Selected Topics in Stochastic Analysis

Lijun Bo

School of Mathematical Sciences
University of Science and Technology of China
lijunbo@ustc.edu.cn

This is a course for Ph.D. candidates majored in probability theory

立志用功如种树然，方其根芽，犹未有干；及其有干，
尚未有枝；枝而后叶，叶而后花实。初种根时，只管栽
培灌溉。勿作枝想。勿作叶想。勿作花想。勿作实想。
悬想何益！但不忘栽培之功，怕没有枝叶花实？
-------王阳明

Course Outline

Course Outline

Treasure Box

- I would recommend you an important reference on a deep introduction of stochastic analysis tools:

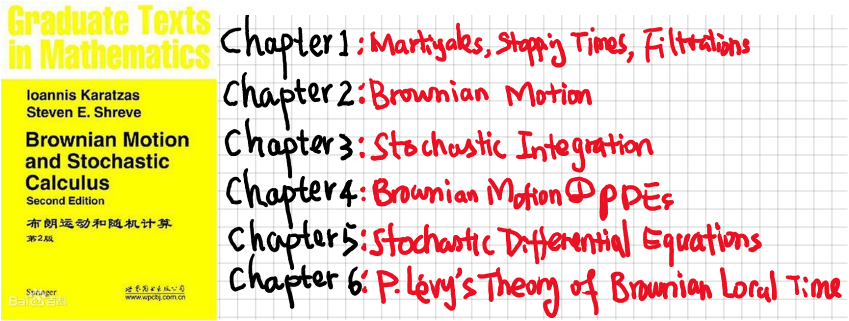


Figure: A Magic Book on Stochastic Analysis

Course Outline

Well-Posedness for SDEs I

- Let $D \subseteq \mathbb{R}^n$ be a domain, i.e., an open connected subset of $\mathbb{R}^n$
- Consider functions $b : D \to \mathbb{R}^{n\times 1}$ and $\sigma : D \to \mathbb{R}^{n\times m}$
- Filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual conditions
- Let $W = (W_t)_{t\in[0,T]}$ be an m-dimensional $(P, \mathbb{F})$-Brownian motion
- An Itô SDE can be described as: for $(t, x) \in [0, T] \times D$,

$$X_t^x = x + \underbrace{\int_0^t b(X_s^x)ds}_{\text{FV part}} + \underbrace{\int_0^t \sigma(X_s^x)dW_s}_{\text{Itô stoch. integral}} \tag{1}$$

- Let $D = \mathbb{R}^n$. Assumption on (b, σ):

($\mathbf{A}_{lip}$) $b : \mathbb{R}^n \to \mathbb{R}^{n\times 1}$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are Lipschitiz continuous with linear growth.

Well-Posedness for SDEs II

- Let $\mathbb{F}^W$ be the filtration generated by Brownain motion W
- Then, the existence and uniqueness of strong solutions of SDE is given by:

Theorem (Well-posedness of SDEs with Strong Solutions I)

Let $(\mathbf{A}_{lip})$ be satisfied. Then, for any $T > 0$, there exists a unique $\mathbb{F}^W$-adapted, continuous solution $X^x = (X_t^x)_{t\in[0,T]}$ satisfying

$$\|X^x\|_T^p := E\left[\sup_{t\in[0,T]} |X_t^x|^p\right] < +\infty, \quad p \geq 1.$$

- Proof. For $p \geq 1$, let $\mathcal{X}_x^p$ be the set of $\mathbb{F}^W$-adapted, continuous processes $X = (X_t)_{t\in[0,T]}$ with $X_0 = x$ satisfying $\|X\|_T < +\infty$.
- Then $(\mathcal{X}_x^p, \|\cdot\|_T)$ is a Banach space.

Well-Posedness for SDEs III

- Define a mapping T on $\mathcal{X}_x^p$ as: for any $X \in \mathcal{X}_x^p$,

$$(TX)_t := x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \ \ t \in [0, T].$$

- Clearly, $(TX)_0 = x$ and $t \to (TX)_t$ is continuous using the continuity of stochastic integrals. Using Itô formula, the linear growth condition of (b, σ), BDG inequality and Grownall's lemma, $\|TX\|_T^p < +\infty$. Hence $TX \in \mathcal{X}_x^p$.
- For any $X, Y \in \mathcal{X}_x^p$, using the assumption $(\mathbf{A}_{lip})$ and the similar argument above, we have

$$\|TX - TY\|_T \leq C_{T,p}\|X - Y\|_T,$$

where $C_{T,p}$ is a positive constant such that $T \to C_{T,p}$ is a non-decreasing function satisfying $\lim_{T\downarrow 0} C_{T,p} = 0$.

