(p-1)}\right] = 0.$$

Cancelling x^p and multiplying by p, this implies

$$f'(t) + \gamma f(t) = 0,$$

where $\gamma = rp - \frac{p(\mu-r)^2}{2\sigma^2(p-1)}$. It follows that $f(t) = e^{\gamma(T-t)}$, since $f(T) = 1$. Therefore $v(t,x) = \frac{1}{p}x^p e^{\gamma(T-t)}$ solves the HJB equation (5.16). Taking derivatives with respect to x, we obtain

$$\begin{aligned}
\partial_x v(t,x) &= x^{p-1}e^{\gamma(T-t)} \\
\partial_{xx} v(t,x) &= (p-1)x^{p-2}e^{\gamma(T-t)},
\end{aligned}$$

and we confirm that indeed $\partial_{xx} v(t,x) < 0$ since $p < 1$. We conclude by computing the optimal control

$$\begin{aligned}
\alpha^*(t,x) &= \widehat{\alpha}\left(x, \partial_x v(t,x), \partial_{xx} v(t,x)\right) \\
&= \frac{\mu - r}{(1-p)\sigma^2},
\end{aligned}$$

which turns out to be constant. This means that the proportion of our wealth invested in the risky asset should always be the same during the entire investment horizon. To implement this trading strategy would of course require continuous rebalancing at each point in time, as the stock price fluctuates.

It is a good exercise to check that the same solution arises in the limiting case $p = 0$, using the utility function $U(x) = \log x$. On the other hand, with an exponential utility $U(x) = -\exp(-px)$, the solution is quite different, and one should work with the *absolute amount of wealth* α_t invested in the stock at each time, as opposed to the *fraction of wealth.* This makes removes the denominator of X_t on the left-hand side of the state equation (5.15), and it is a good exercise to find the optimal control for this model. It turns out again to be constant, but now with the interpretation that a constant *amount* (not proportion) of wealth is invested in the risky asset at any given time.

5.8 The infinite horizon problem

While we focus most of the course on models with a finite time horizon $T > 0$, it is worth explaining what happens in infinite horizon problems. These are popular because the HJB equation loses its dependence on the time variable t, which happens intuitively because at any fixed time t the remaining optimization problem after t "looks the same."

To warm up, let us consider the uncontrolled case. For $x \in \mathbb{R}^d$, consider the SDE

$$dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dW_t,$$

for some nice coefficients b and σ. Define

$$v(x) = E\left[\int_0^\infty e^{-\beta t} f(X_t^x)dt\right],$$

for some $\beta > 0$ and some nice function f. Let L denote the infinitesimal generator of the SDE, which acts on smooth functions φ by

$$L\varphi(x) = b(x)\cdot\nabla\varphi(x) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\intercal(x)\nabla^2\varphi(x)].$$

We claim that if w is a smooth solution of

$$-\beta w(x) + Lw(x) + f(x) = 0,$$

satisfying the growth assumption

$$\lim_{T\to\infty} e^{-\beta T}\mathbb{E}\left[w(X_t^x)\right] = 0,$$

for all $x \in \mathbb{R}^d$, then we have $w(x) = v(x)$. To see this, apply Itô's formula to $e^{-\beta t} w(X_t^x)$ to get

$$d(e^{-\beta t} w(X_t^x)) = \left[-\beta e^{-\beta t} w(X_t^x) + e^{-\beta t} Lw(X_t^x\right])dt \\ + e^{-\beta t} \nabla w(X_t^x) \cdot \sigma(X_t^x) dW_t.$$

For each $T > 0$, we may integrate and take expectations (assuming the stochastic integral is a true martingale) to get

$$\begin{aligned} \mathbb{E}\left[e^{-\beta T} w(X_T^x)\right] &= w(x) + \mathbb{E}\left[\int_0^T e^{-\beta t}(-\beta w(X_t^x) + Lw(X_t^x))dt\right] \\ &= w(x) - \mathbb{E}\left[\int_0^T e^{-\beta t} F(X_t^x) dt\right], \end{aligned}$$

where the last line used the PDE. Send $T \to \infty$, using the growth assumption, to find

$$w(x) = \mathbb{E}\left[\int_0^T e^{-\beta t} F(X_t^x) dt\right] = v(x).$$

In the controlled case, we end up with a similar PDE for the value function, but with a supremum over the controls. Consider the infinite horizon optimal control problem,

$$\begin{cases} \sup_\alpha \mathbb{E}\left[\int_0^\infty e^{-\beta t} f(X_t^x, \alpha_t) dt\right] \\ dX_t^x = b(X_t^x, \alpha_t)dt + \sigma(X_t^x, \alpha_t) dW_t \\ X_0^x = x, \end{cases}$$

where $\beta > 0$ is a constant. Let $J(x, \alpha) = E\left[\int_0^\infty e^{-\beta t} f(X_t^x, \alpha_t) dt\right]$ and define the value function

$$v(x) = \sup_\alpha J(x, \alpha).$$

Now, suppose that w is a smooth solution of the HJB equation

$$-\beta w(x) + \sup_{a \in A}\left\{b(x, a) \cdot \nabla\phi(x) + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^\top(x, a)\nabla^2\phi(x)\right] + f(x, a)\right\} = 0,$$

for all $x \in \mathbb{R}^d$ and all $a \in A$. Furthermore, assume that

$$\lim_{T \to \infty} e^{-\beta T}\mathbb{E}\left[w(X_t^x)\right] = 0,$$

for all $x \in \mathbb{R}^d$ and all controls α. Then, we have $w(x) = v(x)$. It is left as an exercise to sketch the proof, analogous to that of Theorem 5.5.

5.9 Bibliographic notes

TBD

6 Two-player zero-sum games

To warm up for our discussion of n-player stochastic games, we first focus on the important special case of two-player zero-sum games. To fix some ideas, we briefly review two-player zero-sum games in the static, deterministic case.

6.1 Static games

Suppose that we have two players with action sets A and B and an objective function $F : A \times B \to \mathbb{R}$. We say that the game is a *zero-sum* game if the objective of player A is to maximize F and the one of player B is to maximize $-F$ (or equivalently minimize F). In other words, the game is zero-sum if the two players' rewards sum to zero. In this setting, a Nash equilibrium is a pair $(a^*, b^*) \in A \times B$ such that

$$\begin{cases} F(a^*, b^*) = \sup_{a \in A} F(a, b^*) \\ -F(a^*, b^*) = \sup_{b \in B} -F(a^*, b). \end{cases}$$

This condition is equivalent to

$$\inf_{b \in B} F(a^*, b) = F(a^*, b^*) = \sup_{a \in A} F(a, b^*).$$

Observe that we also have the inequalities

$$\begin{aligned} \inf_{b \in B} F(a^*, b) &\leq \sup_{a \in A} \inf_{b \in B} F(a, b) \\ \sup_{a \in A} F(a, b^*) &\geq \inf_{b \in B} \sup_{a \in A} F(a, b). \end{aligned}$$

Since it is also true that $\sup_{a \in A} \inf_{b \in B} F(a, b) \leq \inf_{b \in B} \sup_{a \in A} F(a, b)$, we conclude that (a^*, b^*) satisfies

$$F(a^*, b^*) = \sup_{a \in A} \inf_{b \in B} F(a, b) = \inf_{b \in B} \sup_{a \in A} F(a, b). \tag{6.1}$$

This means that the Nash equilibrium is a *saddle point* of F. In this context, there is an interesting interpretation of the Nash equilibrium for two players. Since $\inf_{b \in B} F(a^*, b) = \sup_{a \in A} \inf_{b \in B} F(a, b)$, we have that a^* maximizes the function $\inf_{b \in B} F(a, b)$. This means that player A maximizes his/her worst case performance againt player B. There is an analogous interpretation for player B.

When a Nash equilibrium exists, the common value in (6.1) is called the *value* of the game. Note that there may be multiple Nash equilibria, but there is only one value. On the contrary, the game *does not have value* if[6]

$$\sup_{a \in A} \inf_{b \in B} F(a, b) < \inf_{b \in B} \sup_{a \in A} F(a, b).$$

These quantities are sometimes called the lower and upper values, respectively. The left-hand side represents the value when player B has the advantage, with player A choosing a first and then player B getting to react to player A's choice.

6.2 Stochastic differential games

Moving on to the stochastic setting, we suppose that two players each control a common d-dimensional state process X, which evolves according to

$$dX_t = b(X_t, \alpha_t, \beta_t)dt + \sigma(X_t, \alpha_t, \beta_t)dW_t,$$

where $X_0 = x$ and W is an m-dimensional Brownian motion. Similarly to the stochastic control framework, the objective function takes the form

$$J(\alpha, \beta) = \mathbb{E}\left[\int_0^T f(X_t, \alpha_t, \beta_t)dt + g(X_T)\right].$$

Players A and B choose control processes α and β taking values in subsets A and B of Euclidean spaces, respectively. As in the deterministic game, the objective of player A is to maximize $J(\alpha, \beta)$, while the objective of player B is to minimize $J(\alpha, \beta)$. We now define the *value* of a two-player game.

Definition 6.1. The game has *value* if

$$\sup_{\alpha \in \mathcal{A}} \inf_{\beta \in \mathcal{B}} J(\alpha, \beta) = \inf_{\beta \in \mathcal{B}} \sup_{\alpha \in \mathcal{A}} J(\alpha, \beta).$$

Intuitively, in a game with *value*, it does not matter which player "plays" first and which one "plays" second.

It is an absolutely crucial point in stochastic differential games to define carefully what one means by an *admissible control*. The three most common choices are:

[6]Note that we always have $\sup\inf \leq \inf\sup$ here.

1. **Open loop**: Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by the Brownian motion. Player A chooses an $\mathcal{A}$-valued $\mathbb{F}$-adapted process $\alpha = (\alpha_t(\omega))$. Similarly for player B.

2. **Closed loop (Markovian)**: Player A chooses a (measurable) function $\alpha : [0,T]\times\mathbb{R}^d \to A$. In this case $\alpha(t, X_t(\omega))$ is the control process. The function $\alpha(t,x)$ is called the *feedback* function. Similarly for player B.

3. **Closed loop (path dependent)**: Player A chooses a (measurable) function $\alpha : [0,T] \times C([0,T];\mathbb{R}^d) \to A$ with the following adaptedness property: For each $t \in [0,T]$ and each $x, y \in C([0,T];\mathbb{R}^d)$ satisfying $x_s = y_s$ for all $s \leq t$, we have $\alpha(t,x) = \alpha(t,y)$. Intuitively, unlike the Markovian case, the control may depend on the entire history of the state process, and this adaptedness constraint simply means that one cannot look into the future. Similarly for player B.

Note that a closed loop Markovian control (and similarly for a closed loop path dependent control) always gives rise to an open loop control. Indeed, if $\alpha : [0,T] \times \mathbb{R}^d \to A$ and $\beta : [0,T] \times \mathbb{R}^d \to B$ are closed loop Markovian controls, then the state process is determined by solving the SDE

$$dX_t = b(X_t, \alpha(t,X_t), \beta(t,X_t))dt + \sigma(X_t, \alpha(t,X_t), \beta(t,X_t))dW_t, \quad X_0 = x,$$

and let us assume for this discussion that the SDE is well-posed. Then $\tilde{\alpha}_t(\omega) := \alpha(t, X_t(\omega)$ and $\tilde{\beta}_t(\omega) := \beta(t, X_t(\omega)$ both define open loop controls.

With this in mind, in a one-player optimal control problem, it typically does not make a difference whether we use open loop or (the smaller class of) closed loop controls, as the optimizer among open loop controls is often representable in closed loop form. However, in games, the choice of admissibility class influences the equilibrium outcome. Intuitively, the key point is that, in the notation of the previous paragraph, the control $\tilde{\alpha}$ *as a process* depends on the choice of the other player! If Player B switches to a different closed loop control $\beta' : [0,T] \times \mathbb{R}^d \to B$ while player A keeps the same control $\alpha : [0,T] \times \mathbb{R}^d \to A$, then we must resolve the state equation

$$dX'_t = b(X'_t, \alpha(t,X'_t), \beta'(t,X'_t))dt + \sigma(X'_t, \alpha(t,X'_t), \beta'(t,X'_t))dW_t, \quad X'_0 = x.$$

This gives rise a different state process, which is then fed into the function α, and the control *process* of player A becomes $\tilde{\alpha}'_t(\omega) = \alpha(t, X'_t(\omega))$. In the open loop regime, this feedback is not present; if player B switches controls, then the control process of player A does not react to this change, because it is the process and not the feedback function which is fixed. This extra layer of feedback in closed loop controls gives rise to a different equilibrium set, and we will see this more clearly in an example.

From now on, assume that we are in the setting of closed loop (Markovian) controls. We will use an optimal-response argument to solve the two-player game. First, suppose that player B chooses $\beta(t,x)$. Then, player A solves the following problem:

$$\begin{cases} \sup_\alpha \mathbb{E}\left[\int_0^T f(X_t,\alpha_t,\beta(t,X_t))dt + g(X_T)\right] \\ dX_t = b(X_t,\alpha_t,\beta(t,X_t)) + \sigma(X_t,\alpha_t,\beta(t,X_t))dW_t. \end{cases}$$

Given β, let $v^\beta(t,x)$ be the value function of player A. In this case, the HJB equation that $v^\beta(t,x)$ should solve is

$$\partial_t v^\beta(t,x) + \sup_{\alpha\in A} h(x,\nabla v^\beta(t,x),\nabla^2 v^\beta(t,x),\alpha,\beta(t,x)) = 0,$$

with terminal $v^\beta(T,x) = g(x)$ and where

$$h(x,y,z,\alpha,\beta) = b(x,\alpha,\beta)\cdot y + \frac{1}{2}\mathrm{Tr}\left[\sigma\sigma^\top(x,\alpha,\beta)z\right] + f(x,\alpha,\beta).$$

We can then find the optimal $\alpha(t,x)$ by maximizing h pointwise above. Similarly, if player A chooses $\alpha(t,x)$, denote the value function of player B as $v_\alpha(t,x)$. The HJB equation for $v_\alpha(t,x)$ is

$$\partial_t v_\alpha(t,x) + \inf_{\beta\in B} h(x,\nabla v_\alpha(t,x),\nabla^2 v_\alpha(t,x),\alpha(t,x),\beta) = 0,$$

with $v_\alpha(T,x) = g(x)$. Again, the optimal $\beta(t,x)$ is the pointwise minimizer of h. Now suppose that the pair (α,β) is Nash. In that case, both $v^\beta(t,x)$ and $v_\alpha(t,x)$ satisfy the same PDE

$$\partial_t v(t,x) + h(x,\nabla v(t,x),\nabla^2 v(t,x),\alpha(t,x),\beta(t,x)) = 0,$$

and thus, by the Feynman-Kac representation, we must have $v \equiv v^\beta \equiv v_\alpha$. We must then have

$$\sup_{\alpha\in A} h(x,\nabla v(t,x),\nabla^2 v(t,x),\alpha,\beta(t,x)) = \inf_{\beta\in B} h(x,\nabla v(t,x),\nabla^2 v(t,x),\alpha(t,x),\beta),$$

which in turn implies that $(\alpha(t,x),\beta(t,x))$ is a *saddle point* for the function $(\alpha,\beta)\to h(x,\nabla v(t,x),\nabla^2 v(t,x),\alpha,\beta)$.

Before stating the verification theorem, we introduce some notation. Define the functions H^+ and H^- as

$$H^+(x,y,z) = \inf_{\beta\in B}\sup_{a\in A} h(x,y,z,\alpha,\beta)$$
$$H^-(x,y,z) = \sup_{\alpha\in A}\inf_{b\in B} h(x,y,z,\alpha,\beta).$$

Suppose also that there are value functions v^+ and v^- that solve

$$\partial_t v^\pm(t,x) + H^\pm(x,\nabla v^\pm(t,x),\nabla^2 v^\pm(t,x)) = 0.$$

The key condition that will ensure existence of value for the game is *Isaacs' condition*, which amounts to requiring that the static two-player zero-sum game determined by the reward function $(\alpha,\beta)\to h(x,y,z,\alpha,\beta)$ has value for each choice of (x,y,z).

Definition 6.2. We say *Isaacs' condition* holds if $H^+ \equiv H^-$.

Theorem 6.3 (Verification theorem)**.** *Assume that Isaacs' condition holds. Assume also that there is a v which is a smooth solution of*

$$\partial_t v(t,x) + H(x,\nabla v(t,x),\nabla^2 v(t,x)) = 0,$$

with terminal condition $v(T,x) = g(x)$. Suppose α and β are measurable functions from $[0,T]\times\mathbb{R}^d$ into A and B, respectively, and that $(\alpha(t,x),\beta(t,x))$ is a saddle point for the function $(\alpha,\beta)\to h(x,\nabla v(t,x),\nabla^2 v(t,x),\alpha,\beta)$, for each $(t,x)\in[0,T]\times\mathbb{R}^d$. If the state equation

$$dX_t = b(X_t,\alpha(t,X_t),\beta(t,X_t))dt + \sigma(X_t,\alpha(t,X_t),\beta(t,X_t))dW_t,$$

is well-posed, then (α,β) is a closed loop Nash equilibrium.

6.3 Buckdahn's example

This section describes a simple two-player zero-sum stochastic differential game which does not have value in open loop controls, but it does have a closed loop Nash equilibrium. Suppose our state process and Brownian motion are each two-dimensional, and write $X = (X^1,X^2)$ and $W = (W^1,W^2)$. The states evolve according to

$$dX_t^1 = \alpha_t dt + \sigma dW_t^1, \quad X_0^1 = x, dX_t^2 \qquad = \beta dt + \sigma dW_t^2, \quad X_0^1 = x.$$

Players each choose controls from the set $A = B = [-1, 1]$. The objective function is

$$J(\alpha, \beta) = \mathbb{E}[|X_T^1 - X_T^2|]$$

Intuitively, the maximizing player A wants to make the difference between states as great as possible, whereas player B wants to keep the states close together.

Let us begin by searching for closed loop equilibrium, following the blueprint of Theorem 6.3. Since volatility is uncontrolled, we define the reduced (unoptimized) Hamiltonian:

$$h(y^1, y^2, \alpha, \beta) = \alpha y^1 + \beta y^2$$

The upper and lower Hamiltonians are

$$\begin{aligned} H^+(x, y, z) &= \inf_{\beta \in [-1,1]} \sup_{\alpha \in [-1,1]} h(y^1, y^2, \alpha, \beta) = |y^1| - |y^2|, \\ H^-(x, y, z) &= \sup_{\alpha \in [-1,1]} \inf_{\beta \in [-1,1]} h(y^1, y^2, \alpha, \beta) = |y^1| - |y^2|. \end{aligned}$$

Hence, Isaacs' condition condition holds, and the game has value as long as the HJB equation is solvable:

$$\begin{aligned} &\partial_t v(t, x_1, x_2) + |\partial_{x_1} v(t, x_1, x_2)| - |\partial_{x_2} v(t, x_1, x_2)| \\ &\qquad + \frac{\sigma^2}{2}\Big(\partial_{x_1 x_1} v(t, x_1, x_2) + \partial_{x_2 x_2} v(t, x_1, x_2)\Big) = 0, \\ &v(T, x_1, x_2) = |x_1 - x_2|. \end{aligned}$$

It is worth noting that the presence of (non-differentiable) absolute value terms suggests that this PDE will not have a classical solution. While we will not go into it, viscosity solution theory will work well here, and this PDE is indeed "solvable enough" that the game can be shown to have value over closed loop controls. Moreover, the saddle point of the function $(\alpha, \beta) \mapsto h(y^1, y^2, \alpha, \beta)$ is clearly

$$\alpha(y^1) = \text{sign}(y^1), \qquad \beta(y^2) = \text{sign}(y^2)$$

Controls of this form are often called *bang-bang* controls, as they jumps to extreme end of control set whenever the sign flips. Bang-Bang controls arise quite often when there is no running cost and there is a bounded action space.

Next, we study the open loop case, and we show that there is no value. This is stated precisely as follows:

Theorem 6.4. *Let*

$$\underline{V}_0 = \sup_\alpha \inf_\beta J(\alpha,\beta), \qquad \overline{V}_0 = \inf_\beta \sup_\alpha J\alpha,\beta.$$

in open-loop controls. If $0 \leq \sigma < \frac{1}{2}\sqrt{\pi T}$*, then* $\underline{V}_0 < \overline{V}_0$

Proof. First we bound $\underline{V}_0$ from above. To do this, fix α. Choose $\beta = \alpha$. Then

$$\begin{aligned}
J(\alpha,\beta) = \mathbb{E}[|X_T^1 - X_T^2|] &= \mathbb{E}\left[\int_0^T (\alpha_t - \beta_t)dt + \sigma(W_T^1 - W_T^1)\right] \\
&= \sigma\mathbb{E}[|W_T^1 - W_T^1|] = \sigma\mathbb{E}[|N(0,2T|] \\
&= 2\sigma\sqrt{T/\pi},
\end{aligned}$$

where the last line is a straightforward calculation. We have shown that for any α we can find β such that $J(\alpha,\beta) \leq 2\sigma\sqrt{T/\pi}$. This shows that $\underline{V}_0 \leq 2\sigma\sqrt{T/\pi}$.

Next, we bound $\overline{V}_0$ from below. Fix β. The crucial point is that, because X^2 depends on β but not on α (as we are dealing with open loop controls), we may define α by

$$\alpha_t = -\frac{\mathbb{E}[X_T^2]}{|\mathbb{E}[X_T^2]|}1_{\{\mathbb{E}[X_T^2]\neq 0\}} + 1_{\{\mathbb{E}[X_T^2]=0\}}.$$

Note that α is admissible, as $|\alpha_t| = 1$. Moreover, $\alpha_t = \alpha_0$ is constant, and by construction $-\mathbb{E}[X_T^2] = \alpha_0|\mathbb{E}[X_T^2]|$. Hence, by Jensen's inequality,

$$\begin{aligned}
J(\alpha,\beta) = \mathbb{E}[|X_T^1 - X_T^2|] &\geq |\mathbb{E}[X_T^1 - X_T^2]| \\
&= \left|\mathbb{E}\left[\int_0^T \alpha_t dt + \sigma W_T^1 - X_T^2\right]\right| \\
&= \left|\mathbb{E}\left[\alpha_0 T - X_T^2\right]\right| \\
&= \left|\alpha_0 T + \alpha_0|\mathbb{E}[X_T^2]|\right| \\
&\geq T.
\end{aligned}$$

In summary, for any β we have chosen α such that $J(\alpha,\beta) \geq T$. Hence $\overline{V}_0 \geq T$. Combining this with our upper bound $\underline{V}_0 \leq 2\sigma\sqrt{T/\pi}$, we find $\underline{V}_0 < \overline{V}_0$ if $\sigma < \frac{1}{2}\sqrt{\pi T}$. ∎

6.4 Bibliographic notes

The literature on two-player zero-sum stochastic differential games is vast. Rufus Isaacs (see, e.g., [72]) kicked off the study of differential games in the mid 1900s, with the seminal paper on *stochastic* differential games due to Fleming and Souganidis [51]. Some research on this subject continuous to this day, and we refer to [16, 110] for some recent work in this area with more comprehensive bibliographies. Notably, the notion of closed loop equilibrium we employed is not the most standard one (with so-called Elliot-Kalton strategies being more common), but it is the one which most readily generalizes to n-player games.

7 n-player stochastic differential games

Here we begin our study of n-player stochastic differential games. The general setup is inevitably notationally cumbersome, but we will do our best. The general setup is as follows. Each of the n-players $i = 1, \ldots, n$ chooses a control $\alpha^i = (\alpha^i_t)_{t\in[0,T]}$ with values in some set A_i (assumed to be a closed subset of a Euclidean space) to influence the state process

$$dX_t = b(X_t, \vec{\alpha}_t)dt + \sigma(X_t, \vec{\alpha}_t)dW_t,$$

where X is a d-dimensional state process, W is an m-dimensional Brownian motion, and $\vec{\alpha}_t = (\alpha^1_t, \ldots, \alpha^n_t)$. The given drift and volatility functions,

$$b : \mathbb{R}^d \times \prod_{i=1}^n A_i \to \mathbb{R}^d, \qquad \sigma : \mathbb{R}^d \times \prod_{i=1}^n A_i \to \mathbb{R}^{d\times m},$$

will be implicitly assumed to be nice enough so that our state SDE above can be solved. For instance, assume they are Lipschitz in x, uniformly in $\vec{\alpha}$.

It is quite important here that we allow the generality of multidimensional state and noise processes. Indeed, in many applications, we have $d = nk$, where k is the dimension of the "private state process" of player i, perhaps influenced only by player i's own choice of control. The current level of generality encompasses many possible frameworks of this nature.

We will work throughout with closed loop Markovian controls. Let $\mathcal{A}_i$ denote the set of measurable functions $\alpha^i : [0,T] \times \mathbb{R}^d \to A_i$. Let $\mathcal{A} \subset \prod_{i=1}^n A_i$ denote the set of $\vec{\alpha} = (\alpha^1, \ldots, \alpha^n)$, such that $\alpha^i \in \mathcal{A}_i$ for each i and the state equation

$$dX_t = b(X_t, \vec{\alpha}(t, X_t))dt + \sigma(X_t, \vec{\alpha}(t, X_t))dW_t$$

has a unique strong solution starting from any initial point. The objective of player i will be to maximize the functional

$$J_i(\vec{\alpha}) = \mathbb{E}\left[\int_0^T f_i(X_t, \vec{\alpha}(t, X_t))dt + g_i(X_T)\right].$$

where we assume the functions f_i and g_i are continuous and bounded from above.

