PROJET DE RECHERCHE (PRE)

There is no overlap gap for the Sherrington Kirkpatrick model at low temperature

Réalisé par:
Victor PRISER

Tuteur de stage:
Pascal Maillard

Enseignante référente:
Laure Giovangigli

Stage effectué du 17/05/2021 au 20/08/202

Organisme d'accueil: Institut de Mathématique de Toulouse

Adresse: 1 R.3, Université Paul Sabatier, 118 Rte de Narbonne, 31400 Toulouse

2ème Année: Majeure mathématiques appliquées (promo 2022)

Non confidentiel

ENSTA
IP PARIS

Attestation de confidentialité / non confidentialité

A remplir par l'entreprise d'accueil et à remettre à la bibliothèque de l'ENSTA Paris[1]

Rapport de Projet de Recherche (PRe)

Par la présente, je soussigné (e) Madame / Monsieur[2] NOM, Prénom
MAILLARD, Pascal

Employé(e) en tant que (qualité)
professeur

dans la société (NOM et ADRESSE de l'établissement) :
Université Toulouse 3 Paul Sabatier

Atteste sur l'honneur que LES DONNÉES contenues dans le rapport de
Madame / Monsieur[1] NOM Prénom PRISER Victor **SONT:**

☐ **CONFIDENTIELLES**

Par conséquent, **l'entreprise** (ou organisme) d'accueil **S'ENGAGE A CONSERVER** le rapport de stage **pendant la Durée d'Utilité Administrative (DUA) soit 5 ans.**

A l'issue de la DUA, l'entreprise s'engage (cocher le choix correspondant) :

☐ A conserver le document

☐ A détruire le document

☐ **NON CONFIDENTIELLES** (**cocher** le choix correspondant)

☒ L'entreprise autorise une **mise en ligne du rapport de stage** (connexion via l'annuaire LDAP de l'école[3]). L'étudiant doit déposer son rapport sur BibNum - **https://bibnum.ensta.fr - (archive institutionnelle de l'Ecole).**

☐ L'entreprise autorise un accès restreint au document (uniquement sur place et au format électronique). L'étudiant doit déposer son rapport sur BibNum - **https://bibnum.ensta.fr - (archive institutionnelle de l'Ecole).**

Un rapport consultable uniquement sur place est stocké sur un PC non connecté au réseau et à internet. Les copies électroniques et imprimées ne sont pas autorisées.

À Toulouse

Le 23 / 08 / 2021

Signature et cachet P. Maillard

---

[1] Possibilité de transmettre cette attestation par mail - à documentation@ensta-paris.fr – si le document est confidentiel. Si le rapport est non confidentiel, elle devra être déposée en même temps que le rapport sur https://bibnum.ensta.fr

[2] Rayer la mention inutile.

[3] La communauté ENSTA Paris, seule, aura accès au texte intégral. Les métadonnées (titre, auteur etc.) du rapport, en revanche, seront accessibles en ligne par tous.

Abstract

The purpose of this internship is to show that there is no overlap gap in the Parisi measure. First, we will describe the SK model so, we can understand the stakes of the problem (section 2). Then we will present a new variational representation of the Parisi functional (section 3). And we will find some conditions that the Parisi measure must respect (section 4). We will also present the link between the new representation and the other one (section 5). In this report, we will present a few of the unsuccessful paths we have explored (section 7) . At the end, you'll see that we have brought a new representation which is strongly linked to the usual one. We haven't managed to solve the problem (section 6). But we hope that we have found useful tools that will help other to find out a solution to the problem.

Keyworld: spin glasses, overlap gap, Parisi measure, Parisi functional, Sherrington and Kirkpatrick, zero temperature, variational representation, control optimal

Contents

1 Introduction **5**

2 The problem **6**
2.1 The Sherrington and Kirkpatrick model 6
2.2 The Parisi functional 7
2.2.1 The recursive definition 7
2.2.2 The Parisi functional defined by a PDE 7
2.3 The Parisi measure 8

3 The start **9**
3.1 Starting idea 9
3.2 The new formulation 9
3.2.1 Lemma 3.2 10
3.3 Fist approach of the problem 11

4 Contribution to the research in this field **13**
4.1 Condition on Θ 13
4.2 Condition on μ 15
4.3 A new PDE 18
4.3.1 A formal way to find the PDE 18
4.3.2 A more rigorous way to see the PDE 19
4.3.3 Different uses of the result 21

5 Parallels between the two representations **22**
5.1 An identification of f 22
5.2 Identification of $\Theta_{\mu\, t}$ 23
5.3 A new formulation of a variation principle 24

6 How to find the solution of our problem **25**
6.1 0 is in the support of μ_P 25
6.2 We are Replica Symmetric when $\beta \leq 1$ 26
6.3 $1 - q_*(\beta)$ and $\mu_P[0,t]$ is like $\mathcal{O}(\frac{1}{\beta})$. 26

7 Some unsuccessful ideas **28**
7.1 Model at zero temperature 28
7.2 Ideas on the PDE in v_μ 30
7.2.1 An explicit solution with the power series 30
7.2.2 Perturbation of the PDE 30

8 Planning of the internship **32**

9 Conclusion **33**

References **34**

1 Introduction

Consider a ferromagnetic solid, then each atom will have a magnetic spin oriented in the same direction and this induce a magnetic field in this same direction. Now, consider a anti ferromagnetic solid, then each atom will have a magnetic spin oriented in purely random directions which will induce no magnetic field at all.

But in reality, there exists no such thing as a "pure" ferromagnetic solid. The solid will have some impurities that will contain magnetic spins not oriented in the same direction. Also there exists some alloys of ferromagnetic and anti ferromagnetic metals (for example gold/iron, copper/magnesium). We call the magnetic state of these solids: a spin glass which refers to "disoriented" magnetic state where magnetic spins are randomly oriented or with no regular patterns.

There are different models of spin glass: the Edwards–Anderson (EA) model, the Sherrington and Kirkpatrick (SK) model, the p-spins model ... But we will only consider the SK model. An important question is to find the ground state of the SK model. Andrea Montanari found in 2019 an algorithm that computes the ground state in a $\mathcal{O}(n^2)$ complexity. But this algorithm is based on an assumption which exists in the literature since about 1980.

Our goal is to find a proof of this assumption.

The results bringing something new to the research in this field will be displayed under the form of lemma and theorem. The other considerations are already known or just give us a better understand ing of the problem (there are sometime not very rigorous).

2 The problem

2.1 The Sherrington and Kirkpatrick model

The SK model was introduced in 1975 by Sherrington and Kirkpatrick [10]. The aim was to simplify the EA model by making all spins interact. Consider the vector $\sigma = (\sigma_1, ..., \sigma_N) \in \Sigma_N = \{-1, 1\}^N$. We define the Hamiltonian as:

$$H_N(\sigma) = \frac{\beta}{\sqrt{2N}} \sum_{1 \leq i,j \leq N} g_{ij}\sigma_i\sigma_j$$

The $\sqrt{2N}$ is a normalisation term. g_{ij} represent the interaction of particles i and j that are i.i.d standard Gaussian random variables. And β is the inverse temperature ($\frac{1}{k_b T}$).

Remark. This Hamiltonian is the opposite of the one we meet in physics (energy). It is a convention that mathematicians often use when they study physics. So, they want to maximize the Hamiltonian while the physicists want to minimize it.

The covariance of the Hamiltonian is a function of the overlap $R_{1,2} = \frac{(\sigma^1, \sigma^2)}{N}$:

$$\mathbb{E}\left[H_N(\sigma^1)H_N(\sigma^2)\right] = N\xi(R_{1,2})$$

For the SK model model, we have $\xi(t) = \beta^2\frac{t^2}{2}$ but we have more generally (for the p-spins model):

$$\xi(t) = \sum_{p \geq 2} \beta_p^2 t^p$$

Let's introduce the partition function:

$$Z_N(\beta) = \sum_{\sigma \in \Sigma_N} \mathrm{e}^{H_N(\sigma)}$$

So, we can define the Gibbs measure on the space Σ_N:

$$G_N(\sigma) = \frac{\mathrm{e}^{H_N(\sigma)}}{Z_N(\beta)}$$

In statistical physics, $G_N(\sigma)$ is the chance to observe the system at the configuration σ. So, we can see that the ground state (state which is observable with most probability) will be the σ that maximizes $\sigma \mapsto H_N(\sigma)$.

To obtain this maximum, it is usual in statistical physics to use the free energy:

$$F_N(\beta) = \frac{1}{N}\log \mathbb{E}\left[Z_N(\beta)\right]$$

And then:

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\max_{\sigma\in\Sigma_N} H_N(\sigma)\right] = \lim_{\beta\to\infty}\lim_{N\to\infty} \frac{F_N(\beta)}{\beta}$$

So, the issue to find the ground state becomes an issue to compute the free energy at all temperatures. A valid computation for high temperatures was first discovered by Sherrington and Kirkpatrick in 1975 [11]. And several years later (1979) Giorgo Parisi proposed a valid formulation for all temperatures. This prediction was confirmed in 2005 by Michel Talagrand [13]:

$$\lim_{N\to\infty} F_N(\beta) = \inf_{\mu\in\mathcal{M}_d} \mathcal{P}(\mu) = \min_{\mu\in\mathcal{M}} \mathcal{P}(\mu)$$

Where $\mathcal{M}_d$ corresponds to the space of probability measures on $[0, 1]$ with a finite number of atoms, $\mathcal{M}$ the space of probability measures on $[0, 1]$. And $\mathcal{P}$ is the Parisi functional that we will define in the next section.

2.2 The Parisi functional

There exists two definitions of the Parisi functional $\mathcal{P}$. One is recursive and the other uses a PDE. The link between those two is well shown in the section 2 of [1].

2.2.1 The recursive definition

Each $\mu \in \mathcal{M}_d$, corresponds to a unique triplet $(k, \mathbf{m}, \mathbf{q})$. Where $\mathbf{m} = (m_i)_{0\leq i\leq k+1}$ and $\mathbf{q} = (q_i)_{0\leq i\leq k+2}$ verifies:

$$\mu = \sum_{i=0}^{k}(m_{i+1} - m_i)\delta_{q_{i+1}}$$

And:

$$0 = m_0 \leq m_1 \leq \cdots \leq m_{k+1} = 1$$

$$0 = q_0 \leq q_1 \leq \cdots \leq q_{k+2} = 1$$

Consider now independent centered Gaussian random variables $(z_j)_{0\leq j\leq k+1}$ with variance $\mathbb{E}\left[z_j^2\right] = \beta^2(q_{j+1} - q_j)$.

We now define recursively $(X_n)_{0\leq n\leq k+2}$ as follows:

$$X_{k+2} = \log\cosh\left(\sum_{j=0}^{k+1} z_j\right)$$

And:

$$X_p = \frac{1}{m_p}\log\mathbb{E}[\exp(m_p X_{p+1})|z_j, j < p]$$

Now we can define the Parisi functional on $\mathcal{M}_d$:

$$\boxed{\mathcal{P}(\mu) = X_0 - \frac{1}{2}\int_0^1 t\beta^2\mu[0,t]\,\mathrm{d}t}$$

But the Parisi functional can be extended continuously on the space of all probability measures on $[0, 1]$: $\mathcal{M}$ following the metric $d(\mu, \mu') = \int_0^1 |\mu'[0,u] - \mu[0,u]|\,\mathrm{d}u$.

2.2.2 The Parisi functional defined by a PDE

Let's define the Parisi PDE for $\mu \in \mathcal{M}$,

$$\begin{cases}\partial_t u_\mu(t,x) + \frac{\beta^2}{2}\left(\partial_{xx}u_\mu(t,x) + \mu[0,t](\partial_x u_\mu(t,x))^2\right) = 0, & (t,x) \in (0,1)\times\mathbb{R} \\ u_\mu(1,x) = \log\cosh(x) \end{cases}.$$

It is proven in [6] that the Parisi PDE admits a unique week solution for all μ which is continuously differentiable in time at continuous point of μ and smooth in space.

