Lecture Notes on "Introduction to Mean Field Games and applications"

Denis Ullmo

May 20, 2025

Contents

A - Introduction to Game Theory

Bibliographic reference: *Game theory*, Maschler, Solan and Zamir (2013)

Game theory is the name given to the methodology of using mathematical tools to model and analyse situations of interactive decision making. These are situations involving several decision makers (called players) with different goals, in which the decision of each affects the outcome for all the decision makers. The foundations of game theory were laid down in the book *The Theory of Games and Economic Behavior*, published in 1944 by the mathematician John von Neumann and the economist Oskar Morgenstern. The theory has been developed extensively since then and today it has applications in a wide range of fields, especially in theoretical economics, networks, political science, military applications, inspection, biology,...

This section, will provide an introduction to the basic concepts of game theory.

1. Extensive and strategic forms of a game

1.1. Tic-tac-toe / Chess games

Every description of a game must include the following elements:

- a set of players (decision makers)
- the possible actions available to each player
- the rules determining the order in which players make their moves
- a rule determining when the game ends
- rule determining the outcome of every possible game ending

For instance, **tic-tac-toe** can be described as paper-and-pencil game for two players who take turns marking the spaces in a three-by-three grid with X or O. The player who succeeds in placing three of their marks in a horizontal, vertical, or diagonal row first is the winner. Fig. 1 illustrates a tic-tac-toe game.

Figure 1: Example, of a tic-tac-toe game where the first player (X) wins the game in seven steps [from https://en.wikipedia.org/wiki/Tic-tac-toe].

This type of game is well described by what we call an "extensive form", which corresponds to a graphical way of describing a game and representing its rules (the order in which players make their moves, the information available to players when they are called to take an action, the termination rules, and the outcome at each endpoint). More precisely, a game in extensive form is given by a *game tree*, which consists of a directed graph in which the set of vertices represents positions in the game and a distinct vertex, called the *root*, represents the starting position of the game. Each player's action is then represented as a transition from one vertex to another. Fig. 2 shows the analogue of Fig. 1 in extended form, and Fig 3 is a representation of the chess game in extended form.

1.2. Rock-paper-scissors

Let us consider now the rock-paper-scissors game illustrated in Fig. 4. Here, we have two players and three different strategies. This kind of representation for a game is called a "*strategic*" (or "*normal*") form.

Are all "extended form" games representable in "strategic form"?

The answer is yes. However, while in the rock-paper-scissor game the strategies are simple: each player has only three strategies $\in \{\text{rock}, \text{paper}, \text{scissor}\}$, for tic-tac-toe, *one* strategy for player I (the X's) implies :

1. A choice (placement of an X on the grid) for her first move
2. One answer to each possible first move of player II
3. For each of these answers, an answer to each possible second move of player II
4. etc...

and in the same way, *one* strategy for player II (the O's) implies :

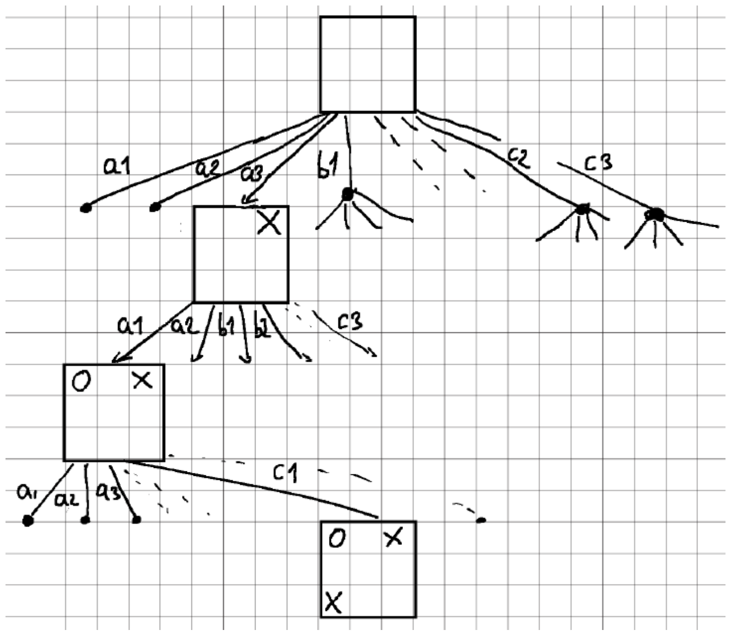


Figure 2: "Extensive form" of the game Tic-Tac-Toe. The part shown in Fig. 1 is one particular route on this tree linking the root to one leaf.

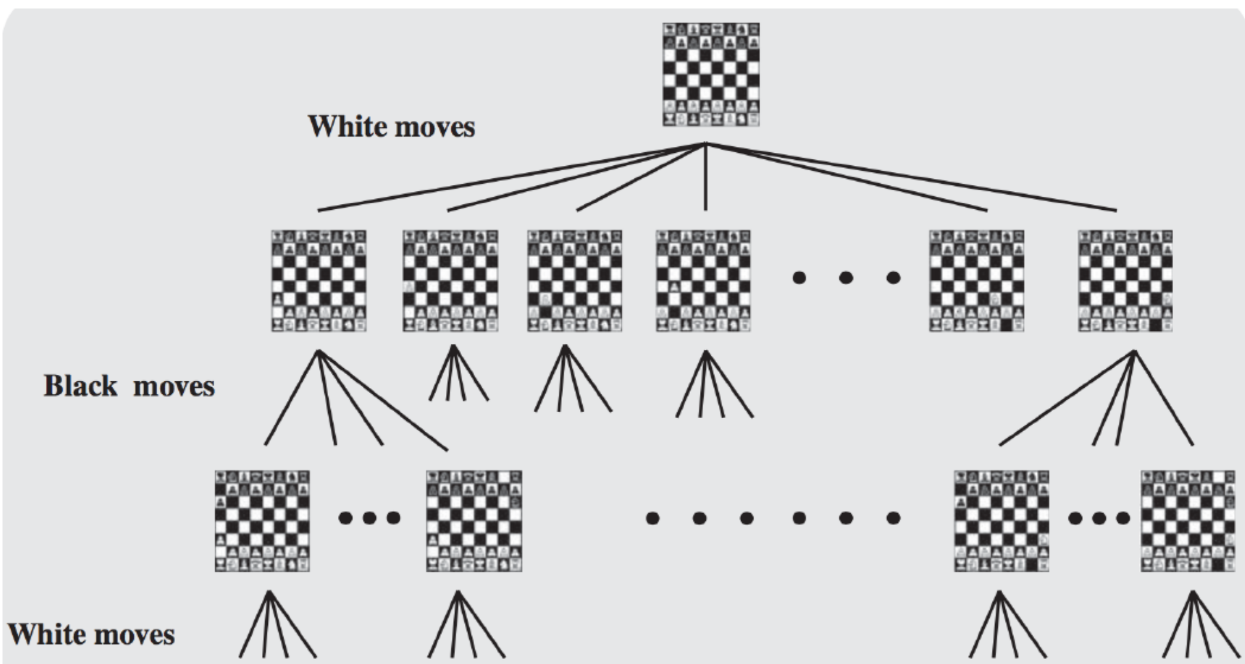


Figure 3: The game of chess presented in extensive form (taken from the book *Game theory*)

1. One answer to each possible first move of player I
2. For each of these answers, an answer to each possible second move of player I

		Player II		
		Rock	Paper	Scissors
	Rock	0, 0	1, –1	1, –1
Player I	Paper	1, –1	0, 0	– 1, 1
	Scissors	–1, 1	–1, 1	0, 0

Figure 4: Rock, paper, scissors as a game in *strategic* (or *normal*) form. In each cell, the left number denotes the outcome for Player I and the right number denotes the outcome for Player II, setting the gain of a player to be 1 for a win, -1 for a loss, and 0 for a draw (equality).

3. etc...

Even for a game as simple as tic-tact-toe, the number of strategies is already quite large (of the order of 9! if we neglect the possibility that some games end before the grid is completely full). An illustration of two strategies (one for player I, one for player II) is given in Fig. 5. For chess, with a larger chessboard, a large option of moves at each turn, and a game duration which is typically much longer, the number of strategies is immensely large.

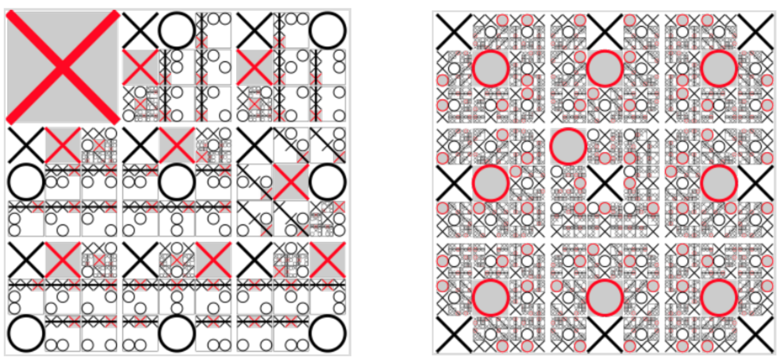


Figure 5: Illustration of *one* strategy for player I (left) and *one* strategy for player II (right) [from https://en.wikipedia.org/wiki/Tic-tac-toe]. These strategies can be proved to be optimal.

It is therefore usually not very convenient to represent games such as chess of even tic-tac-toe in normal form. On the other hand, *if you prove a theorem for games in strategic forms, this theorem will also apply to games in extended forms.*

Are all "strategic form" games representable in "extended form" ?

A priori, the answer is no, since for instance rock-paper-scissor would be a bit boring if players were taking turn to announce their move. One can however introduce the notion of *information set*, which is the set of decision vertices of the player that are indistinguishable by her given her information at that stage of the game. If one enhances the definition of the extended form with this notion of information set, it can be shown that any game in strategic form can be represented in extended form. For rock-paper-scissor, one possible representation in extended form is given in Figure 6. This imply in particular that, *if one proves* *a theorem for all games in strategic forms (with information sets), this theorem will also apply to games in extended forms..*

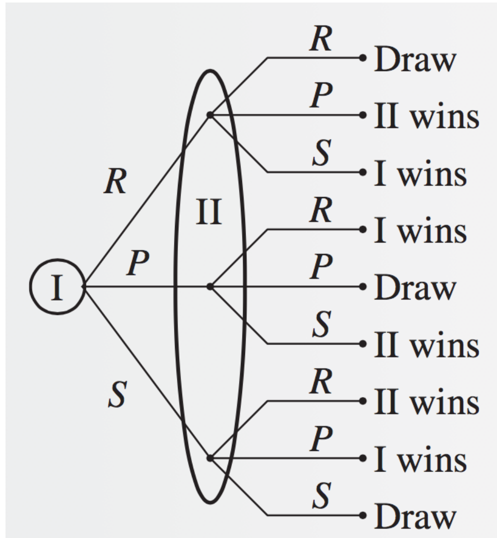


Figure 6: Extended form of the rock-paper-scissors game. The information set of player II before her moves includes the results of all possible moves of player I, giving the same results as if both player decisions were simultaneous.

2. Notion of utility

2.1. Utility as an "ordinal" function

In the examples above the possible outcomes were limited to either winning, losing or drawing. In the most general cases, however, there could be any number of outcomes $O_1, O_2, O_3, \cdots$ (think of the results of an election, for example). The players will choose their strategies depending on their preference about these outcomes.

Definition 1 *A preference relation of player i over a set of outcomes $O = \{O_1, O_2, O_3, \cdots\}$ is a binary relation denoted by $\succsim_i$ such that $O_\alpha \succsim_i O_\beta$ means that O_α is preferred to O_β. We assume*

1. *The preference relation $\succsim_i$ over O is* ***complete****; that is, for any pair of outcomes O_α and O_β in O, either $O_\alpha \succsim_i O_\beta$ or $O_\beta \succsim_i O_\alpha$, or both.*
2. *The preference relation $\succsim_i$ over O is* ***reflexive****; that is, $O_\alpha \succsim_i O_\alpha$ for every $O_\alpha \in O$.*
3. *The preference relation $\succsim_i$ over O is* ***transitive****; that is, for any triple of outcomes O_α, O_β and O_γ in O, if $O_\alpha \succsim_i O_\beta$ and $O_\beta \succsim_i O_\gamma$ then $O_\alpha \succsim_i O_\gamma$.*

Since using real numbers and the $\geq$ ordering relation is convenient for the purposes of conducting analysis, it would be an advantage to be able in general to represent game outcomes by real numbers, and player preferences by the familiar $\geq$ relation. Such a representation of a preference relation is called a *utility function*, and is defined as follows

Definition 2 *Definition: Let O be a set of outcomes and $\succsim$ be a complete, reflexive, and transitive preference relation over O. A function $u : O \to \mathbb{R}$ is called a utility function representing $\succsim$ if for all $O_\alpha, O_\beta, \in O$,*

$$O_\alpha \succsim O_\beta \Longleftrightarrow u(O_\alpha) \geq u(O_\beta) \ . \tag{1}$$

In other words, a utility function u is a function associating each outcome O_α with a real number $u(O_\alpha,)$ in such a way that the more an outcome is preferred, the larger is the real number associated with it. If the set of outcomes is finite, any complete, reflexive, and transitive preference relation can easily be represented by a utility function. A utility function is often called an *ordinal function*, because it represents only the order of preferences between outcomes. The numerical values that a utility function associates with outcomes have a priori no significance, and in principle do not in any way represent the "intensity" of a player's preferences. We will see that this perspective has to be modified a bit when probabilities are involved.

