PDE Methods for Mean Field Games with Non-Separable Hamiltonian: Data in Sobolev Spaces (Continued)

David Ambrose

June 29, 2018

Drexel

Steps of the energy method

- Introduce an approximate problem.
- Prove existence of solutions for the approximate problem.
- **Prove an estimate for these solutions which is uniform in the approximation parameter.**
- Use this estimate to help pass to the limit as the approximation parameter(s) vanish.
- Prove that the limiting solution solves the original problem.
- Make similar estimates for uniqueness.

The mean field games system

- For reference, we (again) write down the mean field games system we are considering:

$$u_t + \varepsilon H(t, x, m, Du) = -\Delta u, \tag{1}$$

$$m_t + \varepsilon \mathrm{div}(m H_p(t, x, m, Du)) = \Delta m, \tag{2}$$

in the context of the planning problem, i.e., with boundary conditions

$$m(0, \cdot) = m_0, \qquad u(T, \cdot) = u_T. \tag{3}$$

The approximate problems

- We had set up the approximate problems with two approximating parameters, n and δ.
- We had let $\mathbb{P}_\delta$ be a mollifier in the x variable; a good choice is truncation of the Fourier series at the level $1/\delta$.
- We mollified the data:

$$\mu^{n,\delta}(0,\cdot) = \mathbb{P}_\delta \mu_0, \qquad w^{n,\delta}(T,\cdot) = \mathbb{P}_\delta w_T.$$

- We had set up an iterative scheme:

$$w_t^{n+1,\delta} + \varepsilon P\Theta(t,x,\mu^{n,\delta},Dw^{n,\delta}) + \Delta w^{n+1,\delta} = 0,$$

$$\mu_t^{n+1,\delta} + \varepsilon \mathrm{div}\left((\mu^{n,\delta}+\bar{m})\Theta_p(t,x,\mu^{n,\delta},Dw^{n,\delta}) \right) - \Delta \mu^{n+1,\delta} = 0.$$

- We initialized with

$$w^{0,\delta} = \mu^{0,\delta} = 0.$$

We skipped the uniform estimate

- The most important step in the energy method is finding an estimate, uniform in the approximation parameter(s).
- We will now establish that

$$\sup_{t\in[0,T]} \|\mu^{n,\delta}\|_{s-1}^2 + \sup_{t\in[0,T]} \|w^{n,\delta}\|_s^2 + \int_0^T \|\mu^{n,\delta}\|_s^2 + \|w^{n,\delta}\|_{s+1}^2 \, d\tau$$

  is bounded by some constant (independent of n and δ).
- To do this, we will introduce a smallness constraint.

The theorem we are proving

- Remember the theorem I stated last time:

Theorem

Let $T > 0$ and $\varepsilon > 0$ be given. Let $s \geq \left\lceil \frac{d+5}{2} \right\rceil$ and let $\mu_0 \in H^{s-1}(\mathbb{T}^d)$ be such that $\bar{m} + \mu_0$ is a probability measure. Let $u_T \in H^s(\mathbb{T}^d)$ be given. Assume that the conditions **(H1)** *and* **(H2)** *are satisfied. Then there exists $\mu \in L^\infty([0,T]; H^{s-1}) \cap L^2([0,T]; H^s)$ and there exists $u \in L^\infty([0,T]; H^s) \cap L^2([0,T]; H^{s+1})$ such that $\bar{m} + \mu$ is a probability measure for all $t \in [0,T]$, and such that $(u, \bar{m} + \mu)$ satisfies (1), (2), (3). Furthermore, for all $s' \in [0, s)$, we have $\mu \in C([0,T]; H^{s'-1})$ and $u \in C([0,T]; H^{s'})$.*

The hypothesis **(H1)**

- **(H1)** The function $\mathcal{H}$ is such that there exists a non-decreasing function $\tilde{F} : [0,\infty) \to [0,\infty)$ such that for all $\beta \in \mathbb{N}^{2d+1}$ with $|\beta| \leq s+2$,

$$\left|\partial^{\beta}\Theta(\cdot,\cdot,\nu,Dy)\right|_{\infty} \leq \tilde{F}\left(|\nu|_{\infty} + |Dy|_{\infty}\right).$$