It follows the definition of the Parisi functional for $\mu \in \mathcal{M}$:

$$\boxed{\mathcal{P}(\mu) = u_\mu(0,0) - \frac{1}{2}\int_0^1 t\beta^2\mu[0,t]\,\mathrm{d}t}$$

2.3 The Parisi measure

It can be shown that the Parisi functional is convex and smooth [2]. And Helly's selection theorem gives us the sequential compacity of $\mathcal{M}$ we need to say that the Parisi functional has a unique minimizer $\mu_P \in \mathcal{M}$ which is called the Parisi measure.

Finally, the ground state can be fond by computing the Parisi measure.

We expect some comportements of the Parisi measure:

1. 0 is contained in the support of μ_P for all beta (demonstrated in [1])

2. Let be $q_*(\beta)$ the last point of the support of μ, then μ admits a point of discontinuity in $q_*(\beta)$ for all β (demonstrated in [1] for some β).

3. μ is expected to be continuous on its support and it is an interval $[0, q_*(\beta)]$ for $\beta > \beta_l$.

The third assumption is the one we are trying to prove. For now, the best we have is that the support of μ has an infinite number of points at zero temperature ($\beta = \infty$) and this was proven in 2017 [3].

And Montanary found in 2019 [9] an algorithm that computes the Parisi measure (so, the ground state energy) only if the third assumption is valid.

We already know that μ_P is continuous on its support by [1] (theorem 2). So, the problem is:

Show that support of μ_P is an interval $[0, q_*(\beta)]$ for $\beta > \beta_l$.

3 The start

The objectif of this section is to give a starting point to the problem and to explain how we manage to find it.

3.1 Starting idea

As we have said before, this problem has been raised around 1980. So, it seams to be a hard problem. But what make it easier now? In 2015, using optimal control, a new formulation of the problem has been found [7]. And my internship supervisor, had an idee to pursue the work in this article. I will describe it here:

In [7], the formulation was (the principle is to apply the Legendre transform on the non-linearity in Parisi PDE: $\partial_u \mu^2$. Then we have a HJB equation that admits a variation representation according to the control optimal theory):

$$\mathcal{P}(\mu) = \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}\left[\log\cosh\left(\int_0^1 \beta^2 \mu[0,t]\alpha_t \,\mathrm{d}t + \int_0^1 \beta \,\mathrm{d}W_t\right) - \beta^2 \frac{1}{2}\int_0^1 \mu[0,t](\alpha_t^2 + t)\,\mathrm{d}t\right] \quad (1)$$

Where $\mathcal{A}_0$ represent the space of progressively adapted processes to $(\mathcal{F}_t)$, the canonical filtration of (W_t), a standard Brownian motion. We see that $\log\cosh$ is convex. So, the idea is to apply the legend transform to $\log\cosh$ and a new formulation appear:

$$\mathcal{P}(\mu) = \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]$$

Where H is the legendre transform of $\log\cosh$. The notation X_t represent $X_t = \mathbb{E}[X|\mathcal{F}_t]$ and $\mathcal{V}$ is a subset of $L^2(\Omega, \mathcal{F}_1, \mathbb{P})$ where the random variables take value in $[-1,1]$. The details will be described in the section bellow.

3.2 The new formulation

Definition 3.1 (Notation)**.** We define:

1. $\mathcal{V} \subset L^2(\Omega, \mathcal{F}_1, \mathbb{P})$ where the random variables take value in $[-1,1]$.
2. $\mathcal{M}$ the space of probability on $[0,1]$.
3. $\mathcal{A}_t$ the space of progressively adapted processes to $(\mathcal{F}_s)_{t \le s \le 1}$,
4. For $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $X_t = \mathbb{E}[X|\mathcal{F}_t]$.
5. For $x \in]-1,1[$, $H(x) = \frac{1}{2}(\log(1+x)(1+x) + \log(1-x)(1-x))$. And $H(1) = H(-1) = \log(2)$.

We want to prove the following theorem:

Theorem 3.1 (Another repésentation)**.** *The Parisi functional can also be defined by,*

$$\mathcal{P}(\mu) = \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]$$

For the proof we will need the help of the the following lemma:

Lemma 3.2 (Legendre transformation of $\log\cosh$)**.** *For all $x \in \mathbb{R}$,*

$$\log\cosh(x) = \sup_{a \in [-1,1]} (ax - H(a))$$

Proof theorem 3.1. With the lemma 3.2 and the result 1 from [7] , we get the first line:

$$\mathcal{P}(\mu) = \sup_{\alpha\in\mathcal{A}_0} \mathbb{E}\left[\sup_{\theta\in[-1,1]} \theta\left(\int_0^1 \beta^2\mu[0,t]\alpha_t\,\mathrm{d}t + \beta W_1\right) - H(\theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t](\alpha_t^2+t)\,\mathrm{d}t\right]$$

We take Θ such that for each ω, $\Theta(\omega) = \arg\max_{\theta\in[-1,1]} \theta\left(\int_0^1 \beta^2\mu[0,t]\alpha_t\,\mathrm{d}t + \beta W_1\right) - H(\theta)$ so, Θ will depend on (α_t) and will take value in $[-1,1]$ that's why we search a sup on $\mathcal{V}$. So, we have the second line:

$$\mathcal{P}(\mu) = \sup_{\alpha\in\mathcal{A}_0} \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\Theta\left(\int_0^1 \beta^2\mu[0,t]\alpha_t\,\mathrm{d}t + \beta W_1\right) - H(\Theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t](\alpha_t^2+t)\,\mathrm{d}t\right]$$

Then by using $\mathbb{E}\left[\Theta_t\alpha_t\right] = \mathbb{E}\left[\mathbb{E}[\Theta\alpha_t|\mathcal{F}_t]\right] = \mathbb{E}\left[\Theta\alpha_t\right]$ we get the third line.

$$\begin{aligned}\mathcal{P}(\mu) &= \sup_{\alpha\in\mathcal{A}_0} \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\int_0^1 \Theta_t\beta^2\mu[0,t]\alpha_t\,\mathrm{d}t + \Theta\beta W_t - H(\Theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t](\alpha_t^2+t)\,\mathrm{d}t\right]\\ &= \sup_{\alpha\in\mathcal{A}_0} \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\int_0^1 \beta^2\mu[0,t](\Theta_t\alpha_t - \frac{1}{2}\alpha_t^2)\,\mathrm{d}t + \Theta\beta\,\mathrm{d}W_1 - H(\Theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t]t\,\mathrm{d}t\right]\end{aligned}$$

Then we switch the two sup and then we can identify the legendre transform of $\frac{\Theta_t^2}{2}$. And we have the fifth line:

$$\begin{aligned}\mathcal{P}(\mu) &= \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\int_0^1 \beta^2\mu[0,t]\frac{\Theta_t^2}{2}\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t]t\,\mathrm{d}t\right]\\ &= \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2\mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]\end{aligned}$$

□

3.2.1 Lemma 3.2

Definition 3.2 (Legendre transformation of a convex a function)**.** Let $I \subset \mathbb{R}$ be an interval and $f: I \to \mathbb{R}$ a convex function. Then, we denote by the Legendre transform of f the function $f^*: I^* \to \mathbb{R}$ defined by:

$$\forall x^* \in I^* \quad f(x^*) = \sup_{a\in I}(ax^* - f(a))$$

Where $I^* = \{x^* \in \mathbb{R} \quad \sup_{a\in I}(ax^* - f(a)) < \infty\}$

Remark. We can define the Legendre transformation for more general functions but it is of no use here.

We can show that f^* is also a convex function. So, we can define $f^{**} = (f^*)^*$. This lead to the following result:

Let $I \subset \mathbb{R}$ be an interval and $f: I \to \mathbb{R}$ a convex function. Then, $f^{**} = f$.

So, we can now give the proof of lemma 3.2:

lemma 3.2. We see that $\log\cosh$ is smooth (composition of smooth functions) and $\log\cosh''(x) = \frac{1}{\cosh^2(x)} > 0$. So, $\log\cosh$ is convex and according to the result above,

$$\log\cosh(x) = \sup_{a^*\in I^*}(a^*x - \log\cosh^*(a^*))$$

With $I^* = \{x^* \in \mathbb{R} \quad \log\cosh^*(x^*) < \infty\}$. Let fixe $x^* \in \mathbb{R}$ and we define $f(a) = ax^* - \log\cosh(a)$. We find that $f'(a) = x^* - \tanh(a)$ and $f''(a) = -\frac{1}{\cosh^2(a)} < 0$. So, f is concave and there is three cases:

1. ($x^* \in]-1,1[$) We have $f'(\text{arc}\tanh(x^*)) = 0$, then $\text{arc}\tanh(x^*)$ maximize f. Because f is concave. We have, $\text{arctanh}(x^*) = \frac{1}{2}(\log(1+x^*) - \log(1-x^*))$ and $\log\cosh(\text{arctanh}(x^*)) = -\frac{1}{2}(\log(1+x^*) + \log(1-x^*))$. So, finally,

$$\sup_{a\in\mathbb{R}} f(a) = f(\text{arctanh}(x^*)) = \frac{1}{2}(\log(1+x^*)(1+x^*) + \log(1-x^*)(1-x^*))$$

2. ($x^* \geq 1$) We have $f' > 0$, so:

$$\sup_{a\in\mathbb{R}} f(a) = \lim_{x\to+\infty} f(x) = \begin{cases} +\infty & \text{if } x^* > 1 \\ \log(2) & \text{else} \end{cases}$$

3. ($x^* \leq -1$) $f' < 0$, so:

$$\sup_{a\in\mathbb{R}} f(a) = \lim_{x\to\infty} f(x) = \begin{cases} +\infty & \text{if } x^* < -1 \\ \log(2) & \text{else} \end{cases}$$

It follows $I^* = [-1,1]$ and for $H(x^*) = \frac{1}{2}(\log(1+x^*)(1+x^*) - \log(1-x^*)(1-x^*))$, we see that $\lim_{x^*\to\pm 1} H(x^*) = \log(2)$. So, we can define H on $[-1,1]$. □

Then the problem we are trying to solve is:

Show that $\text{supp}\,\mu_P = [0, q_*(\beta)]$ for β big enough where μ_P is defined by:

$$\mu_P = \arg\min_{\mu\in\mathcal{M}} \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2\mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]$$

3.3 Fist approach of the problem

Let's define:

$$\mathcal{J}(\mu,\Theta) = \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2\mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]$$

And we want:

$$\inf_{\mu\in\mathcal{M}} \sup_{\Theta\in\mathcal{V}} \mathcal{J}(\mu,\Theta)$$

We have a max and a min. So, the first idea would be to switch the two operators. There exists a min-max theorem that allow us to do that only if the functional is convex in the variable we want to minimize and concave in the variable we want to maximize. But this is not the case here. Indeed, we are linear in μ (so, convex) but we are not concave in Θ because we have a Θ_t^2 that induces convexity.

Before doing anything else we want to understand what happens in this expression that makes the problem not so, easy as it seams. We see that H is bounded by $\log 2$ while the two other terms depend on β and we want β to be big. So, at first we want to neglect the H. That's the problem at zero temperature ($\beta = \infty$). So, there are two terms that are competing with each other. Indeed, $\text{sgn}(W_1)$ will maximize the term $\beta\Theta W_1$ while $\Theta = 1$ will maximize the other term.

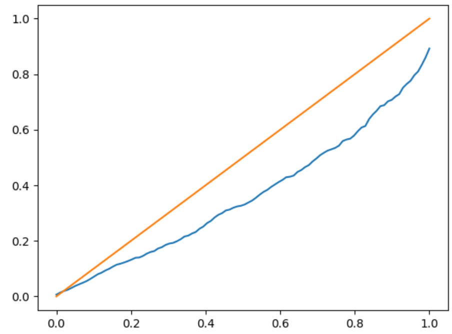

Figure 1: Comparaison between $t \mapsto \Theta_t^2$ (blue) and $t \mapsto t$(red)

The last figure shows that if we take $\Theta = \operatorname{sgn}(W_1)$, we are far from maximizing $\int_0^1 \Theta_t^2 - t \, dt$. We can search for $\Theta = \operatorname{sgn}(W_1 - \alpha)$ or more generally $\Theta = f(W_1)$ but we suspect this is not the case (we found no way to demonstrate this). And worse, even if we find the good Θ, we need a way to verify it. That's the purpose of the next section: finding a condition on the μ and Θ.