Definition 7.1. A *closed loop* (Markovian) Nash equilibrium is defined to be any $\vec{\alpha} \in \mathcal{A}$ such that, for each $i = 1, \ldots, n$ and each $\beta \in \mathcal{A}_i$ such that $(\vec{\alpha}^{-i}, \beta) \in \mathcal{A}$, we have $J_i(\vec{\alpha}) \geq J_i((\vec{\alpha}^{-i}, \beta))$, where we use the standard game-theoretic abbreviation

$$(\vec{\alpha}^{-i}, \beta) := (\alpha^1, \ldots, \alpha^{i-1}, \beta, \alpha^{i+1}, \ldots, \alpha^n).$$

To solve an n-player game, we will get a lot of mileage from thinking of it as a coupled set of stochastic control problems. That is, player i solves a stochastic control problem whose parameters depend on the actions of the other players. Accordingly, define the (unoptimized) Hamiltonians,

$$h_i(x, y, z, \vec{\alpha}) = b(x, \vec{\alpha}) \cdot y + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\intercal(\alpha, \vec{\alpha})z] + f_i(x, \vec{\alpha}).$$

Note h_i is defined on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}_d \times \prod_{i=1}^n A_i$, where we recall that $\mathbb{S}_d$ is the set of symmetric $d \times d$ matrices.

Following the intuition for two-player games developed the previous section, the HJB equation(s) we find for n-player games will ultimately move the Nash equilibrium problem to a static one, using these Hamiltonians h_i as reward functions. However, we must be careful here: Each player will have her own value function $v_i = v_i(t, x)$, and as usual the adjoint variables y and z in $h_i(t, x, y, \vec{\alpha})$ are merely placeholders into which we will eventually substitute the derivatives $\nabla v_i(t, x)$ and $\nabla^2 v_i(t, x)$, respectively. The point is that we must have a different pair of adjoint variables (y, z) for each agent. This leads to the following:

Definition 7.2. We say that the *generalized Isaacs' condition* holds if there exist measurable functions $\alpha_i : [0, T] \times \mathbb{R}^d \times (\mathbb{R}^d)^n \times \mathbb{S}_d^n \to A_i$ such that, for every $(x, \vec{y}, \vec{z}) \in \mathbb{R}^d \times (\mathbb{R}^d)^n \times \mathbb{S}_d^n$, the vector

$$\vec{\alpha}(x, \vec{y}, \vec{z}) := (\alpha_1(x, \vec{y}, \vec{z}), \ldots, \alpha_n(x, \vec{y}, \vec{z})) \in \prod_{i=1}^n A_i$$

is a Nash equilibrium for the static n-player game with reward functions given by

$$\prod_{i=1}^{n} A_i \ni (a_1, \dots, a_n) \mapsto h_i(x, y_i, z_i, a_1, \dots, a_n).$$

Equivalently, for each $(x, \vec{y}, \vec{z}) \in \mathbb{R}^d \times (\mathbb{R}^d)^n \times \mathbb{S}_d^n$ we have

$$h_i(x, y_i, z_i, \vec{\alpha}(x, \vec{y}, \vec{z})) = \sup_{a_i \in A_i} h_i\Big(x, y_i, z_i, (\vec{\alpha}(x, \vec{y}, \vec{z})^{-i}, a_i)\Big). \tag{7.1}$$

Finally, this lets us state and prove the verification theorem, by bootstrapping on the verification Theorem 5.5 that we have already seen for (one-player) stochastic control problems.

Theorem 7.3 (Verification theorem)**.** *Suppose the generalized Isaacs' condition holds. Suppose $\vec{v} = (v_1, \dots, v_n)$, with $v_i : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ for each i, is a $C^{1,2}$ solution of the PDE system*

$$\begin{aligned}
&\partial_t v_i(t,x) + h_i\Big(x, \nabla v_i(t,x), \nabla^2 v_i(t,x), \vec{\alpha}(x, \nabla \vec{v}(t,x), \nabla^2 \vec{v}(t,x))\Big) = 0, \\
&v_i(T,x) = g_i(x),
\end{aligned}$$

where we abbreviate $\nabla \vec{v} = (\nabla v_1, \dots, \nabla v_n)$ and $\nabla^2 \vec{v} = (\nabla^2 v_1, \dots, \nabla^2 v_n)$. Finally, setting $\vec{\alpha}^*(t,x) = \vec{\alpha}(x, \nabla \vec{v}(t,x), \nabla^2 \vec{v}(t,x))$, suppose that the state equation

$$dX_t = b(X_t, \vec{\alpha}^*(t, X_t))dt + \sigma(X_t, \vec{\alpha}^*(t, X_t))dW_t$$

is well-posed. Then $\vec{\alpha}^*$ is a closed loop Nash equilibrium.

Proof. Fix $i \in \{1, \dots, n\}$. Using the notation of $\vec{\alpha}^*(t,x)$, we may rewrite the PDE as

$$\partial_t v_i(t,x) + h_i\Big(x, \nabla v_i(t,x), \nabla^2 v_i(t,x), \vec{\alpha}^*(t,x)\Big) = 0,$$

We may rewrite the PDE using Isaacs' conditionin the form of (7.1) to get

$$\partial_t v_i(t,x) + \sup_{a_i \in A_i} h_i\Big(x, \nabla v_i(t,x), \nabla^2 v_i(t,x), (\vec{\alpha}^*(t,x)^{-i}, a_i)\Big) = 0.$$

Expand the definition of h_i to find

$$\begin{aligned}
\partial_t v_i(t,x) + \sup_{a_i \in A_i} \Big(& b(x, (\vec{\alpha}^*(t,x)^{-i}, a_i)) \cdot \nabla v_i(t,x) \\
&+ \frac{1}{2} \mathrm{Tr}[\sigma\sigma^\intercal(x, (\vec{\alpha}^*(t,x)^{-i}, a_i)) \nabla^2 v_i(t,x)] + f_i(x, (\vec{\alpha}^*(t,x)^{-i}, a_i))\Big) = 0.
\end{aligned}$$

Treating the other functions $(v_k)_{k\neq i}$ as fixed, we use the verification theorem for (one-player) stochastic optimal control (Theorem 5.5) to find that v_i can be represented as the value function

$$v_i(t,x) = \sup_{\alpha^i} \mathbb{E}\left[\int_t^T f_i(X_s^{t,x}, (\vec{\alpha}^*(s, X_s^{t,x})^{-i}, \alpha_s^i))ds + g_i(X_T^{t,x})\right],$$

where given a control α^i the state process $X^{t,x}$ solves

$$\begin{aligned} dX_s^{t,x} &= b(X_s^{t,x}, (\vec{\alpha}^*(s, X_s^{t,x})^{-i}, \alpha_s^i))ds + \sigma(X_s^{t,x}, (\vec{\alpha}^*(s, X_s^{t,x})^{-i}, \alpha_s^i))dW_s, \quad s \in (t,T], \\ X_t^{t,x} &= x. \end{aligned}$$

By the Isaacs' condition, $\alpha_i^*(t,x)$ is the pointwise optimizer in the PDE written above for v_i, and so the verification theorem 5.5 tells us that $\alpha_i^*(t,x)$ is indeed the optimal control. In summary, we have shown that for a given i, if all players $k \neq i$ stick with the (closed loop) controls α_k^*, then the optimal choice for player i is the control α_i^*. This shows that $\vec{\alpha}^*$ is a Nash equilibrium. □

7.1 Private state processes

The general form of Isaacs' condition given in Definition 7.1 is rather cumbersome, and in this section we focus on an important special class of models for which it simplifies. Assume now that the dimension is $d = n = m + 1$.[7] Now, the state process $\vec{X} = (X^1, \ldots, X^n)$ is thought of as a vector of *private state processes*, with player i denoting the state (e.g., wealth, position, etc.) of player i.

The dynamics of each player's private state process are more specific, taking the form

$$dX_t^i = b_i(\vec{X}_t, \alpha_t^i)dt + \sigma_i(\vec{X}_t, \alpha_t^i)dW_t^i + \widehat{\sigma}_i(\vec{X}_t)dB_t.$$

Here, $B, W^1, \ldots, W^n$ are independent Brownian motions. Notice that b_i and σ_i are functions on $\mathbb{R}^n \times A_i$ and depend only on the control of player i, with no direct dependence on the other players' controls. The Brownian motions $W^1, \ldots, W^n$ are interpreted as *independent* or *idiosyncratic noises*, specific to each player, whereas the *common noise* B influences each player equally.

[7] It should be clear from the discussion how we can generalize to $d = nk$ and $m = j_0 + nj$ for some fixed $k, j_0, j \in \mathbb{N}$ with little more than notational changes. This will simply amount to making each of the private state processes X^i have dimension k, each private noise W^i have dimension j, and the common noise B have dimension j_0.

Similarly, the objective of player i is to maximize

$$J_i(\vec{\alpha}) = \mathbb{E}\left[\int_0^T f_i(\vec{X}_t, \alpha_t^i)dt + g_i(\vec{X}_T)\right],$$

where f_i depends only on player i's own control.

In this scenario, the Hamiltonians exhibit a striking simplification. Let $i \in \{1,\ldots,n\}$. For $(x,y,z,\vec{a}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}_n \times \prod_{k=1}^n A_k$, the above definition of h_i specializes to

$$\begin{aligned} h_i(x,y,z,\vec{a}) = &\sum_{k=1}^n b_k(x,a_k)y_k + \frac{1}{2}\sum_{k=1}^n \sigma_k^2(x,a_k)z_{kk} \\ &+ \frac{1}{2}\sum_{k,j=1}^n \widehat{\sigma}_k(x)\widehat{\sigma}_j(x)z_{kj} + f_i(x,a_k). \end{aligned}$$

When optimizing h_i over a_i, keeping $(a_k)_{k\neq i}$ as fixed, we find that all of the terms involving $(a_k)_{k\neq i}$ come out of the supremum. To be clear, for a given i and for $(x,y,z,a_i) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}_n \times A_i$ let us define

$$\widetilde{h}_i(x,y,z,a_i) = b_i(x,a_i)y_i + \frac{1}{2}\sigma_i^2(x,a_i)z_{ii} + f_i(x,a_i).$$

Then, we may write

$$h_i(x,y,z,\vec{a}) = \widetilde{h}_i(x,y,z,a_i) + \sum_{k\neq i}^n b_k(x,a_k)y_k + \frac{1}{2}\sum_{k\neq i}^n \sigma_k^2(x,a_k)z_{kk} \tag{7.2}$$

$$+ \frac{1}{2}\sum_{k,j=1}^n \widehat{\sigma}_k(x)\widehat{\sigma}_j(x)z_{kj}. \tag{7.3}$$

The point is that all of the terms outside of $\widetilde{h}_i(x,y,z,a_i)$ are independent of a_i.

To resolve Isaacs' condition, we need to fix $(x,\vec{y},\vec{z}) \in \mathbb{R}^n \times (\mathbb{R}^n)^n \times \mathbb{S}_n^n$ and find a Nash equilibrium of the static n-player game with reward functions $\vec{a} \mapsto h_i(x,y_i,z_i,\vec{a})$. But, in this case, (7.3) shows that this is equivalent to finding a Nash equilibrium of the static n-player game with reward functions $\vec{a} \mapsto \widetilde{h}_i(x,y_i,z_i,a_i)$. In particular, the action of player i does not depend on the actions of the other players! This n-player game is completely trivial, consisting of n decoupled optimization problems. This observation is summarized in the following lemma:

Lemma 7.4. *In the class of models described above, suppose there exists for each i a measurable function $\alpha_i : [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to A_i$ such that, for every $(x, y, z) \in \mathbb{R}^d \times (R^d \times \mathbb{S}_d$, we have*

$$\alpha_i(x, y, z) \in \arg\max_{a_i \in A_i} \widetilde{h}_i(x, y, z, a_i).$$

Then the Isaacs' condition holds.

Let us check carefully how the PDE system of Theorem 7.3 looks in this case. Define the optimized Hamiltonians for $(x, y, z) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ by

$$H_i(x, y, z) := \sup_{a_i \in A_i} \widetilde{h}_i(x, y, z, a_i). \tag{7.4}$$

We have removed several terms from $\widetilde{h}_i$ which we must add back in, and the HJB system becomes the following: We must find functions $v_1, \ldots, v_n$ on $[0,T] \times \mathbb{R}^n$ satisfying

$$\begin{aligned}
0 = \partial_t v_i(t,x) &+ H_i(x, \partial_{x_i} v_i(t,x), \partial_{x_i x_i} v_i(t,x)) \\
&+ \sum_{k \neq i}^n b_k(x, \alpha_i(x, \partial_{x_k} v_k(t,x), \partial_{x_k x_k} v_k(t,x))) \partial_{x_k} v_i(t,x) \\
&+ \frac{1}{2} \sum_{k \neq i}^n \sigma_k^2(x, \alpha_i(x, \partial_{x_k} v_k(t,x), \partial_{x_k x_k} v_k(t,x))) \partial_{x_k x_k} v_i(t,x) \\
&+ \frac{1}{2} \sum_{k,j=1}^n \widehat{\sigma}_k(x) \widehat{\sigma}_j(x) \partial_{x_k x_j} v_i(t,x),
\end{aligned}$$

with the boundary condition $v_i(T, x) = g_i(x)$.

This may not look like much of an improvement. However, it is it quite valuable that the Isaacs' condition holds *automatically* for this class of models. Moreover, the PDE structure clearly shows where the optimization happens for player i (inside the Hamiltonian H_i), with the other terms coming from drifts and volatility parameters from the other players' prive state processes and from the uncontrolled common noise coefficients $\widehat{\sigma}_k$. It is a good exercise to see what goes wrong with this program if the common noise coefficient $\widehat{\sigma}_k$ is allowed to depend on a_k, and not just x_k.

7.2 Linear-quadratic games

In this section, parallel to Section 5.5, we study a very simple n-player game, to begin to make sense of the general notational mess of the previous section.

We work in a model that fits the structure discussed in Section 7.1. The private state processes $X^1, \ldots, X^n$ evolve according to

$$dX_t^i = \alpha_t^i dt + dW_t^i.$$

The objective of player i is to maximize

$$J_i(\vec{\alpha}) = \mathbb{E}\left[\int_0^T -\frac{1}{2}|\alpha_t^i|^2 dt - \frac{\lambda}{2}|\overline{X}_T - X_T^i|^2\right], \qquad \overline{X}_t = \frac{1}{n}\sum_{k=1}^n X_t^k,$$

where $\lambda > 0$ is constant. This is similar to the one-player control problem of Section 5.5, except that the "target" that each player wants to reach at the final time T is not a given deterministic value but rather the empirical mean; the reward function encourages the players to "flock" together, but using minimal energy or fuel to do so.

In the notation of Section 7.1, we have

$$\begin{aligned} b_i(x, a_i) = a_i, \quad \sigma_i(x, a_i) = 1, \quad \widehat{\sigma} \equiv 0, \\ f_i(x, a_i) = -\frac{1}{2}a_i^2, \quad g_i(x) = -\frac{\lambda}{2}|\bar{x} - x_i|^2, \end{aligned}$$

where we write $\bar{x} = \frac{1}{n}\sum_{k=1}^n x_k$ for any vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. We may define the optimized Hamiltonian H of (7.4) becomes

$$\begin{aligned} H_i(x, y, z) &= \sup_{a_i \in A_i} \widetilde{h}_i(x, y, z, a_i) = \sup_{a_i \in A_i}(a_i y - \frac{1}{2}a_i^2) + \frac{1}{2}z \\ &= \frac{1}{2}y^2 + \frac{1}{2}z. \end{aligned}$$

The optimizer is $\alpha_i(x, y, z) = y$. Hence, the HJB system becomes

$$0 = \partial_t v_i(t, x) + \frac{1}{2}|\partial_{x_i} v_i(t, x)|^2 + \sum_{k \neq i}^n \partial_{x_k} v_k(t, x)\partial_{x_k} v_i(t, x) + \frac{1}{2}\sum_{k=1}^n \partial_{x_k x_k} v_i(t, x),$$

with the boundary condition $v_i(T, x) = -\frac{\lambda}{2}|\bar{x} - x_i|^2$.

As in Section 5.5, we seek a separable solution by making the ansatz

$$v_i(t, x) = g(t) + \frac{1}{2}f(t)(\bar{x} - x_i)^2, \tag{7.5}$$

for some functions g and f to be determined, satisfying the boundary conditions $g(T) = 0$ and $f(T) = -\lambda$. Taking derivatives, we find

$$\begin{aligned} &\partial_{x_i} v_i(t, x) = -\left(1 - \frac{1}{n}\right) f(t)(\bar{x} - x_i), \qquad \partial_{x_k} v_i(t, x) = \frac{1}{n}f(t)(\bar{x} - x_i), \ k \neq i \\ &\partial_{x_i x_i} v_i(t, x) = \left(1 - \frac{1}{n}\right)^2 f(t), \qquad \partial_{x_k x_k} v_i(t, x) = \frac{1}{n^2}f(t), \ k \neq i. \end{aligned}$$

Plug these into the PDE to find

$$
\begin{aligned}
0 =& g'(t) + \frac{1}{2}f'(t)(\bar{x}-x_i)^2 + \frac{1}{2}\left(1-\frac{1}{n}\right)^2(\bar{x}-x_i)^2 f^2(t) \qquad (7.6)\\
&- \sum_{k\neq i}^{n}\frac{1}{n}\left(1-\frac{1}{n}\right)(\bar{x}-x_i)(\bar{x}-x_k)f^2(t) + \frac{1}{2}\left(1-\frac{1}{n}\right)^2 f(t) + \frac{1}{2}\sum_{k\neq i}^{n}\frac{1}{n^2}f(t).
\end{aligned}
$$

We make a couple of simplifications. First, combine the third and fourth terms by writing

$$
\begin{aligned}
&\frac{1}{2}\left(1-\frac{1}{n}\right)^2(\bar{x}-x_i)^2 f^2(t) - \sum_{k\neq i}^{n}\frac{1}{n}\left(1-\frac{1}{n}\right)(\bar{x}-x_i)(\bar{x}-x_k)f^2(t)\\
&= \frac{1}{2}\left(1-\frac{1}{n}\right)^2(\bar{x}-x_i)^2 f^2(t) + \frac{1}{n}\left(1-\frac{1}{n}\right)(\bar{x}-x_i)^2 f^2(t)\\
&= \frac{1}{2}\left(1-\frac{1}{n^2}\right)(\bar{x}-x_i)^2 f^2(t).
\end{aligned}
$$

Indeed, the first equality here follows from the observation that the sum would vanish if it included the $k=i$ term, because $\frac{1}{n}\sum_{k=1}^{n}(\bar{x}-x_k)=0$.

Now, combine the $(\bar{x}-x_i)^2$ terms in (7.6) to find

$$
f'(t) + \left(1-\frac{1}{n^2}\right)f^2(t) = 0.
$$

The other terms in (7.6) become

$$
g'(t) + \frac{1}{2}\left(1-\frac{1}{n}\right)^2 f(t) + \frac{n-1}{2n^2}f(t) = 0,
$$

or equivalently

$$
g'(t) + \frac{1}{2}\left(1-\frac{1}{n}\right)f(t) = 0.
$$

Recalling the boundary condition $f(T) = -\lambda$, we solve the first of these ODEs by the usual method,

$$
\int_t^T \left(1-\frac{1}{n^2}\right)ds = \int_t^T -\frac{f'(t)}{f^2(t)}ds = \frac{1}{f(T)} - \frac{1}{f(t)} = -\frac{1}{\lambda} - \frac{1}{f(t)}.
$$

Rearrange to find

$$
f(t) = -\frac{1}{\frac{1}{\lambda} + \left(1-\frac{1}{n^2}\right)(T-t)}.
$$

Plug this into the equation for g to find

$$g(t) = g(T) + \frac{1}{2}\left(1 - \frac{1}{n}\right)\int_t^T f(s)ds = \frac{1}{2}\left(1 - \frac{1}{n^2}\right)\log\left(1 + \lambda\left(1 - \frac{1}{n^2}\right)(T-t)\right).$$

Plug this back into (7.5) to obtain an expression for the value function for player i, and thus the solution of the HJB system.

To compute the optimal control, we recall that the maximizer of the Hamiltonian was $\alpha_i(x, y, z) = y$. Plugging in $\partial_{x_i} v_i(t, x)$ for y, we find that the Nash equilibrium is

$$\begin{aligned}\alpha_i^*(t, x) = \partial_{x_i} v_i(t, x) &= -\left(1 - \frac{1}{n}\right) f(t)(\bar{x} - x_i) \\ &= \frac{\left(1 - \frac{1}{n}\right)(\bar{x} - x_i)}{\frac{1}{\lambda} + \left(1 - \frac{1}{n^2}\right)(T-t)}.\end{aligned}$$

As in the one-player control problem of Section 5.5, the state process of each player "mean-reverts" toward the target, which in this case is the current empirical average:

$$dX_t^i = \frac{\left(1 - \frac{1}{n}\right)(\overline{X}_t - X_t^i)}{\frac{1}{\lambda} + \left(1 - \frac{1}{n^2}\right)(T-t)}dt + dW_t^i.$$

7.3 Competitive Merton problem

In this section we study a model in which n fund managers (players) aim to maximize not only *absolute* wealth at the time horizon $T > 0$ but also *relative* wealth. The precise setup will be essentially a multi-agent form of Merton's problem, which we studied in Section 5.7. As in Merton's problem, we assume that the risky asset's price is a one-dimensional geometric Brownian motion

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

For simplicity, we further assume $r = 0$, which means the risk free asset has a constant price. The wealth of player $i = 1, ..., n$ is given by

$$\begin{aligned}dX_t^i &= \alpha_t^i X_t^i (\mu dt + \sigma dW_t) \\ X_0^i &= x_i > 0\end{aligned}$$

where α_t^i is the fraction of wealth invested in the stock. We assume that the controls are square integrable, so that the SDE is well defined, and note that $X_t^i > 0$ for all t.

Agent i uses the power utility

$$U_i(x) = \frac{1}{\gamma_i} x^{\gamma_i},$$

where $\gamma_i < 1$, and $\gamma_i \neq 0$.[8] The goal of player i is to maximize his expected utility at time T

$$J_i(\vec{\alpha}) = \mathbb{E}\left[U_i\left(X_T^i |\overline{X}_T^{-i}|^{-\theta_i}\right)\right],$$

where $\theta_i \in [0,1]$, and where $\overline{X}_t^{-i} = \left(\prod_{k\neq i} X_t^k\right)^{\frac{1}{n-1}}$ is the geometric mean wealth of the $n-1$ other agents. We can write this objective function alternatively as

$$J_i(\vec{\alpha}) = \mathbb{E}\left[U_i\left(\left|X_T^i\right|^{1-\theta_i}\left|\frac{X_T^i}{\overline{X}_T^{-i}}\right|^{\theta_i}\right)\right].$$

$\left|X_T^i\right|^{1-\theta_i}$ is the absolute wealth of player i, and the ratio $X_T^i/\overline{X}_T^{-i}$ is relative wealth. The parameter θ_i controls the trade-off between absolute wealth and relative wealth, and we might thus call it the *competitiveness parameter*. As usual, the parameter γ_i captures the gent's risk aversion.

A priori, to find an equilibrium, we should write down the system of n HJB equation and try to solve it. However, this turns out to be rather cumbersome, and we prefer to take an alternative but natural approach: We fix each players' strategies, we compute a single player's best response, and then we resolve the fixed point problem for a Nash equilibrium.

What makes this approach tractable here is that we will search search for a Nash equilibrium in which the investment strategies α^i are all constant (i.e., deterministic and independent of the time parameter). We saw in the one-player Merton problem that the optimal strategy was constant in this sense, so this is a natural starting point in this problem. (Note in particular that an agent with $\theta_i = 0$ is not at all competitive and simply solves Merton's problem.) To find such an equilibrium, we first solve one player's optimization problem given an arbitrary choice of the other players' constant strategies.

Fix $(\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$ and $i \in \{1, ..., n\}$. We will now allow player i to choose a control process α^i as a best response to the fixed constant strategies $(\alpha_j)_{j\neq i}$ of the other players, and it will turn out that this best response is itself a constant. This will give us a map from $\mathbb{R}^n \to \mathbb{R}^n$ for which we find a fixed point, giving us an equilibrium.

[8] One could treat the limit case $\gamma_i = 0$ by setting $U_i(x) = \log(x)$, and all of the final formulas we find are valid in this case.