2.2. Linear utility

We now consider the case where the results of a game are not a specific outcome $A \in O = \{A_1, ..., A_K\}$, but a lottery $L = [p_1(A_1), ..., p_K(A_K)]$, with p_α the probability of outcome A_α ($0 \leq p_\alpha \leq 1$ and $\sum_{\alpha=1}^{K} p_\alpha = 1$). We note $\mathcal{L} = \{L\}$ the set of all lotteries, and we assume agents have a complete and transitive preference order relation on $\mathcal{L}$.

Definition 3 *A Linear utility, for agent i, is a utility function u_i defined on the set of lotteries associated with the outcomes O, thus such that $u_i(L) \leq u_i(L') \iff L \precsim_i L'$, and which verifies*

$$u_i(L) = \sum_\alpha p_\alpha(A_\alpha) u_i(A_\alpha) \tag{2}$$

Linear utilities may appear as a very specific subclass of all utility functions on lotteries. However, the following theorem states that under very natural hypothesis, utility functions for lotteries can always been chosen in this way.

Theorem 1 *If player i's preference relation over the set of all lotteries is complete, transitive, and verifies the four von Neumann-Morgenstern axioms, then* the preference relation can be represented by a linear utility function.

The four von Neumann-Morgenstern axioms are: **Continuity**, **Monotonicity**, **Simplification**, and **Equivalence**:

- <u>Continuity</u>:

  If $A \succsim_i B \succsim_i C$, then $\exists\, \theta_i \in [0,1]$ such that $B \simeq_i [\theta_i(A), (1-\theta_i)(B)]$.

  In words: if the probability changes continuously, the preference changes continuously.

- <u>Monotonicity</u>: Let $(\alpha, \beta) \in [0,1]^2$, assume $A \succsim_i B$.

  Then $[(1-\alpha)(B), \alpha(A)] \succsim_i [(1-\beta)(B), \beta(A)] \iff \alpha \geq \beta$

  In words: it is always preferable to have a higher probability of having an outcome one prefers.

- <u>Simplification</u>: Assume you have a compound lottery $\tilde{L} = [q_1(L_1), ..., q_J(L_J)]$ where q_j are probabilities and the L_j are "simple" lotteries $L_j = [p_1^j(A_1), ..., p_K^j(A_K)]$.

  Then $\tilde{L} \approx_i [r_1(A_1), ..., r_K(A_K)]$ with $r_k = \sum_j q_j p_k^j$.

  In words: the way the outcomes are chosen is irrelevant, all that matters is their effective probability.

- <u>Equivalence</u>: Let two lotteries $L = [p_1(A_1), \cdot\cdot, p_k(A_k), \cdot\cdot, p_K(A_K)]$ and $L' = [p_1(A_1), \cdot\cdot, p_k(A'_k), \cdot\cdot, p_K(A_K)]$ (L and L' thus differ only by the k'th outcome).

  Then, $A'_k \approx_i A_k$ implies $L' \approx_i L$.

Proof of the theorem: Let A_1 the least preferred outcome and A_K the preferred outcome (we assume $A_K \succsim_i A_1$). Because of *continuity*, we can find for every outcome A_k a number $\theta_i^k \in \mathbb{R}$ such that

$$A_k \approx_i [(1-\theta_i^k)(A_1), \theta_i^k(A_K)] \ .$$

For any lottery $L = [p_1(A_1), p_2(A_2), \cdots, p_K(A_K)]$, we define

$$\begin{aligned} u_i &: \mathcal{L} \to \mathbb{R} \\ &L \to u_i(L) = \sum_k p_k \theta_i^k \ . \end{aligned}$$

The function u_i is clearly linear, we want to show this is a utility.

Lemma: For every lottery $L \in \mathcal{L}$, $L \approx_i [(1-u_i(L))(A_1), u_i(L)(A_K)]$

Proof of the lemma: Let's introduce the lotteries $M_k = [(1-\theta_i^k)(A_1) + \theta_i^k(A_K)]$. We have

$$\begin{aligned} L &= [p_1(A_1), ..., p_K(A_K)] \\ &\approx_i [p_1(M_1), ..., p_K(M_K)] && \text{(equivalence)} \\ &\approx_i [(1-\sum_k p_k\theta_i^k)(A_1), \sum_k p_k\theta_i^k(A_K)] && \text{(simplification)} \\ &\approx_i [(1-u_i(L))(A_1), u_i(L)(A_K)] \end{aligned}$$

Now, if $L_1 \succsim_i L_2$, this implies $u_i(L_1) \geq u_i(L_2)$ because of the monotonicity axiom, which complete the proof that u_i is indeed a linear utility.

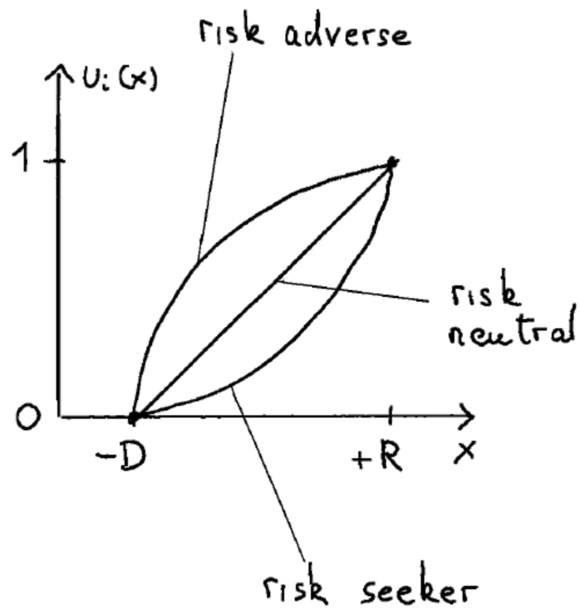


Figure 7: Sketch of linear utilities for risk neutral, risk seeker and risk adverse, agents.

Example: Suppose the outcome of the game is an amount of money $x \in [-D€, +R€]$. Fig. 7 shows sketches of different linear utilities corresponding to risk-averse, risk-neutral or risk-seeking agents. In this figure, $u_i(x)$ is the utility of the "outcome (monetary gain) x", defined by the fact that outcome $x \approx_i [(1-u_i(x))(-D), u_i(x)(R)]]$. For risk-averse agents, $u_i(x) \geq (x+D)/(R+D)$ (they prefer the certainty of x to the possibility of a higher return R but with the risk of a deficit D, even though the expectation is the same). Risk takers, on the other hand, would prefer the possibility of a higher income, even if it involves some risk, to the certainty of the same average return, and so for them $u_i(x) \leq (x+D)/(R+D)$. People are usually risk averse, but there are situations where you can benefit from risk. For example, if you have borrowed money from the mafia and have to pay back the totality of 100€ tonight or risk severe consequences, you might prefer the lottery $A = [0.5(100\$), 0.5(0\$)]$ to $B = [1(50\$)]$.

2.3. Solution concepts: notion of dominated strategy

Here and below, we consider games in *strategic form* involving a set of n players $N = \{1, \cdots, n\}$. We note $S_i = \{s_i\}$ the set of strategies of player $i \in N$, $S = \otimes_{j=1}^{n} S_j$ the set of vectors describing the strategies of all players, $S_{-i} = \otimes_{j \neq i} S_j$ the set of vectors describing the strategies of all players except player i, and s_{-i} an element of S_{-i}. We furthermore assume the existence of n utility functions

$$\begin{aligned} u_i : S &\longrightarrow \mathbb{R} \\ s = (s_i, s_{-i}) &\longmapsto u_i(s) \ . \end{aligned}$$

We begin with the introduction of a little bit of vocabulary:

Definition 4 *A strategy s_i of a player i is* strictly dominated *by another strategy t_i of the same player if*

$$\forall s_{-i} \in S_{-i},\ u_i(s_i, s_{-i}) < u_i(t_i, s_{-i}) \ .$$

Definition 5 *A strategy s_j of a player is* weakly dominated *by t_j if*

$$\begin{aligned} \forall s_{-j} \in S_{-j}, \quad u(s_j, s_{-j}) &\leq u_i(t_j, s_{-j}) \\ \textit{and } \exists\, \tilde{s}_{-j} \in S_{-j} \textit{ s.t. } u(s_j, \tilde{s}_{-j}) &< u_i(t_j, s_{-j}) \end{aligned}$$

If all players are rational and if the fact that they are rational is common knowledge (which mean all players know that all players are rational, and all players know that all of them know, etc.), then dominated strategies can be eliminated. A strategy vector $s \in S$ which is the result iterative elimination of dominated strategies is called a rational solution of a game. Rational solutions do not necessarily exist in all games.

Example: Second price auctions

Rules of the game: An indivisible object if offered for sale. Each of the N players have a private value v_i for the the object that is only known to themselves. A strategy is a bid b_i. All bids are issued at the same time. (The players do not know the bids of the others).

The highest bidder "wins" the auction and takes the object at the price of the second highest bidder (thus the name "second price auction"). If there are ex aequo, the winner is chosen randomly among them. We will try to find the rational solution to this game.

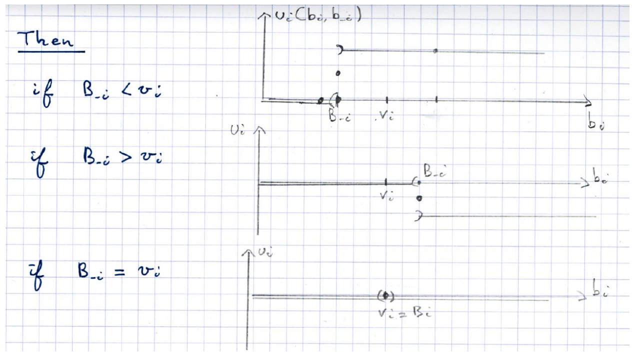
Le us first construct a utility function compatible with the preference relation of the players. For a player i, there are 2 possible outcomes: not winning $(\neg w_i)$, and winning at a cost c_i $((w_i, c_i))$. We assume that $c_i \leq c'_i$, implies $(w_i, c_i) \succsim_i (w_i, c'_i)$, and that the private value v_i can be defined by the fact that winning at a cost v_i is equivalent to not winning $((w_i, v_i) \approx_i \neg w_i)$.

One possible choice for the utility function of player i is therefore

$$u_i(\neg w_i) = u_i(b_i, b_{-i}) = 0 \qquad \text{if } b_i < B_{-i} = \max(b_{-i}) \tag{3}$$
$$u_i(w_i, c_i = B_{-i}) = u_i(b_i, b_{-i}) = v_i - B_{-i} \qquad \text{if } b_i > B_{-i} \tag{4}$$
$$u_i(b_i, b_{-i}) = \alpha(B_{-i}, k)(v_i - B_{-i}) \qquad \text{if } b_i = B_{-i}, \text{ with } k \text{ winners} . \tag{5}$$

(with $\alpha \in [0, 1]$).

Proof



In all three cases, the strategy $b_i = v_i$ is weakly dominant. It is therefore the rational solution of the second price auctions.

3. Nash equilibrium

3.1. Definition of the Nash equilibrium

The concept of dominated strategies plays a very important role in game theory, as it often makes it possible to simplify the analysis by eliminating irrelevant options. However, in most circumstances there are not enough dominated strategies to specify a "rational outcome".

Consider for instance the game in strategic form

		Player II		
		L	**C**	**R**
	T	0 , 6	6 , 0	4 , 3
Player I	**M**	6 , 0	0 , 6	4 , 3
	B	3 , 3	4 , 3	5 , 5

There is no dominated strategy, neither for player I, nor for player II. For instance from the point of view of player II:

- if I plays T: $L \succsim_{II} R \succsim_{II} C$
- if I plays M: $C \succsim_{II} R \succsim_{II} L$
- if I plays B: $R \succsim_{II} C \approx_{II} R$

There is thus no pair of strategies $(A_1, A_2) \in \{L, C, R\}^2$ of player II such that whatever I plays, A_1 is always prefered to A_2, i.e. no dominated strategy.

On the other hand the strategy vector $s^* = (B, R)$ is such if I plays $s_1^* = B$, the best strategy for II is $s_2^* = R$, and reciprocally if II plays $s_2^* = R$, the best strategy for I is $s_1^* = B$. Starting from s^*, neither I nor II can change (alone) their strategy without loosing something. Here $s^* = (B, R)$ is what is called a Nash equilibrium.

