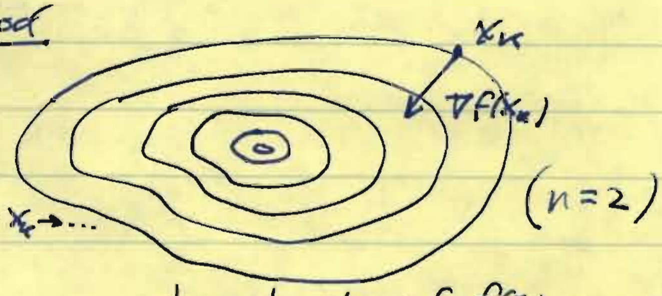
A Variational Principle of the Fokker-Planck Equation

Plan
1. The steepest descent method
2. The variational principle for the FPE
3. The Wasserstein metric
4. Proof of main results

1. The steepest descent method



level sets of $f(x)$

(·) minimize $f(x) : \mathbb{R}^n \to \mathbb{R}$

Iteration: $x_0 \to x_1 \to \cdots \to x_{k-1} \to x_k \to \cdots$

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k \\ d_k = -\nabla f(x_k) / |\nabla f(x_k)| \quad \text{[No need to normalize it in practice.]} \\ f(x_{k+1}) = f(x_k + \lambda_k d_k) = \min\limits_{\lambda} f(x_k + \lambda d_k) \end{cases}$$

(·) ~~Heat flow~~. Why call it the steepest descent?
~~Let $d_k =$~~

$$f(x_k + t d_k) - f(x_k + t d) \le 0 \quad \text{for } \underset{t \sim 0}{t \to 0}? \quad \forall \|d\| = 1.$$

$$\frac{f(x_k + t d_k) - f(x_k + t d)}{t} \le 0 \quad \text{for } t > 0, \ t \ll 1.$$

$$\Longleftrightarrow \quad \frac{[f(x_k + t d_k) - f(x_k)] - [f(x_k + t d) - f(x_k)]}{t} \le 0 \quad \text{fo } 0 < t \ll 1$$

$$\Longleftrightarrow \quad \frac{d}{dt}\Big|_{t=0} f(x_k + t d_k) \le \frac{d}{dt}\Big|_{t=0} f(x_k + t d)$$

$$\nabla f(x_k) \cdot d_k \le \nabla f(x_k) \cdot d$$

Let $a = \frac{\nabla f(x_k)}{|\nabla f(x_k)|}$. $|a|=1$.

$\min_{|\xi|=1} a\cdot\xi$. has a unique sol'n $\xi=-a$

Pf $a\cdot\xi \geq -|a||\xi| = -1$

the Cauchy-Schwarz ineq.

$\xi=-a \Rightarrow a\cdot\xi=-1$.

Uniqueness: Suppose $|\eta|=1$. $a\cdot\eta=-1$.

Then. $|\eta+a|^2 = |\eta|^2+|a|^2+2\eta\cdot a = 2-2=0$

$\Rightarrow \eta=-a$. $\square$

②. ~~Let $I=(0,1)$~~. Consider a nice, bounded domain $\Omega\subset\mathbb{R}^n$. Define

$$E[u] = \frac{1}{2}\int_\Omega |\nabla u|^2 dx$$

Let $\varphi\in C_c^\infty(\Omega)$

$$\begin{aligned}\delta E[u][\varphi] &\equiv \frac{d}{dt}\Big|_{t=0} E[u+t\varphi]\\ &= \frac{1}{2}\int_\Omega \frac{d}{dt}\Big|_{t=0} |\nabla u + t\nabla\varphi|^2 dx\\ &= \int_\Omega \nabla u\cdot\nabla\varphi\, dx\\ &= \int_\Omega (-\Delta u)\varphi\, dx \quad \text{if } u \text{ is smooth.}\end{aligned}$$

Hence $\delta E[u] = -\Delta u$

The steepest descent dynamics for $u=u(x,t)$ is

$$\frac{\partial u}{\partial t} = -\delta E[u] = \Delta u$$

So, the heat eq is the steepest descent w.r.t. the $L^2(\Omega)$-inner product.

Partial check

$$\frac{d}{dt}\int_\Omega \frac{1}{2}|\nabla u(x,t)|^2 dx = \int_\Omega \nabla u(x,t)\cdot \nabla \frac{\partial u}{\partial t}(x,t)\, dx$$

$$= \int_\Omega \nabla u \cdot \nabla(\Delta u)\, dx = -\int_\Omega (\Delta u)^2 dx \le 0.$$

Descent! But, why steepest descent?