4 Contribution to the research in this field

In this section, we present our contributions to the research in this field. We will first present some conditions on μ_P and the Θ optimal. Then we will show that the optimal Θ satisfies an SDE. We were first interested in the problem at zero temperature (without H) and this was harder to find a good condition because we had $\Theta = \pm 1$. Indeed, it was harder to find a good condition by making some variation because the optimal Θ is on the edge of $\mathcal{V}$.

But luckily, with a non zero temperature the H, with infinite derivative in ± 1, allows us to use a variation principle.

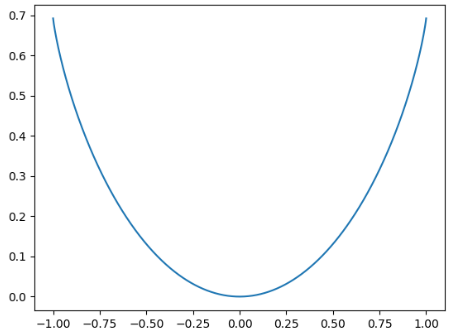

Figure 2: $x \mapsto H(x)$ for $x \in [-1, 1]$

4.1 Condition on Θ

First we need to verify if the optimal Θ is reached. For that we will need a result in [7], the argument used here is a control optimal verification argument (if the sup in the HJB equation is reached then the sup in the variational expression is also reached).

Lemma 4.1 (existence of a unique maximizer Θ_μ)**.** *For $\mu \in \mathcal{M}$, there exists a unique $\Theta_\mu \in \mathcal{V}$ such that:*

$$\sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left(\Theta_t^2 - t\right) \mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right] = \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left(\Theta_{\mu\,t}^2 - t\right) \mathrm{d}t + \Theta_\mu \beta W_1 - H(\Theta_\mu)\right]$$

Proof. We can show ([7] , lemma 4) that the optimal control α_t^* in 1 is reached and is equal to $\partial_x u_\mu(t, X_t)$ (and so, it is unique). With u_μ a weak solution of the Parisi PDE and X solves the SDE:

$$\mathrm{d}X_s = \beta^2 \mu[0,s]\partial_x u_\mu(s, X_s)\,\mathrm{d}s + \beta\,\mathrm{d}W_s$$

With $X_0 = 0$

And so, the optimal Θ_μ would be such that (recall the second line in the proof of the theorem 3.1):

$$\forall \omega \in \Omega \quad \Theta_\mu(\omega) = \mathrm{argsup}_{\theta \in [-1,1]}\ \theta\left(\int_0^1 \beta^2 \mu[0,t]\alpha_t^*(\omega)\,\mathrm{d}t + \beta W_1(\omega)\right) - H(\theta)$$

We remark that $\Theta_\mu(\omega)$ is a maximizer of a continuous, strictly concave ($-H$ is strictly concave) functional over a compact set. So, $\Theta_\mu(\omega)$ is unique and exists. $\square$

Now we can write a lemma that will be useful later:

Lemma 4.2. *For all $\mu \in \mathcal{M}$:*

$$\mathbb{E}\left[\Theta_\mu\right] = 0$$

Proof. We know that Θ is $\mathcal{F}_1$ mesurable. So, $\Theta = \Theta\left((W_t)_{0\leq t\leq 1}\right)$. We use the parity of H and $x \mapsto x^2$. And the fact that a brownian (W_t) is equal in law to $(-W_t)$. And then we see that:

$$\mathcal{J}(\mu, \Theta_\mu(W_t)_{0\leq t\leq 1}) = \mathcal{J}(\mu, -\Theta_\mu(-W_t)_{0\leq t\leq 1})$$

By uniqueness of the maximizer Θ_μ, we get $-\Theta_\mu(-W_t)_{0\leq t\leq 1} = \Theta_\mu(W_t)_{0\leq t\leq 1}$. And:

$$\mathbb{E}\left[\Theta_\mu(-W_t)_{0\leq t\leq 1}\right] = \mathbb{E}\left[\Theta_\mu(W_t)_{0\leq t\leq 1}\right]$$

It follows:

$$\mathbb{E}\left[\Theta_\mu\right] = -\mathbb{E}\left[\Theta_\mu\right]$$

□

Now we need to verify that Θ_μ is not on the edge of $\mathcal{V}$ ($|\Theta_\mu| < 1$):

Lemma 4.3. *For all μ,*

$$|\Theta_\mu| < 1 \quad a.s$$

Proof. We suppose that we have K non negligible such that $\forall \omega \in K \quad \Theta_\mu(\omega) = 1$.

Then, we perturb Θ_μ into $\Theta_\mu - \delta \mathbb{1}_K$ with $\delta > 0$:

$$\begin{aligned}\mathcal{J}(\mu, \Theta_\mu - \delta\mathbb{1}_K) - \mathcal{J}(\mu, \Theta_\mu) &= -\delta\mathbb{E}\left[\mathbb{1}_K\left(\int_0^1 \beta^2\mu[0,t]\Theta_{\mu t}\,\mathrm{d}t + \beta W_1\right)\right] \\ &\quad + \mathbb{P}(K)(H(1) - H(1-\delta)) + \frac{\delta^2}{2}\mathbb{E}\left[\int_0^1 \beta^2\mathbb{1}_{Kt}^2\mu[0,t]\,\mathrm{d}t\right]\end{aligned}$$

H is convex so, H is above all its tangents. So, $H(1) - H(1-\delta) \geq H'(1-\delta)\delta$. And the last term in the equality is positive. And, with Cauchy-Schwartz inequality $\mathbb{E}\left[\mathbb{1}_K W_1\right] \leq \sqrt{\mathbb{P}(K)}||W_1||_{L^2}$. It follows:

$$\mathcal{J}(\mu, \Theta_\mu - \delta\mathbb{1}_K) - \mathcal{J}(\mu, \Theta_\mu) \geq \delta\mathbb{P}(K)\left(H'(1-\delta) - \int_0^1 \beta^2\mu[0,t]\,\mathrm{d}t - \frac{1}{\sqrt{\mathbb{P}(K)}}||W_1||_{L^2}\right)$$

This inequality is true for all $0 < \delta < 2$. And Θ_μ is a maximizer of $\mathcal{J}(\mu, \cdot)$. So:

$$\forall \delta \in]0,2[\quad \delta\mathbb{P}(K)\left(H'(1-\delta) - \int_0^1 \beta^2\mu[0,t]\,\mathrm{d}t - \frac{1}{\sqrt{\mathbb{P}(K)}}||W_1||_{L^2}\right) \leq 0$$

But $H'(x)$ goes to $+\infty$ when x goes to 1. So, for small δ, the inequality is violated.

And then $\mathbb{P}(K)$ must be zero. The same can be done with the case $\Theta_\mu = -1$. □

Remark. We don't have $|\Theta_\mu| \leq C_\mu < 1$

Definition 4.1. For $\Theta \in \mathcal{V}$ such that $|\Theta| < 1$, we define the process:

$$\Xi_\Theta = \int_0^1 \beta^2\Theta_t\mu[0,t]\,\mathrm{d}t + \beta W_1 - H'(\Theta)$$

So, we can now give some properties of Θ_μ,

Lemma 4.4. *For all $\mu \in \mathcal{M}$,*

$$\Xi_{\Theta_\mu} = 0 \quad a.s$$

Proof. For $\varepsilon\Lambda$ such that $\Theta_\mu + \varepsilon\Lambda \in \mathcal{V}$:

$$\begin{aligned}\mathcal{J}(\mu, \Theta_\mu + \varepsilon\Lambda) - \mathcal{J}(\mu, \Theta_\mu) &= \varepsilon\mathbb{E}\left[\Lambda\left(\int_0^1 \beta^2\mu[0,t]\Theta_{\mu t}\,\mathrm{d}t + \beta W_1\right)\right] \\ &+ \mathbb{E}\left[H(\Theta_\mu) - H(\Theta_\mu + \varepsilon\Lambda))\right] + \frac{\varepsilon^2}{2}\mathbb{E}\left[\int_0^1 \beta^2\Lambda_t^2\mu[0,t]\,\mathrm{d}t\right] \\ &= \varepsilon\mathbb{E}\left[\Lambda\Xi_{\Theta_\mu}\right] + \mathcal{O}(\varepsilon^2)\end{aligned}$$

Because $|\Lambda| < 2 \quad a.s$, we have the $\mathcal{O}(\varepsilon^2)$.

Θ_μ maximize $\mathcal{J}(\mu, \cdot)$, then for $\varepsilon\Lambda$ such that $\Theta_\mu + \varepsilon\Lambda \in \mathcal{V}$ and ε is small enough:

$$\mathbb{E}\left[\Lambda\Xi_{\Theta_\mu}\right] \le 0$$

For $\Lambda = \mathbb{1}_{|W_1|<k}\mathbb{1}_{|\Theta_\mu|<1-\delta}\Xi_{\Theta_\mu}$, there exists ε such that $\Theta_\mu + \varepsilon\Lambda \in \mathcal{V}$. Indeed:

$$\mathbb{1}_{|W_1|<k}\mathbb{1}_{|\Theta_\mu|<1-\delta}|\Xi_{\Theta_\mu}| \le \beta k + |H'(1-\delta)| + \int_0^1 \beta^2\mu[0,t] = M_\delta \quad a.s$$

So, we can take $\varepsilon \le \frac{\delta}{M_\delta}$. Then, for all $\delta, k > 0$, $\mathbb{1}_{|W_1|<k}\mathbb{1}_{|\Theta_\mu|<1-\delta}\Xi_{\Theta_\mu} = 0 \quad a.s.$ If we take the limit δ to 0, k to infinity, with lemma 4.3 and the fact that W_1 is finite almost surely, we deduce the result. □

So, we have a condition on Θ_μ. If we knew all the processes that satisfy $\Xi_\Theta = 0$ we would have a good knowledge of Θ_μ. But this is not an easy task due to what we will do bellow:

Formally we get by Itô's formula and the representation theorem ($\Theta_t = \int_0^1 \gamma_t\,\mathrm{d}W_t$ because Θ_t is a continuous $(\mathcal{F}_t)$- martingale and $\Theta_0 = 0$ with lemma 4.2):

$$\Xi_\Theta = \int_0^1 \beta^2\mu[0,t]\Theta_t - \frac{1}{2}H'''(\Theta_t)\gamma_t^2\,\mathrm{d}t + \int_0^1 \beta - H''(\Theta_t)\gamma_t\,\mathrm{d}W_t$$

We are tempted to say that the process $\int_0^t \beta^2\Theta_t\mu[0,t]\,\mathrm{d}t + \beta W_t - H'(\Theta_t)$ is always zero:

$$\frac{1}{2}H'''(\Theta_t)\gamma_t^2 = \beta^2\mu[0,t]\Theta_t \tag{2}$$

And:

$$\beta = H''(\Theta_t)\gamma_t$$

And $\Theta_0 = 0$.

But $H''(x) = \frac{1}{1-x^2}$ so, $H'''(x) = 2x(H''(x))^2$. It follows from 2:

$$\Theta_t(H''(\Theta_t)\gamma_t)^2 = \beta^2\mu[0,t]\Theta_t \tag{3}$$

Then $\mu[0,t] = 1$ which is true only if we are RS (replica symmetric ($\mu = \delta_0$)). But we want to show that the support of μ is an interval $[0, q_*(\beta)]$. So, we shouldn't be in this case. Then we must find a process Θ_t such that the process $\int_0^t \beta^2\Theta_t\mu[0,t]\,\mathrm{d}t + \beta W_t - H'(\Theta_t)$ is zero in 0 and 1 but non-zero elsewhere. It's like a brownian bridge. For this, we might want to look in the direction of conditioned processes [5]. But we haven't tackled this here.