More generally we have

Definition 6 *A strategy vector* $s^* = (s_1^*, s_2^*, \cdots, s_n^*)$ *is a Nash equilibrium*

$$\Leftrightarrow \forall i \in N, \forall s_i \in S_i \quad u(s_i, s^*_{-i}) \leq u(s^*) \ .$$

3.2. Nash equilibrium for two-players zero-sum game

Consider the game

I \ II	L	C	R
T	3 , -3	-5 , 5	-2 , 2
M	1 , -1	4 , -4	1 , -1
B	6 , -6	-3 , 3	-5 , 5

$\Leftrightarrow$

I \ II	L	C	R
T	3	-5	-2
M	1	4	1
B	6	-3	-5

,

which is such that for any strategy vector s, $u_{II}(s) = -u_I(s)$. We assume that the goal of both players is to make "*the best of the worse case scenario*". For instance for player I

if $s_I = T$, worse case $\rightarrow s_{II} = C \quad ; \quad u_I = -5$

if $s_I = M$, worse case $\rightarrow s_{II} = L$ or $R \quad ; \quad u_I = 1$

if $s_I = B$, worse case $\rightarrow s_{II} = R \quad ; \quad u_I = -5$

We thus have $\underline{v}_I \equiv \max_{s_I} \min_{s_{II}} u_I(s_I, s_{II}) = \max(-5, 1, -5) = 1$, and the corresponding MaxMin strategy is $s_I = M$.

In the same way, $\underline{v}_{II} \equiv \max_{s_{II}} \min_{s_I} u_{II}(s_I, s_{II}) = -\min_{s_{II}} \max_{s_I} u_I(s_I, s_{II}) = -\min(6, 4, 1) = -1$, and the corresponding MaxMin strategy is $s_{II} = R$.

One can check here that $s^* = (M, R)$ is a Nash equilibrium, and that $\underline{v}_I = -\underline{v}_{II} = u_I(s^*)$.

More generally

Definition 7 *A two-players zero-sum game is a game such that*

$$\forall s \in S = S_1 \otimes S_2, \qquad u_2(s) = -u_1(s) \ .$$

We note $u(s) = u_1(s) = -u_2(s)$.

For any two-player zero-sum game we can define

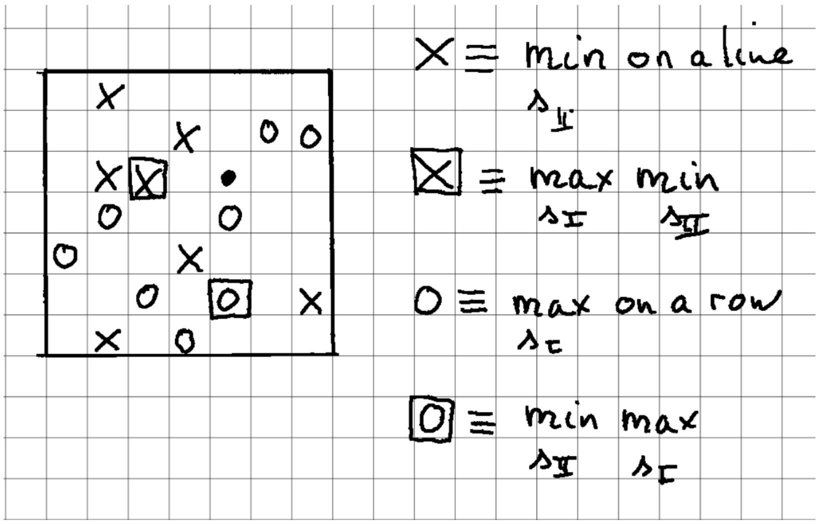
$$\begin{aligned} \underline{v}_1 &= \max_{s_1} \min_{s_2} u(s_1, s_2) \\ \underline{v}_2 &= \max_{s_2} \min_{s_1} u_2(s_1, s_2) = -\min_{s_2} \max_{s_1} u(s_1, s_2) \end{aligned}$$

the results of the MinMax strategies for player one and player two. We further note $\underline{v} = \underline{v}_1$ and $\overline{v} = -\underline{v}_2$.

Theorem 2 *A two-player zero-sum game is said to have* a value *if $\underline{v} = \overline{v} = v$. In that case there is a strategy vector for which this value is attained (i.e. $\exists\, s^*$ such that $u(s^*) = v$) and such that s^* is a Nash equilibrium for the game.*

NB: the reciprocal is not true : a two-player zero-sum game can have a Nash equilibrium and not a value.

Proof



We see that

$$v = \underline{v} = u(\boxed{X}) \leq u(\bullet) \leq u(\boxed{O}) = \overline{v} = v$$

Thus $u(\bullet) = v$, and $\bullet \equiv (s_I^*, s_{II}^*)$ is a Nash equilibrium.

4. Nash Theorem and von Neumann Theorem

4.1. Mixed form extension of strategic form games

Consider a game in strategic form, which is thus defined by:

- A set of players $N = \{1, \cdots, n\}$.
- A set of strategies $S_i = \{s_i\}$ for each player $i \in N$.
- The utility functions

$$u_i : \mathbf{S} = S_1 \times S_2 \times \cdots \times S_n \longrightarrow \mathbb{R}$$
$$(s_i, s_{-i}) \longrightarrow u_i(s_i, s_{-i})$$

Definition 8 *A* mixed *strategy is simply a probability distribution over the set of* pure *strategies S_i. We note Σ_i the set of mixed strategies.*

$$\Sigma_i \equiv \left\{ \sigma_i : S_i \mapsto [0,1] / \sum_{s_i \in S_i} \sigma_i(s_i) = 1 \right\} .$$

Notation If $S_i = \{a, b, c\}$, we note

$$\sigma_i \equiv \left(\frac{1}{3}(a), \frac{1}{3}(b), \frac{1}{3}(c) \right)$$

the mixed strategy corresponding to a probability $1/3$ to follow strategy a, a probability $1/3$ to follow strategy b, and a probability $1/3$ to follow strategy c.

Definition 9 *The* mixed extension *of a game in strategic form is defined by*

- *The set of mixed strategies $\Sigma_i = \{\sigma_i\}$*

- *The utilities*

$$\begin{aligned} u_i : \mathbf{\Sigma} = \Sigma_1 \times \cdots \times \Sigma_n &\longrightarrow \mathbb{R} \\ (\sigma_i, \sigma_{-i}) &\longrightarrow u_i(\sigma_i, \sigma_{-i}) = \mathbb{E}[u_i(s_i, s_{-i})] \\ &= \sum_{(s_1,\cdots,s_n)} u_i(s_1, \cdots, s_n)\sigma_1(s_1)\sigma_2(s_2)\cdots\sigma_n(s_n) \ . \end{aligned}$$

NB1 The mixed extension of a game in strategic form is itself a game in strategic form.

NB2 The form of the utility functions of the mixed extension is justified because we assume the utilities of the players satisfy the von-Neumann Morgenstern axioms, and thus the utility functions are linear.

4.2. Nash Theorem

Theorem 3 **(Nash 1950)** *Every games in strategic form (with a finite number of players, each of them having a finite number of strategies) has (at least) one Nash equilibrium in mixed strategies.*

The derivation of this theorem is a bit technical, and we will not provide it here. We will instead give a derivation of the related result:

4.3. Von Neumann Theorem

Theorem 4 **(von Neumann 1928)** *Every two-player zero-sum games in strategic form in which each player has a finite number of strategies has a value.*

Remark 1 If a 2-player 0-sum game has a "value", this value is associated with a Nash equilibrium. The Von Neumann Theorem is therefore somewhat the equivalent of the Nash theorem for 2-player 0-sum games. Note though that having a value imply a Nash equilibrium, but that the reverse is not true.

Remark 2 Unless it is a constant, the extrema of a linear function are necessarily at the boundary of its domain of definition. Since we assume the von Neumann-Morgenstern axioms, and thus linearity for the utility functions,

$$u_i(\sigma_i, \sigma_{-i}) = \sum_j \sigma_i(s_i^j) u_i(s_i^j, \sigma_{-i}) \ .$$

this implies in particular

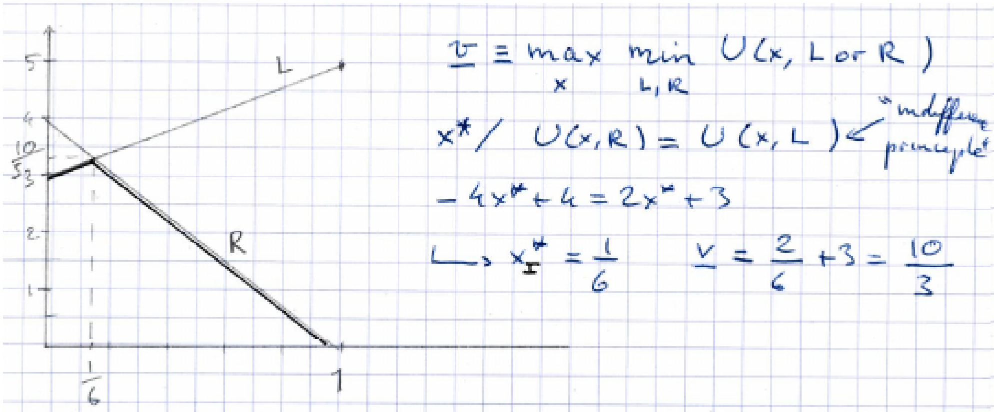
$$\begin{aligned} \underline{v} &= \max_{\sigma_1\in\Sigma_1} \min_{\sigma_2\in\Sigma_2} u(\sigma_1, \sigma_2) = \max_{\sigma_1\in\Sigma_1} \min_{s_s\in S_2} u(\sigma_1, s_2) \\ \overline{v}_2 &= \min_{\sigma_2\in\Sigma_2} \max_{s_1\in S_1} u_(s_1, \sigma_2) \end{aligned}$$

A simple example Consider the 2-players 0-sum game with two pure strategies corresponding to the table

		II	
		L	**R**
I	**T**	5	0
	B	3	4

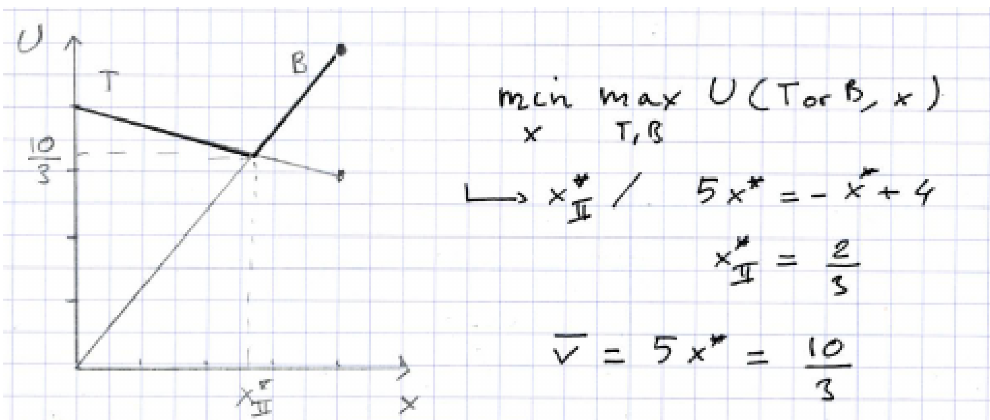
For player one, a mixed strategy corresponds to $[x_1(T), (1-x_1)(B)]$, $x_1 \in [0,1]$. We thus have

$$\begin{aligned} u(x_1, R) &= (1-x_1)\cdot 4 = -4x_1 + 4 \\ u(x_1, L) &= x_1 \cdot 5 + (1-x)\cdot 3 = 2x_1 + 3 \end{aligned}$$



For player two, a mixed strategy corresponds to $[x_2(L), (1-x_2)(R)$, $x_2 \in [0,1]$, and thus

$$u(T, x_2) = 5x_2$$
$$u(B, x_2) = x_2 \cdot 3 + (1-x) \cdot 4 = -x + 4$$



We have thus $\underline{v} = \overline{v} = v = 10/3$, ans $\sigma^* = (1/6, 2/3)$.

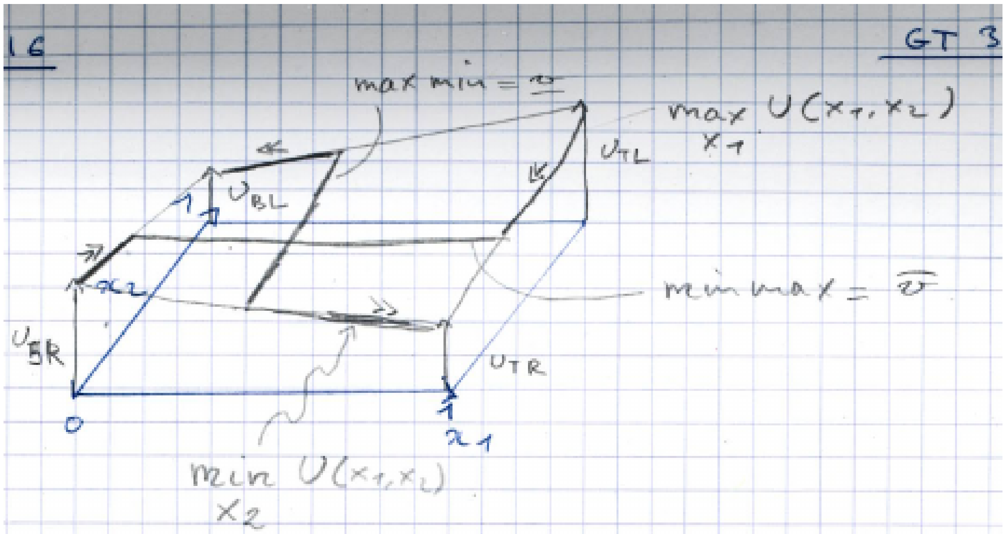
Why is this working ? Consider the general case of a 2-players 0-sum game with two pure strategies

	L	**R**
T	u_{TL}	u_{TR}
B	u_{BL}	u_{BR}

We furthermore assume that there is no dominant strategy, and thus $(u_{TL} > u_{BL}) \Leftrightarrow (u_{TR} < u_{BR})$ and $(u_{TL} > u_{TR}) \Leftrightarrow (u_{BL} < u_{BR})$. The graph of the mixed strategies utility function

$$u(x_1, x_2) = x_1 x_2 u_{BR} + (1-x_1)x_2 u_{TR} + x_1(1-x_2)u_{BL} + (1-x_1)(1-x_2)u_{TL}$$

can be represented by the graph



Since the minmax and maxmin lines cross, they should have the save value, implying $\underline{v} = \overline{v} = v$.