Same argument (modulo some details)

$$(?)\quad \lim_{t\to 0} \frac{E[u+t(-\Delta u)]-E[u]}{t} \le \lim_{t\to 0}\frac{E[u+t\varphi]-E[u]}{t}. \quad \forall \varphi.$$

$$\Leftrightarrow (?)\quad \delta E[u][-\Delta u] \le \delta E[u][\varphi]. \quad \forall \varphi.$$

$$\int_\Omega \nabla u \cdot \nabla(-\Delta u)\, dx \le \int_\Omega \nabla u\cdot \nabla\varphi\, dx$$

$$-\int_\Omega (\Delta u)^2 dx \le \int_\Omega \nabla u \cdot \nabla \varphi\, dx$$

$\varphi = \Delta v.$ $\Rightarrow -\int_\Omega (\Delta u)^2 dx \le \int_\Omega \nabla u \cdot \nabla(\Delta v) = -\int_\Omega \Delta u \cdot \Delta v.$

True! $\int_\Omega (\Delta u)^2 dx = \int_\Omega \Delta u\cdot \Delta u\, dx \le \int_\Omega \Delta u \cdot \Delta v\, dx$

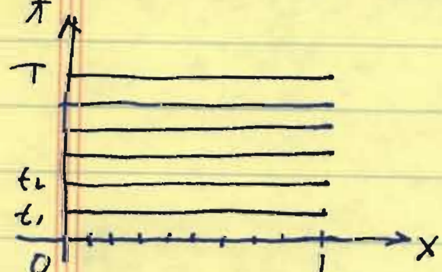
if $\int|\Delta u|^2 = 1.$ $\int |\Delta v|^2 = 1.$

Also: for: $u_t = \Delta u$

$$\frac{d}{dt}\int_\Omega \frac{1}{2}u^2 dx = \int_\Omega u\frac{\partial u}{\partial t}dx = \int_\Omega u\,\Delta u = -\int_\Omega |\nabla u|^2 \le 0.$$

Discretization of $u_t = \Delta u$.

Consider spatial 1-dim problem. $\Omega = (0,1)$.



$u^k(x) \sim u(x, t_k)$

$$\frac{u^{k+1}-u^k}{\Delta t} = \Delta u^{k+1}$$

$$\boxed{u^{k+1} = u^k + \Delta t\, \Delta u^{k+1}}$$ steepest descent!

$-\Delta u^{k+1}$ direction!

$$\min_u \left(\|u-u^k\|^2_{L^2(\Omega)} + \frac{1}{2}\Delta t \int_\Omega |\nabla u|^2 \right)$$

$$\Longrightarrow \quad u^{k+1} = u^k + \Delta t\, \Delta u^{k+1}!$$

2. The variational principle of FPE in $\mathbb{R}^n$

FPE. $\frac{\partial \Psi}{\partial t} = \nabla\cdot(u\nabla\Psi) + \beta^{-1}\Delta\Psi$ $(D=1)$

$\beta^{-1} = k_B T.$

$\Psi : \mathbb{R}\to\mathbb{R}.$ $\left[\Psi\in C^\infty;\ \Psi\geq 0;\ |\Psi'(x)| \leq C(|\Psi(x)|+1)\right]$

$\sim \partial_x \log\Psi(x) \leq C$

$\log\Psi \leq C(|x|+1)$

$\boxed{\Psi\leq C_1 e^{C|x|}}$

Reasonable!

Define $F[u] = \underbrace{\int_{\mathbb{R}^n}\Psi u\,dx}_{E[u]} + \underbrace{\beta^{-1}\int_{\mathbb{R}^n} u\log u}_{\beta^{-1}S(u)}.$

If $u=u(x,t)$ is a soln of the FPE then $\frac{d}{dt}F[u(\cdot,t)]\leq 0.$

Define $K = \{\rho\in L^1(\mathbb{R}^n) : \rho\geq 0.\ \int_{\mathbb{R}^n}\rho\,dx^n = 1,\ M(\rho)<\infty\}$

$M(\rho) = \int_{\mathbb{R}^n}|x|^2\rho(x)\,dx.$

$h=\Delta t.$ $d(\cdot,\cdot)$ some metric — will be described later.