7.3.1 Solution of one player's best response

Define $X_t = X_t^i \left| \overline{X}_t^{-i} \right|^{-\theta_i}$. The idea is to write player i's optimization problem as a stochastic control problem with one-dimensional state process X, and to do this we must first compute the (controlled) dynamics of X. For each k, Itô s formula yields

$$d \log X_t^k = \left(\alpha_k \mu - \frac{\sigma^2 \alpha_k^2}{2} \right) dt + \sigma \alpha_k dW_t.$$

Thus

$$\begin{aligned}
d \log X_t &= d \log X_t^i - \frac{\theta_i}{n-1} \sum_{k \neq i} d \log X_t^k \\
&= \left(\alpha_t^i \mu - \frac{|\alpha_t^i|^2 \sigma^2}{2} - \frac{\theta_i}{n-1} \sum_{k \neq i} \left(\alpha_k \mu - \frac{\sigma^2 \alpha_k^2}{2} \right) \right) dt \\
&\qquad + \left(\sigma \alpha_t^i - \frac{\theta_i}{n-1} \sum_{k \neq i} \sigma \alpha_k \right) dW_t \\
&=: \left(\alpha_t^i \mu - \frac{\alpha_i^2 \sigma^2}{2} - \theta_i \mu \overline{\alpha}^{-i} + \frac{\theta_i \sigma^2}{2} \overline{\alpha^2}^{-i} \right) dt + \left(\sigma \alpha_t^i - \theta_i \sigma \overline{\alpha}^{-i} \right) dW_t
\end{aligned}$$

where we abbreviate $\overline{\alpha}^{-i} = \frac{1}{n-1} \sum_{k \neq i} \alpha_k$ and $\overline{\alpha^2}^{-i} = \frac{1}{n-1} \sum_{k \neq i} \alpha_k^2$. Exponentiating and applying Itô's formula again, our state process is

$$\begin{aligned}
\frac{dX_t}{X_t} &= \left(\alpha_t^i \mu - \frac{|\alpha_t^i|^2 \sigma^2}{2} - \theta_i \mu \overline{\alpha}^{-i} + \frac{\theta_i \sigma^2}{2} \overline{\alpha^2}^{-i} + \frac{\sigma^2}{2} \left(\alpha_t^i - \theta_i \overline{\alpha}^{-i} \right)^2 \right) dt \\
&\qquad + \sigma \left(\alpha_t^i - \theta_i \overline{\alpha}^{-i} \right) dW_t \\
&=: \left[\left(\mu - \sigma^2 \theta_i \overline{\alpha}^{-i} \right) \alpha_t^i + \eta \right] dt + \sigma \left(\alpha_t^i - \theta_i \overline{\alpha}^{-i} \right) dW_t,
\end{aligned}$$

where $\eta = -\theta_i \mu \overline{\alpha}^{-i} + \frac{\theta_i \sigma^2}{2} \overline{\alpha^2}^{-i} + \frac{\theta_i^2 \sigma^2}{2} (\overline{\alpha}^{-i})^2$. The goal of player i is to choose α_i to maximize

$$\mathbb{E} \left[\frac{1}{\gamma_i} X_T^{\gamma_i} \right].$$

This is a standard stochastic control problem, and we are well-versed in solving such things by now. We identify the Hamiltonian H as

$$
\begin{aligned}
H(x,y,z) &= \sup_{a\in\mathbb{R}}\left[xy\left(\mu-\sigma^2\theta_i\overline{\alpha}^{-i}\right)a+\frac{1}{2}\sigma^2x^2z\left(a-\theta_i\overline{\alpha}^{-i}\right)^2\right]+xy\eta \\
&= \sup_{a\in\mathbb{R}}\left[a\left(xy\left(\mu-\sigma^2\theta_i\overline{\alpha}^{-i}\right)-\sigma^2x^2z\theta_i\overline{\alpha}^{-i}\right)+\frac{1}{2}\sigma^2x^2za^2\right] \\
&\qquad +\frac{1}{2}\sigma^2x^2z\theta_i^2\left(\overline{\alpha}^{-i}\right)^2+xy\eta
\end{aligned}
$$

As in the Merton problem, we assume $z<0$ to avoid an infinite Hamiltonian, and we will justify this later by checking that the solution v of our HJB equation satisfies $\partial_{xx}v(t,x)<0$. Assuming $z<0$, the optimizer of the Hamiltonian is

$$
\widehat{\alpha}(x,y,z)=-\frac{xy(\mu-\sigma^2\theta_i\overline{\alpha}^{-i})-\sigma^2x^2z\theta_i\overline{\alpha}^{-i}}{\sigma^2x^2z}, \tag{7.7}
$$

and we can write

$$
H(x,y,z)=-\frac{\left(xy(\mu-\sigma^2\theta_i\overline{\alpha}^{-i})-\sigma^2x^2z\theta_i\overline{\alpha}^{-i}\right)^2}{2\sigma^2x^2z}+\frac{1}{2}\sigma^2x^2z\theta_i^2\left(\overline{\alpha}^{-i}\right)^2+xy\eta.
$$

The HJB equation is then

$$
\begin{aligned}
0 &= \partial_tv(t,x)-\frac{\left(x\partial_xv(t,x)(\mu-\sigma^2\theta_i\overline{\alpha}^{-i})-\sigma^2x^2\partial_{xx}v(t,x)\theta_i\overline{\alpha}^{-i}\right)^2}{2\sigma^2x^2\partial_{xx}v(t,x)} \\
&\qquad +\frac{1}{2}\sigma^2x^2\partial_{xx}v(t,x)\theta_i^2\left(\overline{\alpha}^{-i}\right)^2+x\partial_xv(t,x)\eta, \\
v(T,x) &= \frac{1}{\gamma_i}x^{\gamma_i}.
\end{aligned}
$$

We look for a separable solution by making the ansatz $v(t,x)=\frac{1}{\gamma_i}f(t)x^{\gamma_i}$, where f is a function to be determined and that satisfies the boundary condition $f(T)=1$. The partial derivatives of v are then

$$
\begin{aligned}
\partial_tv(t,x) &= \frac{1}{\gamma_i}f'(t)x^{\gamma_i} \\
\partial_xv(t,x) &= f(t)x^{\gamma_i-1} \\
\partial_{xx}v(t,x) &= -(1-\gamma_i)f(t)x^{\gamma_i-2}.
\end{aligned}
$$

Now note that in the HJB equation, each time we have a partial derivative of v, we either have $\partial_tv(t,x)$ or $x\partial_xv(t,x)$ or $x^2\partial_{xx}v(t,x)$, which all give us

x^{γ_i}. In the fraction in the HJB equation, we get a factor of $(x^{\gamma_i})^2$ in the numerator and x^{γ_i} in the denominator. Hence, the factor x^{γ_i} cancels out of the equation entirely, and the HJB equation becomes a first linear ODE of f

$$\frac{1}{\gamma_i}f'(t)+\left[\frac{\mu-\sigma^2\theta_i\overline{\alpha}^{-i}+\sigma^2\theta_i\overline{\alpha}^{-i}(1-\gamma_i)}{2\sigma^2(1-\gamma_i)}-\frac{1}{2}\sigma^2(1-\gamma_i)\theta_i^2\left(\overline{\alpha}^{-i}\right)^2+\eta\right]f(t)=0.$$

This ODE is of course well-posed, and the terminal condition $f(T)=1$ gives us

$$f(t)=e^{\rho(T-t)},$$

where $\rho\in\mathbb{R}$ is precisely γ_i times the long expression appearing in the bracket above; we will not keep track of this value, as it will appear only in the value function and not in the optimal control. We conclude that our ansatz was valid, and $v(t,x)=\frac{1}{\gamma_i}f(t)x^{\gamma_i}$ indeed solves the HJB equation. Recalling (7.7), the optimal control is then

$$\begin{aligned}\widehat{\alpha}(x,\partial_x v(t,x),\partial_{xx}v(t,x)) &= \frac{\mu-\sigma^2\theta_i\overline{\alpha}^{-i}+\sigma^2\theta_i\overline{\alpha}^{-i}(1-\gamma_i)}{\sigma^2(1-\gamma_i)}\\ &=: \alpha_i^*,\end{aligned}$$

which we notice does not depend on t or x because of the cancellation of x^{γ_i} and $f(t)$.

7.3.2 Solution of equilibrium

We have now determined how one player responds optimally to other players' strategies, and we find our desired Nash equilibrium as a fixed point of this map $(\alpha_1,\dots,\alpha_n)\mapsto(\alpha_1^*,\dots,\alpha_n^*)$. That is, suppose now that $\alpha_i=\alpha_i^*$ for each i. Let

$$\overline{\alpha}=\frac{1}{n}\sum_{k=1}^n\alpha_k,$$

and note that

$$\overline{\alpha}^{-i}=\frac{n\overline{\alpha}-\alpha_i}{n-1}.$$

Then

$$\begin{aligned}\alpha_i=\frac{\mu}{\sigma^2(1-\gamma_i)}-\left(\frac{\theta_i}{1-\gamma_i}-\theta_i\right)\overline{\alpha}^{-i}&=\frac{\mu}{\sigma^2(1-\gamma_i)}-\frac{\theta_i\gamma_i}{1-\gamma_i}\left(\frac{n\overline{\alpha}-\alpha_i}{n-1}\right)\\ \Leftrightarrow\qquad \alpha_i=\frac{\frac{\mu}{\sigma^2(1-\gamma_i)}-\frac{\theta_i\gamma_i}{1-\gamma_i}\frac{n}{n-1}\overline{\alpha}}{1-\frac{\theta_i\gamma_i}{1-\gamma_i}\frac{1}{n-1}}&=:\frac{\mu\phi_i}{\sigma^2}-\psi_i\overline{\alpha},\end{aligned}$$

where we define

$$\phi_i = \frac{1}{1-\gamma_i\left(1+\frac{\theta_i}{n-1}\right)}, \qquad \psi_i = \frac{\theta_i\gamma_i\frac{n}{n-1}}{1-\gamma_i\left(1+\frac{\theta_i}{n-1}\right)}.$$

Note that we are in trouble here if $1 = \gamma_i\left(1+\frac{\theta_i}{n-1}\right)$, and in this case there is no equilibrium! We assume it is not the case to avoid diving by zero, and it is worth noting that for sufficiently large n we can avoid this problem no matter what the values of θ_i are. Assuming ϕ_i and ψ_i are well defined, we may now average over i to get

$$\overline{\alpha} = \frac{\mu\overline{\phi}}{\sigma^2} - \overline{\psi}\overline{\alpha}.$$

Thus

$$\overline{\alpha} = \frac{\mu\overline{\phi}}{\sigma^2(1+\overline{\psi})}.$$

Therefore

$$\alpha_i = \frac{\mu}{\sigma^2}\left(\phi_i - \frac{\psi_i\overline{\phi}}{1+\overline{\psi}}\right) = \frac{\mu}{\sigma^2\left(1-\gamma_i\left(1+\frac{\theta_i}{n-1}\right)\right)}\left(1 - \frac{n}{n-1}\frac{\gamma_i\theta_i\overline{\phi}}{1+\overline{\psi}}\right).$$

Note that $\frac{\mu}{\sigma^2(1-\gamma_i)}$ corresponds to the Merton problem's optimal strategy. Thus, α_i can be interpreted as the Merton's portfolio plus a perturbation. This perturbation tells us that if $\gamma_i > 0$ then the player will invest less in the risky asset if he is competitive, but if $\gamma_i < 0$, on the contrary, the more competitive is the player, the more he will invest in the risky asset. By sending $n \to \infty$, this simplifies nicely to

$$\alpha_i = \frac{\mu}{\sigma^2(1-\gamma_i)}\left(1 - \gamma_i\theta_i\frac{\overline{\left(\frac{1}{1-\gamma}\right)}}{1+\overline{\left(\frac{\gamma\theta}{1-\gamma}\right)}}\right).$$

In fact, it is often convenient to work with the *risk tolerance* parameter $\delta > 0$, defined by $\gamma = 1 - 1/\delta$. Then $\delta = 1/(1-\gamma)$, and we may write the above as

$$\alpha_i = \frac{\mu}{\sigma^2}\left(\delta_i - \frac{\theta_i\overline{\delta}(\delta_i - 1)}{1+\overline{\theta(\delta-1)}}\right),$$

and we can think of the term inside the parentheses as the *modified* or *effective risk tolerance*; player i uses the Merton portfolio $\mu\delta_i/\sigma^2$, except that the risk tolerance δ_i is adjusted due to competition.

7.4 Bibliographic notes

The literature on n-player stochastic differential games is sparse, and it has only recently begun to appear in book form. See the lecture notes [25, Chapter 5] and the more recent tome [30]. Some of the key early (PDE-focused) work was by Bensoussan-Frehse [7, 8, 9] and Friedman [53]. The text of Ladyzhenskaia et al. [86] is a classic reference for the kinds of PDE systems that arise, though there is no discussion of games here. For a probabilistic approach based on the maximum principle, see [64, 65] or the aforementioned books [25, 30].

The somewhat contrived linear quadratic game of Section 7.2 is inspired by (and is a simplification of) the model of [32]. The competitive Merton problem is shamelessly adapted from my own recent paper [85].

8 Stochastic differential mean field games

Here we turn to the final topic of the course, where we look at the mean field or continuum limit of n-player stochastic differential games. This combines essentially all of the material we have seen so far in the course. This might be a good time to review the material of Section 3 on McKean-Vlasov equations.

8.1 Review: Linear-Quadratic Game

We will ease into things by reviewing the linear-quadratic model studied in Section 7.2. In this n-player game, the ith agent's objective and state processes are given by

$$\begin{cases} J_i(\vec{\alpha}) & = \mathbb{E}\left[\int_0^T -\frac{1}{2}\left|\alpha_t^i\right|^2 dt - \frac{\lambda}{2}\left|\overline{X}_T - X_T^i\right|^2\right] \\ dX_t^i & = \alpha_t^i dt + dW_t^i, \qquad \bar{X}_t = \frac{1}{n}\sum_{k=1}^n X_t^k \end{cases}$$

We proved that in this case, at the Nash equilibrium, the state process evolves according to

$$dX_t^i = \frac{\lambda\left(1-\frac{1}{n}\right)\left(\overline{X}_t - X_t^i\right)}{1+\lambda\left(1-\frac{1}{n^2}\right)(T-t)}dt + dW_t^i, \tag{8.1}$$

and that the value function is

$$v_i(t,\vec{x}) = -\frac{\lambda(\bar{x}-x_i)^2}{2(1+\lambda(1-\frac{1}{n^2})(T-t))} + \frac{1}{2}\left(1-\frac{1}{n^2}\right)\log\left(1+\lambda\left(1-\frac{1}{n^2}\right)(T-t)\right)$$

The SDE system (8.1) is not exactly of McKean-Vlasov type because of the factors $(1-1/n)$ and $(1-1/n^2)$. Nonetheless, it is not hard to adapt the arguments of Section 3 (more specifically, Theorem 3.3) to show that, as $n \to \infty$, we have the weak convergence $X^i \Rightarrow Y^i$, where $Y^1, \ldots, Y^n$ are independent copies of the McKean-Vlasov equation

$$dY_t^i = \frac{\lambda(\mathbb{E}(Y_t^i) - Y_t^i)}{1+\lambda(T-t)}dt + dW_t^i, \quad Y_0^i = X_0^i.$$

Moreover, the value functions converge in an "averaged" sense. Define $v : [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$v(t,x,y) = -\frac{\lambda(y-x)^2}{2(1+\lambda(T-t))} + \frac{1}{2}\log\left(1+\lambda(T-t)\right).$$

Then, for each $t \in [0,T]$ and $x_1, x_2, \ldots \in \mathbb{R}$, we have

$$\lim_n \max_{i=1,\ldots,n} \left| v_i^n(t, x_1, ..., x_n) - v\left(t, x_i, \frac{1}{n}\sum_{k=1}^n x_k\right)\right| = 0.$$

This limiting value function v may look familiar. Indeed, recall the (one-player) linear-quadratic control problem studied in Section 5.5:

$$\begin{cases} \sup\limits_{\alpha} & \mathbb{E}\left[\int_0^T -\frac{1}{2}\left|\alpha_t\right|^2 dt - \frac{\lambda}{2}\left|z - \bar{X_T}\right|^2\right] \\ dX_t & = \alpha_t dt + dW_t, \qquad z \in \mathbb{R}, \lambda > 0. \end{cases}$$

Here, $z \in \mathbb{R}$ is some given target the controller wants to reach at the final time. We saw that the optimal state process is

$$dX_t = \frac{\lambda(z - X_t)}{1+\lambda(T-t)}dt + dW_t, \tag{8.2}$$

and the value function is

$$u(t,x) = \frac{\lambda(z-x)^2}{2(1+\lambda(T-t)} + \frac{1}{2}\log(1+\lambda(T-t)),$$

or $u(t,x) = v(t,x,z)$. In fact, we can define a fixed point problem using this "target" z as a surrogate for our population average, and in doing so arrive at the same limit of our n-player equilibria. Precisely, consider the following fixed point problem:

1. Fix a target $z \in \mathbb{R}$.

2. Solve 1-player control problem to get (8.2).

3. Compute $\Phi(z) := \mathbb{E}[X_T]$.

This defines a map $\Phi : \mathbb{R} \to \mathbb{R}$. Let us find the fixed point. Note that

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] + \int_0^t \frac{\lambda(z - \mathbb{E}[X_s])}{1 + \lambda(T - s)} ds.$$

In other words, the function $f(t) = \mathbb{E}[X_t] - z$ solves the ODE

$$f'(t) = -\frac{\lambda f(t)}{1 + \lambda(T - t)}, \qquad f(0) = \mathbb{E}[X_0] - z.$$

The solution is easily seen to be

$$f(t) = (\mathbb{E}[X_0] - z)\frac{1 + \lambda(T - t)}{1 + \lambda T},$$

so that

$$\Phi(z) = \mathbb{E}[X_T] = z + f(T) = z + \frac{\mathbb{E}[X_0] - z}{1 + \lambda T}.$$

Then $\Phi(z) = z$ if and only if $z = \mathbb{E}[X_0]$. It is no coincidence that if one starts with $\mathbb{E}[X_0^i] = z$ in the n-player game, the limiting McKean-Vlasov equation can be solved with $\mathbb{E}[Y_t^i] = z$.

8.2 Setting up a general mean field game

The mean field games we study are to be seen as continuum analogues of n-player games of the following form. This is essentially a controlled version of the n-particle SDE system associated to the McKean-Vlasov limit we studied in Section 3. Consider the n-player stochastic differential game, where each player $i \in \{1, ..., n\}$ has a (private) state process X^i given by

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha_t^i)dt + \sigma(X_t^i, \mu_t^n, \alpha_t^i)dW_t^i,$$

where

$$\mu_t^n = \frac{1}{n}\sum_{k=1}^n \delta_{X_t^k},$$

and α^i is the control of player i. We typically assume the initial states (X_0^i) are i.i.d. with some given distribution $\lambda_0 \in \mathcal{P}(\mathbb{R}^d)$. As usual, we will be fairly loose about the precise assumptions on the coefficients $b : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times A \to$

$\mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times A \to \mathbb{R}^{d\times m}$. Here, the state X_t^i is d-dimensional, and the driving Brownian motions W_t^i is m-dimensional. The objectives are

$$J_i^n(\vec{\alpha}) = \mathbb{E}\left[\int_0^T f(X_t^i, \mu_t^n, \alpha_t^i)dt + g(X_T^i, \mu_T^n)\right].$$

The crucial feature here is that the functions (b, σ, f, g) are the same for each agent, and the Brownian motions W^i are independent. We will see later how to incorporate correlation in the noises and heterogeneity in the coefficients (b, σ, f, g), but for now we stick to the simplest setup.

The goal is to describe the Nash equilibrium for large n, or the mean field game (MFG) limit. We will first do this heuristically. Suppose $(\alpha^{n,1}, \ldots, \alpha^{n,n})$ is a Nash equilibrium for each n. Given the symmetry of the game, it is not unreasonable to suspect that for each n we can find a single function $\widehat{\alpha}_n$ such that $\alpha_t^{n,i} = \widehat{\alpha}(t, X_t^i, \mu_t^n)$. It is less reasonable to expect that $\widehat{\alpha}_n = \widehat{\alpha}$ does not depend on n, but let us make this assumption anyway, as we are doing heuristics. Then, the state processes become

$$dX_t^i = b(X_t^i, \mu_t^n, \widehat{\alpha}(t, X_t^i, \mu_t^n))dt + \sigma(X_t^i, \mu_t^n, \widehat{\alpha}(t, X_t^i, \mu_t^n))dW_t^i.$$

This is precisely an equation of McKean-Vlasov type. Recalling Theorem 3.3, it should hold (if the coefficients and, notably, $\widehat{\alpha}$ are nice enough) that (X^1, μ^n) converges in law to (X, μ), where $\mu = (\mu_t)_{t\in[0,T]}$ is a deterministic measure flow and X^* solves

$$dX_t^* = b(X_t^*, \mu_t, \widehat{\alpha}(t, X_t^*, \mu_t))dt + \sigma(X_t^*, \mu_t, \widehat{\alpha}(t, X_t^*, \mu_t))dW_t,$$

with $\mu_t = \mathcal{L}(X_t^*)$ for all t. On the other hand, because $(\alpha^{n,1}, \ldots, \alpha^{n,n})$ is a Nash equilibrium, for any other control β we have

$$\begin{aligned} &\mathbb{E}\left[\int_0^T f(X_t^1, \mu_t^n, \alpha_t^{n,1})dt + g(X_T^1, \mu_T^n)\right] \\ &\quad \geq \mathbb{E}\left[\int_0^T f(X_t^1, \mu_t^n, \beta_t)dt + g(X_T^1, \mu_T^n)\right]. \end{aligned} \tag{8.3}$$

Here we mean that X^1 on the right-hand side is controlled by β, not by $\alpha^{n,1}$. When only player 1 changes strategy and the remaining $n-1$ players stick with the same $\alpha^{n,i}$, the empirical measure μ^n should not change much and should thus converge to the same limit μ. Hence, the process X^1 controlled by β should converge to the solution X^β of the SDE

$$dX_t^\beta = b(X_t^\beta, \mu_t, \beta_t)dt + \sigma(X_t^\beta, \mu_t, \beta_t)dW_t,$$

with the same measure flow μ as before. Sending $n \to \infty$ on both sides of (8.3), we find

$$\begin{aligned}\mathbb{E}&\left[\int_0^T f(X_t^*, \mu_t, \widehat{\alpha}(t, X_t^*, \mu_t))dt + g(X_T^*, \mu_T)\right] \\ &\geq \mathbb{E}\left[\int_0^T f(X_t^\beta, \mu_t, \beta_t)dt + g(X_T^\beta, \mu_T)\right].\end{aligned}$$

This should hold for every choice of β, and we conclude that the control $\widehat{\alpha}$ is optimal for the control problem

$$\begin{array}{ll}\sup_\alpha & \mathbb{E}\left[\int_0^T f(X_t, \mu_t, \alpha_t)dt + g(X_T, \mu_T)\right] \\ dX_t & = b(X_t, \mu_t, \alpha_t)dt + \sigma(X_t, \mu_t, \alpha_t)dW_t.\end{array}$$

Moreover, μ is precisely the law of the optimally controlled state process X^* above!

In summary the above heuristic leads us to the following fixed point problem:

Definition 8.1. Define a map $\Phi : C([0,T], \mathcal{P}(\mathbb{R}^d)) \to C([0,T], \mathcal{P}(\mathbb{R}^d))$ as follows:

1. Fix a (deterministic) measure flow $\mu = (\mu_t)_{t\in[0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$, to represent a continuum of agents' state processes.

2. Solve the control problem faced by a *typical* or *representative agent*:

$$(P_\mu) \qquad \left\{\begin{array}{ll}\sup_\alpha & \mathbb{E}\left[\int_0^T f(X_t^{\mu,\alpha}, \mu_t, \alpha_t)dt + g(X_T^{\mu,\alpha}, \mu_T)\right] \\ dX_t^{\mu,\alpha} & = b(X_t^{\mu,\alpha}, \mu_t, \alpha_t)dt + \sigma(X_t^{\mu,\alpha}, \mu_t, \alpha_t)dW_t.\end{array}\right.$$

   Here the law of the initial state $X_0^{\mu,\alpha}$ is the given λ_0.

3. Let α^* be the optimally controlled state process, which we assume is unique, and define $\Phi(\mu) = (\mathcal{L}(X_t^{\mu,\alpha^*}))_{t\in[0,T]}$.

We say that $\mu \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ is a *mean field equilibrium* (MFE) if it is a fixed point of Φ, or $\mu = \Phi(\mu)$.

Remark 8.2. Definition 8.1 is central to the rest of the course, so we take the time to clarify it with some comments:

1. While our definition, strictly speaking, refers only to the measure flow μ, it is often useful to include the optimal control. That is, we may refer more descriptively to the pair (μ, α^*) as an MFE, where α^* is optimal for the control problem (P_μ).

2. In general, the optimal control α^* for (P_μ) need not be unique. We could then could define

$$\Phi(\mu) = \left\{ (\mathcal{L}(X_t^*))_{t\in[0,T]} : X^* \text{ is optimal for } (P_\mu) \right\},$$

and then try to find a fixed point for this set-valued map, or $\mu \in \Phi(\mu)$. In this case, it becomes even more appropriate to include the control α^* in the definition of an MFE, as in remark (1) above.

3. It cannot be stressed enough that the measure flow $\mu = (\mu_t)_{t\in[0,T]}$ must be seen as *fixed* when solving the control problem (P_μ). The control problem (P_μ) is a completely standard stochastic optimal control problem, as the measure flow should be seen just as a time-dependence in the coefficient.

4. The fixed point can instead be formulated on the control process itself. That is, suppose we start with a control α, and then solve the McKean-Vlasov equation

$$dY_t = b(Y_t, \mu_t, \alpha_t)dt + \sigma(Y_t, \mu_t, \alpha_t)dW_t, \quad \mu_t = \mathcal{L}(Y_t), \ t \in [0, T].$$

Note that $Y = X^{\mu,\alpha}$. Using this measure flow, we then find an optimal control α^* for the control problem (P_μ) relative to this measure flow μ. If α^* agrees with the α we started with, then μ is a mean field equilibrium.

Theorem 8.3. *Assume the coefficients (b, σ, f, g) are bounded and continuous, the control space A is a closed subset of a Euclidean space, and $b = b(x, m, a)$ is Lipschitz in all variables, using $\mathcal{W}_1$ for the measure argument. Assume for simplicity that σ is constant. Assume the initial states (X_0^i) are i.i.d. and square integrable. Let μ be a Mean Field Equilibrium with corresponding optimal control α^*, which we assume is given in feedback form by a Lipschitz function $\alpha^* : [0, T] \times \mathbb{R}^d \to A$. Let player i use control $\alpha_t^{n,i} = \alpha^*(t, X_t^i)$. Then there exists $(\epsilon_n)_n \searrow 0$ such that the controls $\vec{\alpha}_n = (\alpha^{n,1}, \ldots, \alpha^{n,n})$ form an open-loop ϵ_n-Nash equilibrium for the n-player game, for each n.*

Remark 8.4. In Theorem 8.3, under suitable assumptions, it can be shown that the same controls $\vec{\alpha}_n = (\alpha^{n,1}, \ldots, \alpha^{n,n})$ form a *closed-loop* (Markovian) ϵ_n-Nash equilibrium for the n-player game, for each n. However, the closed-loop case is remarkably much more difficult to prove!