4.4. Indifference principle

We finish this section with one important property of Nash equilibria for mixed strategies, which is characterized by the following result

Theorem 5 **(Indifference principle)** *Let σ^* be a Nash equilibrium in mixed strategies of a strategic form game, and let s_i^α and s_i^β be two pure strategies of player i. Then*

$$\left(\sigma_i^*(s_i^\alpha) > 0 \quad \textit{and} \quad \sigma_i^*(s_i^\beta) > 0\right) \Rightarrow \left(u_i(s_i^\alpha, \sigma_{-i}^*) = u_i(s_i^\beta, \sigma_{-i}^*)\right) .$$

Proof Assume $u_i(s_i^\alpha, \sigma_{-i}^*) > u_i(s_i^\beta, \sigma_{-i}^*)$, and let

$$\sigma_i(t_i) = \begin{cases} \sigma_i^*(t_i) & \text{if } t_i \neq s_i^\alpha \text{ or } s_i^\beta \\ 0 & \text{if } t_i = s_i^\beta \\ \sigma_i^*(s_i^\alpha) + \sigma_i^*(s_i^\beta) & \text{if } t_i = s_i^\alpha \end{cases}$$

Without the $t_i \neq s_i^\alpha, s_i^\beta$, this amount to play a pure strategy s_i^α, which makes sense since it has a higher utility.

We then have

$$\begin{aligned} u_i(\sigma_i, \sigma_{-i}^*) = & \sum_{t_i \neq s_i^\alpha \text{ or } s_i^\beta} \sigma_i^*(t_i) u_i(t_i, \sigma_{-i}^*) + (\sigma_i^*(s_i^\alpha) + \sigma_i^*(s_i^\beta)) u_i(s_i^\alpha, \sigma_{-i}^*) \\ > & \sum_{t_i \neq s_i^\alpha \text{ or } s_i^\beta} \sigma_i^*(t_i) u_i(t_i, \sigma_{-i}^*) + \sigma_i^*(s_i^\alpha) u_i(s_i^\alpha, \sigma_{-i}^*) + \sigma_i^*(s_i^\beta) u_i(s_i^\beta, \sigma_{-i}^*) \\ = & \, u_i(\sigma^*) \end{aligned}$$

This is impossible if σ^* is a Nash equilibrium, and thus the initial hypothesis cannot hold.

	Dove	Hawk
Dove	4, 4	2, 8
Hawk	8, 2	1, 1

Table 1: Hawk and Dove game (lines : player 1, defender; columns : player 2, invader)

5. Evolutionary stable strategies

Consider the "hawk and dove" game characterized in strategic form by the following table

The hawk and dove game is a symmetric game, in the sense that $S_1 = S_2 = \{\text{Dove}, \text{Hawk}\}$ and $\forall (s_a, s_b) \in S, u_1(s_a, s_b) = u_2(s_b, s_a)$.

The mixed strategy of player 1 and 2 are

$$\sigma_1(x_1) = (x_1(\text{Dove}), (1 - x_1)\text{Hawk})$$
$$\sigma_2(x_2) = (x_2(\text{Dove}), (1 - x_2)\text{Hawk})\ .$$

One can check easily that $\sigma^* = (\sigma_1(x^*), \sigma_2(x^*))$ is a Nash equilibrium for $x_1^* = x_2^* = x^* = 1/5$.

Assume now that the game is played repeatedly, and that the outcome of the game is related not to a utility, but to the number of offspring of the animal playing that game, and furthermore that these offspring tend to play the same strategies as their parents (up to some random mutation). An agent (defender) is thus characterized by the mixed strategy "x" that she is following.

Assume now that this agent has a probability y to meet an invader that follows the Dove strategy, and $(1-y)$ that she follows Hawk (this could be the probability that the invader plays one of this pure strategies, or just implies that the invader plays the mixed strategy y, or a mix of the two). Then for the outcome u_D for the defender depending on her strategy choice is given by

$$\begin{aligned} u_1(\text{Dove }, y) =& 4y + 2(1-y) \\ u_1(\text{Hawk}, y) =& 8y + \ (1-y) \\ u_1(x, y) \ =& x(4y + 2(1-y)) + (1-x)(8y + (1-y)) \\ =& x(1-5y) + (1+7y) \end{aligned}$$

Thus :

- if $(y > 1/5)$, the average number of offsprings decreases with $x \quad \longrightarrow \quad$ more Doves
- if $(y < 1/5)$, the average number of offsprings increases with $x \quad \longrightarrow \quad$ more Hawks
- if $(y = 1/5)$, the average number of offsprings is independent of x.

If we assume that agents can alternatively be invaders and defenders, this gives the intuition that a population of agents with strategy $x = 1/5$ should be stable in some sense, and this is clearly related to the fact that $\sigma^* = (1/5, 1/5)$ is a Nash equilibrium.

To define this notion of stability more precisely, one introduces the concept of "evolutionary stable strategy (ESS)", which corresponds to the following question: Consider an initial situation where everyone follows a strategy x^*, and assume that a small group of size ϵ switches to a new strategy $x \neq x^*$. Will this small group become larger or simply disappear ?

This can be phrased more formally as

Definition 10 *A mixed strategy x^* in a two players symmetric game is said to be an evolutionary stable strategy (ESS) if for every mixed strategy $x \neq x*$ $\exists \epsilon_0 > 0$ such that $\forall \epsilon < \epsilon_0$:*

$$(1-\epsilon)u_1(x, x^*) + \epsilon u_1(x, x) < (1-\epsilon)u_1(x^*, x^*) + \epsilon u_1(x^*, x)$$

Returning to our Hawk and Dove example above, we have $u_1(x_1, x_2) = x_1(1 - 5x_2) + (1 + 7x_2)$ and $x^* = 1/5$. Therefore

$$\begin{aligned}
u_1(x, x^*) &= u_1(x^*, x^*) = 1 + 7x^* = \frac{12}{5} \\
u_1(x, x) &= x(1 - 5y) + 1 + 7x = -5x^2 + 8x + 1 \\
u_1(x^*, x) &= x^*(1 - 5x) + (1 + 7x) = 6x + \frac{6}{5}
\end{aligned}$$

What we want to check is:

$$\begin{aligned}
&(1 - \epsilon)u_1(x, x^*) + \epsilon u_1(x, x) < (1 - \epsilon)u_1(x^*, x^*) + \epsilon u_1(x^*, x) \\
&\Leftrightarrow u_1(x, x) < u_1(x^*, x) \\
&\Leftrightarrow 5x^2 - 2x + < u_1(x^*, x) \\
&\Leftrightarrow 5(x - \frac{1}{5})^2 > 0
\end{aligned}$$

which is here indeed the case. So the Nash strategy $x^* = 1/5$ is here indeed evolutionary stable.

As a final remark, note that $x^* =$ ESS can mean either that $u_1(x, x^*) < u_1(x^*, x^*)$ for all other mixed strategies x, or that $u_1(x, x^*) \leq u_1(x^*, x^*)$, but in that case the x's such that $u_1(x, x^*) = u_1(x^*, x^*)$ verify $u_1(x, x) < u_1(x^*, x)$. The first case would correspond to a *strict* Nash equilibrium, which however does not really exist in mixed strategies. In the second case an extra constraint is imposed beyond the Nash equilibrium.

B - Optimal control

Problem: I'm on a beach, I would like to spend some time in the shade while the sun is high, then spend some time on the shore, and eventually go back to the bus station in time to go home. What mathematical framework would be appropriate for this task ?

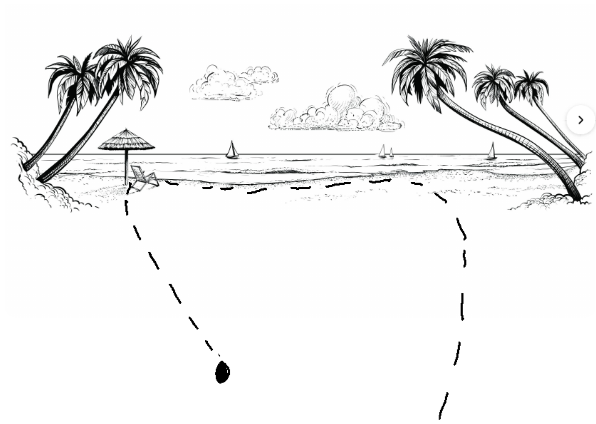


Figure 8: [From https://www.papierpeintpanoramique.be/plage-en-noir-et-blanc.html]

This can be cast in the form of an *optimal control* problem, which is characterised by

- A *state* variable $\mathbf{X}_t$ describing the state of the agent at time t (here $\mathbf{X}_t$ would just be the location of the agent at time t).
- A *control* variable $\mathbf{v}_t$, chosen by the agent, and that will impact the trajectory followed by $\mathbf{X}_t$. For instance here we will assume $\mathbf{v}_t$ is the desired velocity at time t, but that the motion is subject to some randomness, so that $\mathbf{X}_t$ follows a Langevin equation

$$dX_t = \mathbf{v}_t dt + \sigma d\xi_t \qquad , \tag{6}$$

  with ξ a "normalized" white noise ($\mathbb{E}[d\xi_t] = 0$, $\mathbb{E}[d\xi_t^\alpha d\xi_t^\beta] = \delta_{\alpha\beta}\, dt$) and σ the strength of the noise.
- A *cost function*, which is a functional of $\mathbf{v}(\cdot)$, and a function of the starting point x_0 and of the initial time t_0.

$$J_{x_0 t_0}[\mathbf{v}(\cdot)] = \mathbb{E}\left[\int_{t_0}^{T} \underbrace{L(\mathbf{v}_\tau, \mathbf{X}_\tau, \tau)}_{\text{running cost}}\, d\tau + \underbrace{C_T(\mathbf{X}_T)}_{\text{final cost}}\right]_{|\mathbf{X}_{t_0}=\mathbf{x}_0} \qquad . \tag{7}$$

  Here we can assume the simple form

$$L(\mathbf{v}, \mathbf{X}_t, t) = \frac{1}{2} m\, \mathbf{v}^2 - V(\mathbf{x}, t)\ , \tag{8}$$

  where the first term indicates that we would like a desired velocity that is not too high, and the second term implements our preference to be under the beach umbrella when the sun is high, and near the shore later. Finally $C_T(\mathbf{X}_T)$ tells us that we rather be at the bus stop when it is time to leave, i.e. at the end of the optimization time T.

Solving this optimal control problem boils down to determining the optimal control $\mathbf{v}_t^*$ such that the cost $J_{\mathbf{x}_0 t_0}[\mathbf{v}^*(\cdot)]$ is minimal.

Optimal control is the bread and butter of engineering sciences, and can be applied in a vast variety of contexts. The state variable is not necessarily your location on the beach, but can be the speed of an engine, the amount of stock purchased by an investor, or the amount of energy stored in a battery. Similarly, the relationship between the control variable and the dynamics of the state variable may be considerably more complex than the Langevin equation (6), and the cost function may have a more general form than Eq. (7) (and of course than Eq. (8)).

Here, I will limit the discussion to the above framework, which is sufficient to introduce the general philosophy, and which also has the advantage that the formalism we will be dealing with is very reminiscent of something very familiar to physicists.