Algorithm (definition):

$u^{k+1}\in K$ is defined as the minimizer of

$I[u] = \frac{1}{2}d(u^{k-1},u)^2 + h\,F(u).$

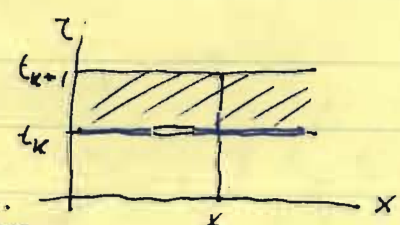
over all $u\in K$

$W(\cdot,\cdot) = d(\cdot,\cdot)$: the Wasserstein metric. $(p=2)$

<u>Lemma</u> $\forall u^{(0)}\in K.\ \exists!\ u\in K$ s.t

$I[u] = \min_{v\in K} I[v].$

Will be proved later



Define $u_h(t) = u^{(k)}(t) = u_h^{(k)}(t)$

piecewise constant in time.

$\Rightarrow$ Define $u_h : \mathbb{R}^n\times(0,\infty)\to[0,\infty).$

Main Theorem

Let $\rho^{(0)} \in K$ w/ $F(\rho^{(0)}) < \infty$. Let $h > 0$ and $\{\rho_h^{(k)}\}$ and ρ_h be constructed as above. Then

$$\rho_h(t) \longrightarrow \rho(t) \text{ in } L^1(\mathbb{R}^n) \quad \forall t (0,\infty),$$

$$\rho_h(t) \longrightarrow \rho \text{ in } L^1((0,T)\times\mathbb{R}^n) \quad \forall T>0,$$

where $\rho \in C^\infty((0,\infty)\times\mathbb{R}^n)$ is the unique sol'n to the FPE

$$\frac{\partial \rho}{\partial t} = \nabla\cdot(\rho\nabla\Psi) + \beta^{-1}\Delta\rho$$

w/ intial condition $\rho(t) \to \rho^{(0)}$ strongly in $L^1(\mathbb{R}^n)$ as $t \to 0$ and $M(\rho), E(\rho) \in L^\infty((0,T))$ $\forall T > 0$.

3. The Wasserstein metric

Assume (Ω, d) is separable and complete.

Let Ω be a metric space with the metric $d(\cdot\cdot)$: $\Omega\times\Omega \to \mathbb{R}$. [(i) $d(x,y) \ge 0$ $\forall (x,y) \in \Omega\times\Omega$, $d(x,y)=0 \iff x=y$. (ii) $d(x,y) = d(y,x)$ $\forall x, y \in \Omega$. (iii) $d(x,y) \le d(x,z) + d(z,y)$ $\forall x,y,z \in \Omega$]. Let Σ be the Borel algebra of Ω. [i.e. Σ is the smallest σ-algebra of Ω that contains all the open subsets of Ω.] Let

$$\mathcal{P} = \{\text{all probability measures } \mu \text{ on } (\Omega,\Sigma) \text{ s.t. } \int_\Omega d(x,y)^2 d\mu(x) < \infty \text{ for some } y \in \Omega \text{ (and hence all)}\}.$$

Let $\mu, \nu \in \mathcal{P}$. Define

$\Pi(\mu,\nu) = \{$all probability measures on $\Omega\times\Omega$ that have marginals μ and ν, resp.$\}$ ($\Sigma\times\Sigma$)

i.e. for any $\pi \in \Pi(\mu,\nu)$,

$$\pi(A\times\Omega) = \mu(A), \quad \pi(\Omega\times A) = \nu(A) \quad \forall A \in \Sigma.$$

[$\Sigma\times\Sigma$ = smallest σ-algebra of $\Omega\times\Omega$ containing all $A\times B$ w/ $A\in\Sigma$ and $B\in\Sigma$.]

Note: The set $\Pi(\mu,\nu) \neq \phi$.

In fact, $\mu \times \nu \in \Pi(\mu,\nu)$ [$\mu\times\nu$: the product measure.]

$$(\mu\times\nu)(Q) = \int_\Omega \nu(Q_x)\,d\mu(x) = \int_\Omega \mu(Q^y)\,d\nu(y)$$

where $Q_x = \{y \in \Omega : (x,y) \in Q\}$

$Q^y = \{x \in \Omega : (x,y) \in Q\}$.