Proof. We provide a sketch of the proof in the n-player equilibrium is interpreted in the open loop sense, as this is much easier to handle. If each player uses α^*, then

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha^*(t, X_t^i))dt + \sigma(X_t^i, \mu_t^n, \alpha^*(t, X_t^i))dW_t^i$$

Because the coefficients are Lipschiz, we know from Theorem 3.3 on McKean-Vlasov equations that

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} |X_t^1 - \overline{X}_t|^2 + \sup_{t\in[0,T]} \mathcal{W}_1^2(\mu_t^n, \mu_t)\right],$$

where $\mu = (\mu_t)_{t\in[0,T]}$ and $\overline{X}$ solve the McKean-Vlasov equation

$$\begin{cases} d\overline{X}_t = & b(\overline{X}_t, \mu_t, \alpha^*(t, \overline{X}_t))dt + \sigma dW_t^1 \\ \mu_t = & \mathcal{L}(\overline{X}_t), \qquad \forall t \in [0, T]. \end{cases}$$

It follows from dominated convergence that the objective of player 1 has the following limit:

$$\lim_{n\to\infty} J_1^n(\vec{\alpha}_n) = \mathbb{E}\left[\int_0^T f(\overline{X}_t, \mu_t, \alpha^*(t, \overline{X}_t))dt + g(\overline{X}_T, \mu_T)\right] =: J^\infty(\alpha^*).$$

We then need to show that player 1 cannot earn much more by deviating his strategy. Define

$$\epsilon_n := \sup_\beta \ J_1^n((\vec{\alpha}_n^{-1}, \beta)) - J_1^n(\vec{\alpha}_n) \geq 0,$$

where $(\vec{\alpha}_n^{-1}, \beta)$ means that player 1 switches to β. Of course, $\epsilon_n \geq 0$, and trivially $\vec{\alpha}_n$ is an ϵ_n-Nash equilibrim for each n. It remains to show that $\epsilon_n \to 0$.

To do this, fix β. We may then write

$$J_1^n((\vec{\alpha}_n^{-1}, \beta)) = \mathbb{E}\left[\int_0^T f(Y_t^1, \nu_t^n, \beta_t)dt + g(Y_T^1, \nu_T^n)\right],$$

where $\alpha_t^i = \alpha^*(t, X_t^i)$ (we are working with open loop here) and

$$\begin{cases} dY_t^i = & b(Y_t^i, \nu_t^n, \alpha_t^i)dt + \sigma dW_t^i \quad \forall i \neq 1 \\ dY_t^1 = & b(Y_t^1, \nu_t^n, \beta_t)dt + \sigma dW_t^1 \\ \nu_t^n = & \frac{1}{n}\sum_{k=1}^n \delta_{Y_t^k}. \end{cases}$$

From the Mean Field Equilibrium property, $J^\infty(\alpha^*) \geq J^\infty(\beta)$ for all β, where

$$J^\infty(\beta) := \mathbb{E}\left[\int_0^T f(Y_t^\beta, \mu_t, \beta_t)dt + g(Y_T^\beta, \mu_T)\right],$$
$$dY_t^\beta = b(Y_t^\beta, \mu_t, \beta_t)dt + \sigma dW_t^1, \quad Y_0^\beta = X_0^1.$$

To prove that $\epsilon_n \to 0$, it suffices to show that

$$J_1^n((\vec{\alpha}_n^{-1}, \beta)) \underset{n\to\infty}{\to} J^\infty(\beta), \tag{8.4}$$

uniformly in β. To do this, we first show $(Y^1, \nu^n) \to (Y^\beta, \mu)$, uniformly in β. To do this, we first show that $\nu^n \to \mu$ uniformly in β. To see this, define

$$\mu_t^{n,-1} = \frac{1}{n-1}\sum_{k=2}^n \delta_{X_t^k}.$$

A standard estimate we have seen before shows that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \mathcal{W}_1(\mu_t^{n,-1}, \mu_t^n)\right] \leq \frac{C}{n}, \tag{8.5}$$

where C depends on the Lipschitz constants of the coefficients and on the second moment of the initial state. On the other hand,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]} \mathcal{W}_1^2(\mu_s^{n,-1}, \nu_s^n)\right] \\
&\quad\leq \mathbb{E}\left[\sup_{s\in[0,t]} \frac{1}{n-1}\sum_{k=2}^n |X_s^k - Y_s^k|^2\right] \\
&\quad\leq \frac{1}{n-1}\sum_{k=2}^n \mathbb{E}\left[\sup_{s\in[0,t]} \int_0^s \left|b(X_r^k, \mu_r^n, \alpha_r^i) - b(Y_r^k, \nu_r^n, \alpha_r^i)\right|^2\right] \\
&\quad\leq \frac{C}{n-1}\sum_{k=2}^n \mathbb{E}\left[\int_0^t \left(\sup_{r\in[0,s]} |X_r^k - Y_r^k|^2 + \sup_{r\in[0,s]} \mathcal{W}_1(\mu_r^n, \nu_r^n)\right) ds\right] \\
&\quad\leq \frac{C}{n} + \frac{C}{n-1}\sum_{k=2}^n \mathbb{E}\left[\int_0^t \left(\sup_{r\in[0,s]} |X_r^k - Y_r^k|^2 + \sup_{r\in[0,s]} \mathcal{W}_1(\mu_r^{n,-1}, \nu_r^n)\right) ds\right].
\end{aligned}$$

Here, the constant $C > 0$ may change from line to line but never depends on n. By Gronwall's inequality,

$$\begin{aligned}\mathbb{E}&\left[\sup_{s\in[0,t]}\frac{1}{n-1}\sum_{k=2}^{n}|X_s^k - Y_s^k|^2\right]\\&\le \frac{C}{n} + C\mathbb{E}\left[\int_0^t \sup_{r\in[0,s]}\mathcal{W}_1(\mu_r^{n,-1},\nu_r^n)ds\right].\end{aligned}$$

Apply Gronwall's inequality again to get

$$\mathbb{E}\left[\sup_{s\in[0,t]}\mathcal{W}_1^2(\mu_s^{n,-1},\nu_s^n)\right] \le \frac{C}{n}.$$

Combine this with (8.5) to get

$$\mathbb{E}\left[\sup_{t\in[0,T]}\mathcal{W}_1(\nu_t^n,\mu_t^n)\right] \le \frac{C}{n}.$$

Finally, we compare Y^β and Y^1,

$$\begin{aligned}|Y_t^1 - Y_t^\beta| &\le \int_0^t \left|b(Y_r^1,\nu_r^n,\beta_r) - b(Y_r^\beta,\mu_r,\beta_r)\right| dr\\&\le C\int_0^t \left(|Y_r^1 - Y_r^\beta| + \mathcal{W}_1(\nu_r^n,\mu_r)\right) dr.\end{aligned}$$

Apply Gronwall's inequality and the triangle inequality to get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t^1 - Y_t^\beta|\right] &\le C\mathbb{E}\left[\sup_{t\in[0,T]}\mathcal{W}_1(\nu_t^n,\mu_t^n)\right]\\&\le \frac{C}{n} + C\mathbb{E}\left[\sup_{t\in[0,T]}\mathcal{W}_1(\mu_t,\mu_t^n)\right].\end{aligned}$$

Now, because $\mu^n \to \mu$, we conclude that

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t^1 - Y_t^\beta| + \sup_{t\in[0,T]}\mathcal{W}_1(\mu_t,\mu_t^n)\right] = 0.$$

Notably, this convergence is uniform in β, as none of our estimates depended on the choice of β. Using dominated convergence, this proves that the limit (8.4) holds, uniformly in β. $\square$

8.3 Controlled Mckean-Vlasov dynamics

This short section serves to highlight the related but crucially distinct problem of optimal control of McKean-Vlasov equations, also known as *mean field control*. Now is a good time to recall the precise statement of Definition 8.1.

Definition 8.5. The general mean field control problem is described as follows:

1. For each control α, let μ^α denote the law of the McKean-Vlasov equation,

$$dX_t^\alpha = b(X_t^\alpha, \mu_t^\alpha, \alpha_t)dt + \sigma(X_t^\alpha, \mu_t^\alpha, \alpha_t)dW_t, \quad \mu_t^\alpha = \mathcal{L}(X_t^\alpha), \ t \in [0,T]. \tag{8.6}$$

2. Solve the following optimization problem:

$$\sup_\alpha \mathbb{E}\left[\int_0^T f(X_t^\alpha, \mu_t^\alpha, \alpha_t)dt + g(X_T^\alpha, \mu_T^\alpha)\right].$$

The key difference between the mean field control problem of Definition 8.5 and the MFG problem of Definition 8.1 is that the former requires that the law μ be matched to the controlled state process *before the optimization*. In other words, in the MFG problem, the measure parameter is not controlled or influenced in any way by the control during the resolution of the optimization problem (P_μ).

Intuitively, the mean field control problem represents a cooperative game with a continuum of agents. The setup is the same as the MFG problem, but the notion of equilibrium is different, and it is natural to ask how these two notions relate. In general, they are completely different, and to see this more clearly it helps to think about what the n-player analogue of the mean field control problem should be. For a fixed α, we know the McKean-Vlasov equation (8.6) arises as the limit of n-particle systems. So, for a vector of controls $\vec{\alpha} = (\alpha^1, \ldots, \alpha^n)$, say in closed-loop form, define the n-particle system

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha^i(t, \vec{X}_t))dt + \sigma(X_t^i, \mu_t^n, \alpha^i(t, \vec{X}_t))dW_t^i, \quad \mu_t^n = \frac{1}{n}\sum_{k=1}^n \delta_{X_t^k}.$$

We have studied competitive (Nash) equilibria in this context, but consider now the following cooperative problem: We now play the role of a central planner, and we simultaneously choose all of the controls for all of the players to optimize the average payoff received by the players. That is, we solve the optimal control problem

$$\sup_{\vec{\alpha}} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\int_0^T f(X_t^i, \mu_t^n, \alpha^i(t, \vec{X}_t))dt + g(X_T^i, \mu_T^n)\right].$$

We will not justify this here or study it any further, but reference will be provided in Section 8.9.

8.4 Analysis of a semi-linear-quadratic mean field game

Going back to our original MFG problem, we will analyze a relatively simple setup to illustrate how a typical uniqueness proof goes. We will see that Banach's fixed point theorem typically does not apply, but for small enough time horizon T one can indeed show the fixed point map to be a contraction. The typical existence proof, with no constraint on the time horizon, boils down to an application of Schauder's fixed point theorem, which we state without proof:

Theorem 8.6 (Schauder's Fixed Point Theorem)**.** *Let K be a convex compact subset of a topological vector space. Suppose $\Psi : K \to K$ is continuous. Then Ψ has a fixed point.*

We will apply Schauder's theorem with a convex compact set $K \subset \mathcal{P}(\mathcal{C}^d)$, where $\mathcal{C}^d = C([0,T];\mathbb{R}^d)$. Note that $\mathcal{P}(\mathcal{C}^d)$ equipped with weak convergence is not itself a topological vector space, but it is a subspace of one, namely the space of bounded signed measures on $\mathcal{C}^d$.

Let us restate our setup from Definition 8.1 in our specific context. We are given functions

$$f, g : \mathbb{R}^d \times \mathcal{P}^1(\mathbb{R}^d) \to \mathbb{R}^d.$$

Controls α_t will take values in $\mathbb{R}^d$. Let $\lambda_0 \in \mathcal{P}^2(\mathbb{R}^d)$ denote a square-integrable initial law.

We define a map $\Phi : \mathcal{P}^1(\mathcal{C}^d) \to \mathcal{P}(\mathcal{C}^d)$ as follows:

1. Fix $\mu \in \mathcal{P}(\mathcal{C}^d)$.

2. Solve the representative agent's optimal control problem:

$$(P_\mu) \qquad \begin{cases} \sup_\alpha & \mathbb{E}\left[\int_0^T \left(f(X_t^\alpha, \mu_t) - \frac{1}{2}|\alpha_t|^2\right) dt + g(X_T^{\mu,\alpha}, \mu_T)\right] \\ dX_t^\alpha & = \alpha_t dt + dW_t, \quad X_0^\alpha \sim \lambda_0. \end{cases}$$

The control α can be any adapted process (relative to the filtration generated by the Brownian motion) with $\mathbb{E}\int_0^T |\alpha_t|^2 dt < \infty$.

3. Let α^* be an optimally controlled state process (which will be unique in this case), and define $\Phi(\mu) = \mathcal{L}(X^{\alpha^*})$.

A mean field equilibrium (MFE) is now a fixed point $\mu \in \mathcal{P}(\mathcal{C}^d)$ of this map, $\mu = \Phi(\mu)$.

Theorem 8.7. *Assume $f = f(x, m)$ and $g = g(x, m)$ have first and second derivatives in x which are jointly continuous functions of (x, m), using the metric $\mathcal{W}_1$ for the variable $m \in \mathcal{P}^1(\mathbb{R}^d)$. Assume also that f and g and these derivatives are all uniformly bounded. Then there exists a MFE. Moreover, if the time horizon $T > 0$ is sufficiently small, and if $f(x, m)$ and $g(x, m)$ are $\mathcal{W}_1$-Lipschitz in m, uniformly in x, then the MFE is unique.*

Proof.
Step 1: We first fix $\mu \in \mathcal{P}^1(\mathcal{C}^d)$ and solve the control problem using the Hopf-Cole transformation. The Hamiltonian is

$$H(x, y) = \sup_{\alpha \in \mathbb{R}^d} \left(a \cdot y - \frac{1}{2}|a|^2 \right) + f(x, \mu_t) = \frac{1}{2}|y|^2 + f(x, \mu_t),$$

and the optimizer is $\widehat{\alpha}(x, y) = y$. The HJB equation is then

$$\begin{aligned} \partial_t v(t, x) + \frac{1}{2}|\nabla v(t, x)|^2 + f(x, \mu_t) + \frac{1}{2}\Delta v(t, x) &= 0 \\ v(T, x) &= g(x, \mu_T) \end{aligned}$$

We now apply the Hopf-Cole transformation: Let $u(t, x) = e^{v(t,x)}$. Then

$$\nabla u = u \nabla v, \qquad \Delta u = u(\Delta v + |\nabla v|^2).$$

Multiply the HJB equation on both sides by $u(t, x)$ to get

$$\begin{aligned} d_t u(t, x) + \frac{1}{2}\Delta u(t, x) + u(t, x) F(x, \mu_t) &= 0 \\ u(T, x) &= e^{g(x, \mu_T)} \end{aligned}$$

From the Feynman-Kac formula we can express the solution in terms of a standard d-dimensional Brownian motion $W = (W_t)_{t \in [0,T]}$, as follows:

$$\begin{aligned} u(t, x) &= \mathbb{E}\left[\exp\left(g(W_T, \mu_T) + \int_t^T f(W_s, \mu_s) ds \right) \,\middle|\, W_t = x \right] \\ &= \mathbb{E}\left[\exp\left(g(W_T - W_t + x, \mu_T) + \int_t^T f(W_s - W_t + x, \mu_s) ds \right) \right]. \end{aligned}$$

Then, $v(t,x) = \log u(t,x)$, and the optimal feedback control is

$$\alpha_\mu(t,x) = \nabla v(t,x) = \frac{\nabla u(t,x)}{u(t,x)}.$$

Step 2: Now that we have identified the (candidate) optimal feedback control, we next establish useful continuity properties of $\alpha_\mu(t,x)$, as a function of $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}^1(\mathcal{C}^d)$. This step of finding enough regularity of the optimal control, particularly as a function of the measure, is often the crux of an MFG fixed point analysis. We will show the following:

(i) $\alpha_\mu(t,x)$ is bounded uniformly in (t,x,μ).

(ii) The map $(t,x,\mu) \mapsto \alpha_\mu(t,x)$ is jointly continuos on $[0,T]\times\mathbb{R}^d\times\mathcal{P}^1(\mathcal{C}^d)$.

(iii) $|\partial_j \alpha^i_\mu(t,x)|$ is bounded uniformly in t,x,μ,i,j, where $\alpha^i_\mu(t,x)$ is the ith component of the vector $\alpha_\mu(t,x)$, and ∂_j denotes the derivatives with respect to the jth component of $x \in \mathbb{R}^d$. In particular, $\alpha_\mu(t,\cdot)$ is Lipshitz, uniformly in (t,μ).

(iv) If $f(x,m)$ and $g(x,m)$ are $\mathcal{W}_1$-Lipschitz in m, uniformly in x, then there exists $C > 0$ such that $|\alpha_\mu(t,x) - \alpha_\nu(t,x)| \le \mathcal{W}_1(\mu,\nu)$ for all t,x,μ,ν.

It will follow from these facts that the state process

$$dX^\mu_t = \alpha_\mu(t,X^\mu_t)dt + dW_t, \quad X^\mu_0 \sim \lambda_0, \tag{8.7}$$

has a unique strong solution, which ensures that we can apply the verification theorem to conclude that X^μ is indeed the optimally controlled state process.

To prove these facts, we first express $\alpha_\mu(t,x)$ and its first derivatives in more convenient forms. First, using the expression for $u(t,x)$ above, we may write

$$\begin{aligned}\alpha^i_\mu(t,x) &= \frac{\partial_i u(t,x)}{u(t,x)} \\ &= \frac{\mathbb{E}\left[Z(t,x,\mu)\left(\partial_i g(W_T - W_t + x,\mu_t) + \int_t^T \partial_i f(W_s - W_t + x,\mu_s)ds\right)\right]}{\mathbb{E}[Z(t,x,\mu)]},\end{aligned}$$

where we define the random variable

$$Z(t,x,\mu) = \exp\left(g(W_T - W_t + x,\mu_T) + \int_t^T f(W_s - W_t + x,\mu_s)ds\right).$$

Because $Z(t,x,\mu)$ is strictly positive, we see easily that α^i_μ is uniformly bounded. Indeed, if $\partial_i g$ and $\partial_i f$ are uniformly bounded by C, then

$$|\alpha^i_\mu(t,x)| \le \frac{\mathbb{E}\left[Z(t,x,\mu)\left(C+(T-t)C\right)\right]}{\mathbb{E}[Z(t,x,\mu)]} \le (1+T)C.$$

This proves (i). Moreover, by invoking the bounded convergence theorem, we see that claim (ii) follows from this expression for $\alpha^i_\mu(t,x)$ and the assumed joint continuity of f and g. To prove (iv), notice that e^h is Lipschitz whenever h is bounded and Lipschitz, simply because the function $x \mapsto e^x$ is locally Lipschitz. It is straightforward to check that $\mathcal{W}_1(\mu_t,\nu_t) \le \mathcal{W}_1(\mu,\nu)$ for all $\mu,\nu \in \mathcal{P}(\mathcal{C}^d)$, and (iv) follows.

Lastly, we prove (iii) by computing

$$\partial_j \alpha^i_\mu = \partial_j\left(\frac{\partial_i u}{u}\right) = \frac{\partial_{ij} u}{u} - \frac{\partial_i u \partial_j u}{u^2} = \frac{\partial_{ij} u}{u} - \alpha^i_\mu \alpha^j_\mu.$$

The first term can be written as

$$\frac{\partial_{ij} u(t,x)}{u(t,x)} = \frac{\mathbb{E}\left[Z(t,x,\mu)\left(S_{ij}(t,x,\mu) + S_i(t,x,\mu) S_j(t,x,\mu)\right)\right]}{\mathbb{E}[Z(t,x,\mu)]},$$

where we define the random variables

$$S_{ij}(t,x,\mu) = \partial_{ij} g(W_T - W_t + x, \mu_T) + \int_t^T \partial_{ij} f(W_s - W_t + x, \mu_s) ds,$$
$$S_i(t,x,\mu) = \partial_i g(W_T - W_t + x, \mu_T) + \int_t^T \partial_i f(W_s - W_t + x, \mu_s) ds,$$

for each $i,j = 1,\dots,d$. Because the first and second derivatives of f and g are uniformly bounded, we conclude that (iii) holds.

Step 3: Now that we have identified the optimal state process X^μ in (8.7), we define $\Phi(\mu) = \mathcal{L}(X^\mu)$. We show now that Φ is a continuous function from $\mathcal{P}^1(\mathcal{C}^d)$ to itself. First, notice that for any $\mu \in \mathcal{P}^1(\mathcal{C}^d)$ we use the boundedness of $\alpha_\mu(t,x)$ from the previous step to find

$$|X^\mu_t| \le Ct + |W_t| + |X_0|,$$

where the constant $C > 0$ may change from line to line but does not depend on the choice of μ. Recalling that $\mathbb{E}[|X_0|^2] < \infty$ and $\Phi(\mu) = \mathcal{L}(X^\mu)$, we find

$$\sup_{\mu \in \mathcal{P}^1(\mathcal{C}^d)} \int_{\mathcal{C}^d} \|x\|_\infty^2 \Phi(\mu)(dx) \le CT + C\mathbb{E}[|X_0|^2] + C\mathbb{E}[\|W\|_\infty^2] < \infty, \quad (8.8)$$

where $\|x\|_\infty = \sup_{t\in[0,T]}|x_t|$ for $x \in \mathcal{C}^d$. In particular, this shows that $\Phi(\mathcal{P}^1(\mathcal{C}^d)) \subset \mathcal{P}^1(\mathcal{C}^d)$.

We next show that Φ is continuous. To this end, let $\mu, \nu \in \mathcal{P}^1(\mathcal{C}^d)$. Use the Lipschitz continuity of $\alpha_\mu(t,\cdot)$ to get

$$\begin{aligned}
|X_t^\mu - X_t^\nu| &\le \int_0^t |\alpha_\mu(s,X_s^\mu) - \alpha_\nu(s,X_s^\nu)|ds \\
&\le \int_0^t |\alpha_\mu(s,X_s^\mu) - \alpha_\mu(s,X_s^\nu)|ds + \int_0^t |\alpha_\mu(s,X_s^\nu) - \alpha_\nu(s,X_s^\nu)|ds \\
&\le C\int_0^t |X_s^\mu - X_s^\nu|ds + \int_0^t |\alpha_\mu(s,X_s^\nu) - \alpha_\nu(s,X_s^\nu)|ds.
\end{aligned}$$

Use Gronwall's inequality to get

$$\|X^\mu - X^\nu\|_\infty \le C\int_0^T |\alpha_\mu(s,X_s^\nu) - \alpha_\nu(s,X_s^\nu)|ds.$$

Hence,

$$\begin{aligned}
\mathcal{W}_1(\Phi(\mu),\Phi(\nu)) &\le \mathbb{E}[\|X^\mu - X^\nu\|] \\
&\le C\mathbb{E}\int_0^T |\alpha_\mu(s,X_s^\nu) - \alpha_\nu(s,X_s^\nu)|ds. \qquad (8.9)
\end{aligned}$$

Now, if $\mu^n \to \mu \in \mathcal{P}^1(\mathcal{C}^d)$, then we use the continuity of α in μ and the bounded convergence theorem to conclude that $\mathcal{W}_1(\Phi(\mu^n),\Phi(\nu)) \to 0$. That is, Φ is continuous. If the additional assumption holds, that $f(x,m)$ and $g(x,m)$ are $\mathcal{W}_1$-Lipschitz in m, uniformly in x, then using property (iv) of Step 2 in (8.9) yields

$$\mathcal{W}_1(\Phi(\mu),\Phi(\nu)) \le CT\mathcal{W}_1(\mu,\nu).$$

If $T < 1/C$, we see that Φ is a contraction on the complete metric space $(\mathcal{P}(\mathcal{C}^d),\mathcal{W}_1)$, and Banach's fixed point theorem proves the claimed existence and uniqueness for small time horizon.

Step 4: Lastly, to apply Schauder's theorem, we find a compact convex set $K \subset \mathcal{P}(\mathcal{C}^d)$ such that Φ maps K into itself. First, recall that we saw at the beginning of Step 3 that

$$M := \sup_{\mu\in\mathcal{P}^1(\mathcal{C}^d)} \int_{\mathcal{C}^d} \|x\|_\infty^2 \, \Phi(\mu)(dx) < \infty.$$

Next, we make use of Aldous' criterion for tightness [75, Theorem 16.11]: Suppose we can show that for any $(\mu^n) \subset \mathcal{P}^1(\mathcal{C}^d)$, and stopping times τ_n is any sequence of stopping times taking values in $[0,T]$, and any $\delta_n \to 0$, we have

$$\lim_{n\to\infty} \mathbb{E}\left[|X^{\mu^n}_{T\wedge(\tau_n+\delta_n)} - X^{\mu^n}_{\tau_n}| \wedge 1\right] = 0.$$

Then $\Phi(\mathcal{P}^1(\mathcal{C}^d)) = \{\mathcal{L}(X^\mu) : \mu \in \mathcal{P}^1(\mathcal{C}^d)\}$ is tight. To check this, note simply that because $|\alpha_\mu(t,x)| \le \|\alpha\|_\infty < \infty$ is bounded, we have

$$\begin{aligned}\mathbb{E}\left[|X^{\mu^n}_{T\wedge(\tau_n+\delta_n)} - X^{\mu^n}_{\tau_n}|\right] &\le \mathbb{E}\left[\int_{\tau_n}^{T\wedge(\tau_n+\delta)} |\alpha_{\mu^n}(t, X^{\mu^n}_t)|dt + |W_{T\wedge(\tau_n+\delta)} - W_{\tau_n}|\right] \\ &= \delta\|\alpha\|_\infty + \sqrt{\delta}.\end{aligned}$$

Now, let $X : \mathcal{C} \to \mathcal{C}$ denote the coordinate map, defined by $X_t(x) = x_t$ for $x \in \mathcal{C}$. Define

$$K = \{P \in \mathcal{P}^1(\mathcal{C}^d) : \mathbb{E}^P[\|X\|^2] \le M, \quad \mathbb{E}^P[|X_{T\wedge(\tau+\delta)} - X_\tau|] \le \delta||\alpha||_\infty + \sqrt{\delta},\ \forall \tau, \delta > 0\},$$

where τ ranges over $[0,T]$-valued stopping times. Clearly K is convex. It is tight thanks to Aldous' criterion. In light of Theorem 2.13, the second moment bound built into the definition of K ensures that K is in fact compact in $\mathcal{P}^1(\mathcal{C}^d)$. Hence, the closure $\overline{K}$ is convex and compact in both $\mathcal{P}^1(\mathcal{C}^d)$ and $\mathcal{P}(\mathcal{C}^d)$. Moreover, the topology generated by the metric $\mathcal{W}_1$ coincides with the weak convergence topology on $\overline{K}$. We saw that Φ is $\mathcal{W}_1$-continuous on $\mathcal{P}^1(\mathcal{C}^d)$, and we conclude that it is weakly continuous on $\overline{K}$. We can finally apply Schauder's Fixed Point Theorem to deduce that an equilibrium exists. $\square$

Remark 8.8. Theorem 8.7 is far from optimal, and it can be easily generalized in many directions. The boundedness constraints on the coefficients can be easily relaxed if one is more careful about integrability when applying the dominated convergence theorem. More importantly, there is no need to restrict one's attention to the semi-linear-quadratic structure; this was merely convenient in obtaining a workable expression for the optimal control using the Hopf-Cole transformation. That said, the structure of this proof is standard: (1) For each measure flow, solve the optimal control problem using your favorite methodology (HJB analysis, probabilistic maximum principle, etc.). (2) Find a way to establish continuity of the optimal control as a function of the input measure flow. (3) Argue that the laws of optimal state processes for various measure flows can be safely confined into a single compact set. (4) Apply Schauder's (or Kakutani's) fixed point theorem.