1. The noiseless case : Lagrange equations

In the absence of noise (ie $\sigma = 0$), it is not necessary to take the expectation value in Eq. (7), and the "desired velocity" $\mathbf{v}_t$ becomes just "the" velocity

$$\mathbf{v}_t = \dot{\mathbf{X}}_t \; . \tag{9}$$

Introducing a Lagrange multiplier $\mathbf{P}_t$ to impose the constraint Eq. (9) (so that $\mathbf{v}_t$ and $\mathbf{X}_t$ can be treated as independent variables), the optimal velocity $\mathbf{v}_t^*$ is such that the modified cost functional

$$\begin{aligned} J_{x_0 t_0}[\mathbf{v}(\cdot), \mathbf{X}(\cdot)] &= \int_{t_0}^{T} \left(L(\mathbf{v}_\tau, \mathbf{X}_\tau, \tau) + \mathbf{P}_\tau(\dot{\mathbf{X}}_\tau - \mathbf{v}_\tau)\right) d\tau \; + \; C_T(\mathbf{X}_T) & (10) \\ &= \int_{t_0}^{T} \left(L(\mathbf{v}_\tau, \mathbf{X}_\tau, \tau) - (\dot{\mathbf{P}}_\tau \mathbf{X}_\tau + \mathbf{P}_\tau \mathbf{v}_\tau)\right) d\tau \; + \; [\mathbf{P}_T \mathbf{x}_T - \mathbf{P}_0 \mathbf{X}_0 + C_T(\mathbf{X}_T)] \; , & (11) \end{aligned}$$

is invariant under small variations $\delta \mathbf{v}_t$ and $\delta \mathbf{X}_t$. We get

$$\frac{\delta J_{\mathbf{x}_0}}{\delta \mathbf{v}_t} = 0 = \frac{\partial L}{\partial \mathbf{v}} - \mathbf{P}_t \tag{12}$$

$$\frac{\delta J_{\mathbf{x}_0}}{\delta \mathbf{X}_t} = 0 = \frac{\partial L}{\partial \mathbf{X}} - \dot{\mathbf{P}}_t \tag{13}$$

which implies

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \mathbf{v}}\right) = \frac{\partial L}{\partial \mathbf{X}} \; . \tag{14}$$

We recognize the usual Lagrange equation of classical mechanics, which is not a surprise since the cost function Eq. (7) has exactly the structure of a classical action derived from the Lagrangian $L(\mathbf{v}, \mathbf{X}, t)$. With the form Eq. (8) of the Lagrangian, Eq. (14) take the even simpler form of Newton second law

$$m\ddot{\mathbf{X}} = -\partial_{\mathbf{X}} V \; .$$

Finally, the variation with respect $\mathbf{X}_T$ leads to $\partial_{\mathbf{X}_T}[\mathbf{P}_T \mathbf{X}_T - \mathbf{P}_0 \mathbf{X}_0 + C_T(\mathbf{X}_T)] = 0 = \mathbf{P}_T + C_T'(\mathbf{X}_T)$ giving the boundary conditions

$$X_{t=0} = x_0 \tag{15}$$

$$P_{t=T} = \frac{\partial L}{\partial \mathbf{v}}(\mathbf{v}_T^*, \mathbf{x}_T^*, T) = -C_T'(\mathbf{X}_T) \; . \tag{16}$$

These "mixed" boundary conditions, for which initial position and final momentum are fixed, are the only significant difference with what we are used to in classical mechanics.

2. The noisy case : Bellman equation and linear programming

In the presence of noise, the above Lagrange approach (where one follows the trajectory of an individual agent) is difficult to generalize, and it turns out to be simpler to adopt an Eulerian approach where one focus on fixed specific location in the space through which the agent goes through as time passes.

For this purpose, we introduce the value function $u(\mathbf{x},t)$ defined as the optimal cost function for an agent located at $\mathbf{x}$ at time t

$$u(\mathbf{x},t) \equiv \inf_{\mathbf{v}(.)} J_{\mathbf{x},t}[\mathbf{v}(.)] \ .$$

We stress again that, contrarily to $\mathbf{X}_t$, here $\mathbf{x}$ is a given location in space, with no time dependence. To compute $u(\mathbf{x},t)$ and finally relate the result to $\mathbf{v}_t^*$, we turn to the notion of dynamic programming, and more specifically to Bellman's optimality principle, which states that

> "An optimal policy has the property that, whatever the initial state and the initial decision, the remaining decisions must constitute an optimal policy with respect to the state resulting from the first decision."

(note this is based under the assumption that the agent is allowed to re-adapt her strategy if, because of the noise, the trajectory actually followed is not the "averaged one" anticipated).

This Bellman optimality principle means that the computation of u can be simplified by discretizing the problem and optimizing on infinitesimally small time frames. If we split the optimization into two intervals,$[t,t+dt]$ and $[t+dt,T]$, we get the Bellman equation

$$u(\mathbf{x},t) = \inf_{\mathbf{v}_t} \mathbb{E}\left[\int_t^{t+dt} L(\mathbf{v}_\tau,\mathbf{x}_\tau,\tau)d\tau \ + \ u(\mathbf{x}{+}d\mathbf{x},t{+}dt)\right] , \tag{17}$$

where now the minimisation is performed only *at time* t and not for the entire function $\mathbf{v}(\cdot)$. Furthermore, because of Eq. (6), $d\mathbf{x}$ should be understood here as $d\mathbf{x} = \mathbf{v}_t dt + \sigma d\xi_t$. Since the noise $|d\xi_t| \sim \sqrt{dt}$, we need to expand u to order 2 in $d\mathbf{x}$

$$u(\mathbf{x}{+}d\mathbf{x},t{+}dt) = u(\mathbf{x},t) + \partial_t u(\mathbf{x},t)dt + \partial_{\mathbf{x}} u(\mathbf{x},t)d\mathbf{x} + \frac{1}{2}\partial^2_{\alpha\beta}u(\mathbf{x},t)dx^\alpha dx^\beta + o(dt) \ .$$

With $\mathbb{E}[d\xi^\alpha]=0$ and $\mathbb{E}[d\xi^\alpha d\xi^\beta] = \delta_{\alpha\beta}dt$, this gives

$$\mathbb{E}[u(\mathbf{x}{+}d\mathbf{x},t{+}dt)] = u(\mathbf{x},t) + \left[\partial_t u(\mathbf{x},t) + \mathbf{v}_t\cdot\partial_{\mathbf{x}}u(\mathbf{x},t) + \frac{1}{2}\sigma^2\partial^2_{\alpha\alpha}u(\mathbf{x},t)\right]dt + o(dt) \ .$$

Keeping only the terms of order dt in Eq. (17), we obtain

$$\inf_{\mathbf{v}_t}\left[L(\mathbf{v}_t,x,t) + \partial_t u(\mathbf{x},t) + \mathbf{v}_t\cdot\partial_{\mathbf{x}}u(\mathbf{x},t) + \frac{1}{2}\sigma^2\partial^2_{\alpha\alpha}u(\mathbf{x},t)\right] = 0 \ .$$

The optimal choice $\mathbf{v}_t^*$ should therefore fulfil the condition $\partial_{\mathbf{v}_t^*}[L(\mathbf{v}_t^*,x,t) + \mathbf{v}_t^*\cdot\partial_{\mathbf{x}}u(\mathbf{x},t)] = 0$, that is

$$\frac{\partial L}{\partial \mathbf{v}_t^*}(\mathbf{v}_t^*,x,t) = -\partial_{\mathbf{x}}u(\mathbf{x},t) \ , \tag{18}$$

and introducing

$$H(\mathbf{p},\mathbf{x},t) = \mathbf{v}_t^*\cdot\mathbf{p} - L(\mathbf{v}_t^*,x,t) \ , \tag{19}$$

the values function should fulfil the PDE

$$\partial_t u(\mathbf{x},t) - H(-\partial_{\mathbf{x}}u,\mathbf{x},t) + \frac{1}{2}\sigma^2\Delta u(\mathbf{x},t) = 0 \ , \tag{20}$$

which is known as the Hamilton-Jacobi Bellman equation.

Noting that at the final time $t = T$ the cost Eq. (7) simplifies to $J_{x,T} = C_T(x)$, and that, since there is nothing left to optimize on, this is also the "optimal cost", the value function need to fulfil the terminal boundary condition

$$u(\mathbf{x}, T) = C_T(\mathbf{x}) \ . \tag{21}$$

Solving the optimal control problem thus amounts to solve *backward* the PDE (20) from the boundary condition (21), and, once $u(\mathbf{x}, t)$ is known, to obtain the optimal control (which is our goal) from (18). (Note the sign in front of the Laplacian also impose that Eq. (21) should be solved backward in time.)

Quadratic running cost

If the running cost $L(\mathbf{v}_t, \mathbf{X}_t, t)$ takes the simple form Eq. (8), the optimality condition Eq. (18) and the function H defined in Eq. (19) read explicitly

$$\mathbf{v}_t = -\frac{1}{m}\frac{\partial u}{\partial \mathbf{x}}(\mathbf{x}, t) \tag{22}$$

$$H(\mathbf{p}, \mathbf{x}, t) = \frac{p^2}{2m} + V(\mathbf{x}) \ , \tag{23}$$

giving for the HJB equation (20)

$$\partial_t u(\mathbf{x}, t) - \frac{(\nabla_{\mathbf{x}} u)^2}{2m} + \frac{\sigma^2}{2}\Delta u(\mathbf{x}, t) = V(\mathbf{x}) \ . \tag{24}$$

Back to the noiseless case

When $\sigma = 0$, one can also follow this Eulerian route, with the only difference that the HJB equation (20) reduces to the Hamilton-Jacoby equation[1]

$$\partial_t u(\mathbf{x}, t) - H(-\partial_{\mathbf{x}} u, \mathbf{x}, t) = 0 \ , \tag{25}$$

(with the same terminal boundary condition $u(x, T) = C_T(x)$). The Hamilton equations

$$\dot{\mathbf{x}} = +\frac{\partial H}{\partial \mathbf{p}}$$
$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{x}}$$

are the equivalent, in the Hamiltonian formalism to the Lagrange equation Eq. (14), and give rise to the same dynamics, and it can be shown using traditional arguments from classical mechanics that the optimal control derived from solving the Hamilton-Jacobi equation (25) and setting the optimal control $\mathbf{v}_t^*$ by Eq. (18) gives the same results as defining it as the velocity Eq. (9) for the trajectories solution of the Lagrange equation (14).

C - Introduction to mean field games

1. From optimal control to differential games

If we define a *game* as a *competitive optimization*, i.e. a situation in which more than one agent is trying to determine a strategy that would yield a profit (or minimize a cost), but for which their own outcome is influenced by the strategies of the other agents, then the optimal control problem discussed above is

[1] Unfortunately, when making the connection with the traditional classical dynamics u is not the action S, but minus the action ($u \equiv -S$). Therefore $p = -\partial_{\mathbf{x}} u$ (instead of $p = \partial_x S$), and the usual HJ equation $\partial_t S + H(\partial_x S, x) = 0$ reads $-\partial_t u + H(-\partial_x S, x) = 0$.

transformed into a game when more than one agent is performing optimal control, and the cost function of any given agent is influenced by the strategy choices of the others.

Let's therefore consider a set of agents $i = 0, 1, \cdots, N$, whose individual states are described by a continuous variable $\mathbf{X}^i \in \mathbb{R}^d$. Depending on the problem at hand, the state variable $\mathbf{X}^i$ can represent a physical location, the amount of resources owned by a company, etc. (To make things more concrete, you might consider the case where these "agents" are tourists arriving at the beach at the same time and competing for space under the umbrellas or near the shore). We will assume that each of these state variables follows a Langevin equation

$$d\mathbf{X}_t^i = \mathbf{v}_t^i dt + \sigma d\xi_t^i \tag{26}$$

with initial condition $\mathbf{X}_{t=t_0}^i = \mathbf{x}_0^i$, and where the Brownian motions (ξ_t^j, ξ_t^i) are independent for $i \neq j$. Specializing to a quadratic Lagrangian as in Eq. (8), each individual will tuned her control parameter $\mathbf{v}_t^i$ to optimize a cost functional

$$J_{x_0^i, t_0}[\mathbf{v}^i(\cdot)] = \mathbb{E}\left[\int_{t_0}^T \left(\frac{(\mathbf{v}_t^i)^2}{2m} - \tilde{V}[\hat{\rho}_t(\cdot)](\mathbf{X}_t^i, t)\right) dt + C_T(\mathbf{X}_T^i)\right]_{|\mathbf{x}_{t=t_0}=\mathbf{x}_0} . \tag{27}$$

The main, but essential, difference from the plain vanilla optimal control considered above is that here the potential $\tilde{V}$ is supposed to include interactions with the other agents. Assuming that the agents are essentially identical (that is, they differ only in their state $\mathbf{X}_t^i$, but otherwise have the same dynamics, optimize the same cost function, and interact symmetrically), the interaction can be expressed in terms of the empirical density of agents

$$\hat{\rho}_t(\mathbf{x}) = \frac{1}{N}\sum_{j\neq i} \delta(\mathbf{x} - \mathbf{X}_t^j) . \tag{28}$$

In other words $\tilde{V}[\hat{\rho}_t(\cdot)](\mathbf{X}_t^i)$ is both a function of the agent's state $\mathbf{X}_t^i$ and a functional of the density of agents $\hat{\rho}_t$ in the state space. A typical example for $\tilde{V}$ would be

$$\tilde{V}[\hat{\rho}_t(\cdot)](\mathbf{x}) = g\hat{\rho}(\mathbf{x}) + V_0(\mathbf{x}, t) , \tag{29}$$

where $V_0(\mathbf{x}, t)$ is a "one body" term playing a role similar as its equivalent in the original optimal control problem, and $g\hat{\rho}(\mathbf{x})$ implement the preference (for $g > 0$) or aversion (for $g < 0$) of crowded places.

If we consider a given agent i and assume the strategies $\mathbf{v}_t^{-i}$ of the other agents fixed, the optimization performed by i is just a standard optimal control problem leading to the HJB equation

$$\partial_t u^i(\mathbf{x}, t) - \frac{(\nabla_\mathbf{x} u^i)^2}{2m} + \frac{\sigma^2}{2}\Delta u^i(\mathbf{x}, t) = \tilde{V}[\hat{\rho}](\mathbf{x}) . \tag{30}$$

with boundary condition $u^i(\mathbf{x}, T) = C_T(\mathbf{x})$ and control parameter fixed as $\mathbf{v}_t^{i*} = -\frac{1}{m}\nabla_\mathbf{x} u$. Finding a *Nash equilibrium* therefore correspond to solving simultaneously the coupled (through $\tilde{V}[\hat{\rho}]$) $(N + 1)$ equation Eq. (30) for $i = 0, \cdots, N$.