Well-Posedness for SDEs IV

- Choose $T = t_0$ small enough such that $C_{T,p} < 1$, we have a unique fixed point $X^* = TX^*$ on $[0, t_0]$.
- Since $C_{T,p}$ depends on T, p only, we can divide $[0, T]$ into infinitely many small time intervals. In each small interval, we have a unique fixed point and fit them together on $[0, T]$.
- The proof of Theorem 1 is complete.
- In many cases, (b, σ) may be *not globally* Lipschitz continuous
- We impose the following conditions:

($\mathbf{A}_{loc}$) $b : \mathbb{R}^n \to \mathbb{R}^{n\times 1}$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are locally Lipschitz continuous.
Remark: **($\mathbf{A}_{loc}$)** implies the pathwise uniqueness of SDE.

($\mathbf{A}_{Lyn}$) There exists a function $V : \mathbb{R}^n \to \mathbb{R}_+$ satisfying

(i) Let $q_R := \inf_{|x|>R} V(x)$, then $\lim_{R\to\infty} q_R = +\infty$;

Well-Posedness for SDEs V

(ii) There exists a constant $C > 0$ such that, for all $t \in [0, T]$,

$$E[V(X^x_{t\wedge\tau_R})] \leq V(x) + C\int_0^t (1 + E[V(X^x_{s\wedge\tau_R})])ds.$$

Here $\tau_R := \inf\{t \in [0, T];\ |X^x_t| \geq R\}$ and $\tau_R = T$ if the set is empty.

Theorem (Well-posedness of SDEs with Strong Solutions II)

Let assumptions $(\mathbf{A}_{loc})$ *and* $(\mathbf{A}_{Lyn})$ *be satisfied. Then, for any* $T > 0$, *there exists a unique* $\mathbb{F}^W$*-adapted, continuous solution of SDE* (1).

- Proof. By Theorem 1, using the assumption $(\mathbf{A}_{loc})$, we have, for any $t \in [0, T]$ and $R > 0$, SDE (1) has a unique continuous strong solution on $[0, t \wedge \tau_R]$.

Well-Posedness for SDEs VI

- We prove $\tau_R \to T$ as $R \to \infty$. In fact, by the condition (ii) of $(\mathbf{A}_{Lyn})$, we have, for all $t \in [0, T]$,

$$E[V(X^x_{t\wedge\tau_R})] \leq e^{Ct}(1 + V(x)).$$

- Therefore, for all $t \in [0, T]$, and $R > 0$,

$$\begin{aligned} P(\tau_R < t) &\leq \frac{1}{q_R} E\left[\mathbb{1}_{\tau_R<t} V(X_{\tau_R})\right] = \frac{1}{q_R} E\left[\mathbb{1}_{\tau_R<t} V(X_{t\wedge\tau_R})\right] \\ &\leq \frac{e^{Ct}}{q_R}(1 + V(x)). \end{aligned}$$

- This yields that $P(\tau_R < t) \to 0$ as $R \to \infty$, using the condition (i) of $(\mathbf{A}_{Lyn})$.
- Then, the proof of Theorem 2 is complete.

Well-Posedness for SDEs VII

- Question: Let conditions of Theorem 1 hold and $Z = (Z_t)_{t\in[0,T]}$ an m-dimensional continuous semimartingale. Prove that the following SDE:
$$X_t^x = x + \int_0^t \sigma(X_s^x)dZ_s, \quad t \in [0, T],$$
admits a unique $\mathbb{F}^Z$-adapted, continuous solution for $x \in \mathbb{R}^n$.
- Hints: Consider the canonical decomposition of the continuous semimartingale Z given by $Z = M + A$, where $M \in \mathcal{M}_{loc}$ and $A \in \mathcal{V}$
  - Firstly, consider $dA << dt$ and $d[M, M] << dt$
  - Secondly, consider using the time change

Examples for SDEs I

- Linear SDE. A general linear SDE can be written as:

$$X_t = Y_t + \int_0^t X_s dZ_s, \quad t \in [0, T], \tag{2}$$

where both $Y = (Y_t)_{t\in[0,T]}$ and $Z = (Z_t)_{t\in[0,T]}$ are one-dimensional continuous semimartingales

Lemma (Closed-Form of Solutions of Linear SDE)

The linear SDE (2) *admits a closed-form solution given by*

$$X_t = \mathcal{E}(Z)_t \left(Y_0 + \int_0^t \mathcal{E}(Z)_s^{-1}(dY_s - d[Y, Z]_s) \right),$$

where $\mathcal{E}(Z)$ is Doléans-Dade exponential of the semimartingale Z.