~~Example~~ Define

$$W(\mu,\nu) = \sqrt{\inf_{\pi\in\Pi(\mu,\nu)} \int_\Omega\int_\Omega d(x,y)^2\,d\pi(x,y)}$$

~~Then~~ (~~Under some conditions, (e.g. (Ω,d) is separable~~)

Theorem. $W: \mathcal{P}\times\mathcal{P} \to \mathbb{R}$ is a metric. It is a complete metric. $\square$ Need (Ω,d) to be complete.

True for W_p with $1\le p<\infty$

~~Remark~~ So, ① $W(\mu,\nu)\ge 0$ $\forall \mu,\nu\in\mathcal{P}$. $W(\mu,\nu)=0 \Leftrightarrow \mu=\nu$

② $W(\mu,\nu) = W(\nu,\mu)$ $\forall \mu,\nu\in\mathcal{P}$.

③ $W(\mu,\nu) \le W(\mu,\xi) + W(\xi,\nu)$ $\forall \mu,\nu,\xi\in\mathcal{P}$

$\lim_{k,j\to\infty} W(\mu_k,\mu_j)=0 \Longrightarrow \exists \mu\in\mathcal{P} \ni \lim_{k\to\infty} W(\mu_k,\mu)=0$.

Call $W(\cdot,\cdot)$ the Wasserstein metric (or distance).

Remark. We can generalize it to $W_p(\cdot,\cdot)$ for $1\le p<\infty$:

$$W_p(\mu,\nu) = \left[\inf_{\pi\in\Pi(\mu,\nu)} \int_\Omega\int_\Omega d(x,y)^p\,d\pi(x,y)\right]^{1/p}.$$

<u>Example</u> $\Omega = \{1, 2, \cdots, N\}$. $d(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y. \end{cases}$

$[\Rightarrow d(x,y)^2 = d(x,y)]$ This is the discrete metric.

$\Sigma = 2^\Omega = \{$all subsets of $\Omega\}$. $|\Sigma| = 2^N$.

$\mathcal{P} = \{$all prob. measures on $(\Omega, \Sigma)\}$.

$= \{\mu = (\mu_1, \cdots, \mu_N) \in \mathbb{R}^N : 0 \le \mu_i \le 1, i = 1, \cdots, N. \ \sum_1^N \mu_i = 1\}$

$= \{$all random $\mathbb{R}^N$ vectors$\}$

$\Pi = \{$all prob. measures on $(\Omega \times \Omega, 2^{\Omega \times \Omega})\}$

$= \{\pi = (\pi_{ij}) \in \mathbb{R}^{N \times N} : \sum_{i,j=1}^N \pi_{ij} = 1\}$.

Let $\mu, \nu \in \mathcal{P}$. Then

$$\Pi(\mu,\nu) = \left\{ \pi = (\pi_{ij}) \in \Pi : \sum_{j=1}^N \pi_{ij} = \mu_i, \ 1 \le i \le N, \ \sum_{i=1}^N \pi_{ij} = \mu_j \ \ 1 \le j \le N \right\}$$

$\Pi(\mu,\nu) \neq \phi$ since $\mu \otimes \nu \in \Pi(\mu,\nu)$

$\mu \otimes \nu = (\mu_i \nu_j)_{i,j=1}^N$.

$$\begin{array}{cc} \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} & \begin{array}{l} \to \mu_1 \\ \to \mu_i \\ \to \mu_N \end{array} \\ \begin{array}{ccc} \downarrow & \downarrow & \downarrow \\ \nu_1 & \nu_i & \nu_N \end{array} & \end{array}$$

The Wasserstein metric is:

$$W(\mu,\nu) = \sqrt{\inf_{\pi \in \Pi(\mu,\nu)} \sum_{i,j=1}^N d(i,j)^2 \pi_{ij}}$$

$$W(\mu,\nu)^2 = \inf_{\pi \in \Pi(\mu,\nu)} \sum_{i \neq j} \pi_{ij}$$

$$= \inf_{\pi \in \Pi(\mu,\nu)} (1 - \operatorname{tr} \pi)$$

Clearly $\exists \pi^{\mu,\nu} \in \Pi(\mu,\nu)$ such that $W(\mu,\nu) = 1 - \operatorname{tr} \pi^{\mu,\nu}$

① $W(\mu,\nu) \geq 0$. $W(\mu,\nu) = 0 \Longrightarrow 1 - \operatorname{tr}\pi^{\mu,\nu} = 0$