4.2 Condition on μ

We want also to find the optimal μ_P by a variation principle. But, fot that, we need to know how Θ_μ is perturbed when μ is perturbed. The following theorem seams to be easy at first but $\mathcal{V}$ is not compact so, we can't conclude quickly by continuity of the applications we are studying.

Lemma 4.5 (variation of a solution)**.** *The application:*

$$T : \mu \mapsto \Theta_\mu$$

is continuous.

Proof. According to lemma 4.1, we can define the application $T : \mu \to \Theta_\mu$.

We know, by recalling a result in the proof of the lemma 4.1, that:

$$\forall \omega \in \Omega \quad T(\mu)(\omega) = \operatorname{argsup}_{\theta \in [-1,1]} \theta \left(\int_0^1 \beta^2 \mu[0,t] \partial_x u_\mu(t, X_t(\omega)) \, \mathrm{d}t + \beta W_1(\omega) \right) - H(\theta)$$

With:

$$\mathrm{d}X_s^\mu = \beta^2 \mu[0,s] \partial_x u_\mu(s, X_s^\mu) \, \mathrm{d}s + \beta \, \mathrm{d}W_s$$

And $X_0^\mu = 0$.

We want to prove that T is continuous.

To do this we want to prove that, $\mu \mapsto \partial_x u_\mu(t, X_t)$ is continuous. Indeed, if we had this, we would have:

$$\forall \omega \in \Omega \quad T(\tilde{\mu})(\omega) = \operatorname{argsup}_{\theta \in [-1,1]} \theta \left(\int_0^1 \beta^2 \mu[0,t] \partial_x u_\mu(t, X_t(\omega)) \, \mathrm{d}t + \beta W_1(\omega) \right) - H(\theta) + \theta \underset{\tilde{\mu} \to \mu}{o}(1)$$

Because the functional to maximize is strictly concave, the maximizer won't change by a small modification of the functional to maximize.

According to the lemma 8 in [7], we know that $||\partial_x u_\mu - \partial_x u_{\tilde{\mu}}||_\infty < C(\beta) d(\mu, \tilde{\mu})$ with $d(\mu, \nu) = \int_0^1 |\mu[0,t] - \nu[0,t]| \, \mathrm{d}t$. So, we just have to prove the continuity of $\mu \mapsto X_t^\mu$.

We have:

$$\begin{aligned} \mathrm{d}(X_t^\mu - X_t^{\tilde{\mu}}) &= \beta^2 \mu[0,t] \left(\partial_x u_\mu(t, X_t^\mu) - \partial_x u_{\tilde{\mu}}(t, X_t^{\tilde{\mu}}) \right) + \beta^2 (\tilde{\mu} - \mu)[0,t] \partial_x u_{\tilde{\mu}}(t, X_t^{\tilde{\mu}}) \, \mathrm{d}t \\ &= \beta^2 \mu[0,t] \left(\partial_x u_\mu(t, X_t^\mu) - \partial_x u_\mu(t, X_t^{\tilde{\mu}}) \right) + \beta^2 \mu[0,t] \left(\partial_x u_{\tilde{\mu}}(t, X_t^{\tilde{\mu}}) - \partial_x u_\mu(t, X_t^{\tilde{\mu}}) \right) \\ &+ \beta^2 (\tilde{\mu} - \mu)[0,t] \partial_x u_{\tilde{\mu}}(t, X_t^{\tilde{\mu}}) \, \mathrm{d}t \end{aligned}$$

So:

$$\begin{aligned} \left| \frac{\mathrm{d}\left(X_t^\mu - X_t^{\tilde{\mu}} \right)}{\mathrm{d}t} \right| &\leq \beta^2 (d(\mu, \tilde{\mu}) + ||\partial_x u_\mu - \partial_x u_{\tilde{\mu}}||_\infty) + \beta^2 \mu[0,t] \left| \partial_x u_\mu(t, X_t^\mu) - \partial_x u_\mu(t, X_t^{\tilde{\mu}}) \right| \\ &\leq C(\beta) d(\mu, \tilde{\mu}) + \beta^2 \mu[0,t] \left(|X_t^\mu - X_t^{\tilde{\mu}}| + o(||X_t^\mu - X_t^{\tilde{\mu}}||) \right) \end{aligned}$$

The majoration are obtained with the utilisation of some other results in [7]: $||\partial_x u_\mu||_\infty, ||\partial_{xx} u_\mu||_\infty < 1$. Because the derivative of $X_t^\mu - X_t^{\tilde{\mu}}$ is small when $X_t^\mu - X_t^{\tilde{\mu}}$ is small and $X_0^\mu - X_0^{\tilde{\mu}} = 0$ we get that $X^\mu \underset{\tilde{\mu} \to \mu}{\to} X^{\tilde{\mu}}$

To to this more rigorously, we say that $|X_t^\mu - X_t^{\tilde{\mu}}| < Y_t$ where:

$$\begin{cases} \frac{\mathrm{d}Y_t}{\mathrm{d}t} = C d(\mu, \tilde{\mu}) + C' Y_t \\ Y_0 = 0 \end{cases}$$

So: $Y_t = \frac{d(\mu, \tilde{\mu}) C}{C'} (\mathrm{e}^{C't} - 1) \underset{\tilde{\mu} \to \mu}{\to} 0$. □

The last lemma is really important for the next lemma:

Lemma 4.6 (Condition on Θ_{μ_P})**.** *There exists $\varepsilon > 0$ such that for all μ that satisfies $d(\mu, \mu_P) < \varepsilon$ we have:*

$$\mathbb{E} \left[\int_0^1 (\mu - \mu_P)[0,t] (\Theta_t^2 - t) \, \mathrm{d}t \right] \geq 0$$

Proof. For $\mu \in \mathcal{M}$, we must have

$$\mathcal{J}(\mu, \Theta_\mu) - \mathcal{J}(\mu_P, \Theta_{\mu_P}) \geq 0$$

And:

$$\mathcal{J}(\mu, \Theta_\mu) - \mathcal{J}(\mu_P, \Theta_{\mu_P}) = \mathcal{J}(\mu_P, \Theta_\mu) - \mathcal{J}(\mu_P, \Theta_{\mu_P}) + \frac{1}{2}\mathbb{E}\left[\int_0^1 \beta^2(\mu - \mu_P)[0,t](\Theta^2_{\mu\, t} - t)\,\mathrm{d}t\right]$$

But Θ_{μ_P} maximizes $\Theta \mapsto \mathcal{J}(\mu_P, \Theta)$. So, we must have: For all $\mu \in \mathcal{M}$:

$$\frac{1}{2}\mathbb{E}\left[\int_0^1 \beta^2(\mu - \mu_P)[0,t](\Theta^2_{\mu\, t} - t)\,\mathrm{d}t\right] \geq 0$$

And now if we use the continuity of $\mu \mapsto \Theta_\mu$ (lemma 4.5), the result follows. □

Let's define $f : t \mapsto \mathbb{E}\left[\Theta^2_{\mu_P\, t}\right] - t$. We can now have some conditions on f:

Theorem 4.7 (Condtion on f)**.** *For all $s \in [0,1]$,*

$$\int_0^s f(t)\,\mathrm{d}t \leq 0$$

And equal to zero on the support of μ_P.

Proof.

$$\int_0^1 (\mu - \mu_P)[0,t]f(t) = \int_0^1 \int_0^t f(t)\,\mathrm{d}\mu(s)\,\mathrm{d}t - \int_0^1 \int_0^t f(t)\,\mathrm{d}\mu_P(s)\,\mathrm{d}t$$

Then we use Fubini's theorem and:

$$\int_0^1 (\mu - \mu_P)[0,t]f(t) = \int_0^1 \int_s^1 f(t)\,\mathrm{d}t\,\mathrm{d}\mu(s) - \int_0^1 \int_s^1 f(t)\,\mathrm{d}t\,\mathrm{d}\mu_P(s)$$

Let's define:

$$K : \mu \mapsto \int_0^1 \int_s^1 f(t)\,\mathrm{d}\mu(s)$$

According to the last lemma, μ_P must be a local minimizer of K.

We remark that K admits at least one global minimizer. We can take μ such that μ is supported on $\arg\min_{s\in[0,1]} \int_s^1 f(t)\,\mathrm{d}t$ (this minimizer exists by the extreme value theorem).

K is convex (even linear) so, all local minimizers are global minimizers. So, μ_P is a global minimizer of K.

Now let's define $F(s) = \int_0^s f(t)\,\mathrm{d}t$. F will be maximized on the support of μ_P. And by [1], we know that 0 belongs to the support of μ_P. So, $F(0) = 0$ maximize F.

Then we will have $F \leq 0$ and equal to zero on the support of μ_P.

□

From the last theorem, we know that f is equal to zero on the support of μ_P. Indeed, suppose there is a point t_0 in the support of μ_P such that $f(t_0) \neq 0$. There are two cases:

1. If $f(t_0) > 0$, we have, by continuity of f, F strictly increasing in t_0. And that's impossible (or $t_0 = 1$ but we know that 1 is not in the support of μ_P) because $F(t_0) = 0$ and $F \leq 0$.

2. If $f(t_0) < 0$, we have by continuity of f, F strictly decreasing in t_0. And that's impossible (or $t_0 = 0$ but we already know that $f(0) = 0$ (lemma 4.2)) with the same arguments as above.

Then we can think that we have enough information on μ_P and do one of the thing bellow:

1. We can consider μ_* such that $\operatorname{supp}\mu_* = [0, q_\beta]$ and then we can show that the optimal Θ is such that $\int_0^t f(s)\,\mathrm{d}s$ is zero on $[0, q_\beta]$ and negative outside. That would means that $\mu_* = \mu_P$.
2. We can also show that if there exists (μ_*, Θ) such that $\operatorname{supp}\mu_* = [0, t_1] \cup [t_2, q_\beta]$ and Θ is optimal with respect of F zero on $[0, t_1] \cup [t_2, q_\beta]$ and negative outside. Then there is something wrong (for example: there exists another (μ_*, Θ) or the Θ is not defined).

But for now we have not enough information on Θ, we will see in the next section that Θ_μ can be described by a PDE.

4.3 A new PDE

In this section, we will introduce a new PDE which will permit to find an SDE satisfied by Θ and also define the Parisi measure. The first subsection will present an intuitive way to find the PDE. While the second section will give a more rigorous proof using the Parisi PDE.

We have to transform the expression:

$$\mathcal{P}(\mu) = \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta\beta W_1 - H(\Theta)\right]$$

In a form usually seen in optimal control theory:

$$v(x,t) = \inf_{\gamma \in \mathcal{A}_t} \mathbb{E}\left[\int_t^1 J(\gamma_s, X_s^\gamma, s)\,\mathrm{d}s + U(X_1^\gamma)\middle| X_t^\gamma = x\right]$$

For this we need to use the Itô's representation theorem on (Θ_t) which is a continuous martingale on the brownian filtration:

$$\Theta_t^\gamma = \int_0^t \gamma_t\,\mathrm{d}W_t + h$$

With γ an adapted process from 0 to t and $h = 0$ according to lemma 4.2. And so, we can take the sup on the γ and the expression is transformed into:

$$\sup_{\gamma \in \mathcal{A}_0} \mathbb{E}\left[\frac{1}{2}\int_0^1 \beta^2 \mu[0,t]\left({\Theta_t^\gamma}^2 - t\right) + \beta\gamma_t\,\mathrm{d}t - H(\Theta_1^\gamma)\right]$$

Because $\mathbb{E}\left[\Theta_1 W_1\right] = \mathbb{E}\left[\int_0^1 \gamma_t\,\mathrm{d}W_t \int_0^1 \mathrm{d}W_t\right] = \mathbb{E}\left[\int_0^1 \gamma_t\,\mathrm{d}t\right]$.

Now we can define:

$$v_\mu(x,t) = \sup_{\gamma \in \mathcal{A}_t} \mathbb{E}\left[\frac{1}{2}\int_t^1 \beta^2 \mu[0,s]\left({\Theta_s^\gamma}^2 - s\right) + \beta\gamma_s\,\mathrm{d}s - H(\Theta_1^\gamma)\middle|\Theta_t^\gamma = x\right]$$

And we have $v_\mu(0,0) = \mathcal{P}(\mu)$.