- This has the immediate consequence that Θ and its derivatives are bounded in terms of the norm of the solution:

$$\sum_{|\beta|\leq 2} \left\|(\partial^{\beta}\Theta)(t,\cdot,\mu,Dw)\right\|_{s-1}$$
$$\leq \bar{C}F\left(\|\mu\|^{2}_{\lceil \frac{d+1}{2}\rceil} + \|Dw\|^{2}_{\lceil \frac{d+1}{2}\rceil}\right)(1+\|\mu\|_{s-1}+\|w\|_{s})^{s-1}.$$

Our norms to estimate

- For all $n \in \mathbb{N}$, we define M_n and N_n to be

$$M_n = \sup_{t\in[0,T]} \left(\|Dw^n\|_{s-1}^2\right) + \sup_{t\in[0,T]} \left(\|\mu^n\|_{s-1}^2\right),$$

$$N_n = \sum_{1\le|\alpha|\le s} \int_0^T \|\partial^\alpha D w^n\|_0^2 \; d\tau + \sum_{0\le|\alpha|\le s-1} \int_0^T \|\partial^\alpha D\mu^n\|_0^2 \; d\tau.$$

- Our solutions actually depend on δ as well, but we will suppress this dependence for now.
- The N_n characterizes the parabolic gain of regularity; we will focus on estimating M_n and we will get an estimate for N_n as we go.

We begin with μ^n :

$$\frac{d}{dt}\frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)^2\ dx$$

$$=\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\left(\partial^\alpha\Delta\mu^{n+1}\right)\ dx-\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left(D\mu^n\cdot\Theta_p(\cdot,x,\mu^n,Dw^n)\right)\ dx$$

$$-\,\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^{d}\Theta_{x_ip_i}(\cdot,x,\mu^n,Dw^n)\right)\ dx$$

$$-\,\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^{d}\left[\left(\Theta_{qp_i}(\cdot,x,\mu^n,Dw^n)\right)\left(\partial_{x_i}\mu^n\right)\right]\right)\ dx$$

$$-\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^{d}\sum_{j=1}^{d}\left[\left(\Theta_{p_ip_j}(\cdot,x,\mu^n,Dw^n)\right)\left(\partial^2_{x_ix_j}w^n\right)\right]\right)\ dx.$$

We integrate with respect to time and rearrange:

$$
\begin{aligned}
&\frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}(t,x)\right)^2\ dx-\frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}(0,x)\right)^2\ dx+\int_0^t\int_{\mathbb{T}^d}\left|D\partial^\alpha \mu^{n+1}\right|^2\ dxd\tau\\
&\quad=-\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}\right)\partial^\alpha\left(D\mu^n\cdot\Theta_p(\tau,x,\mu^n,Dw^n)\right)\ dxd\tau\\
&\quad-\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\tau,x,\mu^n,Dw^n)\right)\ dxd\tau\\
&\quad-\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\tau,x,\mu^n,Dw^n)\right)\left(\partial_{x_i}\mu^n\right)\right]\right)\ dxd\tau\\
&\quad-\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\sum_{j=1}^d\left[\left(\Theta_{p_ip_j}(\tau,x,\mu^n,Dw^n)\right)\left(\partial^2_{x_ix_j}w^n\right)\right]\right)\ dxd\tau\\
&\quad=I+II+III+IV.
\end{aligned}
$$

Work to do:

- We need to estimate these terms, I, II, III, and IV in terms of M_n and N_n.
- We will actually do some further decomposing: $I = I_A + I_B$, $III = III_A + III_B$, and $IV = IV_A + IV_B$.
- Each of these further decompositions involves separating out the leading term, as in an example in the last lecture.
- This is for μ; we have corresponding work to do for w as well.