$\Longrightarrow \sum_{i=1}^{N} \pi_{ii}^{\mu,\nu} = 1$. Since $\sum_{i,j=1}^{N} \pi_{ij}^{\mu\nu} = 1$, all $\pi_{ij}^{\mu\nu} \geq 0$.

we have all $\pi_{ij}^{\mu,\nu} = 0$ if $i \neq j$

$$\begin{array}{ll} \begin{bmatrix} * & & \\ & * & \\ & & * \end{bmatrix} & \to \mu_1 \\ \downarrow & \\ \nu_1 & \end{array} \qquad \begin{aligned} &\mu_i = \pi_{ii}^{\mu,\nu} = \nu_i \quad \forall i: 1 \leq i \leq N \\ &\Longrightarrow \mu = \nu. \end{aligned}$$

If $\mu = \nu$. then $\pi^{\mu,\nu}$ can be chosen as

$$\pi_{ij}^{\mu\nu} = \delta_{ij}\mu_i \quad (i,j = 1, \cdots, N)$$
$$= \operatorname{diag}(\mu_1, \cdots, \mu_N)$$

Then. $W(\mu,\nu) = 1 - \operatorname{tr}\pi^{\mu,\nu} = 0$.

Hence $W(\mu,\nu) = 0 \Longleftrightarrow \mu = \nu$.

② $W(\mu,\nu) = W(\nu,\mu)$. $\pi^{\nu,\mu} = (\pi^{\mu,\nu})^T$ ← transpose.

③ $W(\mu,\nu) \leq W(\mu,\xi) + W(\xi,\nu)$ — not easy.

Homework ?

<u>Example 2</u>. More realistic and useful

$\Omega = \mathbb{R}^n$. d = Euclid distance. $\mathcal{B}$ or Σ: Borel σ-alg.

$\mathcal{P}$, Π: same as in the general setting.

$\mathcal{P} = \{$ all prob. Borel measures μ on $\mathbb{R}^n$ s.t. $\int_{\mathbb{R}^n} |x|^2 d\mu(x) < \infty \}$

$\Pi = \{$ all prob. Borel measures on $\mathbb{R}^n \times \mathbb{R}^n \}$

$\forall \mu, \nu \in \mathcal{P}$. The Wasserstein distance is

$$W(\mu,\nu)^2 = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |x-y|^2 d\pi(x,y).$$

<u>Thm</u> $W(\cdot,\cdot)$ is a complete metric of $\mathcal{P}$. $\square$

<u>Some basic properties</u>

Let $\mu, \nu \in \mathcal{P}$.

① $\exists \pi^{\mu,\nu} \in \Pi(\mu,\nu)$ s.t.

$$W(\mu,\nu) = \sqrt{\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |x-y|^2 d\pi^{\mu,\nu}(x,y)}$$

② Let $\mu_k \in \mathcal{P}$ $(k=1,2,\cdots)$, $\mu \in \mathcal{P}$. Then

$$W(\mu_k,\mu) \to 0 \iff \begin{cases} \mu_k \xrightarrow{*} \mu \\ \int_{\mathbb{R}^n} |x|^2 d\mu_k \to \int_{\mathbb{R}^n} |x|^2 d\mu . \end{cases}$$

Here,

$\mu_k \xrightarrow{*} \mu$ means: $\int_{\mathbb{R}^n} \varphi \, d\mu_k \to \int_{\mathbb{R}^n} \varphi \, d\mu \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$.

③ Let $\mu_k, \nu_k, \mu, \nu \in \mathcal{P}$ $(k=1,2,\cdots)$. Then

$\mu_k \xrightarrow{*} \mu$, $\nu_k \xrightarrow{*} \nu \Rightarrow W(\mu,\nu) \le \liminf_{k\to\infty} (\mu_k, \nu_k)$.

i.e. $W(\cdot,\cdot)$ is weak-* lower semicontinuous.

④ Let $\mu \in \mathcal{P}$, $a \in \mathbb{R}^n$. Then

$$W(\mu, \delta_a) = \int_{\mathbb{R}^n} |x-a|^2 d\mu(x)$$

⑤ Let $\theta > 0$, $f \in L^1(\mathbb{R}^n)$, $0 \le f$, $\int_{\mathbb{R}^n} f\,dx = 1$, $\int |x|^2 f\,dx < \infty$ and

$S_\theta[f] = \theta^{n/2} f(\sqrt{\theta}\, x)$. Then

$$W(S_\theta[f], S_\theta[g]) = \frac{1}{\sqrt{\theta}} W(f,g).$$

Here $W(f,g) = W(f\,dx, g\,dx)$.