Examples for SDEs II

- Proof. Let us first consider the simple linear case:

$$X_t^1 = 1 + \int_0^t X_s^1 dZ_s, \quad t \in [0, T].$$

Then, the solution is Doléans-Dade exponential of Z, i.e., $X_t^1 = \mathcal{E}(Z)_t$

- Question: Write the expression of $\mathcal{E}(Z)$.
- Consider the solution of the linear SDE admitting the form:

$$X_t = X_t^1 L_t, \quad t \in [0, T], \tag{3}$$

where $L = (L_t)_{t\in[0,T]}$ is a continuous semimartingale with $L_0 = Y_0$.

Examples for SDEs III

- By integration by parts, we have

$$\begin{aligned} dX_t = d(X_t^1 L_t) &= L_t dX_t^1 + X_t^1 dL_t + d[X^1, L]_t \\ &= X_t^1 L_t dZ_t + X_t^1 dL_t + d[X^1, L]_t \\ &= X_t dZ_t + X_t^1 dL_t + d[X^1, L]_t \end{aligned}$$

- Compare it with (2), i.e., $dX_t = X_t dZ_t + dY_t$, we obtain

$$dY_t = X_t^1 dL_t + d[X^1, L]_t, \quad L_0 = Y_0.$$

- This gives that

$$dL_t = (X_t^1)^{-1} dY_t - \underbrace{(X_t^1)^{-1} d[X^1, L]_t}_{\text{FV part of } L}, \quad L_0 = Y_0. \tag{4}$$

Examples for SDEs IV

- Therefore, it holds that

$$d[X^1, L]_t = d\left[\int_0^\cdot X_s^1 dZ_s, \int_0^\cdot (X_s^1)^{-1} dY_s\right]_t = d[Y, Z]_t.$$

- Using (4), we get

$$\begin{aligned} dL_t &= (X_t^1)^{-1}(dY_t - d[Y, Z]_t) = \mathcal{E}(Z)_t^{-1}(dY_t - d[Y, Z]_t) \\ L_0 &= Y_0. \end{aligned}$$

- Then, we arrive at

$$L_t = Y_0 + \int_0^t \mathcal{E}(Z)_s^{-1}(dY_s - d[Y, Z]_s), \quad t \in [0, T].$$

- Thus, the proof of Lemma 3 is complete.

Examples for SDEs V

- The above linear SDE includes many important examples of stochastic models in practice
- Let $W = (W_t)_{t\in[0,T]}$ be a scalar Brownian motion, i.e., $m = 1$
- Ornstein-Uhlenbeck (OU) process. The OU process can be described as follows: for $x \in \mathbb{R}$,

$$X_t^x = x + \int_0^t \alpha(\beta - X_s^x)ds + \sigma W_t, \quad t \in [0, T]$$

where $\alpha, \sigma > 0$ and $\beta \in \mathbb{R}$.

- OU process is a class of important stochastic models in physics (Langevin Equation) and finance (Vasicek Model)

Examples for SDEs VI

- The discretization of OU processes is a AR(1) model: for $\beta = 0$,

$$\begin{aligned} dX_t^x &= -\alpha X_t^x dt + \sigma dW_t \\ X_{t+1}^x - X_t^x &= -\alpha X_t^x + \sigma(W_{t+1} - W_t) \\ X_{t+1}^x &= (1-\alpha)X_t^x + \xi, \quad \xi \sim N(0, \sigma^2) \end{aligned}$$

- The OU admits a closed-form solution:

$$X_t^x = xe^{-\alpha t} + \beta(1 - e^{-\alpha t}) + \int_0^t \sigma e^{\alpha(s-t)} dW_s$$

- The OU process is both continuous Gaussian process and semimartingale:
  - Mean function:

$$E[X_t^x] = xe^{-\alpha t} + \beta(1 - e^{-\alpha t}).$$

  - OU process is mean-reverting since $\lim_{t\to\infty} E[X_t^x] = \beta \in \mathbb{R}$.