8.5 MFG PDE System

In this section we derive a PDE formulation for the MFG model studied in the previous section. In short, this will be a forward-backward PDE system consisting of an HJB equation for the value function of the control problem and a Fokker-Planck equation for the measure flow associated with the SDE. These two PDEs are coupled: the measure flow depends on the choice of control, which is obtained by maximizing the Hamiltonian along the solution of the HJB equation, and in turn the HJB equation depends on the choice of measure flow.

First, we recall how to derive the Fokker-Planck (a.k.a. Kolmogorov forward) equation associated with the SDE

$$dX_t = \alpha(t, X_t)dt + dW_t, \quad X_0 \sim \mu_0(dx) = m_0(x)dx,$$

where α is bounded and measurable, and the initial distribution μ_0 has a density m_0. Apply Itô's formula for a smooth function φ to get

$$\varphi(X_t) = X_0 + \int_0^t \left(\alpha(s, X_s) \cdot \nabla\varphi(X_s) + \frac{1}{2}\Delta\varphi(X_s)\right) ds + \int_0^t \nabla\varphi(X_s) \cdot dW_s.$$

As α is bounded, Girsanov's theorem shows that the law of X is equivalent to the law of Brownian motion started from initial law μ_0. Since the latter process has a density at each time t, so does X_t. That is, the law of X_t admits a probability density function, which we denotes $\mu(t, x)$. Taking expectations in the above equation yields

$$\int_{\mathbb{R}^d} \varphi(x)\mu(t, x)dx = \int_{\mathbb{R}^d} \varphi(x)m_0(x)dx + \int_0^t \int_{\mathbb{R}^d} \left(\alpha(s, x) \cdot \nabla\varphi(x) + \frac{1}{2}\Delta\varphi(x)\right) \mu(s, x)dxds.$$

If $\mu \in C^{1,2}\left([0, T] \times \mathbb{R}^d\right)$, then we may differentiate the above to find

$$\begin{aligned}
\int \varphi(x)\partial_t\mu(t, x)dx &= \frac{d}{dt}\int_{\mathbb{R}^d} \varphi(x)\mu(t, x)dx \\
&= \int_{\mathbb{R}^d} \left(\alpha(t, x) \cdot \nabla\varphi(x) + \frac{1}{2}\Delta\varphi(x)\right) \mu(t, x)dx \\
&= \int \left[-\mathrm{div}(\mu(t, x)\alpha(t, x)) + \frac{1}{2}\Delta\mu(t, x)\right] \varphi(x)dx,
\end{aligned}$$

where div acts on the x variable, and the last step follows from integration by parts. As this holds for all smooth φ, we find that μ solves the PDE

$$\begin{aligned}
\partial_t\mu(t, x) + \mathrm{div}(\mu(t, x)\alpha(t, x)) - \frac{1}{2}\Delta\mu(t, x) &= 0 \\
\mu(0, x) &= m_0(x)
\end{aligned}$$

With this in mind, we next write down the HJB equation. For a fixed measure flow $(\mu_t)_{t\in[0,T]}$, the HJB equation should encode the following control problem:

$$\begin{aligned}
&\sup_{\alpha} \mathbb{E}\left[\int_0^t (f(X_t^\alpha, \mu_t) - \frac{1}{2}|\alpha_t|^2)dt + g(X_T^\alpha, \mu_T)\right] \\
&dX_t^\alpha = \alpha_t dt + dW_t.
\end{aligned}$$

This HJB equation, as we saw in the previous section, is simply

$$\begin{aligned}
\partial_t v(t,x) + \frac{1}{2}|\nabla v(t,x)|^2 + \frac{1}{2}\Delta v(t,x) + f(x,\mu_t) = 0 \\
v(T,x) = g(x,\mu_T)
\end{aligned}$$

The optimal control is $\alpha(t,x) = \nabla v(t,x)$, so we now plug this into the Fokker-Planck equation above to attain our coupled MFG PDE system:

$$\begin{aligned}
\partial_t v(t,x) + \frac{1}{2}|\nabla v(t,x)|^2 + \frac{1}{2}\Delta v(t,x) + f(x,\mu_t) = 0 \\
\partial_t \mu(t,x) + \mathrm{div}(\mu(t,x)\nabla v(t,x)) - \frac{1}{2}\Delta \mu(t,x) = 0 \\
v(T,x) = g(x,\mu_T), \quad \mu(0,x) = m_0(x).
\end{aligned}$$

This is a *forward-backward* PDE system because one equation has a terminal condition (time T) while the other has an initial condition (time zero).

Forward-backward systems in general, whether ODEs, SDEs, or PDEs, tend to be quite hard to analyze, and the example below sheds some light on why. Mean field games are inherently forward-backward systems, and our existence Theorem 8.7 is somehow typical for forward-backward systems: Existence and uniqueness by contraction arguments only work on small time horizons, because the forward-backward structure impedes the usual Picard iteration argument. On the other hand, existence can be established more easily on arbitrary time horizon using a compactness-based fixed point theorem (like Schauder's) rather than a contraction-based fixed point theorem (Banach's).

Example 8.9. To understand the difficulty of forward-backward systems, consider the following ODE system. Let $T > 0$ and $a \in \mathbb{R}$, and consider:

$$\begin{aligned}
x'(t) = y(t), \quad x(0) = 0 \\
y'(t) = -x(t), \quad y(T) = a.
\end{aligned}$$

This implies $x''(t) = y'(t) = -x(t)$. Solve this to get $x(t) = c_1 \sin(t) + c_2 \cos(t)$ for some $c_1, c_2 \in \mathbb{R}$. The initial condition implies $0 = x(0) = c_2$. Thus $x(t) = c_1 \sin(t)$ and $y(t) = x'(t) = c_1 \cos(t)$. The terminal condition entails $a = y(T) = c_1 \cos(T)$. Now, there are two possibilities:

1. Suppose $\cos(T) = 0$. If $a = 0$, there are infinitely many solutions (one for each choice of $c_1 \in \mathbb{R}$). If instead $a \neq 0$, then there are no solutions.

2. Suppose $\cos(T) \neq 0$. Then there exists a unique solution, with $c_1 = a/\cos(T)$.

As we can see, if we are flexible about the value of T, we can guarantee uniqueness.

8.6 Uniqueness

As we know, we cannot expect mean field equilibria to be unique in general. However, there is a well known monotonicity assumption, often called the *Lasry-Lions monotonicity condition*, which does ensure uniqueness. This is very similar in spirit to Theorem 4.6 of Section 4.1.

Suppose the following hold, in addition to the usually omitted potpourri of technical assumptions:

(1) There is no mean field interaction in the state dynamics: $b = b(x, a)$ and $\sigma = \sigma(x, a)$.

(2) The running objective function is separable, in the sense that $f(x, m, a) = f_1(x, a) + f_2(x, m)$ for some functions f_1 and f_2.

(3) For each measure flow μ, there is a unique optimal control for the problem

$$(P_\mu) \qquad \begin{cases} \sup_\alpha & \mathbb{E}\left[\int_0^T \left(f_1(X_t^\alpha, \alpha_t) + f_2(X_t^\alpha, \mu_t)\right) dt + g(X_T^\alpha, \mu_T)\right] \\ dX_t^\alpha & = b(X_t^\alpha, \alpha_t)dt + \sigma(X_t^\alpha, \alpha_t)dW_t. \end{cases}$$

(4) Monotonicity: For each $m_1, m_2 \in \mathcal{P}(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \left(g(x, m_1) - g(x, m_2)\right)(m_1 - m_2)(dx) \leq 0, \tag{8.10}$$

$$\int_{\mathbb{R}^d} \left(f_2(x, m_1) - f_2(x, m_2)\right)(m_1 - m_2)(dx) \leq 0.$$

Again, the most important of these assumptions is the monotonicity condition (4). Another way of writing it is

(4') For any $\mathbb{R}^d$-valued random variables X and Y, we have

$$\mathbb{E}\left[g(X, \mathcal{L}(X)) + g(X, \mathcal{L}(X)) - g(X, \mathcal{L}(Y)) - g(Y, \mathcal{L}(X))\right] \leq 0,$$
$$\mathbb{E}\left[f_2(X, \mathcal{L}(X)) + f_2(X, \mathcal{L}(X)) - f_2(X, \mathcal{L}(Y)) - f_2(Y, \mathcal{L}(X))\right] \leq 0.$$

Some examples will follow the uniqueness theorem:

Theorem 8.10. *Under the above conditions, there is at most one mean field equilibrium.*

Proof. Suppose $\mu = (\mu_t)_{t\in[0,T]}$ and $\nu = (\nu_t)_{t\in[0,T]}$ are two MFE, and suppose they are distinct. Let α and β be the optimal controls for (P_μ) and (P_ν), respectively. Note that α and β must be distinct, because if $\alpha_t = \beta_t$ a.s. for each t then we would have $\mu_t = \mathcal{L}(X_t^\alpha) = \mathcal{L}(X_t^\beta) = \nu_t$ for each t, contradicting the assumption that μ and ν are distinct. Now, because α is optimal for (P_μ), it certainly outperforms β, and we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \left(f_1(X_t^\alpha, \alpha_t) + f_2(X_t^\alpha, \mu_t)\right) dt + g(X_T^\alpha, \mu_T)\right] \\
&\quad > \mathbb{E}\left[\int_0^T \left(f_1(X_t^\beta, \alpha_t) + f_2(X_t^\beta, \mu_t)\right) dt + g(X_T^\beta, \mu_T)\right],
\end{aligned}$$

with strict equality thanks to assumption (3). Similarly, β is optimal for (P_ν), and so

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \left(f_1(X_t^\beta, \beta_t) + f_2(X_t^\beta, \nu_t)\right) dt + g(X_T^\beta, \mu_T)\right] \\
&\quad > \mathbb{E}\left[\int_0^T \left(f_1(X_t^\alpha, \alpha_t) + f_2(X_t^\alpha, \nu_t)\right) dt + g(X_T^\alpha, \nu_T)\right].
\end{aligned}$$

Adding these two inequalities, we see that the f_1 terms cancel out, leaving

$$\begin{aligned}
0 <& \mathbb{E}\left[\int_0^T \left(f_2(X_t^\alpha, \mu_t) + f_2(X_t^\beta, \nu_t) - f_2(X_t^\alpha, \nu_t) - f_2(X_t^\beta, \mu_t)\right) dt\right] \\
&+ \mathbb{E}\left[g(X_T^\alpha, \mu_T) + g(X_T^\beta, \nu_T) - g(X_T^\alpha, \nu_T) - g(X_T^\beta, \mu_T)\right].
\end{aligned}$$

From assumption (4) above, or rather (4'), the right-hand side is ≤ 0, which is a contradiction. □

A few examples illustrate the kinds of objective functions covered by the monotonicity assumption.

Example 8.11. If $g(x,m) = g(x)$ or $g(x,m) = g(m)$ depends on only one of the variables, then the inequality (8.10) holds.

Example 8.12. Suppose $g(x,m) = (\varphi(x) - \langle m, \varphi\rangle)^2$, for some function φ, where we abbreviate $\langle m, \varphi\rangle = \int_{\mathbb{R}^d} \varphi \, dm$. This objective function encourages agents to make $\varphi(X_t^i)$ as distant as possible from the average $\frac{1}{n}\sum_{k=1}^n \varphi(X_t^k)$. For example, if $\varphi(x) = x$, then this encourages agents to spread out their state processes. Then

$$g(x,m) = \varphi^2(x) - 2\varphi(x)\langle m, \varphi\rangle + \langle m, \varphi\rangle^2.$$

With the previous example in mind, we compute

$$\begin{aligned}
&\int_{\mathbb{R}^d} (g(x,m_1) - g(x,m_2))\,(m_1 - m_2)(dx) \\
&\quad = -2\int_{\mathbb{R}^d} \varphi(x)\,(\langle m_1, \varphi\rangle - \langle m_2, \varphi\rangle)\,(m_1 - m_2)(dx) \\
&\quad = -2\,(\langle m_1, \varphi\rangle - \langle m_2, \varphi\rangle)^2 \\
&\quad \le 0.
\end{aligned}$$

Example 8.13. Suppose $g(x,m) = \int_{\mathbb{R}^d} \varphi(x,y)\, m(dy)$ for some bounded continuous function φ, which we assume for simplicity is symmetric in the sense that $\varphi(x,y) = \varphi(y,x)$. Then

$$\begin{aligned}
&\int_{\mathbb{R}^d} (g(x,m_1) - g(x,m_2))\,(m_1 - m_2)(dx) \\
&\quad = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \varphi(x,y)(m_1 - m_2)(dy)(m_1 - m_2)(dx).
\end{aligned}$$

This can be shown to be nonnegative for all m_1, m_2 if and only if $-\varphi$ is a *positive definite kernel*. This means that for each $n \in \mathbb{N}$ the $n \times n$ matrix $(-\varphi(x_i, x_k))_{i,k=1,\ldots,n}$ is positive semidefinite. To see why this is equivalent to the inequality (8.10), suppose that we choose two discrete measures supported on the same n points, namely $m_1 = \sum_{k=1}^n p_k \delta_{x_k}$ and $m_2 = \sum_{k=1}^n q_k \delta_{x_k}$ for some $n \in \mathbb{N}$ $x_1, \ldots, x_n \in \mathbb{R}^d$. Then we get

$$\int_{\mathbb{R}^d} (g(x,m_1) - g(x,m_2))\,(m_1 - m_2)(dx) = \sum_{i=1}^n \sum_{k=1}^n (p_i - q_i)(p_k - q_k)\varphi(x_i, x_k).$$

This must be nonpositive, for every choice of the p_k's, q_k's, and x_k's. Any measure can be approximated by discrete measures of this form.

Example 8.14. To make the above example more concrete, suppose $g(x,m) = \int_{\mathbb{R}^d} h(|x-y|)\, m(dy)$ for some bounded continuous function h on $\mathbb{R}_+$. This is a special case of the previous example, with $\varphi(x,y) = h(|x-y|)$. To mention the requisite jargon, to say that φ is a *positive definite kernel* is equivalent to saying that *positive definite function*, and this is a somewhat better understood class of functions. Bochner's theorem characterizes such functions, but let us just mention a sufficient condition: h is positive definite if it is continuous, nonincreasing, and convex.

8.7 Mean field games with different types of agents

We will not spend much time on this, but in applications it is useful to know that there is a fairly straightforward way to adapt the MFG paradigm to model more heterogeneous agents. The idea is to introduce a *type parameter*, similar to what we did for static games in Section 4.4.

In the n-player game, we suppose that each agent $i = 1, \ldots, n$ is assigned a type parameter $\theta_i \in \Theta$, where Θ is some Polish space. The state process of agent i is then

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha_t^i, \theta_i)dt + \sigma(X_t^i, \mu_t^n, \alpha_t^i, \theta_i)dW_t^i, \quad X_0^i = x_0^i,$$

where the initial states $x_0^i \in \mathbb{R}^d$ are given and where now

$$\mu_t^n = \frac{1}{n}\sum_{k=1}^n \delta_{(X_t^k, \theta_k)}$$

is the empirical joint distribution of states and types. To be clear, b and σ are now functions on the space $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d \times \Theta) \times A \times \Theta$, where A is as usual the action space. Similarly, the type parameter enters into the objective functions, and agent i seeks to maximize

$$J_i(\vec{\alpha}) = \mathbb{E}\left[\int_0^T f(X_t^i, \mu_t^n, \alpha_t^i, \theta_i)dt + g(X_T^i, \mu_T^n, \theta_i)\right].$$

We can expect to have a meaningful mean field limit if the empirical type distribution converges, in the sense that we have the weak convergence

$$\frac{1}{n}\sum_{k=1}^n \delta_{(x_0^k, \theta_k)} \to M = M_0(dx, d\theta) \in \mathcal{P}(\mathbb{R}^d \times \Theta).$$

(Of course, the x_0^k and θ_k can depend on n, but we suppress this from the notation.) To adapt mean field game problem of Definition 8.1, the idea is to assign initial states and type parameters independently to each of the continuum of agents, according to the distribution M_0, at the beginning of the game. In other words, at time zero each agent is independently given a type, and then the game is played. Agents now interact not only through the state distribution but also through the state-type distribution, and it is this distribution that should solve a fixed point problem.

Definition 8.15. As usual, let $\mathcal{C}^d = C([0,T];\mathbb{R}^d)$ denote the path space. Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, supporting an $\mathbb{F}$-Brownian motion W as well as $\mathcal{F}_0$-measurable random variables ξ and θ, with values in $\mathbb{R}^d$ and Θ, respectively, and with $(\xi,\theta) \sim M_0$, where $M_0 \in \mathcal{P}(\mathbb{R}^d \times \Theta)$ is a given distribution. Define a map $\Phi : \mathcal{P}(\mathcal{C}^d \times \Theta) \to \mathcal{P}(\mathcal{C}^d \times \Theta)$ as follows:

1. Fix a (deterministic) measure $\mu \in \mathcal{P}(\mathcal{C}^d \times \Theta)$, to represent the state-type distribution of a continuum of agents.

2. Solve the control problem faced by a typical agent:

$$(P_\mu) \qquad \begin{cases} \sup_\alpha & \mathbb{E}\left[\int_0^T f(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dt + g(X_T^{\mu,\alpha}, \mu_T, \theta)\right] \\ dX_t^{\mu,\alpha} & = b(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dt + \sigma(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dW_t, \quad X_0^{\mu,\alpha} = \xi. \end{cases}$$

3. Let α^* be the optimal control, which we assume is unique, and define $\Phi(\mu) = \mathcal{L}(X^{\mu,\alpha^*}, \theta)$.

We say that μ is a *mean field equilibrium* (MFE) if it is a fixed point of Φ, or $\mu = \Phi(\mu)$.

There is an alternative, equivalent formulation that can be more intuitive.

Definition 8.16. As usual, let $\mathcal{C}^d = C([0,T];\mathbb{R}^d)$ denote the path space. Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, supporting an $\mathbb{F}$-Brownian motion W. We are given also a type distribution $M_0 \in \mathcal{P}(\mathbb{R}^d \times \Theta)$ is a given distribution. Define a map $\Phi : \mathcal{P}(\mathcal{C}^d \times \Theta) \to \mathcal{P}(\mathcal{C}^d \times \Theta)$ as follows:

1. Fix a (deterministic) measure $\mu \in \mathcal{P}(\mathcal{C}^d)$, to represent the state-type distribution of a continuum of agents.

2. Fix a type parameter $(x_0, \theta) \in \mathbb{R}^d \times \Theta$.

3. Solve the control problem faced by a typical agent:

$$(P_\mu) \quad \begin{cases} \sup_\alpha & \mathbb{E}\left[\int_0^T f(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dt + g(X_T^{\mu,\alpha}, \mu_T, \theta)\right] \\ dX_t^{\mu,\alpha} & = b(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dt + \sigma(X_t^{\mu,\alpha}, \mu_t, \alpha_t, \theta)dW_t, \quad X_0^{\mu,\alpha} = x_0. \end{cases}$$

   Note that here we use the deterministic initial state x_0 and type parameter θ.

4. Let α^* be the optimal control, which we assume is unique, and define a new probability measure $\nu_{x_0,\theta} = \mathcal{L}(X^{\mu,\alpha^*})$.

5. Define a new measure $\nu \in \mathcal{P}(\mathcal{C}^d)$ by $\nu = \int_{\mathbb{R}^d \times \Theta} \nu_{x_0,\theta} M_0(dx_0, d\theta)$. More explicitly, this is the (mean) measure defined by

$$\nu(S) = \int_{\mathbb{R}^d \times \Theta} \nu_{x_0,\theta}(S) M_0(dx_0, d\theta), \quad \text{for Borel sets } S \subset \mathcal{C}^d,$$

   which is a well defined probability measure as long as the map $(x_0, \theta) \mapsto \nu_{x_0,\theta}$ is measurable.

6. Let $\Phi(\mu) = \nu$.

We say that μ is a *mean field equilibrium* (MFE) if it is a fixed point of Φ, or $\mu = \Phi(\mu)$.

In the second definition, we solve the control problem (P_μ) of a typical agent for each possible type parameter (x_0, θ) that the agent may be assigned. The idea is that the agent arrives at time zero and is randomly assigned a type parameter from the distribution M_0, and once this assignment is made they solve a control problem. The measures $\nu_{x_0,\theta}$ represent the law of the optimal state process *given the assignment of the type parameter*. Since these type parameters are assigned independently to all agents, we aggregate these conditional measures into the unconditional one ν, and this is what should agree with the starting measure μ which was assumed to represent the (unconditional) state-type distribution of the population.

Example 8.17. One natural specification of the above setup is when there finitely many types, or $\Theta = \{\theta_1, \ldots, \theta_K\}$. Intuitively, we think of the population as consisting of K different types of agents, with type-k agents all having type parameter θ_k. For any mean field term $\mu \in \mathcal{P}(\mathbb{R}^d \times \Theta)$, one may define the *mean field of subpopulation k* as the conditional law $\mu_k \in \mathcal{P}(\mathbb{R}^d)$, defined by

$$\mu_k(\cdot) = \frac{\mu(\cdot \cap \{\theta_k\})}{\mu(\mathbb{R}^d \times \{\theta_k\})}.$$

This is well defined as long as $\mu(\mathbb{R}^d \times \{\theta_k\}) > 0$. One can then allow the coefficients and objective functions to explicitly depend in different ways on the mean field of each subpopulation. That is, we could work with a running objective function of the form

$$f(x, \mu_1, \ldots, \mu_K, \alpha).$$

In the analysis of an MFG system with different types, one will typically need to assume the coefficients in Definition 8.15 or 8.16 are continuous in the $\mathcal{P}(\mathbb{R}^d \times \Theta)$ variable. In this finite-types example, it can be shown that the ma $\mathcal{P}(\mathbb{R}^d \times \Theta) \ni \mu \mapsto (\mu_1, \ldots, \mu_K) \in \mathcal{P}(\mathbb{R}^d)^K$ is (weakly) continuous, as long as we restrict to the set of μ satisfying $\mu(\mathbb{R}^d \times \{\theta_k\}) > 0$ for all $k = 1, \ldots, K$. This useful continuity property fails miserably when the type space Θ is uncountable; the operation of conditioning a measure is not very well-behaved topologically.

8.7.1 A linear-quadratic mean field game with different types of agents

In this section we revisit the simple linear-quadratic n-player game of Section 7.2, and we now allow each agent to have a different parameter $\lambda_i \geq 0$. The n-player game consists of state processes $X^1, \ldots, X^n$ given by

$$n - player : dX_t^i = \alpha_t^i dt + dW_t^i, \quad X_0^i = x_0^i$$

with each agent i choosing α^i to try to maximize

$$\mathbb{E}\left[-\int_0^T \frac{1}{2}|a_t^i|^2 dt - \frac{\lambda_i}{2}|\overline{X}_T - X_T^i|^2\right], \qquad \overline{X}_t := \frac{1}{n}\sum_{k=1}^n X_t^k.$$

To formulate the corresponding mean field game, we imagine that the empirical measure of initial states and types is approximately some $M \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}_+)$. That is, $\frac{1}{n}\sum_{k=1}^n \delta(x_0^k, \lambda_k) \approx M$.

We proceed as in Definition 8.16, defining a map $\Phi : \mathcal{P}(\mathcal{C}^d \times \Theta) \to \mathcal{P}(\mathcal{C}^d \times \mathbb{R}_+)$ as follows:

1. Fix a target $z \in \mathbb{R}^d$ and a type parameter $(x_0, \lambda) \in \mathbb{R}^d \times \mathbb{R}_+$.

2. Solve the control problem faced by a typical agent:

$$(P[z, x_0, \lambda]) \qquad \begin{cases} \sup_\alpha & \mathbb{E}\left[\int_0^T -\frac{1}{2}|\alpha_t|^2 dt - \frac{\lambda}{2}|z - X_T^\alpha|^2\right] \\ dX_t^\alpha & = \alpha_t dt + dW_t, \quad X_0^\alpha = x_0. \end{cases}$$

3. Let α^* be the optimal control, which we will show is unique, and calculate the resulting expected final value of the state process, $\widehat{z}[z, x_0, \lambda] = \mathbb{E}[X_T^{\alpha^*}]$

4. A mean field equilibrium is now any $z \in \mathbb{R}^d$ satisfying the fixed point equation

$$z = \int_{\mathbb{R}^d \times \mathbb{R}_+} \widehat{z}[z, x_0, \lambda] \, M(dx_0, d\lambda). \tag{8.11}$$

For fixed (z, x_0, λ) we have already seen how to solve the control problem $(P[z, x_0, \lambda])$. Specifically, from Section 5.5, the optimal control in step 3 above is

$$\alpha^*(t, x) = \frac{z - x}{\frac{1}{\lambda} + T - t}.$$

Note that the optimal control does not depend on the initial state x_0. The optimal state process is then

$$dX_t^{\alpha^*} = \frac{z - X_t^{\alpha^*}}{\frac{1}{\lambda} + T - t} dt + dW_t \,, \quad X_0^{\alpha^*} = x_0$$

Taking expectations, we find

$$\mathbb{E}[X_t^{\alpha^*}] = \int_0^t \frac{z - \mathbb{E}[X_s^{\alpha^*}]}{\frac{1}{\lambda} + T - s} ds + x_0.$$

We can compute this as follows. Define $f(t) = \mathbb{E}[X_t^{\alpha^*}] - z$. Then $f(\cdot)$ solves the ODE

$$f'(t) = -\frac{f(t)}{\frac{1}{\lambda} + T - t} \,, \quad f(0) = x_0 - z.$$

The unique solution is quickly found to be

$$\begin{aligned} f(t) &= (x_0 - z) \exp\left(- \int_0^t \frac{1}{\frac{1}{\lambda} + T - s} ds \right) \\ &= (x_0 - z) \frac{\frac{1}{\lambda} + T - t}{\frac{1}{\lambda} + T} \end{aligned}$$

Hence

$$\widehat{z}[z, x_0, \lambda] = \mathbb{E}[X_T^{\alpha^*}] = f(T) + z = z + \frac{x_0 - z}{1 + \lambda T}$$

To resolve the fixed point equation in step 4 above, note that

$$\begin{aligned}\int \widehat{z}[z, x_0, \lambda]\, M(dx_0, d\lambda) &= z + \int \frac{x_0 - z}{1 + \lambda T}\, M(dx_0, d\lambda) \\ &= z + \int \frac{x_0}{1 + \lambda T}\, M(dx_0, d\lambda) - z \int \frac{1}{1 + \lambda T}\, M(dx_0, d\lambda).\end{aligned}$$

Hence, $z \in \mathbb{R}^d$ solves the fixed point equation (8.11) if and only if

$$z = \frac{\int \frac{x_0}{1+\lambda T}\, M(dx_0, d\lambda)}{\int \frac{1}{1+\lambda T}\, M(dx_0, d\lambda)}.$$

This is precisely the mean of the terminal position X_T, which is exactly the target toward which each player tries to steer.