⑥ $\forall \theta > 0$. $M_\theta := \{\mu \in \mathcal{P} : \int_{\mathbb{R}^n} |x|^2 d\mu = \theta\}$ is closed w.r.t. $W(\cdot,\cdot)$.

Remarks on proofs of these results.

⊙ Symmetry: $W(\mu,\nu)$.

For any $\pi \in \Pi(\mu,\nu)$, we can construct $\pi^T \in \Pi(\nu,\mu)$ by $\pi^T(A\times B) = \pi(B\times A)$ $\forall A, B \in \mathcal{B}$.

⊙ $\mu=\nu \Rightarrow W(\mu,\nu)=0$.

Define $F[\phi] = \int_{\mathbb{R}^n} \phi(x,x)\,d\mu(x)$ $\forall \phi \in C_c^\infty(\mathbb{R}^n\times\mathbb{R}^n)$.

Clearly, F is linear and

$$|F[\phi]| \le \|\phi\|_\infty$$

Riesz's Thm $\Rightarrow \exists\, \pi \in \mathcal{P}(\mathbb{R}^n\times\mathbb{R}^n)$ s.t.

$$F[\phi] = \iint \phi(x,y)\,d\pi(x,y)$$

Hence. $\iint \phi(x,y)\,d\pi(x,y) = \iint \phi(x,y)\,d\pi$ $\forall \phi$

$\Rightarrow \pi$ is concentrated on $\{x=y\}$.

Hence ~~$\Rightarrow$~~ $W(\mu,\nu)^2 \le \iint |x-y|^2\,d\pi = 0$.

Also, $\phi(x,y)=f(x)$

$$F[\phi] = \iint f(x)\,d\pi$$

$$F[\phi] = \int \phi(x,x)\,d\mu(x) = \int f(x)\,d\mu$$

$$\Rightarrow \iint f(x)\,d\pi = \int f(x)\,d\mu$$

$\Rightarrow$ projection of π is μ.

⊙ The triangle inequalty is nontrivial.

[cf. P. Clement & W. Desch Proc. AMS 2008, pp 333–339.] for an elementary proof of the triangle inequality]

⊙ Attainment of $\pi \in \Pi(\mu,\nu)$.

Let $\mu,\nu \in \mathcal{P}$ Then $\exists\, \pi \in \Pi(\mu,\nu)$ s.t.

$$W(\mu,\nu)^2 = \iint |x-y|^2\,d\pi(x,y)$$

<u>Pf</u>. $\exists\, \pi_k \in \Pi(\mu,\nu)$ s.t.

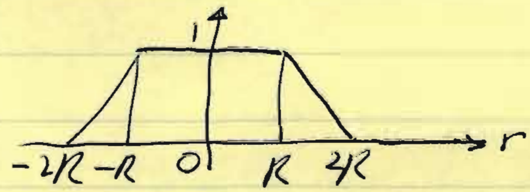
~~$W(\mu,\nu)^2 = \lim_{k\to\infty} \iint |x-y|^2\,d\pi_k(x,y)$~~

In ~~fact~~

$$W(\mu,\nu)^2 \le \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\, d\pi_k(x,y) \le W(\mu,\nu)^2 + \frac{1}{k}$$
$(k=1,\cdots)$

Up to a subseq. $\pi_k \xrightarrow{*} \pi \in \Pi(\mu,\nu)$

Let $R>0$. $\eta_R(r) = \eta_R(|x|)$:



$$W(\mu,\nu)^2 + \frac{1}{k} \ge \int |x-y|^2\, d\pi_k \to \int \cancel{|x-y|^2\, d\pi}$$

$$\ge \int \eta_R(x)\,\eta_R(y)\,|x-y|^2\, d\pi_k \;\cancel{\to \int \eta_R(x)\,\eta_R(y)\,|x-y|^2\, d\pi}$$

$k\to\infty \implies \quad W(\mu,\nu)^2 \ge \int \eta_R(x)\,\eta_R(y)\,|x-y|^2\, d\pi$

$R\to\infty \implies \quad W(\mu,\nu)^2 \ge \int |x-y|^2\, d\pi$.