4.3.1 A formal way to find the PDE

Now we will do things formally and admit that we have all the continuity we need, we will see after that we can make a link between the Parisi PDE and the one we will find formally.

By Itô's formula:

$$\forall\gamma \quad v_\mu(x,t) = \mathbb{E}\Big[-v(\Theta_1^\gamma,1) - \int_t^1 \partial_t v_\mu(\Theta_s^\gamma,s) + \frac{1}{2}\partial_{xx}v_\mu(\Theta_s^\gamma,s)\gamma_s^2\,\mathrm{d}s\Big|\Theta_t^\gamma = x\Big]$$

But:

$$\forall\gamma \quad v_\mu(x,t) \geq \mathbb{E}\Big[\frac{1}{2}\int_t^1 \beta^2\mu[0,s]\left(\Theta_s^{\gamma\,2} - s\right) + \beta\gamma_s\,\mathrm{d}s - H(\Theta_1^\gamma)\Big|\Theta_t^\gamma = x\Big]$$

And if we combine the two last lines with $v(x,1) = -H(x)$ and use them for all x and t:

$$\forall\gamma,x,t \quad \partial_t v_\mu(x,t) + \partial_{xx}v_\mu(x,t)\frac{\gamma^2}{2} + \beta\gamma + \frac{\beta^2}{2}\mu[0,t](x^2-t) \leq 0$$

Then for all x,t:

$$\partial_t v_\mu(x,t) + \sup_{\gamma\in\mathbb{R}}\left(\partial_{xx}v_\mu(x,t)\frac{\gamma^2}{2} + \beta\gamma\right) + \frac{\beta^2}{2}\mu[0,t](x^2-t) \leq 0 \tag{4}$$

But as the optimal γ_t is reached in the variational expression, we can say that:

$$\partial_t v_\mu(x,t) + \sup_{\gamma\in\mathbb{R}}\left(\partial_{xx}v_\mu\frac{\gamma^2}{2} + \beta\gamma\right) + \frac{\beta^2}{2}\mu[0,t](x^2-t) = 0$$

And then, if $\partial_{xx}v_\mu < 0$, the maximizer γ would be $-\frac{\beta}{\partial_{xx}v_\mu(x,t)}$. And then v_μ is a (weak) solution of:

$$\begin{cases} \partial_t v_\mu(x,t) - \frac{\beta^2}{2\partial_{xx}v_\mu(x,t)} + \frac{\beta^2}{2}\mu[0,t](x^2-t) = 0 & \forall(t,x)\in[0,1)\times\mathbb{R} \\ v_\mu(x,1) = -H(x) \quad \forall x\in\mathbb{R} \end{cases} \tag{5}$$

All that is not very rigorous but we have an easier way to find this PDE. We do this by identifying a relation between u_μ and v_μ. The intuition is given by the final condition on v_μ which is the Legendre transform of the final condition on u_μ.

4.3.2 A more rigorous way to see the PDE

So, we have the theorem:

Theorem 4.8 (Link between u_μ and v_μ)**.** *For all $x\in\mathbb{R}$ and $t\in[0,1]$:*

$$u_\mu(x,t) = \sup_{h\in\mathbb{R}}\left(xh + v_\mu(h,t)\right) + \frac{1}{2}\int_t^1 \beta^2\mu[0,s]s\,\mathrm{d}s$$

Proof. In [7], we have that:

$$u_\mu(x,t) = \sup_{\alpha\in\mathcal{A}_t}\mathbb{E}\left[\log\cosh\left(\int_t^1 \beta^2\mu[0,s]\alpha_s\,\mathrm{d}s + \beta(W_1-W_t) + x\right) - \frac{1}{2}\int_t^1 \beta^2\mu[0,s]\alpha_s^2\,\mathrm{d}s\right]$$

And then we do the same as in theorem 3.1:

$$\begin{aligned}
u_\mu(x,t) &= \sup_{\alpha\in\mathcal{A}_t}\mathbb{E}\left[\sup_{\theta\in[-1,1]}\theta\left(\int_t^1 \beta^2\mu[0,s]\alpha_s\,\mathrm{d}s + \beta(W_1-W_t) + x\theta\right) - H(\theta) - \frac{1}{2}\int_t^1 \beta^2\mu[0,s](\alpha_s^2)\,\mathrm{d}s\right] \\
&= \sup_{\alpha\in\mathcal{A}_t}\sup_{\Theta\in\mathcal{V}}\mathbb{E}\left[\Theta\left(\int_t^1 \beta^2\mu[0,s]\alpha_s\,\mathrm{d}s + \beta(W_1-W_t) + x\Theta\right) - H(\Theta) - \frac{1}{2}\int_t^1 \beta^2\mu[0,s](\alpha_s^2)\,\mathrm{d}s\right] \\
&= \sup_{\alpha\in\mathcal{A}_t}\sup_{\Theta\in\mathcal{V}}\mathbb{E}\left[\int_t^1 \Theta_s\beta^2\mu[0,s]\alpha_s\,\mathrm{d}s + \Theta\beta(W_1-W_t) + x\Theta - H(\Theta) - \frac{1}{2}\int_t^1 \beta^2\mu[0,s](\alpha_s^2)\,\mathrm{d}s\right] \\
&= \sup_{\alpha\in\mathcal{A}_t}\sup_{\Theta\in\mathcal{V}}\mathbb{E}\left[\int_t^1 \beta^2\mu[0,s](\Theta_s\alpha_s - \frac{1}{2}\alpha_s^2)\,\mathrm{d}s + \Theta\beta(W_1-W_t) + x\Theta - H(\Theta)\right] \\
&= \sup_{\Theta\in\mathcal{V}}\mathbb{E}\left[\int_t^1 \beta^2\mu[0,s]\frac{\Theta_s^2}{2}\,\mathrm{d}s + \Theta\beta(W_1-W_t) + x\Theta - H(\Theta)\right]
\end{aligned}$$

Now we use Itô's representation theorem:

$$\Theta_s^\gamma = \int_t^s \gamma_k \,\mathrm{d}W_k + h$$

And:

$$\begin{aligned}
u_\mu(x,t) &= \sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\int_t^1 \beta^2\mu[0,s]\frac{\Theta_s^2}{2}\,\mathrm{d}s + \Theta\beta(W_1-W_t) + x\Theta - H(\Theta)\right] \\
&= \sup_{\gamma\in\mathcal{A}_t}\sup_{h\in[-1,1]} \mathbb{E}\left[\int_t^1 \beta^2\mu[0,s]\frac{\Theta_s^{\gamma\,2}}{2}\,\mathrm{d}s + \Theta_1^\gamma\beta(W_1-W_t) + xh - H(\Theta_1^\gamma)\middle|\Theta_t^\gamma = h\right] \\
&= \sup_{\gamma\in\mathcal{A}_t}\sup_{h\in[-1,1]} \mathbb{E}\left[\int_t^1 \beta^2\mu[0,s]\frac{\Theta_s^{\gamma\,2}}{2} + \beta\gamma_s\,\mathrm{d}s + xh - H(\Theta_1^\gamma)\middle|\Theta_t^\gamma = h\right] \\
&= \sup_{h\in[-1,1]} (xh + v(h,t)) + \frac{1}{2}\int_t^1 s\beta^2\mu[0,s]\,\mathrm{d}s
\end{aligned}$$

$v_\mu(t,h) = -\infty$ for $|h|>1$ so, we can take the sup on $\mathbb{R}$. □

Now from the Parisi PDE, we can deduce a PDE in v_μ:

Theorem 4.9 (PDE for v_μ). *For all $\mu\in\mathcal{M}$, v_μ is a weak solution of the following PDE:*

$$\begin{cases} \partial_t v_\mu(x,t) - \frac{\beta^2}{2\partial_{xx}v_\mu(x,t)} + \frac{\beta^2}{2}\mu[0,t](x^2-t) = 0 & \forall (t,x)\in[0,1)\times\mathbb{R} \\ v_\mu(x,1) = -H(x) \quad \forall x\in\mathbb{R} \end{cases}$$

Proof. The double legendre transform of a convex function is the same function. So:

$$v(x,t) = -\sup_{h\in\mathbb{R}}(xh - u(h,t)) - \int_t^1 \frac{\beta^2}{2}\mu[0,s]s\,\mathrm{d}s$$

We know by [7] that u_μ is smooth in x and continuous differentiable in time at point of continuity of μ. So, we can define $(x,t)\mapsto g(x,t)$ such that:

$$\partial_x u_\mu(g(x,t),t) = x$$

Because u_μ is strictly convexe. So, ∂u_μ is strictly increasing so, bijective. g will be also smooth in space by the implicit function theorem. And:

$$\partial_x g(x,t) = \frac{1}{\partial_{xx}u_\mu(g(x,t),t)}$$

We have:

$$v_\mu(x,t) = u_\mu(g(x,t),t) - xg(x,t) - \frac{1}{2}\int_t^1 \beta^2\mu[0,s]s\,\mathrm{d}s$$

So:

$$\partial_x v_\mu(x,t) = \partial_x g(x,t)\partial_x u(g(x,t),t) - x\partial_x g(x,t) - g(x,t) = -g(x,t)$$

And:

$$\partial_{xx}v_\mu(x,t) = -\frac{1}{\partial_{xx}u_\mu(g(x,t),t)}$$

But u_μ is not differentiable in t. That's why we have weak solution but we will suppose that the results are the same if we do as if u_μ was smooth in time and so, g is also smooth in time. So:

$$\partial_t v_\mu(x,t) = \partial_t g(x,t)\partial_x u_\mu(g(x,t),t) + \partial_t u_\mu(g(x,t),t) - x\partial_t g(x,t) + \frac{1}{2}\beta^2\mu[0,t]t = \partial_t u_\mu(g(x,t),t) + \frac{1}{2}\beta^2\mu[0,t]t$$

So:

$$\begin{aligned}
&\partial_t v_\mu(x,t) - \frac{\beta^2}{2\partial_{xx}v_\mu(x,t)} + \frac{\beta^2}{2}\mu[0,t](x^2-t) \\
&= \partial_t u_\mu(g(x,t),t) + \frac{1}{2}\beta^2\mu[0,t]t + \frac{\beta^2}{2}\left(\partial_{xx}u_\mu(g(x,t)) + \mu[0,t]\left(\partial_x u_\mu(g(x,t),t)^2\right) - t\right) \\
&= \partial_t u_\mu(g(x,t),t) + \frac{\beta^2}{2}\left(\partial_{xx}u_\mu(g(x,t)) + \mu[0,t]\partial_x u_\mu(g(x,t),t)^2\right) = 0
\end{aligned}$$

And $v_\mu(x,1) = -\sup_{h\in\mathbb{R}}(hx - \log\cosh(h)) = -H(x)$ □

And now if we do the inverse of what we have done at the beginning of the section (control optimal "verification argument"), and that's what have been done in lemma 4 of [7]. Θ_μ solves the SDE (it corresponds to the sup in the HJB equation 4):

Lemma 4.10. *For all $\mu \in \mathcal{M}$:*

$$\mathrm{d}\Theta_{\mu\,t} = -\frac{\beta}{\partial_{xx}v_\mu(\Theta_{\mu\,t},t)}\,\mathrm{d}W_t$$

4.3.3 Different uses of the result

We can remark that $p_\mu = \partial_x v_\mu$ solves the PDE:

$$\begin{cases} \partial_t p_\mu(x,t) + \frac{\beta^2\partial_{xx}p_\mu(x,t)}{2\partial_x p_\mu(x,t)^2} + \beta^2\mu[0,t]x = 0 & \forall(t,x)\in[0,1)\times\mathbb{R} \\ p_\mu(x,1) = -H'(x) \quad \forall x\in\mathbb{R} \end{cases}$$

And by Itô's formula:

$$\begin{aligned}
p_\mu(\Theta_{\mu\,1},1) &= p_\mu(0,0) + \int_0^1 \partial_t p_\mu(\Theta_{\mu\,t},1) + \partial_{xx}p_\mu(\Theta_{\mu\,t},t)\frac{\beta^2}{2\partial_x p_\mu(\Theta_{\mu\,t},t)^2}\,\mathrm{d}t - \int_0^1 \partial_x p_\mu(\Theta_{\mu\,t},t)\frac{\beta}{\partial_x p_\mu(\Theta_{\mu\,t},t)}\,\mathrm{d}W \\
&= -\left(\int_0^1 \beta^2\mu[0,t]\Theta_{\mu\,t}\,\mathrm{d}t + \beta W_1\right)
\end{aligned}$$

And if we notice $-H(\Theta_{\mu\,1}) = p_\mu(\Theta_{\mu\,1},1)$, we have the result we had in the lemma 4.1.