The term I_A

- We have the definition

$$I_A = -\varepsilon \int_0^t \int_{\mathbb{T}^d} (\partial^\alpha \mu^{n+1})(\partial^\alpha D\mu^n) \cdot \Theta_p(\tau, x, \mu^n, Dw^n)\ dxd\tau$$

- We pull Θ_p through the integral:

$$I_A \leq \varepsilon |\Theta_p(t, \cdot, \mu^n, Dw^n)|_\infty \int_0^t \int_{\mathbb{T}^d} |\partial^\alpha \mu^{n+1}||\partial^\alpha D\mu^n|\ dxd\tau.$$

- We use the assumption **(H1)**:

$$I_A \leq \varepsilon F(M_n) \int_0^T \int_{\mathbb{T}^d} |\partial^\alpha \mu^{n+1}||\partial^\alpha D\mu^n|\ dxd\tau.$$

- Next we use Young's inequality, with parameter $\sigma_1 > 0$.

The term I_A (continued)

- Using Young's inequality yields the following:

$$I_A \leq \varepsilon F(M_n)\left(\frac{1}{2\sigma_1}\int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \, d\tau + \frac{\sigma_1}{2}\int_0^T \|\partial^\alpha D\mu^n\|_0^2 \, d\tau\right)$$

$$\leq \varepsilon F(M_n)\left(\frac{1}{2\sigma_1}\int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \, d\tau + \frac{\sigma_1}{2}N_n\right).$$

- We choose $\sigma_1 = 28T\varepsilon F(M_n)$. This implies

$$I_A \leq \frac{1}{56T}\int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \, d\tau + 14\varepsilon^2 T(F(M_n))^2 N_n$$

$$\leq \frac{1}{56}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + 14\varepsilon^2 T(F(M_n))^2 N_n.$$

The term I_B

- We have the definition

$$I_B = \varepsilon \int_0^t \int_{\mathbb{T}^d} (\partial^\alpha \mu^{n+1}) \cdot$$
$$[(\partial^\alpha D\mu^n) \cdot \Theta_p(\tau, x, \mu^n, Dw^n) - \partial^\alpha (D\mu^n \cdot \Theta_p(\tau, x, \mu^n, Dw^n))] \; dx d\tau.$$

- We use Young's inequality with parameter $\sigma_2 = 28T\varepsilon$, finding

$$I_B \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right)$$
$$+14\varepsilon^2 T \int_0^T \left\| (\partial^\alpha D\mu^n) \cdot \Theta_p(\tau, \cdot, \mu^n, Dw^n) - \partial^\alpha (D\mu^n \cdot \Theta_p(\tau, \cdot, \mu^n, Dw^n)) \right\|_0^2 d\tau.$$

- We then use our lemma about derivatives of products, and Sobolev embedding, and **(H1)**, finding

$$I_B \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + c\varepsilon^2 T^2 (F(M_n))^2 M_n (1 + M_n)^{s-1}.$$

The term II

- We have the definition

$$II = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) \partial^\alpha \left((\mu^n + \bar{m}) \sum_{i=1}^{d} \Theta_{x_i p_i}(\tau, x, \mu^n, Dw^n) \right) \, dx d\tau$$

- There is no need to extract a leading-order contribution this time.
- We use Young's inequality, the Sobolev algebra property, and **(H1)**, finding

$$II \leq \frac{1}{56} \left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + c\varepsilon^2 T^2 (F(M_n))^2 (1 + M_n)^s.$$

The term III_A

- We have the definition

$$III_A = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right)(\mu^n + \bar{m}) \cdot$$
$$\sum_{i=1}^{d} \left[\left(\Theta_{qp_i}(\tau, x, \mu^n, Dw^n)\right)\left(\partial^\alpha \partial_{x_i} \mu^n\right)\right] \, dxd\tau.$$

- We estimate this by putting μ^n and Θ_{qp_i} in L^∞ using Sobolev embedding and **(H1)**, and we use N_n for $\partial^\alpha \partial_{x_i} \mu^n$.
- The result is

$$III_A \leq \frac{1}{56}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + c\varepsilon^2 T(1+M_n)(F(M_n))^2 N_n.$$