This correspond to a N-body game, which quickly becomes intractable as N increase.

2. Intermezzo : the mean field approximation for the Ising model (a brief recap)

The Ising model describes a system of spins $\{S_i = \pm 1\}$ with a configurational energy

$$E(\{S_i\}) = -h\sum_i S_i - J\sum_{<ij>} S_i S_j . \tag{31}$$

It is one of the simplest examples of many-body problems, and remain nevertheless in general intractable except in low dimensions ($d = 1, 2$).

In higher dimensions, a very good approximation is however provided by a mean-field approach. Consider all contributions involving spin j

$$\epsilon(S_j) = -hS_j - JS_j \sum_{k \in \text{n.n}} S_k \;, \tag{32}$$

where the sum is over nearest neighbours (n.n.) k of site j. The mean field approximation consists in approximating this contribution by replacing the S_k by their mean value

$$\epsilon_{\text{mf}}(S_j) = -hS_j - JS_j \sum_{k \in \text{n.n.}} \langle S_k \rangle = -h_{\text{mf}} S_j \tag{33}$$

where

$$h_{\text{mf}} = h + Jzm \tag{34}$$

with z the number of neighbours, and m, the magnetisation per spin, which is just the mean value of any given spin

$$m = \frac{1}{N} \sum_i \langle S_i \rangle = \langle S_k \rangle \; \forall k \;. \tag{35}$$

Thus, in the mean field approximation, one replaces the many-body problem governed by the configuration energy Eq. (31) by a one-body problem governed by Eq. (32), i.e. a non-interacting spin experiencing a field h_{mf}. At finite temperature, with $\beta = 1/k_B T$, we can write down the single-spin Boltzmann distribution straight away

$$p(S_j) = \frac{e^{-\beta \epsilon_{\text{mf}}(S_j)}}{\sum_{S_j = \pm 1} e^{-\beta \epsilon_{\text{mf}}}(S_j)} = \frac{e^{\beta h_{\text{mf}} S_j}}{e^{\beta h_{\text{mf}}} + e^{-\beta h_{\text{mf}}}} \;. \tag{36}$$

However, we still have a consistency condition to fulfil: the value of the magnetisation m predicted by (6) should be equal to the value of m used in the expression for h_{mf}. Thus we require

$$m = \sum_{S_j = \pm 1} p(S_j) S_j = \frac{e^{\beta h_{\text{mf}}} - e^{-\beta h_{\text{mf}}}}{e^{\beta h_{\text{mf}}} + e^{-\beta h_{\text{mf}}}} = \tanh(\beta h_{\text{mf}}) \;, \tag{37}$$

and we arrive at the mean-field equation for the magnetisation

$$m = \tanh(\beta h + \beta J z m) \;. \tag{38}$$

3. Mean Field Games

For the Ising model, the mean-field approximation consists in replacing a complex many-body problem by a much simpler one-body problem coupled to a “mean-field” (here the magnetization m), with the only complication that a self-consistent condition must be fulfilled.

The “mean-field game” approximation to the “many-body Nash equilibrium" problem above follow exactly the same logic. For large N, one assumes that the details of the empirical density $\hat{\rho}(\mathbf{x})$ Eq. (28) are irrelevant, and one replace this empirical density by its average

$$\rho(\mathbf{x}) = \mathbb{E}\left[\frac{1}{N} \sum_j \delta(\mathbf{x} - \mathbf{X}^j)\right] \;. \tag{39}$$

(NB: one important difference with Ising is that here the average is taken *over the noise* and not over space, so the average density keep an $\mathbf{x}$ dependence.) Contrary to the empirical density $\hat{\rho}(\mathbf{x})$ which is intrinsically a stochastic function, the averaged $\rho(\mathbf{x})$ is deterministic, and, starting from from the initial distribution $\rho^0(\mathbf{x})$, follow the Fokker-Planck/Kolmogorov equation [11]

$$\begin{cases} \partial_t \rho_t(\mathbf{x}) + \boldsymbol{\nabla}.(\rho_t(\mathbf{x})\, \mathbf{v}_t^*(\mathbf{x})) - \dfrac{\sigma^2}{2} \Delta \rho_t(\mathbf{x}) = 0 \\ \rho_{t=0}(\mathbf{x}) = \rho^0(\mathbf{x}) \end{cases} , \tag{40}$$

where the optimal velocity

$$\mathbf{v}_t^*(\mathbf{x}) = -\frac{\nabla u_t(\mathbf{x})}{m}$$

derives from the Hamilton-Jacobi-Bellman equation

$$\begin{cases} \partial_t u_t(\mathbf{x}) - \dfrac{(\nabla u_t(\mathbf{x}))^2}{2m} + \dfrac{\sigma^2}{2}\Delta u_t(\mathbf{x}) = \tilde{V}[\rho_t(.)](\mathbf{x}) \\ u_T(\mathbf{x}) = C_T(\mathbf{x}) \end{cases} \tag{41}$$

The above two equations (41) and (40) are coupled together (through the terms $\tilde{V}[\rho_t]$ and $\mathbf{v}_t^*$) and both of diffusion type, but respectively backward and forward in time; they form together the MFG system of equations [8]. As for the Ising problem, we have transformed an very complicated many-body problem into a much simpler one body problem (an optimization problem for a single agent), with however a "self-consistent condition" stating that the time evolution of the mean field $\rho_t(\mathbf{x})$ predicted by Eq. (40) should be the one entering or the r.h.s. of Eq. (41).

4. Permanent regime

In the limit of a large optimization period $T \to \infty$, we expect that under reasonable conditions (no explicit time dependence in $\tilde{V}$, etc ...), the system will, beyond a short transient near $t = 0$ and $t = T$, stay in a *permanent regime* for most of the optimization time.

One must keep in mind, however, that by "permanent regime", what we mean is that the physical (observable) quantities, namely the density and desired velocity are time independent

$$\begin{aligned} \rho_t(\mathbf{x}) &\longrightarrow \rho^{\mathrm{er}}(\mathbf{x}) \\ \mathbf{v}_t(\mathbf{x}) &\longrightarrow \mathbf{v}^{\mathrm{er}}(\mathbf{x}) \;. \end{aligned}$$

The value function $u_t(\mathbf{x})$, which is not observable, may however keep a trivial time dependence, which is however constrained by the fact that $\nabla u_t(\mathbf{x}) = -m\mathbf{v}^{\mathrm{er}}(\mathbf{x})$ is time independent. This impose that in the permanent regime, the value function should take the form

$$u_t(\mathbf{x}) = u^{\mathrm{er}}(\mathbf{x}) - \lambda t \;, \tag{42}$$

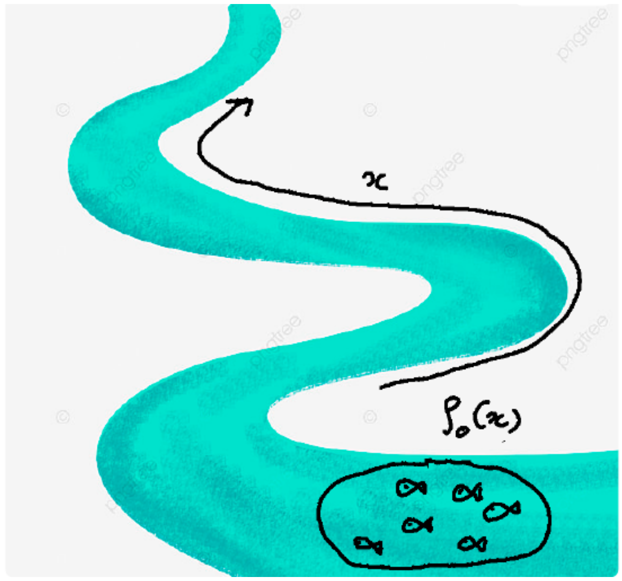
with the constant λ usually fixed by the (spatial) boundary conditions. From Eqs. (40)-(41), we see immediately that the pair $(\rho^{\mathrm{er}}(\mathbf{x}), \mathbf{v}^{\mathrm{er}}(\mathbf{x}))$ characterizing the permanent regime should be solution of the (time independent) couple system of equations

$$\begin{cases} \boldsymbol{\nabla}(\rho^{\mathrm{er}}(\mathbf{x})\mathbf{v}^{\mathrm{er}}(\mathbf{x})) - \dfrac{\sigma^2}{2}\Delta\rho^{\mathrm{er}}(\mathbf{x}) = 0 \\ -\lambda^{\mathrm{er}} - \dfrac{(\nabla u^{\mathrm{er}}(\mathbf{x}))^2}{2m} + \dfrac{\sigma^2}{2}\Delta u^{\mathrm{er}}(\mathbf{x}) = \tilde{V}[\rho^{\mathrm{er}}](\mathbf{x}) \end{cases} . \tag{43}$$

See e.g. P. Cardialaguet et al. [1], for a more mathematical discussion.

D - Study of a toy model: the fish-school problem

As an illustration, let us consider the following problem :



A school of fish live in a (1d) river :

- They start in the morning with an initial distribution $\rho_0(x)$.
- During the day they try to gather food, while staying close together to protect themselves from predators (and the latter concern take priority on the former).

$$\tilde{V}[\hat{\rho}_t(\cdot)](\mathbf{x}) = \underbrace{V_0(x)}_{\text{Intrinsic quality of the place}} + \underbrace{g\rho(x)}_{g>0 \text{ favour dense groups}} . \tag{44}$$

- At the end of the day they find shelter in the best possible place (quality of shelter described by $C_T(x)$).

(NB: do not try to use this theoretical approach to develop your fish farming business.)
All this to say we consider the (1d) MFG problem

$$\partial_t \rho_t + \partial_x(v_t \rho_t) - \frac{\sigma^2}{2} \partial^2_{xx} \rho_t = 0 \tag{45}$$

$$\partial_t u_t - \frac{(\partial_x u_t)^2}{2m} + \frac{\sigma^2}{2} \partial^2_{xx} u_t = \tilde{V}[\rho_t(.)](x) = V_0(x) + g\rho(x) , \tag{46}$$

with the desired velocity given by $v_t(x) = -\frac{1}{m}\partial_x u_t$, the initial/terminal boundary conditions $\rho_{t=0}(\cdot) = \rho_0(\cdot)$, $u_T(\cdot) = C_T(\cdot)$, and the assumption that $g > 0$ is large (in a sense to be defined more precisely later). (NB: the limitation to 1d and to the form (44) of $\tilde{V}$ are not very important for most of the discussion.)

1. Mapping to the non-linear Schrödinger equation

1.1. Bellman

As a first step, we use the well known fact that the Hamilton-Jacobi-Bellman equation for the value function $u(x,t)$ in (46) can be cast into a standard heat equation using a Cole-Hopf transformation [3]

$$\Phi_t(x) = \exp\left(-u_t(x)/m\sigma^2\right) . \tag{47}$$

Indeed, noting that $m\sigma^2$ has the dimension of an action, and using the notation

$$\hbar_{\text{eff}} \equiv m\sigma^2 \ , \tag{48}$$

we have $u_t = -\hbar_{\text{eff}} \log \Phi_t$, and thus

$$\partial_t u_t = -\hbar_{\text{eff}} \frac{\partial_t \Phi}{\Phi} \tag{49}$$

$$\partial_x u_t = -\hbar_{\text{eff}} \frac{\partial_x \Phi}{\Phi} \tag{50}$$

$$\partial^2_{xx} u_t = -\hbar_{\text{eff}} \frac{\partial^2_{xx} \Phi}{\Phi} + \hbar_{\text{eff}} \frac{(\partial_x \Phi)^2}{\Phi^2} \ , \tag{51}$$

which, inserted in Eq. (46) leads for Φ to the time-backwards diffusion equation,

$$-\hbar_{\text{eff}} \partial_t \Phi_t = +\frac{\hbar^2_{\text{eff}}}{2m} \partial^2_{xx} \Phi_t + \tilde{V}[\rho_t]\Phi_t \ , \tag{52}$$

$$\text{with} \quad \Phi_T(\cdot) = \exp(-C_T(\cdot)/\hbar_{\text{eff}}) \ . \tag{53}$$

Note that it follows from equation (53) that $\Phi_T(\cdot) > 0$, and that (52) then implies $\Phi_t(x) > 0$ for all times.