We can also remark that if $\mu = \delta_0$, we find a solution of the new PDE:

$$v_{\delta_0}(x,t) = -H(x) + \frac{\beta^2}{2}\left(\frac{t^2}{2} - t\right) + \frac{\beta^2}{4}$$

And so, Θ_{δ_0} solves the SDE:

$$\begin{cases} \mathrm{d}\Theta_{\delta_0\,t} = \beta(1-\Theta_{\mu\,t}^2)\,\mathrm{d}W_t \\ \Theta_{\delta_0\,t} = 0 \end{cases} \tag{6}$$

And we find the result we had with the equation 6 in the section 4.1.

We have seen in this section that, we had a strong link between the variational representation with α and the one with Θ. So, we will see in the next section that we can do other parallels between the two representations.

5 Parallels between the two representations

In this section we will identify f to an other quantity that was already known. Then will identify $\Theta_{\mu\,t} = \alpha_t^*$.

5.1 An identification of f

In the section 3 in [2], we saw a quantity like $\mathbb{E}\left[\Theta_t^2\right]$ that is equal to t when we are on the support of μ_P. That's the quantity:

$$\Gamma_\mu(t) = \mathbb{E}\left[\partial_x u_\mu(\beta W_t, t)^2 \mathrm{e}^{W_\mu(t)}\right]$$

With:

$$W_\mu(t) = \int_0^t \left(u_\mu(\beta W_t, t) - u_\mu(\beta W_s)\right) \mathrm{d}\mu(s)$$

And now if we differentiate with Itô's formula:

$$\mathrm{d}W_\mu(t) = \mu[0,t] \left(\partial_t u_\mu(\beta W_t, t) + \frac{\beta^2}{2}\partial_{xx} u_\mu(\beta W_t, t)\,\mathrm{d}t + \beta\partial_x u_\mu(\beta W_t, t)\,\mathrm{d}W_t\right)$$

And now we use the Parisi PDE:

$$\mathrm{d}W_\mu(t) = -\mu[0,t]^2\frac{\beta^2}{2}(\partial_x u_\mu(\beta W_t, t))^2\,\mathrm{d}t + \mu[0,t]\beta\partial_x u_\mu(\beta W_t, t)\,\mathrm{d}W_t$$

So, if we define:

$$Y_t = \beta\mu[0,t]\partial_x u_\mu(\beta W_t, t)$$

Because we have, by definition, $W_\mu(0) = 0$:

$$W_\mu(t) = \int_0^t Y_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t Y_s^2\,\mathrm{d}s$$

By Girsanov theorem, we can define the probability measure $\mathbb{Q}$: $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}|\mathcal{F}_t = \mathrm{e}^{W_\mu(t)}$.

And then $\tilde{W}_t = W_t - \int_0^t Y_s\,\mathrm{d}s$ is a brownian motion under $\mathbb{Q}$. So, βW_t is a solution of the SDE:

$$\mathrm{d}X_t = \beta Y_t\,\mathrm{d}t + \beta\,\mathrm{d}\tilde{W}_t = \beta^2\mu[0,t]\partial_x u_\mu(X_t, t)\,\mathrm{d}t + \beta\,\mathrm{d}\tilde{W}_t$$

With $X_0 = 0$.

So, with the Girsanov theorem:

$$\Gamma_\mu(t) = \mathbb{E}\left[\partial_x u(X_t, t)^2\right]$$

And if we recall what we have said in 4.1, $\partial_x u(X_t, t)$ correspond to the maximizer α^* in 1. So,

$$\Gamma_\mu(t) = \mathbb{E}\left[\alpha_t^{*\,2}\right]$$

This was not unknown, indeed, Montanari used $\mathbb{E}\left[\alpha_t^{*\,2}\right]$ instead of $\Gamma_\mu(t)$ in the lemma 3.3 of [9]. But we haven't seen if the link has been made. And we have a funny way to show it.

Because $\mathbb{E}\left[\alpha_t^{*\,2}\right]$ and $\mathbb{E}\left[\Theta_{\mu\,t}^2\right]$ have the same properties on the Parisi measure μ_P. We must verify if $\alpha_t^* = \Theta_{\mu\,t}$.

5.2 Identification of $\Theta_{\mu\, t}$

We differentiate α_t^* and we will see that α_t^* will satisfy the same SDE as $\Theta_{\mu\, t}$.

$$\begin{aligned} \mathrm{d}\alpha_t^* &= \mathrm{d}\partial_x u(X_t,t) \\ &= \partial_{tx} u(X_t,t) + \frac{\beta^2}{2}\Big(\partial_{xxx} u_\mu(X_t,t) + 2\beta^2\mu[0,t]\partial_x u_\mu(X_t,t)\partial_{xx} u_\mu(X_t,t)\Big)\,\mathrm{d}t + \beta\partial_{xx} u_\mu(X_t,t)\,\mathrm{d}W_t \end{aligned}$$

We recall the Parisi PDE:

$$\begin{cases} \partial_t u_\mu(t,x) + \frac{\beta^2}{2}\left(\partial_{xx} u_\mu(t,x) + \mu[0,t](\partial_x u_\mu(t,x))^2\right) = 0, & (t,x) \in (0,1)\times\mathbb{R} \\ u_\mu(1,x) = \log\cosh(x) \end{cases}.$$

And now, if we differentiate the Parisi PDE in x,

$$\begin{cases} \partial_{tx} u_\mu(t,x) + \frac{\beta^2}{2}\left(\partial_{xxx} u_\mu(t,x) + 2\mu[0,t]\partial_x u_\mu(t,x)\partial_{xx} u_\mu(t,x)\right) = 0, & (t,x) \in (0,1)\times\mathbb{R} \\ u_\mu(1,x) = \log\cosh(x) \end{cases}.$$

So, we have:

$$\mathrm{d}\alpha_t^* = \beta\partial_{xx} u_\mu(X_t,t)\,\mathrm{d}W_t$$

Now we do as in the proof of theorem 4.9, we define g such that $\partial_x u_\mu(g(x,t),t) = x$ and:

$$\partial_{xx} u_\mu\left(g(x,t),t\right) = -\frac{1}{\partial_{xx} v_\mu(x,t)}$$

We just have to see that $X_t = g(\alpha_t^*,t)$ because $\alpha_t^* = \partial_x u_\mu(X_t,t)$. And:

$$\mathrm{d}\alpha_t^* = -\frac{\beta}{\partial_{xx} v(\alpha_t^*,t)}\,\mathrm{d}W_t$$

And the initial condition is:

$$\alpha_0^* = 0$$

Because $X_0 = 0$ and $\partial_x u_\mu(0,0)$ (see that $u_\mu(-x,t)$ is also a solution of the Parisi PDE so, u_μ is symmetric).

And if remember the SDE satisfied by $\Theta_{\mu\, t}$ in the last section:

$$\begin{cases} \mathrm{d}\Theta_{\mu\, t} = -\frac{\beta}{v_{xx}(\Theta_{\mu\, t},t)}\,\mathrm{d}W_t \\ \Theta_{\mu\, t} = 0 \end{cases}$$

We have that $\frac{1}{\partial_{xx} v_\mu}$ is bounded by 1 and smooth in x. So, $\frac{1}{\partial_{xx} v_\mu}$ has linear growth and is Lipschitz in x uniformally in t. Then, we have uniqueness of the SDE's solution. So, we have $\alpha_t^* = \Theta_{\mu\, t}$.

It is a result we should have predicted earlier. Indeed, in the proof of the theorem 3.1, we had the result:

$$\mathcal{P}(\mu) = \sup_{\alpha\in\mathcal{A}_0}\sup_{\Theta\in\mathcal{V}} \mathbb{E}\left[\left(\int_0^1 \beta^2\mu[0,t](\Theta_t\alpha_t - \frac{1}{2}\alpha_t^2)\,\mathrm{d}t + \Theta\beta\,\mathrm{d}W_1\right) - H(\Theta) - \frac{1}{2}\int_0^1 \beta^2\mu[0,t]t\,\mathrm{d}t\right]$$

And then, we said we recognized the Legendre transform of $x \mapsto \frac{x^2}{2}$. But the maximizer in $\sup_{h\in\mathbb{R}} xh - \frac{h^2}{2}$ is $h_{\max} = x$. So, the maximizer $\alpha(\Theta)$ for a given Θ respects $\alpha_t = \Theta_t$. And finally,

$$\alpha_t^* = \alpha_t(\Theta_\mu) = \Theta_{\mu\, t}$$

5.3 A new formulation of a variation principle

The quantity has been found by a variation principle (likeα_t^* us with the Θ). Indeed, we have, according to the lemma 3.7 in [13]:

$$\left.\frac{\mathrm{d}\mathcal{P}(\mu_s)}{\mathrm{d}s}\right|_{s=0} = \frac{1}{2}\int_0^1 \beta^2(t-\Gamma_\mu(t))a(t)\,\mathrm{d}\mu(t)$$

This result is true for all $\mu \in \mathcal{M}$. And the continuous function a on $[0,1]$ satisfying $0 \le t + a(t) \le 1$ for $t \in [0,1]$ and $|a(t) - a(\tilde{t})| \le |t - \tilde{t}|$ for $t, \tilde{t} \in [0,1]$.

And for $s \in [0,1]$, we define $\mu_s \in \mathcal{M}$ as the probability measure induced by the mapping $t \mapsto t + sa(t)$, that is, $\mu_s([0, t+sa(t)]) = \mu([0,t])$ for $t \in [0,1]$. This definition may be hard to understand at first. But it is as if we make a dilatation of time on μ. The important condition on a is $|a(t) - a(\tilde{t})| \le |t - \tilde{t}|$ which means that $t \mapsto t + sa(t)$ is strictly increasing for all $s \in [0,1]$ and so, the change of time is a bijection.

And so, μ_P must respect for all a:

$$\frac{1}{2}\int_0^1 \beta^2(t-\Gamma_{\mu_P}(t))a(t)\,\mathrm{d}\mu_P(t) \ge 0$$

The variations they have use are not as general as the ones we have used so, we could use the lemma 4.6 to find this one. So, we recall our result in the proof of (lemma 4.6):

There exists $\varepsilon > 0$ such that for all μ that satisfy $d(\mu, \mu_P) < \varepsilon$ we have (if we remember the definition $f(t) = \Theta^2_{\mu_P\,t} - t$):

$$\int_0^1 (\mu - \mu_P)[0,t] f(t)\,\mathrm{d}t \ge 0$$

Let's fixe a,

There exists η such that for all $s < \eta$, $d(\mu_s, \mu_P) < \eta$. Then, for all $s < \eta$,

We define $\kappa_s(t) = t + sa(t)$. Wiht the definition of a, κ_s is bijective so, we can define the inverse function. And

$$(\mu_s - \mu_P)[0,t] = \mu_s[0, \kappa_s^{-1}(t)] - \mu_P[0,t] = \int_t^{\kappa_s^{-1}(t)} \mathrm{d}\mu_P(k)$$

Then:

$$\begin{aligned}
\int_0^1 (\mu - \mu_P)[0,t] f(t)\,\mathrm{d}t &= \int_0^1 \int_t^{\kappa_s^{-1}(t)} f(t)\,\mathrm{d}\mu_P(k)\,\mathrm{d}t \\
&= \int_0^1 \int_{\kappa_s(k)}^{k} f(t)\,\mathrm{d}\mu_P(k)\,\mathrm{d}t \\
&\simeq \int_0^1 (k - \kappa_s(k))\, f(k)\,\mathrm{d}\mu_P(k) \\
&\simeq \int_0^1 -sa(k) f(k)\,\mathrm{d}\mu_P(k)
\end{aligned}$$

Because $-f(k) = k - \Gamma_{\mu_P}$, we fall into the same result.