The term III_B

- We have the definition

$$III_B = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) \left\{ \partial^\alpha \left((\mu^n + \bar{m}) \sum_{i=1}^{d} \left[\left(\Theta_{qp_i}(\tau, x, \mu^n, Dw^n) \right) \left(\partial_{x_i} \mu^n \right) \right] \right) - (\mu^n + \bar{m}) \sum_{i=1}^{d} \left[\left(\Theta_{qp_i}(\tau, x, \mu^n, Dw^n) \right) \left(\partial^\alpha \partial_{x_i} \mu^n \right) \right] \right\} dx d\tau.$$

- We estimate this similarly to how we estimated I_B :

$$III_B \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + cT^2 \varepsilon^2 (F(M_n))^2 (1 + M_n)^{s+1}.$$

The term IV_A

- We have the definition

$$IV_A = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) (\mu^n + \bar{m}) \cdot \sum_{i=1}^{d} \sum_{j=1}^{d} \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n)\right) \left(\partial^\alpha \partial^2_{x_i x_j} w^n\right) \right] \, dx d\tau.$$

- As in III_A, we put μ^n and $\Theta_{p_i p_j}$ in L^∞, using Sobolev embedding and **(H1)**, and we use N_n to bound $\partial^\alpha \partial^2_{x_i x_j} w^n$:

$$IV_A \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + c\varepsilon^2 T (1 + M_n)(F(M_n))^2 N_n.$$

The term IV_B

- We have the definition

$$IV_B = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) \left\{ \partial^\alpha \left((\mu^n + \bar{m}) \sum_{i=1}^d \sum_{j=1}^d \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n) \right) \left(\partial^2_{x_i x_j} w^n \right) \right] \right) \right.$$
$$\left. - (\mu^n + \bar{m}) \sum_{i=1}^d \sum_{j=1}^d \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n) \right) \left(\partial^\alpha \partial^2_{x_i x_j} w^n \right) \right] \right\} dx d\tau.$$

- We estimate this similarly to I_B and III_B :

$$IV_B \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + cT^2 \varepsilon^2 (F(M_n))^2 (1 + M_n)^{s+1}.$$

What have we been doing?

- We need to bound the growth of the approximate solutions.
- We have shown how to bound $\mu^{n+1,\delta}$, mainly in terms of the previous iterates.
- The structure in the equations that we have used: considering the MFG system as a system for (m, Du), other than the linear terms, the equations only have first spatial derivatives. The linear terms are parabolic, giving a gain of regularity of one derivative. We have used this gain of regularity to control the first spatial derivatives present in the nonlinear terms.

Using our estimate for $\mu^{n+1,\delta}$

- We have shown

$$I + II + III + IV \leq \frac{1}{8}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + c\varepsilon^2 T(F(M_n))^2\left((1+T)(1+N_n)(1+M_n)^{s+1}\right).$$

- This has as one consequence

$$\frac{1}{2}\|\partial^\alpha \mu^{n+1}(t,\cdot)\|_0^2 \leq \frac{1}{2}\|\partial^\alpha \mu^{n+1}(0,\cdot)\|_0^2 + \frac{1}{8}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + c\varepsilon^2 T(F(M_n))^2\left((1+T)(1+N_n)(1+M_n)^{s+1}\right).$$

Using our estimate for $\mu^{n+1,\delta}$ (continued)

- Another consequence is

$$\int_0^t \|D\partial^\alpha \mu^{n+1}\|_0^2 \, d\tau \leq \frac{1}{2}\|\partial^\alpha \mu^{n+1}(0,\cdot)\|_0^2 + \frac{1}{8}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + c\varepsilon^2 T(F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1}\right).$$

- We take the supremum in time of both of these consequences, and we add, finding

$$\frac{1}{4}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + \int_0^T \|D\partial^\alpha \mu^{n+1}\|_0^2 \, d\tau \leq \|\partial^\alpha \mu^{n+1}(0,\cdot)\|_0^2 + c\varepsilon^2 T(F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1}\right).$$