This formalism enables a streamlined "extreme case" analysis. Imagine, for example, that $c \gg 0$ is very large, and suppose $M = \mathcal{L}(X_0, \lambda)$ where λ is the random variable

$$\lambda = \begin{cases} 0\,, & \text{with probability } p \\ c\,, & \text{with probability } 1 - p. \end{cases}$$

Define the condition mean starting positions:

$$y_0 = \mathbb{E}[X_0 | \lambda = 0], \qquad y_\infty = \mathbb{E}[X_0 | \lambda = c].$$

Intuitively, this distribution M means that a fraction of p of the population has no need to steer toward anything (and thus will choose control $\alpha \equiv 0$), whereas the other $1-p$ of the population have a very large penalty for missing the target. We will take $c \to \infty$ soon, to enforce the binding constraint that this $1 - p$ fraction of the population *must reach the target.* The values y_0 is the average initial state of the agents with no penalty, while y_∞ is the average initial state of the agents with the high penalty. Using the above formula, the equilibrium is

$$z = \frac{p \cdot y_0 + \frac{y_\infty}{1+c\cdot T} \cdot (1 - p)}{p + \frac{1}{1+c\cdot T} \cdot (1 - p)}.$$

As $c \to \infty$, this converges exactly to y_0. In other words, equilibrium the agents with the high penalty all steer toward the average initial state of the sub-population of agents that has no terminal cost.

8.8 Mean Field Games with common noise

This section explains how to adapt the mean field game problem to allow for *common noise*, or what economists may call *aggregate shocks*. It may be helpful to review Section 3.6, which explains how the McKean-Vlasov limit reacts to a common noise term.

On the level of the n-player game, the idea is now to model the state processes by

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha_t^i)dt + \sigma(X_t^i, \mu_t^n, \alpha_t^i)dW_t^i + \gamma(X_t^i, \mu_t^n, \alpha_t^i)dB_t,$$
$$\mu^n t_t = \frac{1}{n}\sum_{k=1}^n \delta_{X_t^k},$$

with B and $W^1, \ldots, W^n$ independent Brownian motions. Again, the objective of agent i is to choose α^i to try to maximize

$$\mathbb{E}\left[\int_0^T f(X_t^i, \mu_t^n, \alpha_t^i)dt + g(X_T^i, \mu_T^n)\right].$$

Notably, the Brownian motion B influences all of the state processes, whereas W^i is specific to agent i.

We learned in our study of McKean-Vlasov equations that, unlike the independent noises, the influence of the common noise on the empirical measure should not average out as $n \to \infty$. As a result, an equilibrium measure flow $(\mu_t)_{t\in[0,T]}$ in the MFG should be stochastic, but adapted just to the filtration generated by the common noise.

These considerations lead to the following fixed point problem, which we formulate on some given filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. This space must support two independent $\mathbb{F}$-Brownian motions W and B, along with an $\mathcal{F}_0$-measurable initial state ξ. Let $\mathbb{F}^B = (\mathcal{F}_t^B)_{t\in[0,T]}$ denote the filtration generated by the common noise, i.e., $\mathcal{F}_t^B = \sigma(B_s : s \le t)$. Consider the following procedure:

1. Fix an $\mathbb{F}^B$-adapted $\mathcal{P}(\mathbb{R}^d)$-valued process $\mu = (\mu_t)_{t\in[0,T]}$.

2. Solve the control problem :
   $(P_\mu) : \begin{cases} \sup_\alpha \mathbb{E}[\int_0^T f(X_t, \mu_t, \alpha_t)dt + g(X_T, \mu_T)] \\ dX_t = b(X_t, \mu_t, \alpha_t)dt + \sigma(X_t, \mu_t, \alpha_t)dW_t + \gamma(X_t, \mu_t, \alpha_t)dB_t \end{cases}$
   With the process μ fixed, this is a stochastic optimal control problem *with random coefficients*.

3. Find optimal state process X^*, and set $\Phi(\mu)_=\mathcal{L}(X_t^*|\mathcal{F}_t^B)$.

This defines a map Φ from the set $\mathbb{F}^B$-adapted $\mathcal{P}(\mathbb{R}^d)$-valued processes into itself. A fixed point $\mu = \Phi(\mu)$ is a mean field equilibrium.

Remark 8.18. Stochastic control problems with random coefficients are notably more difficult than their nonrandom counterpart; by "random coefficients" we mean that the functions $(b, \sigma, \gamma, f, g)$ may depend rather generically on the underlying $\omega \in \Omega$. In this course we have not studied such problems. PDE methods become more difficult, as the HJB equation becomes a (typically much less tractable) *backward stochastic PDE*. However, probabilistic methods based on Pontryagin's maximum principle and FBSDEs is still just as feasible here.

8.8.1 Optimal investment revisited

We next illustrate an example of a MFG with common noise which we can solve explicitly. This is the continuum analogue of the n-player game of optimal investment studied in Section 7.3, and in addition it features different types of agents.

Recall that there is a single stock following the dynamics

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t.$$

When the typical agent invests a fraction α_t of his wealth in the stock, his wealth process becomes

$$dX_t = \alpha_t X_t(\mu dt + \sigma dB_t), \quad X_0 = \xi,$$

where $\xi > 0$. The objective is to maximize

$$\sup_{\alpha} \mathbb{E}\left[\frac{1}{\gamma}(X_T \overline{X}^{-\theta})^{\gamma}\right] = \sup_{\alpha} \mathbb{E}\left[\frac{1}{\gamma}(X_T^{1-\theta}(\frac{X_T}{\overline{X}})^{\theta})^{\gamma}\right].$$

where $\gamma < 1$, $\gamma \neq 0$, $\theta \in [0,1]$, and $\overline{X}$ captures the geometric mean wealth.

More precisely, we define a mean field equilibrium as follows. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a probability space that supports an $\mathbb{F}$-Brownian motion B and a random variable (ξ, γ, θ) with a given distribution $M \in \mathcal{P}(\mathbb{R}_+ \times \mathbb{R} \times [0,1])$. The filtration$\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ we take to be minimal in the sense that $\mathcal{F}_t = \sigma(\xi, \gamma, \theta, B_s : s \le t)$, and in particular B and (ξ, γ, θ) are independent. We define a fixed point problem as follows:

1. Fix a random variable $\overline{X} > 0$, to represent the geometric mean wealth.

2. Solve the problem : $\begin{cases} \sup_\alpha \mathbb{E}\left[\frac{1}{\gamma}(X_T\overline{X}^{-\theta})^\gamma\right] \\ \\ dX_t = \alpha_t X_t(\mu dt + \sigma dB_t, \quad X_0 = \xi.) \end{cases}$

3. Find optimal state process X^*. Compute the conditional geometric mean terminal wealth, $\Phi(\overline{X}) = \exp \mathbb{E}[\log X_T^* | \mathcal{F}_T^B]$.[9]

To solve this problem, we follow a similar strategy to Section 7.3. We look for a solution in which the optimal control is constant as a function of time, i.e., $\alpha_t = \alpha$ is $\mathcal{F}_0$-measurable. For a given $\mathcal{F}_0$-measurable random variable β, we write X^β for the process given by

$$dX_t^\beta = \beta X_t^\beta(\mu dt + \sigma dB_t, \quad X_0^\beta = \xi.$$

The key point that makes this problem tractable is that we do not need to solve for the player's best response to *every possible* $\mathcal{F}_T^B$*-measurable random variable* $\overline{X}$, but only those which arise from the terminal value of a constant-in-time control. That is, we will only consider $\overline{X}$ of the form $\overline{X} = \exp \mathbb{E}[\log X_T^\beta \,|\, \mathcal{F}_T^B]$, for $\mathcal{F}_0$-measurable random variables β.

To make this work, we compute the dynamics of $\exp \mathbb{E}[\log X_t^\beta \,|\, \mathcal{F}_t^B]$. To do this, first use Itô's formula to compute

$$d \log X_t^\beta = (\beta\mu - \frac{\beta^2\sigma^2}{2})dt + \beta\sigma dB_t.$$

Taking conditional expectations, we find

$$d\mathbb{E}[\log X_t^\beta \,|\, \mathcal{F}_t^B] = (\overline{\beta}\mu - \frac{\overline{\beta^2}\sigma^2}{2})dt + \overline{\beta}\sigma dB_t,$$

where we write $\overline{Z} := \mathbb{E}[Z]$ for the expectation of an $\mathcal{F}_0$-measurable random variable. Apply Itô's formula once again to get, with $\overline{X}_t := \exp \mathbb{E}[\log X_t^\beta \,|\, \mathcal{F}_t^B]$,

$$d\overline{X}_t = \overline{X}_t\left[\left(\overline{\beta}\mu - \frac{\sigma^2}{2}(\overline{\beta}^2 - \overline{\beta^2})\right)dt + \overline{\beta}\sigma dB_t\right], \quad \overline{X}_0 = \exp \mathbb{E}[\log \xi].$$

For an alternative control α, setting $Y_t = X_t^\alpha \overline{X}_t^{-\theta}$, we find

$$\frac{dY_t}{Y_t} = [(\mu - \sigma^2\theta\bar{\beta})a + \eta]dt + \sigma(a - \theta\bar{\beta})dB_t,$$

[9]Note: For $x_1, \ldots, x_n > 0$, the geometric mean can be written as $(\Pi_{i=1}^n x_i)^{\frac{1}{n}} = \exp(\frac{1}{n}\sum_{i=1}^n \log x_i)$. Hence, the natural extension of the *geometric mean* of a measure m on $\mathbb{R}_+$ is $\exp \int \log x \, m(dx)$.

where$\eta = -\mu\theta\bar{\beta} + \frac{\sigma^2}{2}(\overline{\beta^2} + \theta^2\bar{\beta}^2)$. The goal is now to treat this process Y as the state process, and optimize over controls α (which now do not need to be constant in time). That is, optimize

$$\sup_\alpha \mathbb{E}[\frac{1}{\gamma}Y_T^\gamma].$$

In solving this (standard!) optimal control problem, we end up with a PDE like that of Section 7.3. For a given β, the best response turns out to be

$$\alpha = \frac{\mu - \sigma^2\gamma\theta\bar{\beta}}{\sigma^2(1-\gamma)}.$$

To have an equilibrium, we should have $\alpha = \beta$. This requires

$$\overline{\alpha} = \overline{\beta} = \frac{\mu}{\sigma^2}\mathbb{E}\left[\frac{1}{1-\gamma}\right] - \overline{\alpha}\mathbb{E}\left[\frac{\gamma\theta}{1-\gamma}\right],$$

which we solve to get

$$\overline{\alpha} = \frac{\mu}{\sigma^2}\mathbb{E}\left[\frac{1}{1-\gamma}\right]\Big/\,\mathbb{E}\left[\frac{1-\gamma(1-\theta)}{1-\gamma}\right],$$

and

$$\alpha = \frac{\mu}{\sigma^(1-\gamma)}\left(1 - \gamma\theta\frac{\mathbb{E}\left[\frac{1}{1-\gamma}\right]}{\mathbb{E}\left[\frac{1-\gamma(1-\theta)}{1-\gamma}\right]}\right).$$

We leave it to the reader to fill in the details, following essentially the same argument of Section 7.3.

8.9 Bibliographic notes

Stochastic differential mean field games originate from the simultaneous works of Lasry-Lions [90, 91, 87, 88] and Huang-Malhamé-Caines [70, 92]; the former introduced the MFG PDE system of Section 8.5. The theory grew rapidly since its inception. Applications are by now too many to enumerate, including but not limited to economics and finance [55, 2, 1, 36, 71, 22, 59, 32, 85], electrical engineering [70, 92, 117], and crowd dynamics [81, 18].

The theoretical literature on mean field games can mostly be categorized as PDE-based or probabilistic. For a PDE perspective, see [90, 20, 60], and keep in mind that most PDE papers on the subject take the forward-backward PDE system of Section 8.5 as the starting point. The most common probabilistic approach, due to Carmona-Delarue [27], is based on an application of the stochastic (Pontryagin) maximum principle, which ultimately recasts the MFG problem as a McKean-Vlasov forward-backward SDE [26]. It was in this stream of literature that the difference between mean field games and the related but distinct problem of controlled McKean-Vlasov dynamics (discussed in Section 8.3) was first clarified [31, 29]. Lastly, a more abstract probabilistic approach based on compactification based on weak or martingale solutions of SDEs has gained steam recently [82, 6, 54], and this seems to be a promising way to prove rather general theorems about existence and convergence of equilibria.

The above discussion mostly pertains to solvability theory (existence and uniqueness) for MFGs, and it is safe to say that existence of equilibria is well understood to hold quite generally. The $n \to \infty$ limit theory, on the other hand, is still in its infancy and is much more challenging than that of static mean field games. Half of the picture is well understood: The MFG equilibrium can be used to build approximate equilibria for n-player games, as was observed in the original work of Huang-Malhamé-Caines [92]. This strategy has been extending in many contexts since then, particularly in the probabilistic literature, and it serves as a good justification for the use of a mean field approximation.

The other, less well-understood half of the picture pertains to the limits of the true n-player equilibria. That is, if we are given for each n a Nash equilibrium, what can be said of the limit(s) as $n \to \infty$, say of the empirical measure? Early results [90] required agents to be rather naive in the sense that agent i can only use a feedback control of the form $\alpha(t, X_t^i)$, based on his own state process and no others. More recently the picture has become clear when the n-player equilibria are taken to be open-loop [83, 49]; it can be shown that limit points of n-player approximate equilibria are a kind of "weak MFE." For closed-loop n-player equilibria, a breakthrough came in [23] with the discovery that the master equation, studied in the next section, can be used to prove this convergence, at least when it admits a smooth solution.

9 The master equation

This section develops an interesting analytic approach to mean field games, based on what is known as the *master equation* for mean field games. In some ways this plays the role of an HJB equation, describing how the value function depends on the starting time and state (t, x). For mean field games, the value function should depend on $(t, x, m) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$. The value $V(t, x, m)$ should be read as "the remaining value at time t for an agent with current state x given that the distribution of the continuum of other agents is m." To write a PDE for such a function, we must first understand how to take derivatives of functions on $\mathcal{P}(\mathbb{R}^d)$.

9.1 Calculus on $\mathcal{P}(\mathbb{R}^d)$

Let us say that a function $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is $\mathscr{C}^1$ if there exists a bounded continuous function $\frac{\delta U}{\delta m} : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ such that, for every $m, m' \in \mathcal{P}(\mathbb{R}^d)$,

$$U(m') - U(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}((1-t)m + tm', v)\,(m' - m)(dv)\,dt. \tag{9.1}$$

Equivalently,

$$\frac{d}{dh}\Big|_{h=0} U(m + h(m' - m)) = \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, v)\,(m' - m)(dv). \tag{9.2}$$

Only one function $\frac{\delta U}{\delta m}$ can satisfy (9.1), up to a constant shift; that is, if $\frac{\delta U}{\delta m}$ satisfies (9.1) then so does $(m, v) \mapsto \frac{\delta U}{\delta m}(m, v) + c(m)$ for any function $c : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$. For concreteness we always choose the shift $c(m)$ to ensure

$$\int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m, v)\,m(dv) = 0. \tag{9.3}$$

If $\frac{\delta U}{\delta m}(m, v)$ is continuously differentiable in v, we define its *intrinsic derivative* $D_m U : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$D_m U(m, v) = D_v \left(\frac{\delta U}{\delta m}(m, v) \right), \tag{9.4}$$

where we use the notation D_v for the gradient in v.

Remark 9.1. One may develop a similar theory of differentiation for functions on $\mathcal{P}^p(\mathbb{R}^d)$, for any exponent $p \geq 1$, as long as one is careful to require that the derivative $\delta U/\delta m$, if unbounded, satisfies some kind of growth conditions to ensure that the integral on the right-hand side of (9.1) is well-defined.

Before developing any further theory, we mention a few examples:

Example 9.2. Suppose $U(m) = \langle m, \varphi \rangle$ for $\varphi : \mathbb{R}^d \to \mathbb{R}$ bounded and continuous, where $\langle m, \varphi \rangle = \int \varphi \, dm$. Then, for any $m, m' \in \mathcal{P}(\mathbb{R}^d)$,

$$\frac{d}{dh}\Big|_{h=0} U(m + h(m' - m)) = \frac{d}{dh}\Big|_{h=0} \big(\langle m, \varphi \rangle + h\langle m' - m, \varphi \rangle\big) = \int \varphi \, d(m' - m).$$

Using (9.2), this shows that

$$\frac{\delta U}{\delta m}(m, v) = \varphi(v).$$

If φ is continuously differentiable, then

$$D_m U(m, v) = D_v \varphi(v).$$

Example 9.3. Suppose $U(m) = F(\langle m, \varphi \rangle)$ for smooth bounded functions $\varphi : \mathbb{R}^d \to \mathbb{R}$ and $F : \mathbb{R} \to \mathbb{R}$. Using the chain rule and the calculation of the previous example, we find

$$\frac{\delta U}{\delta m}(m, v) = F(m)\varphi(v), \qquad D_m U(m, v) = F(m) D_v \varphi(v).$$

Example 9.4. Suppose $U(m) = \iint \phi(x, y) m(dx) m(dy)$ for some smooth bounded function ϕ on $\mathbb{R}^d \times \mathbb{R}^d$. First observe

$$\begin{aligned} U(m + h(\tilde{m} - m)) = h^2 \iint \phi d(\tilde{m} - m)^2 + h \iint \phi(x, y) m(dx)(\tilde{m} - m)(dx, dy) \\ + h \iint \phi(x, y)(\tilde{m} - m)(dx) m(dy) + \iint \phi dm^2 \end{aligned}$$

and therefore

$$\frac{d}{dh}|_{h=0} U(m + h(\tilde{m} - m)) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} [\phi(x, y) + \phi(y, x)] m(dy)(\tilde{m} - m)(dx).$$

From this we conclude

$$\frac{\delta U}{\delta m}(m, v) = \int [\phi(x, v) + \phi(v, x)] m(dx).$$

We also make use of second derivatives of functions on $\mathcal{P}(\mathbb{R}^d)$. If, for each $v \in \mathbb{R}^d$, the map $m \mapsto \frac{\delta U}{\delta m}(m, v)$ is $\mathscr{C}^1$, then we say that U is $\mathscr{C}^2$ and let $\frac{\delta^2 U}{\delta m^2}$ denote its derivative, or more explicitly,

$$\frac{\delta^2 U}{\delta m^2}(m, v, v') = \frac{\delta}{\delta m}\left(\frac{\delta U}{\delta m}(\cdot, v)\right)(m, v').$$

We will also make some use of the derivative

$$D_v D_m U(m, v) = D_v[D_m U(m, v)],$$

when it exists, and we note that $D_v D_m U$ takes values in $\mathbb{R}^{d\times d}$. Finally, if U is $\mathscr{C}^2$ and if $\frac{\delta^2 U}{\delta m^2}(m, v, v')$ is twice continuously differentiable in (v, v'), we let

$$D_m^2 U(m, v, v') = D_{v,v'}^2 \frac{\delta^2 U}{\delta m^2}(m, v, v')$$

denote the $d \times d$ matrix of partial derivatives $(\partial_{v_i}\partial_{v'_j}[\delta^2 U/\delta m^2](m, v, v'))_{i,j}$. We have the following lemma, whose proof we leave as an exercise:

Lemma 9.5. *If U is $\mathscr{C}^2$, then we have*

$$D_m^2 U(m, v, v') = D_m^2 U(m, v', v).$$

Moreover, we may write

$$D_m^2 U(m, v, v') = D_m[D_m U(\cdot, v)](m, v').$$

That is, if we fix v and apply the operator D_m to the function $m \mapsto D_m U(m, v)$, then the resulting function is $D_m^2 U$.

The most important result for our purposes is the following, which shows how these derivatives interact with empirical measures.

Proposition 9.6. *Given $U : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, define $u_n : (\mathbb{R}^d)^n \to \mathbb{R}$ by $u_n(\boldsymbol{x}) = U(m_{\boldsymbol{x}}^n)$ for some fixed $n \geq 1$.*

(i) If U is $\mathscr{C}^1$ and if $D_m U$ exists and is bounded and jointly continuous, then u_n is continuously differentiable, and

$$D_{x_j} u_n(\boldsymbol{x}) = \frac{1}{n} D_m U(m_{\boldsymbol{x}}^n, x_j), \text{ for } j = 1, \dots, n. \tag{9.5}$$

(ii) If U is $\mathscr{C}^2$ and if $D_m^2 U$ exists and is bounded and jointly continuous, then u is twice continuously differentiable, and

$$D_{x_k} D_{x_j} u_n(\boldsymbol{x}) = \frac{1}{n^2} D_m^2 U(m_{\boldsymbol{x}}^n, x_j, x_k) + \delta_{j,k} \frac{1}{n} D_v D_m U(m_{\boldsymbol{x}}^n, x_j),$$

where $\delta_{j,k} = 1$ if $j = k$ and $\delta_{j,k} = 0$ if $j \neq k$.

Proof. Let $m \in \mathcal{P}(\mathbb{R}^d)$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$. By continuity, it suffices to prove the claims assuming the points $x_1, \ldots, x_n \in \mathbb{R}^d$ are distinct. Fix an index $j \in \{1, \ldots, n\}$ and a bounded continuous function $\phi : \mathbb{R}^d \to \mathbb{R}^d$, to be specified later. We claim that, under the assumptions of part (i),

$$\lim_{h \downarrow 0} \frac{U(m \circ (\mathrm{Id} + h\phi)^{-1}) - U(m)}{h} = \int_{\mathbb{R}^d} D_m U(m, v) \cdot \phi(v) m(dv) \tag{9.6}$$

holds, where Id denotes the identity map on $\mathbb{R}^d$. Once (9.6) is proven, we complete the proof as follows. For a fixed vector $v \in \mathbb{R}^d$ we may choose ϕ such that $\phi(x_j) = v$ while $\phi(x_i) = 0$ for $i \neq j$. Let $\boldsymbol{v} \in (\mathbb{R}^d)^n$ have j^{th} coordinate equal to v and i^{th} coordinate zero for $i \neq j$. Then $u_n(\boldsymbol{x}) = U(m^n_{\boldsymbol{x}})$ satisfies

$$\begin{aligned}
\lim_{h \downarrow 0} \frac{u_n(\boldsymbol{x} + h\boldsymbol{v}) - u_n(\boldsymbol{x})}{h} &= \lim_{h \downarrow 0} \frac{U(m^n_{\boldsymbol{x}} \circ (\mathrm{Id} + h\phi)^{-1}) - U(m^n_{\boldsymbol{x}})}{h} \\
&= \frac{1}{n} \sum_{k=1}^n D_m U(m^n_{\boldsymbol{x}}, x_k) \cdot \phi(x_k) \\
&= \frac{1}{n} D_m U(m^n_{\boldsymbol{x}}, x_j) \cdot v.
\end{aligned}$$

This proves (i). Under the additional assumptions, (ii) follows by applying (i) again.

It remains to prove (9.6). For $h > 0$, $t \in [0,1]$, and $m \in \mathcal{P}(\mathbb{R}^d)$, let $m_{h,t} = tm \circ (\mathrm{Id} + h\phi)^{-1} + (1-t)m$. Then, using (9.1) and (9.4), respectively, in the first and third equalities below, we obtain

$$\begin{aligned}
U(m \circ (\mathrm{Id} + h\phi)^{-1} - U(m) &= \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(m_{h,t}, v)\, (m \circ (\mathrm{Id} + h\phi)^{-1} - m)(dv)\, dt \\
&= \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta U}{\delta m}(m_{h,t}, v + h\phi(v)) - \frac{\delta U}{\delta m}(m_{h,t}, v) \right) m(dv)\, dt \\
&= h \int_0^1 \int_{\mathbb{R}^d} \int_0^1 D_m U(m_{h,t}, v + sh\phi(v)) \cdot \phi(v)\, ds\, m(dv)\, dt.
\end{aligned}$$

As $h \downarrow 0$ we have $m_{h,t} \to m$ and $sh\phi(v) \to 0$, and we deduce (9.6) from the bounded convergence theorem and continuity of D_m. ☐

Lastly, and in part for more practice working with this notion of derivative, we prove the following technical result, which will be used implicitly. It simply says that, for a function $U = U(x, m)$ on $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, the derivatives in x and in m commute.

Lemma 9.7. *For any function $U = U(x,m) : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, we have*

$$D_x D_m U(x,m,v) = D_m D_x U(x,m,v), \qquad \text{for all } v \in \mathbb{R}^d,$$

as long as the derivatives on both sides exist and are bounded and jointly continuous.