Now we have a good knowledge of the quantity we are using. So, can we find a solution of our problem? The answer is no. We will describe what we can deduce in the next section:

6 How to find the solution of our problem

We suppose that supp$\mu_p = [0, a] \cup [b, q_*(\beta)]$. And we want to prove that there is something wrong. But as we have said before, we haven't managed to find a contradiction. We managed to show somethings already known:

- 0 is in the support of μ_P (I didn't show it exactly but I think I'm really near (It's not really important so I haven't spent a lot of time on it)).
- $\mu_P = \delta_0$ for $\beta \leq 1$.
- $1 - q_*(\beta)$ and $\mu_P[0, t]$ is like $\mathcal{O}(\frac{1}{\beta})$.

Let's recall the important things that we know:

We have a function f defined by $f(t) = \mathbb{E}\left[\Theta^2_{\mu_P\,t}\right] - t$. With, $\Theta_{\mu_P\,t}$ the solution of the SDE:

$$\begin{cases} \mathrm{d}\Theta_{\mu_P\,t} = -\frac{\beta}{\partial_{xx} v_\mu(\Theta_{\mu_P}, t)} \mathrm{d}W_t \\ \Theta_{\mu_P\,0} = 0 \end{cases}$$

With respect of the PDE:

$$\begin{cases} \partial_t v_\mu(x, t) - \frac{\beta^2}{2\partial_{xx} v_\mu(x,t)} + \frac{\beta^2}{2}\mu[0, t](x^2 - t) = 0 \quad \forall (t, x) \in [0, 1) \times \mathbb{R} \\ v_\mu(x, 1) = -H(x) \quad \forall x \in \mathbb{R} \end{cases}$$

And f must be equal to zero on $[0, a] \cup [b, q_*(\beta)]$ and we have $\int_a^t f(t)\,\mathrm{d}t \leq 0$ for $t \leq b$ with equality in $t = b$.

6.1 0 is in the support of μ_P

First we want to differentiate f, the first derivative is simple:

$$f'(t) = \mathbb{E}\left[\frac{\beta^2}{\partial_{xx} v_\mu(\Theta_{\mu_P\,t}, t)^2}\right] - 1$$

But for the second derivative, by application of the Itô's formula:

$$f''(t) = \mathbb{E}\left[\partial_t g(\Theta_{\mu_P\,t}, t) + \frac{1}{2}\partial_{xx} g(\Theta_{\mu_P}, t) g(\Theta_{\mu_P}, t)\right]$$

With:

$$g(x, t) = \frac{\beta^2}{\partial_{xx} v_\mu(x, t)^2}$$

And if differentiate the PDE in v_μ two times in x, we find that g satisfies the PDE:

$$-\partial_t g + \frac{1}{4}\partial_x g^2 - \frac{1}{2} g \partial_{xx} g = \beta\mu[0, t] g^{\frac{3}{2}}$$

So, we have,

$$f''(t) = \mathbb{E}\left[\frac{1}{4}\partial_x g(\Theta_{\mu_P\,t}, t)^2 - \beta\mu[0, t] g(\Theta_{\mu_P\,t}, t)^{\frac{3}{2}}\right]$$

We manage to deduce from this that μ_P is supported in 0 if $x \mapsto \partial_x g(x, t)$ is equal to zero on a finite number of point for $t \leq r$ with respect of $\mu[0, r] = 0$.

Indeed, if it's not the case, we would have a r (the first point of the support of μ_P) such that $\mu_P[0, r] = 0$. And so, $f''(t) \geq 0$ for $t < r$. And we can't have a strict inequality because it would means that $\int_0^r f(t)\,\mathrm{d}t < 0$ (a function strictly convex with $f(0) = f(r)$). So, $f''(t) = 0$ for $t \leq r$ which implies that $\partial_x g(\Theta_{\mu_P\,t}, t) = 0 \quad a.s.$ This means that $\Theta_{\mu_P\,t}$ takes a finite number of values which contradicts the SDE solved by $\Theta_{\mu_P\,t}$.

We are very closed to showing that μ_P is supported to zero. Because the condition seams always true.

6.2 We are Replica Symmetric when $\beta \leq 1$

We can also say that the model is RS ($\mu_P = \delta_0$) for $\beta \leq 1$ because (with the equation 6):

$$\mathrm{d}\Theta_{\delta_0\,t} = \beta(1 - \Theta_{\delta_0\,t}^2)\,\mathrm{d}W_t$$

This means that:

$$f'(t) = \mathbb{E}\left[\beta^2(1 - \Theta_{\delta_0\,t}^2)^2\right] - 1 \leq \beta^2 - 1 \leq 0$$

So, the condition on f is respected. Then δ_0 is a local minimizer of $\mathcal{P}$ so, by convexity of $\mathcal{P}$, it is a global minimizer and $\delta_0 = \mu_P$.

We manage to do one last thing:

6.3 $1 - q_*(\beta)$ and $\mu_P[0,t]$ is like $\mathcal{O}(\frac{1}{\beta})$.

We obtain by using the PDE in v_μ:

$$\frac{1}{2}\beta^2\mu[0,t]f(t) = \frac{1}{2}\mathbb{E}\left[\frac{\beta^2}{\partial_{xx}v_\mu(t, \Theta_{\mu_P\,t})}\right] - \mathbb{E}\left[\partial_t v_\mu(t, \Theta_{\mu_P\,t})\right] \tag{7}$$

We know that for $q > q_*(\beta)$, $\mu[0,q] = 1$, so, we can see that:

$$v_\mu(x,t) = -H(x) + \frac{\beta^2}{2}\left(\frac{t^2}{2} - t\right) + \frac{\beta^2}{4} \quad \forall t > q_*(\beta)$$

We have from Itô's formula:

$$\mathbb{E}\left[\partial_t v_\mu(\Theta_{\mu_P\,t}, t)\right] = \mathbb{E}\left[v_t(\Theta_{\mu_P\,q}, q)\right] + \frac{\beta^2}{2}(\mu[0,q]f(q) - \mu[0,t]f(t)) - \int_t^q \frac{\beta^2}{2}\mu[0,s](f'(s)+1)\,\mathrm{d}s$$

Because:

$$\begin{aligned}
\mathrm{d}\partial_t v_\mu(t, \Theta_{\mu_P\,t}) &= \left(\partial_{tt}v_\mu + \frac{\beta^2}{2}\frac{\partial_{xxt}v_\mu}{\partial_{xx}v_\mu^2}\right)(\Theta_{\mu_P\,t}, t)\,\mathrm{d}t + \cdots\,\mathrm{d}W_t \\
&= \frac{\partial}{\partial t}\left(\frac{\beta^2}{2}\mu[0,t](t - x^2)\right)(\Theta_{\mu_P\,t}, t)\,\mathrm{d}t + \cdots\,\mathrm{d}W_t
\end{aligned}$$

So,

$$\mathbb{E}\left[v_t(\Theta_q, q)\right] - \mathbb{E}\left[v_t(\Theta_{\mu_P\,t}, t)\right] = \int_t^q \frac{\beta^2}{2}\mu[0,s](f'(s)+1)\,\mathrm{d}s - \frac{\beta^2}{2}(\mu[0,q]f(q) - \mu[0,t]f(t))$$

If we evaluate in $q_*(\beta)$ using (7) and $f(q_*(\beta)) = 0$:

$$\frac{\beta^2}{2}(q_*(\beta) - 1) - \int_t^{q_*(\beta)} \frac{\beta^2}{2}\mu[0,s](f'(s)+1)\,\mathrm{d}s = \frac{1}{2}\mathbb{E}\left[\frac{\beta^2}{\partial_{xx}v_\mu(\Theta_{\mu_P\,t}, t)}\right]$$

But with Cauchy-Schwartz inequality, we get:

$$0 > \mathbb{E}\left[\frac{\beta^2}{\partial_{xx}v_\mu(\Theta_{\mu_P\,t}, t)}\right] \geq -\beta\sqrt{f'(t)+1}$$

So, when β grows big $q_*(\beta) - 1 = \mathcal{O}(\beta)$ and $\mu[0,t] = \mathcal{O}(\beta)$ for $t < q_*(\beta) - \varepsilon$.

Also, by using Itô's formula we can prove the same result:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\frac{\beta^2}{\partial_{xx} v_\mu(t, \Theta_{\mu_P\, t})}\right] = \beta^2 \mu[0,t]\mathbb{E}\left[\frac{\beta^2}{\partial_{xx} v_\mu(t, \Theta_{\mu_P\, t})^2}\right] = \beta^2 \mu[0,t](f'(t) + 1)$$

7 Some unsuccessful ideas

Here, we will present some other ideas we have explored. But either we had no useful result either we had no result at all. We will not present all the unsuccessful ideas (because there were a lot), but only the most interesting.

7.1 Model at zero temperature

We have first looked for the model at zero temperature, because it seamed to be easier. The model was:

Let $\mathcal{U}$ be the collection of all cumulative distributions γ on $[0,1)$ induced by a mesure on $[0,1)$ and satisfying $\int_0^1 \gamma(t)\,\mathrm{d}t < \infty$ and $\gamma(1) = \infty$.

For each $\gamma \in \mathcal{U}$:

$$\begin{cases} \partial_t \psi_\gamma(t,x) = -\frac{1}{2}\left(\partial_{xx}\psi_\gamma(t,x) + \gamma(t)\left(\partial_x \psi_\gamma(t,x)\right)^2\right) & \forall (t,x) \in [0,1) \times \mathbb{R} \\ \psi_\gamma(1,x) = |x| \quad \forall x \in \mathbb{R} \end{cases}$$

Then we get the Parisi functional:

$$\mathcal{P}(\gamma) = \psi_\gamma(0,0) - \frac{1}{2}\int_0^1 t\gamma(t)\,\mathrm{d}t$$

And the Parisi measure (here we have the distribution of the μ_P: $\gamma_P(t) = \mu_P[0,t]$) at zero temperature is given by (according to [3]):

$$\gamma_P = \underset{\gamma \in \mathcal{U}}{\arg\min} \mathcal{P}(\gamma)$$

With control optimal theory (as done in [7]), we can show that:

$$\gamma_P = \underset{u \in \mathcal{U}}{\arg\min} \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}\left[\left|\int_0^1 \gamma(t)\alpha_t\,\mathrm{d}t + W_1 + x\right| - \frac{1}{2}\int_0^1 \gamma(t)(\alpha_t^2 + t)\,\mathrm{d}t\right]$$

We apply the Legendre transformation:

$$|x| = \sup_{\theta \in [-1,1]} (x\theta)$$

Then (as we have done a lot of times):

$$\begin{aligned} \mathcal{P}(\gamma) &= \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}\left[\sup_{\theta \in [-1,1]} \theta\left(\int_0^1 \gamma(t)\alpha_t\,\mathrm{d}t + W_1\right) - \frac{1}{2}\int_0^1 \gamma(t)(\alpha_t^2 + t)\,\mathrm{d}t\right] \\ &= \sup_{\alpha \in \mathcal{A}_0} \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\Theta\left(\int_0^1 \gamma(t)\alpha_t\,\mathrm{d}t + W_1\right) - \frac{1}{2}\int_0^1 \gamma(t)(\alpha_t^2 + t)\,\mathrm{d}t\right] \\ &= \sup_{\alpha \in \mathcal{A}_0} \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\int_0^1 \Theta_t\gamma(t)\alpha_t\,\mathrm{d}t + \Theta W_1 - \frac{1}{2}\int_0^1 \gamma(t)(\alpha_t^2 + t)\,\mathrm{d}t\right] \\ &= \sup_{\alpha \in \mathcal{A}_0} \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\int_0^1 \gamma(t)\left(\alpha_t\Theta_t - \frac{1}{2}\alpha_t^2\right)\,\mathrm{d}t + \Theta W_1 - \frac{1}{2}\int_0^1 \gamma(t)t\,\mathrm{d}t\right] \\ &= \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \gamma(t)\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta W_1\right] \end{aligned}$$

So, we want to show that γ_p defined by:

$$\gamma_p = \underset{\gamma \in \mathcal{U}}{\arg\min} \sup_{\Theta \in \mathcal{V}} \mathbb{E}\left[\frac{1}{2}\int_0^1 \gamma(t)\left(\Theta_t^2 - t\right)\,\mathrm{d}t + \Theta W_1\right]$$

Is strictly increasing.