1.2. Fokker-Planck

In the same way, we can introduce the function $\Gamma_t(x) = \rho_t(x)/\Phi_t(x)$ [2], which is therefore such that

$$\rho_t(x) = \Gamma_t(x)\Phi_t(x) \ . \tag{54}$$

We thus have

$$\partial_t \rho_t = \Gamma_t \, \partial_t \Phi_t + \partial_t \Gamma_t \, \Phi_t \ , \tag{55}$$

$$\partial^2_{xx} \rho_t = \Gamma_t \, \partial^2_{xx} \Phi_t + \partial^2_{xx} \Gamma_t \, \Phi_t + 2\partial_x \Gamma_t \, \partial_x \Phi_t \ , \tag{56}$$

and, using that $v_t = -\frac{1}{m}\partial_x u_t = \sigma^2(\partial_x \Phi_t/\Phi_t)$,

$$\partial_x(v_t \rho_t) = \sigma^2 \partial_x \left(\frac{\partial_x \Phi_t}{\Phi_t} (\Phi_t \Gamma_t) \right) = \sigma^2 \partial_x \left(\partial^2_{xx} \Phi_t \, \Gamma_t + \partial_x \Phi_t \, \partial_x \Gamma_t \right) \ . \tag{57}$$

Inserting these equations into (45), we get

$$\partial_t \Gamma_t - \frac{\sigma^2}{2} \partial^2_{xx} \Gamma_t = \frac{\Gamma_t}{\Phi_t} \underbrace{\left(-\hbar_{\text{eff}} \partial_t \Phi_t - \frac{\hbar^2_{\text{eff}}}{2m} \partial^2_{xx} \Phi_t \right)}_{\hbar^{-1}_{\text{eff}} \Phi_t \tilde{V}[\rho_t](x)} = -\Gamma_t \, \tilde{V}[\rho_t](x) \ . \tag{58}$$

This second variable thus follows a heat equation almost identical to Eq. (52), but forward in time. Specializing to the form (44) of $\tilde{V}[\rho_t](x)$, the MFG system of equations (46)-(45) thus amount to the system

$$+\hbar_{\text{eff}} \partial_t \Gamma_t(x) = +\frac{\hbar^2_{\text{eff}}}{2m} \partial^2_{xx} \Gamma_t(x) + (V_0(x) + g\rho_t(x))\Gamma_t(x) \tag{59}$$

$$-\hbar_{\text{eff}} \partial_t \Phi_t(x) = +\frac{\hbar^2_{\text{eff}}}{2m} \partial^2_{xx} \Phi_t(x) + (V_0(x) + g\rho_t(x))\Phi_t(x) \ , \ , \tag{60}$$

with the understanding that $\rho \equiv \Gamma\Phi$, and initial/terminal conditions (53) and $\Gamma_{t=0}(\cdot) = \rho^0(\cdot)/\Phi_{t=0}(\cdot)$. Again, as both m^0 and $\Phi_{t=0}$ are strictly positive functions, we have $\Gamma_{t=0}(\cdot) > 0$, and Eq. (59) then implies this is the case for all times.

Under these transformations, the MFG system has been recast in a pair of non-linear heat equations, differing only by the sign on the left hand side and by their asymmetric boundary conditions.

Let us now consider the scalar the Schrödinger equation describing the quantum evolution of a wave amplitude $\psi_t(x)$ in a reversed potential $-V_0(x)$

$$i\hbar\,\partial_t\psi_t(x) = -\frac{\hbar^2}{2m}\partial^2_{xx}\psi_t(x) - V_0(x)\psi_t(x)\;. \tag{61}$$

If now a large number of (bosonic) particles are interacting together through a local two-body interaction $V_{\text{int}}(x,x') = -g\,\delta(x-x')$, the mean field approximation to this many-body problem amounts to the substitution $V_0(x) \to \tilde{V}[\rho^{\text{qu}}](x) = V_0(x) + g\rho^{\text{qu}}(x)$ with $\rho^{\text{qu}} \equiv \psi^*\psi$ and $\int_{\mathbb{R}}\rho^{\text{qu}}(x) = 1$. This leads to the non-linear Schrödinger equation

$$i\hbar\,\partial_t\psi_t(x) = -\frac{\hbar^2}{2m}\partial^2_{xx}\psi_t(x) - (V_0(x) + \rho^{\text{qu}}_t(x))\psi_t(x) \tag{62}$$

$$-i\hbar\,\partial_t\psi^*_t(x) = -\frac{\hbar^2}{2m}\partial^2_{xx}\psi^*_t(x) - (V_0(x) + \rho^{\text{qu}}_t(x))\psi^*_t(x)\;. \tag{63}$$

We thus see that Eqs. (62)-(63) are the equivalent to Eqs. (59)–(60) under the formal correspondence $\mu\sigma^2 \to \hbar$, $\phi(\mathbf{x},t) \to \psi(\mathbf{x},it)$ and $\Gamma(\mathbf{x},t) \to [\psi(\mathbf{x},it)]$.

In the same way, starting from the system (43) describing the permanent regime, and introducing the new variables, $\Phi^{\text{er}}_t(x) = \exp\left(-u^{\text{er}}_t(x)/\hbar_{\text{eff}}\right)$ and $\Gamma^{\text{er}}_t(x) = \rho^{\text{er}}_t(x)/\Phi^{\text{er}}_t(x)$, we see that both have to follow the same equation

$$\lambda^{\text{er}}\psi^{\text{er}} = -\frac{\mu\sigma^4}{2}\Delta\psi^{\text{er}} - \tilde{V}[m^{\text{er}}](\mathbf{x})\psi^{\text{er}}\;, \tag{64}$$

with either $\psi^{\text{er}}_t(x) = \Phi^{\text{er}}$ or $\psi^{\text{er}}_t(x) = \Gamma^{\text{er}}_t(x)$. Note that if $\psi^{\text{er}}(\mathbf{x})$ is a solution of Eq. (64), then the two time dependent functions

$$\Gamma_t(x) = \exp\bigl\{-\frac{1}{\hbar_{\text{eff}}}\lambda t\bigr\}\psi^{\text{er}}(x)$$

$$\phi_t(x) = \exp\bigl\{+\frac{1}{\hbar_{\text{eff}}}\lambda t\bigr\}\psi^{\text{er}}(x)\;,$$

(with again $\hbar_{\text{eff}} \equiv m\sigma^2$), are solutions of, respectively Eqs. (59)-(60), in the very same way as

$$\psi_t(x) = \exp\bigl\{-\frac{i}{\hbar}\lambda t\bigr\}\psi^{\text{er}}(x)$$

$$\psi^*_t(x) = \exp\bigl\{+\frac{i}{\hbar}\lambda t\bigr\}\psi^{\text{er}}(x)$$

is solution of both Eq. (62) and Eq. (64).

The non linear Schrödinger equation has been used for decades to describe systems of interacting bosons in the mean field approximation (see e.g. [6, 7, 4, 9, 10]), and for almost a century the context of fluid mechanics ([5]). Physicists therefore know a lot about this equation and its solutions, and some of this knowledge can be transferred to the context of (quadratic) mean field games.

2. Tools of quantum mechanics ($\hbar_{\text{eff}} \equiv m\sigma^2$)

2.1. Heisenberg picture and Ehrenfest's relations

Quantum mechanics 101 In the introductory class of quantum mechanics, the "system" under consideration is described by a wavefunction $\psi_t(x)$ which evolves according to the Schrödinger equation

$$i\hbar\,\partial_t\psi_t = \hat{H}\psi_t \tag{65}$$

$$-i\hbar\,\partial_t\psi^*_t = \hat{H}\psi^*_t\;. \tag{66}$$

with

$$
\begin{aligned}
\hat{H} &\equiv \frac{\hat{p}^2}{2m} + V(\hat{x}) \\
\hat{p}\,(\cdot) &\equiv -i\hbar\partial_x\,(\cdot) \\
\hat{x}\,(\cdot) &\equiv x \times (\cdot)\ .
\end{aligned}
$$

An "observable" is typically an operator $O(\hat{p},\hat{x})$ constructed as a function of $\hat{p}$ and $\hat{x}$. It's expectation value is defined as

$$O(t) = \langle \psi_t | \hat{O} | \psi_t \rangle = \int dx\, \psi_t^*(x) \hat{O} \psi_t(x)$$

and is therefore such that it's time evolution is given by $i\hbar\partial_t O = \langle -\hat{H}\psi_t|\hat{O}|\psi_t\rangle + \langle\psi_t|\hat{O}\hat{H}|\psi_t\rangle = \langle\psi_t|[\hat{O},\hat{H}]|\psi_t\rangle$, that is

$$\frac{d}{dt}O = \frac{i}{\hbar}\langle\psi_t|[\hat{H},\hat{O}]|\psi_t\rangle\ .$$

If for instance we choose $\hat{O} \equiv \hat{x}$, we have

$$[\hat{H},\hat{x}] = [\frac{\hat{p}^2}{2m},\hat{x}]\frac{1}{2m} = (\hat{p}\ \underbrace{[\hat{p},\hat{x}]}_{-i\hbar} + \underbrace{[\hat{p},\hat{x}]}_{-i\hbar}\ \hat{p}) = -\frac{i\hbar}{m}\hat{p}$$

so that

$$\frac{d}{dt}\langle\hat{x}\rangle = \frac{\langle\hat{p}\rangle}{m}\ . \tag{67}$$

In the same way, with $\hat{O} \equiv \hat{p}$, we have $[\hat{x}^n,\hat{p}] = i\hbar\, n\, \hat{x}^{n-1}$ (e.g. by recurrence : $[\hat{x}^n,\hat{p}] = (\hat{x}^{n-1}[\hat{x},\hat{p}] + [\hat{x}^{n-1},\hat{p}]\,\hat{x}) = i\hbar x^{n-1} + i\hbar(n-1)\hat{x}^{n-2}\hat{x})$, and therefore

$$\frac{d}{dt}\langle\hat{p}\rangle = \langle -\partial_x V(\hat{x})\rangle\ . \tag{68}$$

Eqs. (67)-(68) are the Heisenberg equations, which state that if a wavefunction is localized enough on the classical scale, it behaves essentially as a classical particle.

Mean Field Games Because of the similarities between Eqs. (60)-(59) and Eqs. (62)-(63), we can follow the same philosophy as for standard quantum mechanics and introduce an operator formalism on some appropriate functional space. We thus define the operator position and momentum as

$$
\begin{aligned}
\hat{X}\,(\cdot) &\equiv x \times (\cdot)\ , \\
\hat{\Pi}\,(\cdot) &\equiv -\hbar_{\text{eff}}\,\partial_x\,(\cdot) \\
[\hat{X},\hat{\Pi}] &= +\hbar_{\text{eff}}\ ,
\end{aligned}
$$

where again $\hbar_{\text{eff}} \equiv m\sigma^2$. The mean filed game equations Eqs. (60)-(59) can then be expressed as

$$+\hbar_{\text{eff}}\,\partial_t\Gamma_t = \hat{H}\,\Gamma_t \tag{69}$$

$$-\hbar_{\text{eff}}\,\partial_t\Phi_t = \hat{H}\,\Phi_t\ , \tag{70}$$

with[2]

$$\hat{H} = \frac{\hat{\Pi}^2}{2m} + \tilde{V}[\rho](\hat{X})\ .$$

[2]Note that he sign convention for $\hat{H}$ is not the same as in [12]

For an arbitrary operator $\hat{O}$ defined in terms of $\hat{X}$ and $\hat{\Pi}$, we define its average

$$\langle \hat{O} \rangle(t) \equiv \langle \Phi_t | \hat{O} | \Gamma_t \rangle = \int dx\, \Phi_t(x)\; \hat{O}\; \Gamma_t(x)\;, \tag{71}$$

where the couple (Φ_t, Γ_t) defines the state of the system and evolves according to Eqs (69)-(70). Note that, as $\Gamma_t(x)\Phi_t(x) \equiv \rho_t(x)$, when $\hat{O}$ depends only on the position: $\hat{O} = O(\hat{X})$, the latter average reduces to the usual mean value with respect to the density:

$$\langle \hat{O} \rangle(t) = \int dx\, \rho_t(x) O(x)\;. \tag{72}$$

Following the same logic as in the quantum case, we can express the time evolution of the mean value of an observable in terms of a commutator:

$$\frac{d}{dt}\langle \hat{O} \rangle = \langle \frac{\partial \hat{O}}{\partial t} \rangle - \frac{1}{\hbar_{\text{eff}}} \langle [\hat{H}, \hat{O}] \rangle\;, \tag{73}$$

and in particular, we get

$$\frac{d}{dt}\langle \hat{X} \rangle = \frac{\langle \hat{\Pi} \rangle}{m}\;, \tag{74}$$

$$\frac{d}{dt}\langle \hat{\Pi} \rangle = \langle \hat{F}[\rho_t] \rangle\;, \tag{75}$$

where $\hat{F}[m_t] \equiv -\partial_x \tilde{V}[\rho_t](\hat{X})$ is named by analogy the "force" operator.

Furthermore, when local interactions are assumed,

$$\tilde{V}[\rho_t](x) = V_0(x) + f[\rho(x)]\;, \tag{76}$$

then

$$\mathbb{E}[F_{\text{int}}] \equiv \int \rho(x)(-\partial_x(f[\rho(x)]))dx = -\int \rho f'(\rho)\rho' dx = F(-\infty) - F(+\infty) = 0\;,$$

where $F(z) = \int z f'(z) dz$. Therefore the mean force depends only on the external potential

$$\langle \hat{F} \rangle = \langle \hat{F}_0 \rangle \qquad \text{where} \qquad \hat{F}_0 \equiv -\partial_x V_0(\hat{X}) \tag{77}$$

and Eq. (75) can be simplified accordingly.