We do the same thing for $w^{n+1,\delta}$

- We go through the same steps for the w equation, and we estimate V, VI_A, VI_B, VII_A, and VII_B.
- We get a corresponding conclusion for w and we add it to the conclusion for μ, finding

$$\begin{aligned}
&\frac{1}{4}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2+\sup_{t\in[0,T]}\|\partial^\alpha\partial_{x_j}w^{n+1}\|_0^2\right)\\
&\quad+\int_0^T\|D\partial^\alpha\mu^{n+1}(\tau,\cdot)\|_0^2\,d\tau+\int_0^T\|D\partial^\alpha\partial_{x_j}w^{n+1}(\tau,\cdot)\|_0^2\,d\tau\\
&\qquad\le\|\partial^\alpha\mu^{n+1}(0,\cdot)\|_0^2+\|\partial^\alpha\partial_{x_j}w^{n+1}(T,\cdot)\|_0^2\\
&\qquad\quad+c\varepsilon^2T(F(M_n))^2\left((1+T)(1+N_n)(1+M_n)^{s+1}\right).
\end{aligned}$$

Ready for our induction

- We sum over multi-indices and whatnot, we multiply by 4, and we substitute the boundary conditions:

$$M_{n+1} + 4N_{n+1} \leq 4\|\mathbb{P}_{\delta}\mu_0\|_{s-1}^2 + 4\|D\mathbb{P}_{\delta}w_T\|_{s-1}^2 + c\varepsilon^2 T(F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1} \right). \quad (4)$$

- Let $\mathcal{S}$ be a real number such that

$$4\|\mu_0\|_{s-1}^2 + 4\|Dw_T\|_{s-1}^2 \leq \mathcal{S}.$$

- We state our smallness assumption:
  (H2) The function F and the constants c, ε, T, and $\mathcal{S}$ satisfy

$$c\varepsilon^2 T(F(2\mathcal{S}))^2 \left((1+T)(1+2\mathcal{S})^{s+2} \right) \leq \mathcal{S}.$$

Induction proof: the end of the estimate

- Claim: for all n, we have $M_n + 4N_n \leq 2\mathcal{S}$.
- Base case: $\mu^0 = w^0 = 0$, so $M_0 + N_0 = 0$.
- Inductive hypothesis: Assume $M_n + 4N_n \leq 2\mathcal{S}$.
- Notice that $M_n \leq 2\mathcal{S}$ and $N_n \leq 2\mathcal{S}$.
- Also, recall that $\|\mathbb{P}_\delta f\|_{s-1} \leq \|f\|_{s-1}$, for any f.
- We then have

$$M_{n+1} + 4N_{n+1} \leq 4\|\mu_0\|_{s-1}^2 + 4\|Dw_T\|_{s-1}^2 + c\varepsilon^2 T(F(2\mathcal{S}))^2\left((1+T)(1+2\mathcal{S})^{s+2}\right). \quad (5)$$

- So, $M_{n+1} + 4N_{n+1} \leq \mathcal{S} + \mathcal{S} = 2\mathcal{S}$.
- This completes the proof.

About the smallness constraint

- Let's state that smallness constraint again: $\mathcal{S}$ is bigger than a multiple of the data, and we assume

$$c\varepsilon^2 T(F(2\mathcal{S}))^2 \left((1+T)(1+2\mathcal{S})^{s+2} \right) \leq \mathcal{S}.$$

- Given a Hamiltonian and some data, and a time horizon T, this can be enforced by taking ε small.
- Given a Hamiltonian and some data and a value of ε, then this can be enforced by taking T small.
- For some Hamiltonians, it could be the case that $(F(2\mathcal{S}))^2 \to 0$ as $\mathcal{S} \to 0$, faster than $\mathcal{S}$. Then, given values of ε and T, the smallness constraint can be enforced by taking the data sufficiently small.
- So, the smallness constraint simultaneously considers the length of the time horizon, the strength of the coupling in the system, and the size of the data.

Thanks for your attention.