So, the model seams easier because we have not the H. But here we have a constraint $\Theta \in [-1, 1]$. In the case of non zero temperature, we could ignore the constraint because $H(x) = +\infty$ for $x \notin [-1, 1]$

Here, we have $\Theta = \pm 1 \quad a.s.$ We can see that in the variational representation. But an easier way is to see that the sup in the Legendre transformation of $x \mapsto |x|$ is always 1 or -1.

So, it is harder to do a variation method (we have to use Lagrange's multiplier) on Θ.

But we know that $\Theta = \pm 1 \quad a.s.$ So, There exists $A \in \mathcal{F}_1$ such that:

$$\Theta = 2\mathbb{1}_A - 1$$

If we note $\mathbb{E}[\mathbb{1}_A|\mathcal{F}_t] = A_t$, the problem can be rewrite as:

$$\mathcal{P}(\gamma) = \sup_{A \in \mathcal{F}_1} \mathbb{E}\left[\frac{1}{2}\int_0^1 \gamma(t)(4A_t^2 - 4A_t + 1 - t)\,\mathrm{d}t\right] + 2\mathbb{E}\left[\mathbb{1}_A W_1\right]$$

If we define:

$$f(A) = \mathbb{E}\left[\int_0^1 \gamma(t)(A_t^2 - A_t)\,\mathrm{d}t\right] + \mathbb{E}\left[\mathbb{1}_A W_1\right]$$

We want to maximize f. Because the maximizer of f is the maximizer in the expression of $\mathcal{P}(\gamma)$. So, we search for $\tilde{A}$ which verify:

$$\forall B \in \mathcal{F}_1 \quad f(B) - f(\tilde{A}) \leq 0$$

We fixe a little ($\mathbb{P}(H) << 1$) $H \in \mathcal{F}_1$. Let be $H_1 = \tilde{A} \cap H$, $H_2 = H \backslash H_1$. Then, we can look for the variation in f:

$$\begin{aligned} f\left(\left(\tilde{A}\backslash H_1\right) \cup H_2\right) - f(\tilde{A}) &\simeq \mathbb{E}\left[\int_0^1 \gamma(t)(2\tilde{A}_t(H_{2t} - H_{1t}) - (H_{2t} - H_{1t}))\,\mathrm{d}t\right] + \mathbb{E}\left[(\mathbb{1}_{H_2} - \mathbb{1}_{H_1})W_1\right] \\ &\simeq \mathbb{E}\left[(\mathbb{1}_{H_2} - \mathbb{1}_{H_1})\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)\right] \leq 0 \end{aligned}$$

So, $\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)$ is positive in $\tilde{A}$ and negative in $\Omega\backslash\tilde{A}$.

Indeed:

- For $\omega \in \tilde{A}$, we take a small $H \in \mathcal{F}_1 \subset \tilde{A}$ such that $\omega \in H$ and $\mathbb{P}H > 0$. Then we make a variation of f by H.

$$\mathbb{E}\left[-\mathbb{1}_H\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)\right] \leq 0$$

  Because H is small, it means:

$$\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)(\omega) \geq 0$$

- For $\omega \notin \tilde{A}$, we do the same and this time:

$$\mathbb{E}\left[\mathbb{1}_H\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)\right] \leq 0$$

  So:

$$\left(\int_0^1 \gamma(t)(2\tilde{A}_t - 1)\,\mathrm{d}t + W_1\right)(\omega) \leq 0$$

7.2 Ideas on the PDE in v_μ

We have the PDE, for all μ:

$$\begin{cases} \partial_t v_\mu(x,t) - \frac{\beta^2}{2\partial_{xx} v_\mu(x,t)} + \frac{\beta^2}{2}\mu[0,t](x^2-t) = 0 \quad \forall (t,x) \in [0,1) \times \mathbb{R} \\ v_\mu(x,1) = -H(x) \quad \forall x \in \mathbb{R} \end{cases}$$

We have tried to find an explicit solution with the help of power series. And we tried to perturb the PDE in μ.

7.2.1 An explicit solution with the power series

Because of the $t - x^2$ in the equation, we might want to look for a solution of the form:

$$v(x,s) = \sum_{n=1}^{\infty} s^n P_n(x^2-1) + cH(x)$$

With:

$$P_n = \sum_{k=0}^{n} a_k^n X^k$$

And:

$$\gamma(s) = \sum_{n=0}^{\infty} \gamma_n s^n$$

This leads to a tedious calculous. And we find the horrible result:

$$\begin{cases} p = 0 & \sum_{k=0}^{n-2} \gamma_k c_0^{n-k-1} + \sum_{k=0}^{n-1} (k+1) a_0^{k+1} c_0^{n-k} + 2(n+1) c a_1^{n+1} = \gamma_n 2c \\ 1 \le p \le n-2 & \sum_{k=0}^{n-p-2} \gamma_k c_p^{n-k-1} + \sum_{k=0}^{n-p} \gamma_k c_{p-1}^{n-k} + \sum_{k=0}^{n-1} \sum_{i=0}^{p} (k+1) a_i^{k+1} c_{p-i}^{n-k} + 2(n+1) c a_{p+1}^{n+1} = 0 \\ n-1 \le p \le n & \sum_{k=0}^{n-p} \gamma_k c_{p-1}^{n-k} + \sum_{k=0}^{n-1} \sum_{i=0}^{p} (k+1) a_i^{k+1} c_{p-i}^{n-k} + 2(n+1) c a_{p+1}^{n+1} = 0 \end{cases}$$

With:

$$\begin{cases} & c_0^n = 2a_1^n + 8a_2^n \\ 1 \le k \le n-2 & c_k^n = \left(2a_{k+1}^n(k+1)\right)(1+2k) + a_{k+2}(k+1)(k+2) \\ & c_{n-1}^n = (2na_n^n)(1+2(n-1)) \end{cases}$$

We found no way to use this result. That's why we have forsaken this idea.

7.2.2 Perturbation of the PDE

We wanted also to perturb the PDE in μ. We define $w_\mu^{\varepsilon\,\nu} = \frac{1}{\varepsilon}\left(v_{\mu+\varepsilon\nu} - v_\mu\right)$. We find that, $w_\mu^{\varepsilon\,\nu}$ solves the PDE (for small ε):

$$\partial_t w + \frac{\beta^2 w}{2\partial_{xx} v_\mu^2} + \frac{\beta^2}{2}\nu[0,t](x^2-t) = 0$$

With $w(x,1) = 0$.

And we find that $w_\mu^{\varepsilon\,\nu}$ and $-w_\mu^{\varepsilon\,-\nu}$ solves the same PDE (for small ε). Then:

$$w_\mu^{\varepsilon\,\nu} \simeq -w_\mu^{\varepsilon\,\nu} \tag{8}$$

Now we thought we had a contradiction. Because we wanted to say that:

$$\inf_{\mu \in \mathcal{M}} v_\mu = v_{\mu_P} \tag{9}$$

So, $w_{\mu_P}^{\varepsilon\,\nu} \geq 0$ for ν, ε such that: $\mu + \varepsilon\nu \in \mathcal{M}$. And if μ_p is supported on $[0, q]$. We can take whatever ν such that $\nu[0,1] = 0$ and ν supported on $[0, q]$. Which means we have:

$$w_{\mu_P}^{\varepsilon\,\nu}, w_{\mu_P}^{\varepsilon\,-\nu} \geq 0$$

And so, by 8:

$$w_{\mu_P}^{\varepsilon\,\nu} = 0$$

This is absurd so, μ_p can't be supported on $[0, q]$.

But it is false because 9 is wrong. Actually, we have only:

$$\inf_{\mu \in \mathcal{M}} v_\mu(0,0) = v_{\mu_P}(0,0)$$

8 Planning of the internship

Apart from trying to solve this problem, I had the opportunity to follow a lot of conferences about various subjects. But I will just describe bellow the planning of my work on the internship subject.

date	What I have done
17/05 to 28/05	First look to the context of the problem, trying to understand what the difficulties are
31/05 to 11/06	Trying to solve the problem in the case of zero temperature
14/06 to 25/06	Understanding that the problem at non-zero temperature is actually easier and doing somme variation methods to fine conditions on μ_p
28/06 to 02/07	Finding the condition on f and trying to solve the problem.
05/07 to 16/07	Finding the duality between the representations
19/07 to 30/07	Trying to see what the others have done and seeing if there are still some links to be made
02/08 to 13/08	Trying one last time to solve the problem
16/08 to 20/08	Writing the internship report

9 Conclusion

I have really enjoyed my internship. It was hard to find good ideas, but it's pleasant to see my comprehension of the problem increasing with the time. And my internship supervisor always found ideas of how to pursue the work when I had none, so, I was never blocked too long. We haven't solved the problem. Nevertheless, my supervisor and I, have good understanding of the problem. And we hope, the tools we have brought might provide some help to others.

As I wanted, this internship has brought me a good view of what research is. Indeed, I had the opportunity to work, speak, eat with other researchers for three months. The thing I would remember about the researcher's work, it is better to work formally (even if we know it's wrong) in order to find some ideas first, and then we do things rigorously.

This internship gave to me an opportunity to extend my knowledge in stochastic calculus and learn a bit of control optimal stochastic. And due to the various conferences that I could follow seen, I learned a lot about other topics.

To conclude, my supervisor has recently found the articles [12],[4] and [8]. Those are articles that tackle the duality between a control optimal problem and its Legendre transform (that's what we have done). To pursue the work I think it would be interesting to read those articles in order to find other ideas.

References

[1] Antonio Auffinger and Wei Kuo Chen. On properties of Parisi measures. *Probability Theory and Related Fields*, 161(3-4):817–850, 2015.

[2] Antonio Auffinger and Wei-Kuo Chen. The Parisi Formula has a Unique Minimizer. *Communications in Mathematical Physics*, 335(3):1429–1444, 2015.

[3] Antonio Auffinger, Wei Kuo Chen, and Qiang Zeng. The SK model is Full-step Replica Symmetry Breaking at zero temperature. *arXiv*, pages 1–17, 2017.

[4] Jean Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[5] Steven N. Evans and Alexandru Hening. Markov processes conditioned on their location at large exponential times. *Stochastic Processes and their Applications*, 129(5):1622–1658, 2019.

[6] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proceedings of the American Mathematical Society*, 144(7):3135–3150, 2015.

[7] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proceedings of the American Mathematical Society*, 144(7):3135–3150, 2015.

[8] Ioannis Karatzas, John P. Lehoczky, Steven E. Shreve, and Gan Lin Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal on Control and Optimization*, 29(3):702–730, 1991.

[9] Andrea Montanari. Optimization of the Sherrington-Kirkpatrick Hamiltonian. *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, 2019-Novem:1417–1433, 2019.

[10] Dmitry Panchenko. Introduction to the SK model. *Current Developments in Mathematics*, 2014(1):231–291, 2014.

[11] David Sherrington and Scott Kirkpatrick. Solvable Model of a Spin-Glass. *Phys. Rev. Lett.*, 35(26):1792–1796, 1975.

[12] Mathematical Statistics. The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets Author (s): D . Kramkov and W . Schachermayer Source : The Annals of Applied Probability , Aug ., 1999 , Vol . 9 , No . 3 (Aug ., 1999), pp . Published by : In. 9(3):904–950, 1999.

[13] Michel Talagrand. Parisi measures. *Journal of Functional Analysis*, 231(2):269–286, 2006.