2.2. Conservation laws

For local interactions of the form (76) we can introduce an action functional

$$S[\Phi,\Gamma] \equiv \int_0^T dt \int_{\mathbb{R}} dx \left[-\frac{\hbar_{\text{eff}}}{2}\big(\Phi(\partial_t \Gamma) - (\partial_t \Phi)\Gamma\big) - \frac{\hbar_{\text{eff}}^2}{2m}(\partial_x \Phi)(\partial_x \Gamma) + \Phi\, V_0(\mathbf{x})\, \Gamma + F[\Phi\, \Gamma] \right]\;, \tag{78}$$

where $F(m) = \int^m f(m')dm'$, which extremals are solutions of the system (59, 60). Indeed, the variational equation $\delta S/\delta \Gamma = 0$ (respectively $\delta S/\delta \Phi = 0$) is equivalent to Eq. (59) (respectively Eq. (60)) with $\tilde{V}[\rho]$ in the form Eq. (76). This property will provide us with a variational approximation scheme for the solutions of the MFG system. Finally, for a pair (Φ,Γ) solving the MFG equations, the action Eq. (78) can be rewritten as

$$S[\Phi,\Gamma] = -\int dt\, dx \left[\frac{\hbar_{\text{eff}}}{2}\big(\Phi(\partial\Gamma) - (\partial\Phi)\Gamma\big) \right] \;+\; \int dt\, E_{\text{tot}}(t)$$

where we have introduced the "total energy",

$$E_{\text{tot}}(t) \equiv E_{\text{kin}} + E_{\text{pot}} + E_{\text{int}} \tag{79}$$

with "kinetic", "potential", and "interaction" energies defined respectively as

$$E_{\text{kin}} = \frac{1}{2m}\langle\hat{\Pi}^2\rangle, \tag{80}$$

$$E_{\text{pot}} = \langle V_0(\hat{X})\rangle, \tag{81}$$

$$E_{\text{int}} = \int dx F[\rho(x)]. \tag{82}$$

The integrand in the action functional (78) does not depend explicitly on time, so that by Noether theorem, there is a conserved quantity along the trajectories. This conserved Noether charge is (due to the sign conventions here) minus the previously defined total energy, so that

$$\frac{dE_{\text{tot}}(t)}{dt} = 0 \tag{83}$$

and for any pair $(\Phi(x), \Gamma(x))$ solving the MFG equations, one has

$$S[\Phi, \Gamma] = -\frac{\hbar_{\text{eff}}}{2}\int dt\, dx\, [\Phi(\partial\Gamma) - (\partial\Phi)\Gamma] \; + \; E_{\text{tot}}T\;. \tag{84}$$

2.3. Solitons

Let us come back to the time independent non-linear Schrödinger equation with $V_0(x) \equiv 0$

$$\frac{\hbar^2}{2m}\partial^2_{xx}\psi + |\psi|^2\psi = \lambda\psi\;.$$

It can be shown relatively easily that the lowest energy solution of this equation takes the form

$$\psi(x) = \frac{1}{\sqrt{2\eta}}\frac{1}{\cosh\left(\frac{x}{\eta}\right)}$$

with $\eta = 2\hbar^2/mg$ and $\lambda = g/4\eta$. We observe that quite intuitively, $\eta \to 0$ as $g \to \infty$. The size of the soliton is given by a balance between the attractive interaction energy and the quantum effects which implies that if the wave function is too localized in position, it has to have a big extension in momentum, and thus a large kinetic energy. In the MFG context the role of this "quantum pressure" is played by the noise, which prevents the density of agents to collapse completely.

By analogy, we can expect that in the limit of large g (which we can define as the fact that the variations of the external potential around any point, $\delta U_0 = \partial_x V_0 \delta x + (\partial^2_{xx}V_0)\delta x^2 + \cdots$, are dominated by the first term $\partial V_0 \delta x$ for a displacement of order $|\delta x| \sim \eta$), the MFG problem will be characterized by a density of agent localized on a distance of scale η.

3. Generic scenario for the fish-school problem

Going back to our fish-school problem, we want therefore to solve the system of coupled equations (45)-(46) describing respectively the dynamics of the fish-school and the optimization of individual fishes, with some initial fish density $\rho_{t=0}(\cdot) = \rho_0(\cdot)$ and terminal cost for the location at the end of the day $C_T(\cdot)$. We furthermore assume that the first concern of the fished it to avoid being eaten by predators, so that g is positive and large (in the precise sense defined above).

3.1. Dynamical equations

In this context, we expect the following generic scenario for full duration of the game

- **Near** $t = 0$**:** Formation of the soliton (a very short time process).

- **Second stage:** Propagation of the soliton (for most of the interval $[0,T]$).
- **Near $t = T$:** Destruction of the soliton (again a very short time process).

Let us focus on the second stage, where a soliton of size $\eta \sim \hbar_{\text{eff}}^2/mg = m\sigma^4/g$ is formed and not yet destroyed. The assumption that g is large, and more precisely that $\eta|\partial_{xx}^2 V_0| \ll \partial_x V_0$ so that implies that $\langle \partial_x V_0(\hat{X})\rangle = \partial_x V_0(\langle \hat{X}\rangle)$. Noting $X = \langle \hat{X}\rangle$ the mean position of the fish-school and $P = \langle \hat{\Pi}\rangle$ its mean momentum, the system Eqs. (74)-(75) thus reads

$$\frac{d}{dt}X = \frac{P}{m} \; , \tag{85}$$

$$\frac{d}{dt}P = -\partial_x V_0(X) \; , \tag{86}$$

in other words, the centre of mass of the fish-school just follow classical dynamics of a point particle in the potential $V_0(x)$.

The initial position of the soliton is given by

$$X_0 = \langle \hat{X}\rangle_{\rho_0} = \int_{\mathbb{R}} x\rho_0(x)dx \; , \tag{87}$$

and its final momentum by

$$\begin{aligned} P_T &\equiv \int_{\mathbb{R}} \Phi_T(x)\,\hat{\Pi}\,\Gamma_T(x)\,dx \\ &= -\int_{\mathbb{R}} \Gamma_T(x)\,\hat{\Pi}\,\Phi_T(x)\,dx \\ &= +\int_{\mathbb{R}} (\Gamma_T(x)\Phi_T(x))\,\frac{\hbar_{\text{eff}}\partial_x\Phi_T(x)}{\Phi_T(x)}\,dx \\ &= \int_{\mathbb{R}} \rho_T(x)\partial_x(-C_T(x))\,dx = -\langle \partial_x C_T(\hat{X})\rangle \; , \end{aligned}$$

where we have used that the terminal condition Eq. (53) reads $\hbar_{\text{eff}} \log \Phi_T(\cdot) = -u_T(\cdot) = -C_T(\cdot)$. For large g, where we assume that $C_T(x)$ do not vary much on the scale η, this reads simply

$$-P_T = \partial_x C_T(X_T) \; . \tag{88}$$

An illustration of the corresponding dynamics is given in Fig. 9.

3.2. Large T limit

The mixed initial/terminal boundary conditions Eqs. (87)-(88) require to find a trajectory starting from the vertical line $X = X_0$ and ending on the curve $P_T = -\partial_x C_T(X_T)$. In the large T limit, most of the trajectories simply wander off to infinity, and the only way to satisfy this constraint as $T \to \infty$ is that to spend most of the time at, or very near, an unstable fixed point (X^*, P^*) ($= (0,0)$ for the choice of parameters in Fig. 9).

Let $f_t : (X_0, P_0) \mapsto (X_t, P_t)$ the Hamiltonian flow corresponding to the dynamics Eqs. (85)-. (86), and

$$\begin{aligned} W^s &\equiv \{(X,P)/ \lim_{t\to+\infty} f_t(X,P) = (X^*,P^*)\} \\ W^u &\equiv \{(X,P)/ \lim_{t\to-\infty} f_t(X,P) = (X^*,P^*)\} \; , \end{aligned}$$

the stable and unstable manifold associated with the fixed point (X^*, P^*), we see that in the large T limit, the generic scenario described at the beginning of this section can be made more precise:

- **First stage:** Form the soliton (a very short time process close to $t = 0$).

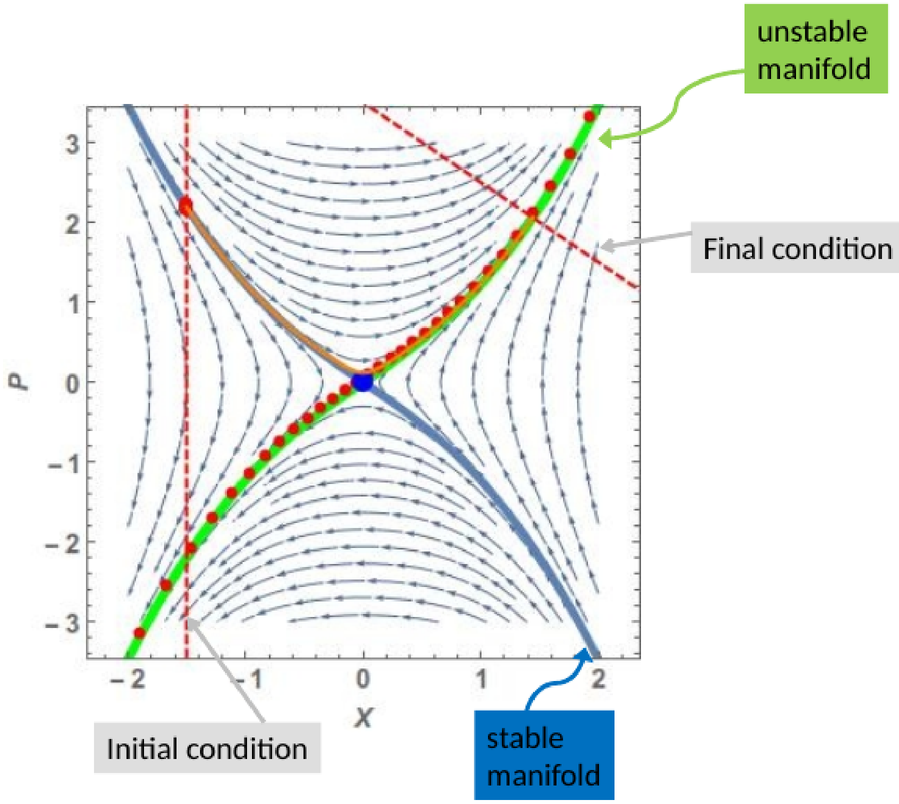


Figure 9: Illustration of the dynamics of the school-fish soliton, for the various parameters of the problem taken as $V_0(x) = -\frac{1}{4}x^4 - \frac{1}{2}x^2$, $C_T(x) = -\frac{7}{2}x + \frac{1}{2}x^2$ (so that $P_T = \frac{7}{2} - X_T$) and $X_0 = -1.5$.

- **Second stage:** Follow the *stable manifold* W^s until the fixed point (X^*, P^*) is reached.
- **Third stage:** Spend most of the time in $[0, T]$ waiting in the close neighbourhood of (X^*, P^*).
- **Fourth stage:** Leave the stable fixed point following the *unstable manifold* W^u, just in time to reach the curve $P_t = -\partial_x C_T(X_T)$ at time T.
- **Fifth stage** Destroy the soliton (again a very short time process).

The permanent regime of section 4. is thus seen here to correspond to a soliton waiting on an unstable fixed point, and can be understood in a more general context as a high dimensional generalisation of the concept of unstable fixed point.

References

[1] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta. Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.*, 51(5):3558–3591, 2013.

[2] O. Guéant. Mean field games equations with quadratic hamiltonian: a specific approach. *Math. Models Methods Appl. Sci.*, 22:1250022, 2012.

[3] E. Hopf. The partial differential equation $u_t + uu_x = u_{xx}$. *Comm. Pure Appl. Math.*, 3:201–230, 1950.

[4] D. J. Kaup. Perturbation theory for solitons in optical fibers. *Phys. Rev. A*, 42:5689–5694, 1990.

[5] C. Kharif, E. Pelinovsky, and A. Slunyaev. *Rogue Waves in the Ocean.* Advances in Geophysical and Environmental Mechanics and Mathematics. Springer, 2008.

[6] Y. S. Kivshar and B. A. Malomed. Dynamics of solitons in nearly integrable systems. *Rev. Mod. Phys.*, 61:763–915, 1989.

[7] A. M. Kosevich. Particle and wave properties of solitons. *Physica D*, 41(2):253 – 261, 1990.

[8] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II – Horizon fini et contrôle optimal. *C. R. Acad. Sci. Paris, Ser. I*, 343(10):679 – 684, 2006.

[9] V. M. Pérez-García, H. Michinel, J. I. Cirac, M. Lewenstein, and P. Zoller. Dynamics of Bose-Einstein condensates: Variational solutions of the Gross–Pitaevskii equations. *Phys. Rev. A*, 56:1424–1432, 1997.

[10] L. Pitaevskii and S. Stringari. *Bose-Einstein Condensation.* Clarendon Press, Oxford, 2003.

[11] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications.* Springer Series in Synergetics. Springer, Heidelberg, 1996.

[12] Denis Ullmo, Igor Swiecicki, and Thierry Gobron. Quadratic mean field games. *Phys. Rep.*, 799:1–35, 2019.