Proof. Fix $x \in \mathbb{R}^d$, $m, m' \in \mathcal{P}(\mathbb{R}^d)$, and for $t \in [0,1]$ let $m^t = (1-t)m + tm'$. Then by (9.1),

$$U(x,m') - U(x,m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta U}{\delta m}(x, m^t, v)(m' - m)(dv)dt.$$

Apply D_x to both sides and use dominated convergence to interchange the order of D_x and the integral to get

$$D_x U(x,m') - D_x U(x,m) = \int_0^1 \int_{\mathbb{R}^d} D_x \frac{\delta U}{\delta m}(x, m^t, v)(m' - m)(dv)dt.$$

Another application of (9.1) shows that for every $v \in \mathbb{R}^d$,

$$\frac{\delta D_x U}{\delta m}(x,m,v) = D_x \frac{\delta U}{\delta m}(x,m,v).$$

To conclude the proof, apply D_v to both sides, commute D_v and D_x on the right-hand side, and use the definition of D_m from (9.4). □

9.2 Ito's formula for $F(X_t, \mu_t)$

Using the calculus developed in the previous section, we will now see how to derive an Itô's formula for smooth functions of an Itô process and a (conditional) measure flow associated to a potentially different Itô process:

Theorem 9.8. *Consider processes X and $\bar{X}$ that solve the SDE:*

$$\begin{aligned} dX_t &= b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \gamma(X_t, \mu_t)dB_t, \qquad X_0 = \xi \\ d\bar{X}_t &= \bar{b}(\bar{X}_t, \mu_t)dt + \bar{\sigma}(\bar{X}_t, \mu_t)dW_t + \bar{\gamma}(\bar{X}_t, \mu_t)dB_t, \qquad \bar{X}_0 = \bar{\xi} \end{aligned} \tag{9.7}$$

where W, B are independent Brownian motions and $\mu_t = \mathcal{L}(\bar{X}_t | \mathcal{F}_t^B)$, where $(\mathcal{F}_t^B)_{t\in[0,T]}$ is the filtration generated by B. Assume for simplicity that the coefficients are all bounded, uniformly Lipschitz in the x variable, and continuous in the measure variable.

Then for $F : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ smooth we have

$$\begin{aligned} dF(X_t, \mu_t) = &[D_x F(X_t, \mu_t) b(X_t, \mu_t) + \frac{1}{2}\mathrm{Tr}[D_x^2 F(X_t, \mu_t)(\sigma\sigma^T + \gamma\gamma^T)(X_t, \mu_t)]]dt \\ &+ \int_{\mathbb{R}^d} D_m F(X_t, \mu_t, v)\bar{b}(v, \mu_t)\mu_t(dv)dt \\ &+ \int_{\mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_v D_m F(X_t, \mu_t, v)(\bar{\sigma}\bar{\sigma}^T + \bar{\gamma}\bar{\gamma}^T)(v, \mu_t)]\mu_t(dv)dt \\ &+ \iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_m^2 F(X_t, \mu_t, v, \tilde{v})\bar{\gamma}(v, \mu_t)\bar{\gamma}^T(\tilde{v}, \mu_t)]\mu_t(dv)\mu_t(d\tilde{v})dt \\ &+ \int_{\mathbb{R}^d} \mathrm{Tr}[D_x D_m F(X_t, \mu_t, v)\gamma(X_t, \mu_t)\bar{\gamma}^T(v, \mu_t)]\mu_t(dv)dt \\ &+ dM_t, \end{aligned}$$

where M_t is a martingale given by

$$\begin{aligned} dM_t = &D_x F(X_t, \mu_t)\Big(\sigma(X_t, \mu_t)dW_t + \gamma(X_t, \mu_t)dB_t\Big) \\ &+ \int_{\mathbb{R}^d} D_m F(X_t, \mu_t, v)\bar{\gamma}(v, \mu_t)\mu_t(dv)dB_t. \end{aligned}$$

Remark 9.9. There are multiple ways to show that μ_t is well defined. One is to set $\tilde{\mu} = \mathcal{L}(\bar{X}|B) = \mathcal{L}(\bar{X}|\mathcal{F}_T^B)$ as the conditional measure on the path space $C([0,T];\mathbb{R}^d)$ and then notice that for its marginal $\tilde{\mu}_t$ we have

$$\tilde{\mu}_t = \mathcal{L}(\bar{X}_t|\mathcal{F}_T^B) = \mathcal{L}(\bar{X}_t|\mathcal{F}_t^B \vee \sigma(B_s - B_t, s \geq t)) = \mathcal{L}(\bar{X}_t|\mathcal{F}_t^B)$$

by independence of increments of Brownian motion.

Proof of Theorem 9.8. We want to approximate the above given formula by empirical measures. To this end, we take $(W^i)_{i\in\mathcal{N}}$ to be iid Brownian motions, independent of W and B. Also consider iid random variables $(\bar{\xi}^i)_{i\in\mathcal{N}}$ such that $\bar{\xi}^i \overset{d}{=} \bar{\xi}$.

Define $\bar{X}^i$ as the solution to the system of SDE's given by

$$d\bar{X}_t^i = \bar{b}(\bar{X}_t^i, \mu_t)dt + \bar{\sigma}(\bar{X}_t^i, \mu_t)dW_t^i + \bar{\gamma}(\bar{X}_t^i, \mu_t)dB_t, \qquad \bar{X}_0^i = \bar{\xi}^i.$$

Let $\mu_t^n = \frac{1}{n}\sum_{k=1}^n \delta_{\bar{X}_t^k}$ denote the empirical measure. Because $\bar{X}^i$ are conditionally i.i.d. given B, it holds that $\mu^n \to \mu$ in probability as $n \to \infty$.

We now apply the standard Itô's formula to $F(X_t, \mu_t^n)$, which is simply a function of $n+1$ Itô processes:

$$
\begin{aligned}
dF(X_t,\mu_t^n) =& [D_xF(X_t,\mu_t^n)b(X_t,\mu_t) + \frac{1}{2}\mathrm{Tr}[D_x^2F(X_t,\mu_t^n)(\sigma\sigma^T + \gamma\gamma^T)(X_t,\mu_t)]]dt \\
&+ \frac{1}{n}\sum_{k=1}^n D_mF(X_t,\mu_t^n,\bar{X}_t^k)\bar{b}(\bar{X}_t^k,\mu_t)dt \\
&+ \frac{1}{n^2}\sum_{k=1}^n \frac{1}{2}\mathrm{Tr}[D_m^2F(X_t,\mu_t^n,\bar{X}_t^k,\bar{X}_t^k)\bar{\sigma}\bar{\sigma}^T(\bar{X}_t^k,\mu_t)]dt \\
&+ \frac{1}{n}\sum_{k=1}^n \frac{1}{2}\mathrm{Tr}[D_vD_mF(X_t,\mu_t^n,\bar{X}_t^k)(\bar{\sigma}\bar{\sigma}^T + \bar{\gamma}\bar{\gamma}^T)(\bar{X}_t^k,\mu_t)]dt \\
&+ \frac{1}{n^2}\sum_{j,k=1}^n \frac{1}{2}\mathrm{Tr}[D_m^2F(X_t,\mu_t^n,\bar{X}_t^k,\bar{X}_t^j)\bar{\gamma}(\bar{X}_t^k,\mu_t)\bar{\gamma}^T(\bar{X}_t^j,\mu_t)]dt \\
&+ \frac{1}{n}\sum_{k=1}^n \mathrm{Tr}[D_xD_mF(X_t,\mu_t^n,\bar{X}_t^k)\gamma(X_t,\mu_t)\bar{\gamma}^T(\bar{X}_t^k,\mu_t)]dt + dM_t^n
\end{aligned}
$$

and the martingale part is given by

$$
\begin{aligned}
dM_t^n =& d\tilde{M}_t^n + D_xF(X_t,\mu_t^n)(\sigma(X_t,\mu_t)dW_t + \gamma(X_t,\mu_t)dB_t) \\
&+ \frac{1}{n}\sum_{k=1}^n D_mF(X_t,\mu_t^n,\bar{X}_t^k)\bar{\gamma}(\bar{X}_t^k,\mu_t)dB_t
\end{aligned}
$$

where we define the martingale $\tilde{M}_t^n$ by setting $\tilde{M}_0^n = 0$ and

$$
d\tilde{M}_t^n = \frac{1}{n}\sum_{k=1}^n D_mF(X_t,\mu_t^n,\bar{X}_t^k)\bar{\sigma}(\bar{X}_t^k,\mu_t)dW_t^k.
$$

Note next that $\tilde{M}_t^n \to 0$ in the sense that $\mathcal{E}[\sup|\tilde{M}_t^n|^2] \to 0$. Indeed, this follows by Doob's inequality after we notice that

$$
\begin{aligned}
|[\tilde{M}_T^n]| &\leq \frac{1}{n^2}\sum_{k=1}^n \int_0^T |D_mF(X_t,\mu_t^n,\bar{X}_t^k)\bar{\sigma}(\bar{X}_t^k,\mu_t)|^2 dt \\
&\leq \frac{T}{n}\|D_mF\|_\infty^2\|\bar{\sigma}\|_\infty^2 \to 0.
\end{aligned}
$$

The term in the third line of the expression for $dF(X_t,\mu_t^n)$ is order $O(n^{-1})$ and thus converges to zero, as D_m^2 and $\bar{\sigma}$ are bounded. All the other terms converge to the corresponding terms because $\mu_t^n \to \mu_t$. Indeed, the term on the second line in the expression for $dF(X_t, \mu_t^n)$ can be written as

$$\int_{\mathbb{R}^d} D_m F(X_t, \mu_t^n, v)\bar{b}(v, \mu_t)\, \mu_t^n(dv),$$

which converges to

$$\int_{\mathbb{R}^d} D_m F(X_t, \mu_t, v)\bar{b}(v, \mu_t)\, \mu_t(dv)$$

because $D_m F$ and $\bar{b}$ are bounded and continuous in all variables. □

Remark 9.10. The key example to keep in mind where $\bar{b} = b$, $\bar{\sigma} = \sigma$, $\bar{\gamma} = \gamma$, and $\xi = \bar{\xi}$. Then $X = \bar{X}$, and our Itô's formula describes the dynamics of functions of the pair (X_t, μ_t) coming from a McKean-Vlasov equation.

To simplify notation, we denote the generator for X by L and the generator for μ by $\bar{L}$. That is, for smooth functions F as above, we set

$$LF(x,m) := D_x F(x,m) b(x,m) + \frac{1}{2}\mathrm{Tr}[D_x^2 F(x,m)(\sigma\sigma^T + \gamma\gamma^T)(x,m)], \tag{9.8}$$

and and $\bar{L}F(x,m)$ is given by the other dt terms in Itô's formula. Precisely,

$$\begin{aligned} \bar{L}F(x,m) = &\int_{\mathbb{R}^d} D_m F(x,m,v)\bar{b}(v,m)m(dv) \\ &+ \int_{\mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_v D_m F(x,m,v)(\bar{\sigma}\bar{\sigma}^T + \bar{\gamma}\bar{\gamma}^T)(v,m)]m(dv) \\ &+ \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_m^2 F(x,m,v,\tilde{v})\bar{\gamma}(v,m)\bar{\gamma}^T(\tilde{v},m)]m(dv)m(d\tilde{v}) \\ &+ \int_{\mathbb{R}^d} \mathrm{Tr}[D_x D_m F(x,m,v)\gamma(x,m)\bar{\gamma}^T(v,m)]m(dv) \end{aligned} \tag{9.9}$$

Itô's formula then reads

$$dF(X_t, \mu_t) = (LF(X_t, \mu_t) + \bar{L}F(X_t, \mu_t))dt + \text{martingale}.$$

To be careful, we should not really think of $\bar{L}$ itself as the generator for the process (μ_t), as the final term in the definition of $\bar{L}$ is really a cross-variation term between (X_t) and (μ_t).

9.3 Feynman-Kac formula

With an analogue of Itô's formula in hand, we can next derive PDEs for expectations of functionals of the pair (X_t, μ_t). Define the operators L and $\bar{L}$ as in (9.8) and (9.9), respectively.

Theorem 9.11. *Let $(t, x, m) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, and consider processes X and $\bar{X}$ that solve the SDEs*

$$
\begin{aligned}
dX_s^{t,x,m} &= b(X_s^{t,x,m}, \mu_s^{t,m})ds + \sigma(X_s^{t,x,m}, \mu_s^{t,m})dW_s + \gamma(X_s^{t,x,m}, \mu_s^{t,m})dB_s, \\
d\bar{X}_t^{t,m} &= \bar{b}(\bar{X}_s^{t,m}, \mu_s^{t,m})ds + \bar{\sigma}(\bar{X}_s^{t,m}, \mu_s^{t,m})dW_s + \bar{\gamma}(\bar{X}_s^{t,m}, \mu_s^{t,m})dB_s,
\end{aligned}
$$

on $s \in (t, T]$, with initial conditions $X_t^{t,x,m} = x$ and $\bar{X}_t^{t,m} \sim m$, and with $\mu_s^{t,m} = \mathcal{L}(\bar{X}_s^{t,m} \,|\, (B_r - B_t)_{r \in [t,s]})$. Define the value function

$$
V(t, x, m) = \mathbb{E}\left[g(X_s^{t,x,m}, \mu_s^{t,m}) + \int_t^T f(X_s^{t,x,m}, \mu_s^{t,m})ds\right],
$$

for some given nice functions g and f. Suppose $U : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is smooth and satisfies

$$
\begin{aligned}
\partial_t U(t, x, m) + LU(t, x, m) + \bar{L}U(t, x, m) + f(x, m) = 0 \\
U(T, x, m) = g(x, m).
\end{aligned}
$$

Then $U \equiv V$.

Proof. Fix (t, x, m), and abuse notation by omitting the superscripts. That is, write $X_s = X_s^{t,x,m}$ and $\bar{X}_s = \bar{X}_s^{t,m}$. Apply Itô's formula to get

$$
dU(X_t, \mu_t) = (\partial_t U(t, X_t, \mu_t) + LU(t, X_t, \mu_t) + \bar{L}U(t, X_t, \mu_t))dt + dM_t,
$$

where M is a martingale. Imposing the terminal condition $U(T, x, m) = g(x, m)$ and the initial condition $U(t, X_t, \mu_t) = U(t, x, m)$, we may write the above in integral form as

$$
\begin{aligned}
&g(X_T, \mu_T) - U(t, x, m) \\
&\quad = \int_t^T \big(\partial_t U(s, X_s, \mu_s) + LU(s, X_s, \mu_s) + \bar{L}U(s, X_s, \mu_s)\big)\, ds + M_T - M_t.
\end{aligned}
$$

Take expectations to get

$$
\begin{aligned}
&\mathbb{E}[g(X_T, \mu_T)] - U(t, x, m) \\
&\quad = \mathbb{E}\int_t^T \big(\partial_t U(s, X_s, \mu_s) + LU(s, X_s, \mu_s) + \bar{L}U(s, X_s, \mu_s)\big)\, ds \\
&\quad = -\mathbb{E}\int_t^T f(s, X_s, \mu_s)ds.
\end{aligned}
$$

□

9.4 Verification theorem

We now consider a version of Theorem 9.11 in which the first process is controlled, but not in the common noise volatility coefficient γ.

Let $(t,x,m) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)$, and consider the controlled processes X and $\bar{X}$ that solve the SDEs

$$\begin{aligned}
dX_s^{t,x,m} &= b(X_s^{t,x,m},\mu_s^{t,m},\alpha_s)ds + \sigma(X_s^{t,x,m},\mu_s^{t,m},\alpha_s)dW_s + \gamma(X_s^{t,x,m},\mu_s^{t,m})dB_s, \\
d\bar{X}_t^{t,m} &= \bar{b}(\bar{X}_s^{t,m},\mu_s^{t,m})ds + \bar{\sigma}(\bar{X}_s^{t,m},\mu_s^{t,m})dW_s + \bar{\gamma}(\bar{X}_s^{t,m},\mu_s^{t,m})dB_s,
\end{aligned}$$

on $s \in (t,T]$, with initial conditions $X_t^{t,x,m} = x$ and $\bar{X}_t^{t,m} \sim m$, and with $\mu_s^{t,m} = \mathcal{L}(\bar{X}_s^{t,m} \,|\, (B_r - B_t)_{r\in[t,s]})$. Define the value function

$$V(t,x,m) = \sup_\alpha \mathbb{E}\left[g(X_s^{t,x,m},\mu_s^{t,m}) + \int_t^T f(X_s^{t,x,m},\mu_s^{t,m},\alpha_s)ds\right],$$

for some given nice functions g and f. Define the Hamiltonian

$$H(x,m,y,z) = \sup_{a\in A}\left[y\cdot b(x,m,a) + \frac{1}{2}\mathrm{Tr}[z(\sigma\sigma^\top + \gamma\gamma^\top)(x,m,a)] + f(x,m,a)\right].$$

Theorem 9.12. *Suppose $U : [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$ is smooth and satisfies*

$$\begin{aligned}
&\partial_t U(t,x,m) + H(x,m,D_xU(t,x,m),D_x^2U(t,x,m)) + \bar{L}U(t,x,m) = 0 \\
&U(T,x,m) = g(x,m).
\end{aligned}$$

Suppose also that there exists a measurable function $\alpha : [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d) \to A$ such that $\alpha(t,x,m)$ attains the supremum in $H(x,m,D_xU(t,x,m),D_x^2U(t,x,m))$ for each (t,x,m) and also the SDE

$$dX_t = b(X_t,\mu_t,\alpha(t,X_t,\mu_t))dt + \sigma(X_t,\mu_t,\alpha(t,X_t,\mu_t))dW_t + \gamma(X_t,\mu_t)dB_t$$

is well-posed. Then $U \equiv V$, and $\alpha(t,X_t,\mu_t)$ is an optimal control.

Proof. Fix $(t,x,m) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)$, and for ease of notation omit the superscript by writing $X = X^{t,x,m}$. Fix a control α, apply Itô's formula to $U(s,X_s,\mu_s)$, take expectations to get

$$\begin{aligned}
&\mathbb{E}g(X_T,\mu_T) - U(t,x,m) \\
&\quad = \mathbb{E}\int_t^T \Big(\partial_t U(s,X_s,\mu_s) + \bar{L}U(s,X_s,\mu_s) \\
&\qquad\qquad + D_xU(s,X_s,\mu_s)\cdot b(X_s,\mu_s,\alpha_s) \\
&\qquad\qquad + \frac{1}{2}\mathrm{Tr}[D_x^2U(s,X_s,\mu_s)(\sigma\sigma^\top + \gamma\gamma^\top)(X_s,\mu_s,\alpha_s)]\Big)ds.
\end{aligned}$$

Using the PDE for U, we immediately find

$$\mathbb{E}g(X_T, \mu_T) - U(t, x, m) \le -\mathbb{E}\int_t^T f(X_s, \mu_s, \alpha_s)ds.$$

As this holds for each choice of α, we conclude that $U \ge V$. On the other hand, if we choose $\alpha_s = \alpha(s, X_s, \mu_s)$, then the above inequality becomes an equality, and we see that $U(t, x, m)$ is the value corresponding to this particular control; hence, $U \le V$. □

9.5 The master equation for mean field games

As a corollary of the verification theorem of the previous section, we can finally derive the *master equation* for mean field games. We now specialize the previous discussion to the case where $\bar{b} = b$, $\bar{\sigma} = \sigma$, $\bar{\gamma} = \gamma$, so that the dynamics of X and $\bar{X}$ are the same.

In order to define the value function for the mean field game, we must take care that the equilibrium is unique. We assume that for each $(t, m) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$ there is a unique MFE $(\mu_s^{t,m})_{s\in[t,T]}$ starting from (t, m). More precisely, suppose that for each (t, m), the following map $\mu \mapsto \mu^*$ has a unique fixed point:

1. Let $(\mu_s)_{s\in[t,T]}$ be a measure flow adapted to filtration generated by the common noise $(B_r - B_t)_{r\in[t,T]}$.

2. Solve the optimal control problem

$$\sup_\alpha \mathbb{E}\left[g(\bar{X}_T, \mu_T) + \int_t^T f(\bar{X}_s, \mu_s, \alpha_s)ds\right],$$
$$d\bar{X}_s = b(\bar{X}_s, \mu_s, \alpha_s)ds + \sigma(\bar{X}_s, \mu_s, \alpha_s)dW_s + \gamma(\bar{X}_s, \mu_s)dB_s, \ s \in [t, T]$$

   with initial state $\bar{X}_t \sim m$.

3. Let $\mu_s^* = \mathcal{L}(\bar{X}_s^* \,|\, (B_r - B_t)_{r\in[t,s]})$ denote the conditional measure flow of the optimal state process.

This unique fixed point $(\mu_s^{t,m})_{s\in[t,T]}$ is indeed best interpreted as a MFE starting from (t, m). Intuitively, the distribution of players at time t is given by m, and we optimize only our remaining reward after time t.

We can then define the value function as

$$V(t, x, m) = \sup_\alpha \mathbb{E}\left[g(X_T, \mu_T^{t,m}) + \int_t^T f(X_s, \mu_s^{t,m}, \alpha_s)ds\right],$$

where

$$
\begin{aligned}
dX_s &= b(X_s, \mu_s^{t,m}, \alpha_s)ds + \sigma(X_s, \mu_s^{t,m}, \alpha_s)dW_s + \gamma(X_s, \mu_s^{t,m})dB_s, \ \ s \in [t,T] \\
X_t &= x
\end{aligned}
$$

This quantity $V(t,x,m)$ is the remaining in-equilibrium value, after time t, to a player starting from state x at time t, given that the distribution of other agents starts at m at time t.

Define the Hamiltonian

$$
H(x,m,y,z) = \sup_{a \in A} \left[y \cdot b(x,m,a) + \frac{1}{2}\mathrm{Tr}[z(\sigma\sigma^\top + \gamma\gamma^\top)(x,m,a)] + f(x,m,a) \right].
$$

Let $\alpha(x,m,y,z)$ denote a maximizer. Define

$$
\begin{aligned}
\widehat{b}(x,m,y,z) &= b(x,m,\alpha(x,m,y,z)), \\
\widehat{\sigma}(x,m,y,z) &= \sigma(x,m,\alpha(x,m,y,z)).
\end{aligned}
$$

The master equation then takes the form

$$
\begin{aligned}
0 =& \partial_t U(t,x,m) + H(x,m,D_xU(t,x,m), D_x^2U(t,x,m)) \\
&+ \int_{\mathbb{R}^d} D_mU(t,x,m,v) \cdot \widehat{b}(v,m,D_xU(t,v,m), D_x^2U(t,v,m))m(dv) \\
&+ \int_{\mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_vD_mU(t,x,m,v)(\widehat{\sigma}\widehat{\sigma}^T + \gamma\gamma^T)(v,m,D_xU(t,v,m), D_x^2U(t,v,m))]m(dv) \\
&+ \iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2}\mathrm{Tr}[D_m^2U(t,x,m,v,\tilde{v})\gamma(v,m)\gamma^T(\tilde{v},m)]m(dv)m(d\tilde{v}) \\
&+ \int_{\mathbb{R}^d} \mathrm{Tr}[D_xD_mU(t,x,m,v)\gamma(x,m)\gamma^T(v,m)]m(dv), \qquad (9.10)
\end{aligned}
$$

with terminal condition $U(T,x,m) = g(x,m)$.

Theorem 9.13. *Suppose $U = U(t,x,m)$ is a smooth solution of the master equation* (9.10)*. Then $U \equiv V$. Moreover, the equilibrium control is given by $\alpha(t,x,m) = \alpha(x,m,D_xU(t,x,m), D_x^2U(t,x,m))$, assuming this function is nice enough for the McKean-Vlasov SDE*

$$
\begin{aligned}
dX_t &= b(X_t, \mu_t, \alpha(t,X_t,\mu_t))dt + \sigma(X_t, \mu_t, \alpha(t,X_t,\mu_t))dW_t + \gamma(X_t,\mu_t)dB_t, \\
\mu_t &= \mathcal{L}(X_t \,|\, \mathcal{F}_t^B),
\end{aligned}
$$

to be well-posed. Finally, the unique solution $\mu = (\mu_t)_{t \in [0,T]}$ of this McKean-Vlasov equation is the unique mean field equilibrium.

Proof. Apply Theorem 9.12 with $\bar{\gamma} = \gamma$ and with the coefficients $\bar{b}$ and $\bar{\sigma}$ given by

$$\bar{b}(x,m) = \widehat{b}(x,m,D_xU(t,x,m),D_x^2U(t,x,m)),$$
$$\bar{\sigma}(x,m) = \widehat{\sigma}(x,m,D_xU(t,x,m),D_x^2U(t,x,m)).$$

□

The master equation is very challenging to analyze, for a number of reasons. First, its state space is infinite-dimensional. Second, it is nonlinear in the spatial derivatives D_xU and D_x^2U. Last but not least, it is *nonlocal*, in the sense that the equation involves both $U(t,x,m)$ at a given point as well as the integral of functions of $U(t,v,m)$ and its derivatives over $m(dv)$.

9.6 Simplifications of the master equation

This section is devoted to describing some of the specializations that lead to simpler, and occasionally solvable master equations. The first basic observation is that if there is no common noise, $\gamma \equiv 0$, then the two last terms of the master equation (9.10) vanish.

9.6.1 Drift control and constant volatility

Let us assume that the MFG is as follows:

$$\begin{aligned}\alpha^* \in \arg\max_{\alpha} \mathbb{E}\left[g(X_T,\mu_T) + \int_0^T f_1(X_t,\mu_t) - f_2(X_t,\alpha_t)dt\right]\\ dX_t = &\alpha_t dt + \sigma dW_t + \gamma dB_t, \;\; \mu_t = \mathcal{L}(X_t|\mathcal{F}_t^B), \, \forall t \in [0,T]\end{aligned}$$

In this example we have that $A = \mathbb{R}^d$, and the Hamiltonian is

$$H(x,y) = \sup_{a \in \mathbb{R}^d} (a \cdot y - f_2(x,a)).$$

Basic results of convex analysis ensure that $a \mapsto D_af_2(x,a)$ and $y \mapsto D_yH(x,y)$ are inverse functions. Hence, the optimizer in the Hamiltonian is $\alpha(x,y) = D_yH(x,y)$.