$\odot$ $W(\mu,\nu)=0 \implies \exists \pi\in\Pi(\mu,\nu) \ni \iint |x-y|^2\, d\pi = 0$

$\implies \pi$ is concentrated on $\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n : \cancel{\phi} x=y\}$.

$\implies \iint \phi(x,y)\, d\pi = \int \phi(x,x)\, d\pi_x$

In particular, $\phi(x,y) = f(x)$

$\implies \iint f(x)\, d\pi = \int f(x)\, d\pi_x \implies \pi_x = \mu$.

Also, $\pi_y = \nu$. $\implies \mu = \nu$.

$\odot$ Also: $\iint |x-y|^2 d\pi = \iint |x|^2 d\pi + \iint |y|^2 d\pi - 2\iint x\,y\, d\pi$

$= \int |x|^2 d\mu + \int |y|^2 d\nu - 2\left(\int x\, d\mu\right)\left(\int y\, d\nu\right) \ge 0$

or: $\mu(A) = \mu(A\times\mathbb{R}^n) = \int_{\mathbb{R}^n} \chi_A(x)\, d\mu(x)$

$= \int_{\mathbb{R}^n\times\mathbb{R}^n} \chi_A(x)\, d\pi(x,y) = \int_{\{x=y\}^c} + \int_{\{x=y\}}$

$= \int_{\{x=y\}} \chi_A(x)\, d\pi(x,y) = \int_{\{x=y\}} \chi_A(y)\, d\pi(x,y)$

$= \iint_{\mathbb{R}^n\mathbb{R}^n} \chi_A(y)\, d\pi(x,y) = \int_{\mathbb{R}^n} \chi_A(y)\, d\nu(y) = \nu(A)$. $\forall A$.

① $\mu_k, \nu_k, \mu, \nu \in \mathcal{P}$ ~~$\lim W(\mu_k, \nu_k)$~~. $\mu_k \xrightarrow{*} \mu$, $\nu_k \xrightarrow{*} \nu$

$$W(\mu,\nu) \le \liminf_{k\to\infty} W(\mu_k,\nu_k)$$

Pf. Let $\alpha = \liminf_{k\to\infty} W(\mu_k,\nu_k) = \lim_{k\to\infty} W(\mu_k,\nu_k)$ (up to a subseq.).

$\exists \pi_k \in \Pi(\mu_k,\nu_k) \ni W(\mu_k,\nu_k)^2 = \iint |x-y|^2 d\pi_k$

Now, up to a subseq. $\pi_k \xrightarrow{*} \pi \in \Pi(\mu,\nu)$.

$$\left[\begin{array}{l} \pi_k(A\times\mathbb{R}^n) = \mu_k(A) \longrightarrow \mu(A) \Longrightarrow \pi(A\times\mathbb{R}^n) = \mu(A) \\ \quad\searrow \pi(A\times\mathbb{R}^n) = \end{array}\right]$$

$$W(\mu,\nu)^2 \le \iint |x-y|^2 d\pi.$$

Only need:
$$\liminf_{k\to\infty} \iint |x-y|^2 d\pi_k \ge \iint |x-y|^2 d\pi.$$

Cut-off η_R.

$$\iint |x-y|^2 d\pi_k \ge \iint \eta_R(x)\eta_R(y)|x-y|^2 d\pi_k$$
$$\longrightarrow \iint \eta_R(x)\eta_R(y)|x-y|^2 d\pi$$
$$\Rightarrow \liminf_k \iint |x-y|^2 d\pi_k \ge \iint \eta_R(x)\eta_R(y)|x-y|^2 d\pi$$

Let $R\to\infty$.
$$\liminf_k W(\mu_k,\nu_k)^2 \ge W(\mu,\nu). \quad \square$$

4. Proof of the main theorem on the convergence to the solution to the PPE.

Proposition Given $\rho^{(0)} \in K$. $h > 0$. $\Rightarrow \exists! \rho \in K$ s.t.

$$I[\rho] = \min_{\tilde{\rho} \in K} I[\tilde{\rho}]$$

$$I[\tilde{\rho}] = \frac{1}{2} W(\tilde{\rho}, \rho^{0})^2 + h F[\rho]$$

$$F[\rho] = \int \Psi \rho + \beta^{-1} \int \rho \log \rho.$$

Note: Replace μ, ν by $f dx$, $g dx$.

$f \in L^1(\mathbb{R}^n)$, $f \geq 0$, $\int f dx = 1$, $\int |x|^2 f(x) dx < \infty$.