Substituting the values just computed, the master equation becomes

$$
\begin{aligned}
&\partial_t U(t,x,m) + H(x, D_x U(t,x,m)) + f_1(x,m) + \\
&+\frac{1}{2} Tr[(\sigma\sigma^T + \gamma\gamma^T) D_x^2 U(t,x,m)] + \\
&+\int_{\mathbb{R}^d} D_m U(t,x,m,v) \cdot D_y H(v, D_x U(t,v,m)) m(dv) + \\
&+\frac{1}{2}\int_{\mathbb{R}^d} Tr[D_v D_m U(t,x,m,v)(\sigma\sigma^\top + \gamma\gamma^\top)] m(dv) + \\
&+\frac{1}{2}\iint_{\mathbb{R}^d\times\mathbb{R}^d} Tr[D_m^2 U(t,x,m,v,v')\gamma\gamma^\top] m(dv) m(dv') + \\
&+\int_{\mathbb{R}^d} Tr[D_x D_m U(t,x,m,v)\gamma\gamma^\top] m(dv) = 0 \\
&U(T,x,m) = g(x,m)
\end{aligned}
$$

The dynamics of the equilibrium state process becomes

$$
\begin{aligned}
dX_t &= D_y H(X_t, D_x U(t, X_t, \mu_t))dt + \sigma dW_t + \gamma dB_t \\
\mu_t &= \mathcal{L}(X_t|\mathcal{F}_t^B),\ \forall t \in [0,T]
\end{aligned}
$$

If we specialize now to the case where $\sigma = 1$ and $\gamma = 0$, the master equation reduces further to

$$
\begin{aligned}
&\partial_t U(t,x,m) + H(x, D_x U(t,x,m)) + f_1(x,m) + \\
&+\frac{1}{2}\Delta_x U(t,x,m)] + \\
&+\int_{\mathbb{R}^d} D_m U(t,x,m,v) D_y \mathcal{H}(v, D_x U(t,v,m)) m(dv) + \\
&+\frac{1}{2}\int_{\mathbb{R}^d} \mathrm{div}_v(D_m U(t,x,m,v)) m(dv) = 0 \\
&U(T,x,m) = g(x,m)
\end{aligned}
$$

and the dynamics of the equilibrium state process are

$$
\begin{aligned}
dX_t &= D_y H(X_t, D_x U(t, X_t, \mu_t))dt + dW_t \\
\mu_t &= \mathcal{L}(X_t),\ \forall t \in [0,T]
\end{aligned}
$$

9.6.2 A linear quadratic model

We return now to our favorite solvable MFG, covered for example in Section 8.7.1:

$$
\begin{aligned}
\alpha^* \in & \arg\max_{\alpha} \mathbb{E}\left[-\frac{\lambda}{2}|\bar{X}_T - X_T| - \int_0^T \frac{1}{2}|\alpha_t|^2 dt\right] \\
dX_t = & \alpha_t dt + dW_t, \quad \bar{X}_t = \mathbb{E}[X_t]
\end{aligned}
$$

In this example, we have that the Hamiltonian is $H(x,y) = \sup_a(a \cdot y - \frac{1}{2}|a|^2) = \frac{1}{2}|y|^2$ with optimizer $\alpha(x,y) = y$. The master equation becomes

$$
\begin{aligned}
&\partial_t U(t,x,m) + \frac{1}{2}|D_x U(t,x,m)|^2 + \frac{1}{2}\Delta_x U(t,x,m)]+ \\
&+ \int_{\mathbb{R}^d} D_m U(t,x,m,v) D_x U(t,v,m) m(dv) + \\
&+ \frac{1}{2}\int_{\mathbb{R}^d} div_v(D_m U(t,x,m,v)) m(dv) = 0 \\
&U(T,x,m) = -\frac{\lambda}{2}|\bar{m} - x|^2, \quad \text{where} \quad \bar{m} = \int_{\mathbb{R}^d} y\, m(dy).
\end{aligned}
$$

In order to try to get a solution, we start making a first ansatz, which consists of assuming that the solution depends only on the mean of the distribution and not on the whole probability distribution, i.e. there exists a function $F : [0,T] \times (\mathbb{R}^d)^2$ such that $U(t,x,m) = F(t,x,\bar{m})$.

Then, using our new developed calculus, we have that

$$
\begin{aligned}
\partial_t U(t,x,m)) =& \partial_t F(t,x,\bar{m}) \\
D_x U(t,x,m) =& D_x F(t,x,\bar{m}) \\
D_x^2 U(t,x,m) =& D_x^2 F(t,x,\bar{m}) \\
D_m U(t,x,m) = D_{\bar{m}} F(t,x,\bar{m}) & \cdot D_m G(m,v) = D_{\bar{m}} F(t,x,\bar{m})
\end{aligned}
$$

where we have used that if we define $G(m) = \bar{m}$, then

$$
\begin{aligned}
\frac{\delta G}{\delta m}(m,v) &= v \\
D_m G(m,v) &= 1.
\end{aligned}
$$

For notational simplicity, in the future we will refer to the third variable of F as y instead of $\bar{m}$. Since $D_y F(t,x,y)$ does not depend on v, the PDE simplifies to

$$\begin{aligned}
\partial_t F(t,x,y) + \frac{1}{2}|D_x F(t,x,y)|^2 + \frac{1}{2}\Delta_x F(t,x,y)]+ \\
+\int_{\mathbb{R}^d} D_y F(t,x,y) D_x F(t,v,y) m(dv) = 0 \\
F(T,x,y) = -\frac{\lambda}{2}|y-x|^2.
\end{aligned}$$

The second idea we postulate is that the derivative of the solution with respect to the spatial variable is affine in that same variable, i.e. $D_x F(t,x,y) = A(t,y)x + B(t,y)$. Then the PDE simplifies to

$$\begin{aligned}
\partial_t F(t,x,y) + \frac{1}{2}|D_x F(t,x,y)|^2 + \frac{1}{2}\Delta_x F(t,x,y)]+ \\
+D_y F(t,x,y) \cdot D_x F(t,y,y) = 0 \\
F(T,x,y) = -\frac{\lambda}{2}|y-x|^2.
\end{aligned}$$

Our final Ansatz is to separate the spatial and temporal variables, i.e. assume that $F(t,x,y) = \frac{1}{2}f(t)|y-x|^2 + g(t)$. Then $D_x F(t,x,y) = f(t)|x-y|$ (so affine in x and y) and $\Delta_x F(t,x,y) = d \cdot f(t)$. In particular, $D_x F(t,y,y) = 0$. The PDE then becomes

$$\begin{aligned}
\frac{1}{2}f'(t)|x-y|^2 + g'(t) + \frac{1}{2}f^2(t)|x-y|^2 + \frac{d}{2}f(t) = 0 \\
g(T) = 0, \ f(T) = -\lambda
\end{aligned}$$

The equations must hold for any x and y in $\mathbb{R}^d$, and matching coefficients shows that the PDE is equivalent to the following system of ODE's

$$\begin{aligned}
f'(t) + f^2(t) = 0 \\
g'(t) + \frac{d}{2}f(t) = 0 \\
g(T) = 0, \ f(T) = -\lambda
\end{aligned}$$

As in Section 5.5, we can explicitly solve these ODEs, yielding

$$U(t,x,m) = F(t,x,\bar{m}) = -\frac{(x-\bar{m})^2}{2(\frac{1}{\lambda} + T - t)} + \frac{d}{2}\log(1 + \lambda(T-t))$$

9.7 Bibliographic notes

The master equation was understood on some level in the lectures of Lions [91], and the first reasonably general accounts of the master equation and its verification theorem appeared in [10, 28]. Analysis of the master equation is difficult, perhaps unsurprisingly. Some first well-posedness theory for classical solutions can be found in [37, 56, 21]. We know from classical control theory that one should not expect classical solutions in general, and this problem is resolved using the theory of viscosity solutions. It is not clear yet if there is a satisfactory viscosity theory for the master equation, with a key challenge posed by the the infinite-dimensional state space $\mathcal{P}(\mathbb{R}^d)$. However, mean field games of potential type can be reformulated as controlled McKean-Vlasov problems, and for such problems the master equation reads as a Hamilton-Jacobi(-Bellman) equation on the space of probability measures; some progress on a viable viscosity theory for such equations appeared already in [20] for first-order systems, and the recent developments of [116] seem promising.

The seminal paper [23] showed not only the strongest well-posedness results to date but also explained how the master equation, when it admits a smooth solution, can be used to prove the convergence of the n-player Nash equilibria to the (unique, in this case) mean field equilibrium. This insight was recently extended in [46, 47] to derive a central limit theorem, a large deviations principle, and non-asymptotic concentration (rate) estimates associated to this law of large numbers-type limit. This kind of analysis is extremely fruitful when it works, but it notably requires uniqueness for the MFE and a quite regular solution of the master equation. See also [34, 35, 5] for a similar analysis for MFGs with finite state space.

The notion of derivative for functions on $\mathcal{P}(\mathbb{R}^d)$ was introduced by Lions in [91], written also in [20] and in [30]. This calculus can be approached from a number of angles, and the exposition of Section 9.1 follows more closely the presentation of [21, Section 2.2].

References

[1] Y. Achdou, F. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll, *Partial differential equation models in macroeconomics*, Phil. Trans. R. Soc. A **372** (2014), no. 2028, 20130397.

[2] Y. Achdou, J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll, *Heterogeneous agent models in continuous time*, Preprint (2014).

[3] C. Aliprantis and K. Border, *Infinite dimensional analysis: A hitchhiker's guide*, 3 ed., Springer, 2007.

[4] R.J. Aumann, *Markets with a continuum of traders*, Econometrica: Journal of the Econometric Society (1964), 39–50.

[5] E. Bayraktar and A. Cohen, *Analysis of a finite state many player game using its master equation*, arXiv preprint arXiv:1707.02648 (2017).

[6] C. Benazzoli, L. Campi, and L. Di Persio, *Mean-field games with controlled jumps*, arXiv preprint arXiv:1703.01919 (2017).

[7] A. Bensoussan and J. Frehse, *Nonlinear elliptic systems in stochastic game theory*, Universität Bonn. SFB 72. Approximation und Optimierung, 1983.

[8] ———, *Stochastic games for N players*, Journal of optimization theory and applications **105** (2000), no. 3, 543–565.

[9] ———, *Regularity results for nonlinear elliptic systems and applications*, vol. 151, Springer Science & Business Media, 2013.

[10] A. Bensoussan, J. Frehse, and S.C.P. Yam, *The master equation in mean field theory*, Journal de Mathématiques Pures et Appliquées **103** (2015), no. 6, 1441–1474.

[11] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons, 2013.

[12] A. Blanchet and G. Carlier, *From Nash to Cournot-Nash equilibria via the Monge-Kantorovich problem*, Phil. Trans. R. Soc. A **372** (2014), no. 2028, 20130398.

[13] ———, *Optimal transport and Cournot-Nash equilibria*, Mathematics of Operations Research **41** (2015), no. 1, 125–145.

[14] F. Bolley, *Separability and completeness for the Wasserstein distance*, Séminaire de probabilités XLI, Springer, 2008, pp. 371–377.

[15] F. Bolley, A. Guillin, and C. Villani, *Quantitative concentration inequalities for empirical measures on non-compact spaces*, Probability Theory and Related Fields **137** (2007), no. 3, 541–593.

[16] R. Buckdahn and J. Li, *Stochastic differential games and viscosity solutions of Hamilton–Jacobi–Bellman–Isaacs equations*, SIAM Journal on Control and Optimization **47** (2008), no. 1, 444–475.

[17] A. Budhiraja, P. Dupuis, and M. Fischer, *Large deviation properties of weakly interacting processes via weak convergence methods*, The Annals of Probability **40** (2012), no. 1, 74–102.

[18] M. Burger, M. Di Francesco, P. Markowich, and M.-T. Wolfram, *Mean field games with nonlinear mobilities in pedestrian dynamics*, arXiv preprint arXiv:1304.5201 (2013).

[19] M. Cáceres, J. Carrillo, and B. Perthame, *Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states*, The Journal of Mathematical Neuroscience **1** (2011), no. 1, 7.

[20] P. Cardaliaguet, *Notes on mean field games*, Tech. report, Technical report, 2010.

[21] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games*, arXiv preprint arXiv:1509.02505 (2015).

[22] P. Cardaliaguet and C.-A. Lehalle, *Mean field game of controls and an application to trade crowding*, arXiv preprint arXiv:1610.09904 (2016).

[23] ———, *Mean field game of controls and an application to trade crowding*, Mathematics and Financial Economics (2016), 1–29.

[24] G. Carmona, *Nash equilibria of games with a continuum of players*, (2004).

[25] R. Carmona, *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*, vol. 1, SIAM, 2016.

[26] R. Carmona and F. Delarue, *Mean field forward-backward stochastic differential equations*, Electronic Communications in Probability **18** (2013).

[27] ———, *Probabilistic analysis of mean-field games*, SIAM Journal on Control and Optimization **51** (2013), no. 4, 2705–2734.

[28] ———, *The master equation for large population equilibriums*, Stochastic Analysis and Applications 2014, Springer, 2014, pp. 77–128.

[29] ———, *Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics*, The Annals of Probability **43** (2015), no. 5, 2647–2700.

[30] ———, *Probabilistic theory of mean field games with applications i-ii*, Springer, 2017.

[31] R. Carmona, F. Delarue, and A. Lachapelle, *Control of mckean–vlasov dynamics versus mean field games*, Mathematics and Financial Economics **7** (2013), no. 2, 131–166.

[32] R. Carmona, J.-P. Fouque, and L.-H. Sun, *Mean field games and systemic risk*, (2013).

[33] J. Carrillo, R.J. McCann, and C. Villani, *Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates*, Revista Matematica Iberoamericana **19** (2003), no. 3, 971–1018.

[34] A. Cecchin and M. Fischer, *Probabilistic approach to finite state mean field games*, arXiv preprint arXiv:1704.00984 (2017).

[35] A. Cecchin and G. Pelino, *Convergence, fluctuations and large deviations for finite state mean field games via the master equation*, arXiv preprint arXiv:1707.01819 (2017).

[36] P. Chan and R. Sircar, *Bertrand and Cournot mean field games*, Applied Mathematics & Optimization **71** (2015), no. 3, 533–569.

[37] J.-F. Chassagneux, D. Crisan, and F. Delarue, *A probabilistic approach to classical solutions of the master equation for large population equilibria*, arXiv preprint arXiv:1411.3009 (2014).

[38] M.G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bulletin of the American mathematical society **27** (1992), no. 1, 1–67.

[39] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Transactions on automatic control **52** (2007), no. 5, 852–862.

[40] D. Dawson and J. Vaillancourt, *Stochastic McKean-Vlasov equations*, NoDEA: Nonlinear Differential Equations and Applications **2** (1995), no. 2, 199–229.

[41] D.A. Dawson, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior*, Journal of Statistical Physics **31** (1983), no. 1, 29–85.

[42] D.A. Dawson and J. Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics: An International Journal of Probability and Stochastic Processes **20** (1987), no. 4, 247–308.

[43] ———, *Large deviations, free energy functional and quasi-potential for a mean field model of interacting diffusions*, vol. 78, American Mathematical Soc., 1989.

[44] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré, *Global solvability of a networked integrate-and-fire model of McKean–Vlasov type*, The Annals of Applied Probability **25** (2015), no. 4, 2096–2133.

[45] ———, *Particle systems with a singular mean-field self-excitation. Application to neuronal networks*, Stochastic Processes and their Applications **125** (2015), no. 6, 2451–2492.

[46] F. Delarue, D. Lacker, and K. Ramanan, *From the master equation to mean field game limit theory: A central limit theorem*, arXiv preprint arXiv:1804.08542 (2018).

[47] ———, *From the master equation to mean field game limit theory: Large deviations and concentration of measure*, arXiv preprint arXiv:1804.08550 (2018).

[48] N. El Karoui, D. Nguyen, and M. Jeanblanc-Picqué, *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, Stochastics: an international journal of probability and stochastic processes **20** (1987), no. 3, 169–219.

[49] M. Fischer, *On the connection between symmetric n-player games and mean field games*, The Annals of Applied Probability **27** (2017), no. 2, 757–810.

[50] W.H. Fleming and H.M. Soner, *Controlled Markov processes and viscosity solutions*, vol. 25, Springer Science & Business Media, 2006.

[51] W.H. Fleming and P.E. Souganidis, *On the existence of value functions of two-player, zero-sum stochastic differential games*, Indiana University Mathematics Journal **38** (1989), no. 2, 293–314.

[52] H. Föllmer and A. Schied, *Stochastic finance: an introduction in discrete time*, Walter de Gruyter, 2011.

[53] A. Friedman, *Stochastic differential games*, Journal of differential equations **11** (1972), no. 1, 79–108.

[54] G. Fu and U. Horst, *Mean field games with singular controls*, SIAM Journal on Control and Optimization **55** (2017), no. 6, 3833–3868.

[55] X. Gabaix, J.-M. Lasry, P.-L. Lions, and B. Moll, *The dynamics of inequality*, Econometrica **84** (2016), no. 6, 2071–2111.

[56] W. Gangbo and A. Swiech, *Existence of a solution to an equation arising from the theory of mean field games*, Journal of Differential Equations **259** (2015), no. 11, 6573–6643.

[57] J. Garnier, G. Papanicolaou, and T.-W. Yang, *Large deviations for a mean field model of systemic risk*, SIAM Journal on Financial Mathematics **4** (2013), no. 1, 151–184.

[58] J. Gärtner, *On the McKean-Vlasov limit for interacting diffusions*, Mathematische Nachrichten **137** (1988), no. 1, 197–248.

[59] R. Gayduk and S. Nadtochiy, *Endogenous formation of limit order books: dynamics between trades*, arXiv preprint arXiv:1605.09720 (2016).

[60] D.A. Gomes and J. Saúde, *Mean field games models—a brief survey*, Dynamic Games and Applications **4** (2014), no. 2, 110–154.

[61] C. Graham and S. Méléard, *Stochastic particle approximations for generalized Boltzmann models and convergence estimates*, The Annals of probability **25** (1997), no. 1, 115–132.

[62] E.J. Green, *Continuum and finite-player noncooperative models of competition*, Econometrica: Journal of the Econometric Society (1984), 975–993.

[63] S.-Y. Ha, K. Lee, and D. Levy, *Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system*, Communications in Mathematical Sciences **7** (2009), no. 2, 453–469.

[64] S. Hamadene, *Backward–forward SDE's and stochastic differential games*, Stochastic processes and their applications **77** (1998), no. 1, 1–15.

[65] ______, *Nonzero sum linear–quadratic stochastic differential games and backward–forward equations*, Stochastic Analysis and Applications **17** (1999), no. 1, 117–130.

[66] U.G. Haussmann and J.P. Lepeltier, *On the existence of optimal controls*, SIAM Journal on Control and Optimization **28** (1990), no. 4, 851–902.

[67] S. Herrmann and J. Tugaut, *Non-uniqueness of stationary measures for self-stabilizing processes*, Stochastic Processes and their Applications **120** (2010), no. 7, 1215–1246.

[68] M. Hitsuda and I. Mitoma, *Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions*, Journal of Multivariate Analysis **19** (1986), no. 2, 311–328.

[69] D. Housman, *Infinite player noncooperative games and the continuity of the Nash equilibrium correspondence*, Mathematics of Operations Research **13** (1988), no. 3, 488–496.

[70] M. Huang, R. Malhamé, and P. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Communications in Information & Systems **6** (2006), no. 3, 221–252.

[71] X. Huang, S. Jaimungal, and M. Nourian, *Mean-field game strategies for a major-minor agent optimal execution problem*, Available at SSRN 2578733 (2017).

[72] R. Isaacs, *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*, Courier Corporation, 1999.

[73] M. Kac, *Foundations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, vol. 1955, 1954, pp. 171–197.

[74] E. Kalai, *Large robust games*, Econometrica **72** (2004), no. 6, 1631–1665.

[75] O. Kallenberg, *Foundations of modern probability*, Springer Science & Business Media, 2006.

[76] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, vol. 113, Springer Science & Business Media, 2012.

[77] T.G. Kurtz and P. Protter, *Weak limit theorems for stochastic integrals and stochastic differential equations*, The Annals of Probability (1991), 1035–1070.

[78] ______, *Weak limit theorems for stochastic integrals and stochastic differential equations*, The Annals of Probability (1991), 1035–1070.

[79] T.G. Kurtz and J. Xiong, *Particle representations for a class of nonlinear SPDEs*, Stochastic Processes and their Applications **83** (1999), no. 1, 103–126.

[80] ______, *A stochastic evolution equation arising from the fluctuations of a class of interacting particle systems*, Communications in Mathematical Sciences **2** (2004), no. 3, 325–358.

[81] A. Lachapelle and M.-T. Wolfram, *On a mean field game approach modeling congestion and aversion in pedestrian crowds*, Transportation research part B: methodological **45** (2011), no. 10, 1572–1589.

[82] D. Lacker, *Mean field games via controlled martingale problems: existence of Markovian equilibria*, Stochastic Processes and their Applications **125** (2015), no. 7, 2856–2894.

[83] ______, *A general characterization of the mean field limit for stochastic differential games*, Probability Theory and Related Fields **165** (2016), no. 3-4, 581–648.

[84] D. Lacker and K. Ramanan, *Rare Nash equilibria and the price of anarchy in large static games*, arXiv preprint arXiv:1702.02113 (2017).

[85] D. Lacker and T. Zariphopoulou, *Mean field and n-agent games for optimal investment under relative performance criteria*, arXiv preprint arXiv:1703.07685 (2017).

[86] O.A. Ladyzhenskaia, V.A. Solonnikov, and N.N. Ural'tseva, *Linear and quasi-linear equations of parabolic type*, vol. 23, American Mathematical Soc., 1988.

[87] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. i–le cas stationnaire*, Comptes Rendus Mathématique **343** (2006), no. 9, 619–625.

[88] ———, *Jeux à champ moyen. ii–horizon fini et contrôle optimal*, Comptes Rendus Mathématique **343** (2006), no. 10, 679–684.

[89] ———, *Mean field games*, Japanese journal of mathematics **2** (2007), no. 1, 229–260.

[90] ———, *Mean field games*, Japanese journal of mathematics **2** (2007), no. 1, 229–260.

[91] P.-L. Lions, *Cours au Collège de France.*

[92] Huang M, R. Malhamé, and P. Caines, *Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε-Nash equilibria*, IEEE transactions on automatic control **52** (2007), no. 9, 1560–1571.

[93] A. Mas-Colell, *On a theorem of Schmeidler*, Journal of Mathematical Economics **13** (1984), no. 3, 201–206.

[94] H.P. McKean, *A class of Markov processes associated with nonlinear parabolic equations*, Proceedings of the National Academy of Sciences **56** (1966), no. 6, 1907–1911.

[95] ———, *Propagation of chaos for a class of non-linear parabolic equations*, Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967) (1967), 41–57.

[96] S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, Probabilistic models for nonlinear partial differential equations, Springer, 1996, pp. 42–95.

[97] S. Méléard and S. Roelly-Coppoletta, *A propagation of chaos result for a system of particles with moderate interaction*, Stochastic processes and their applications **26** (1987), 317–332.

[98] D. Monderer and L. Shapley, *Potential games*, Games and economic behavior **14** (1996), no. 1, 124–143.

[99] J.F. Nash, *Equilibrium points in n-person games*, Proceedings of the national academy of sciences **36** (1950), no. 1, 48–49.

[100] N. Nisan, R. Roughgarden, E. Tardos, and V.V. Vazirani, *Algorithmic game theory*, vol. 1, Cambridge University Press Cambridge, 2007.

[101] K. Oelschläger, *A martingale approach to the law of large numbers for weakly interacting stochastic processes*, The Annals of Probability (1984), 458–479.

[102] K. Oelschläger, *A law of large numbers for moderately interacting diffusion processes*, Probability Theory and Related Fields **69** (1985), no. 2, 279–322.

[103] K.R. Parthasarathy, *Probability measures on metric spaces*, vol. 352, American Mathematical Soc., 2005.

[104] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, vol. 61, Springer Science & Business Media, 2009.

[105] R. Rao, *Relations between weak and uniform convergence of measures with applications*, The Annals of Mathematical Statistics (1962), 659–680.

[106] R.T. Rockafellar, *Convex analysis*, Princeton university press, 2015.

[107] R.W. Rosenthal, *A class of games possessing pure-strategy Nash equilibria*, International Journal of Game Theory **2** (1973), no. 1, 65–67.

[108] D. Schmeidler, *Equilibrium points of nonatomic games*, Journal of Statistical Physics **7** (1973), no. 4, 295–300.

[109] M. Shkolnikov, *Large systems of diffusions interacting through their ranks*, Stochastic Processes and their Applications **122** (2012), no. 4, 1730–1747.

[110] M. Sirbu, *Stochastic Perron's method and elementary strategies for zero-sum differential games*, SIAM Journal on Control and Optimization **52** (2014), no. 3, 1693–1711.

[111] G. Barles P.E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic analysis **4** (1991), no. 3, 271–283.

[112] D.W. Stroock and S.R.S. Varadhan, *Multidimensional diffusion processes*, Springer, 2007.

[113] A.S. Sznitman, *Topics in propagation of chaos*, Ecole d'eté de Probabilités de Saint-Flour XIX - 1989, Lecture Notes in Mathematics, vol. 1464, Springer Berlin / Heidelberg, 1991, pp. 165–251.

[114] J. Tugaut, *Phase transitions of McKean–Vlasov processes in double-wells landscape*, Stochastics An International Journal of Probability and Stochastic Processes **86** (2014), no. 2, 257–284.

[115] C. Villani, *Topics in optimal transportation*, no. 58, American Mathematical Soc., 2003.

[116] C. Wu and J. Zhang, *Viscosity solutions to master equations and McKean-Vlasov SDEs with closed-loop controls*, arXiv preprint arXiv:1805.02639 (2018).

[117] H. Yin, P.G. Mehta, S.P. Meyn, and U.V. Shanbhag, *Synchronization of coupled oscillators is a game*, IEEE Transactions on Automatic Control **57** (2012), no. 4, 920–935.

[118] J. Yong and X.Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43, Springer Science & Business Media, 1999.