Examples for SDEs VII

- Covariance function:

$$\mathrm{Cov}(X_t^x, X_s^x) = \frac{\sigma^2}{2\alpha}\left(e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\right).$$

- The sample paths of OU-processes show a mean-reverting property:

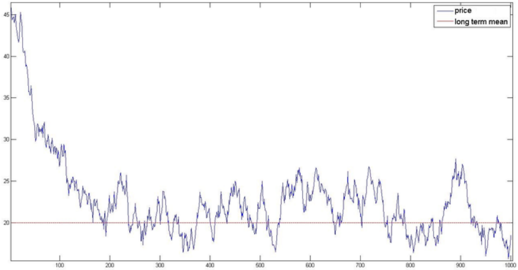


Figure: Sample path of OU processes

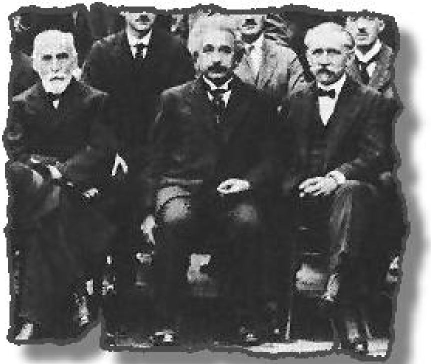
Examples for SDEs VIII

- Questions: Provide an example which is a continuous Gaussian process, but *not* a semimartingale
- Langevin equation. $dV_t = -\alpha V_t dt + dW_t$. The closed-form solution is given by

$$V_t = e^{-\alpha t}\left(V_0 + \int_0^t e^{\alpha s} dW_s\right)$$

  - $\int_0^t V_s ds$: it is used by physicists as a model of physical Brownian motion
  - Recently, Langevin equation and its variation are used to improve the performance of SGD algorithm in Machine Learning and Non-Convex Optimization Problem.
- Paul Langevin (1872-1946): French physicist, Student of French physicist Pierre Curie (1859-1906):

Examples for SDEs IX



Lorentz, Einstein and Langevin in 1927

Figure: Left: Hendrik Antoon Lorentz (1853-1928); Middle: Albert Einstein (1879-1955); Right: P. Langevin

Examples for SDEs X

- Geometric Brownian motion. $dS_t = \mu S_t dt + \sigma S_t dW_t$, $S_0 = x \in D = (0, \infty)$.
- The GBM admits a closed-form solution given by

$$S_t = x \exp\left(\int_0^t (\mu - \frac{\sigma^2}{2}) ds + \sigma W_t\right) = x e^{\mu t} \mathcal{E}(\sigma W)_t$$

- In mathematical finance, GBM is called continuous time Black-Scholes stock model.
- Myron Samuel Scholes (1941-): Canadian-American financial economist, Frank E. Buck Professor of Finance, Emeritus, at the Stanford Graduate School of Business, 1997 Nobel Prize Winner in Economics:

Examples for SDEs XI

Figure: Left: Robert C. Merton (1944-); Right: M. Scholes

Examples for SDEs XII

- The discretization of BS model has a well-known explanation in finance:

$$\underbrace{\frac{S_{t_{i+1}}-S_{t_i}}{S_{t_i}}}_{\text{Stock Return}} = \underbrace{\mu\Delta t_i+\sigma\xi_i}_{\text{Return+Vol. Risk}}, \quad \xi_i := W_{\Delta t_i} \sim N(0,\Delta t_i)$$

  where $\Delta t_i := t_{i+1} - t_i$
- Inhomogeneous GBM. $dX_t = (\theta - aX_t)dt + \sigma X_t dW_t$
- It is also referred to GARCH model (see Lewis (2000))
- Question: Write the closed-form solution of GARCH model.
- Inverse GARCH model. $dX_t = (\theta - aX_t)X_t dt + \sigma X_t dW_t$ where $a, \sigma > 0$ and $\theta > \sigma^2$

Examples for SDEs XIII

- Question: In the above Inverse GARCH model, note that $b(x) := (\theta - ax)x$ is only locally Lipschitz continuous. Prove the existence and uniqueness of strong solutions of the above SDE.
- Question: Let $Y_t = (X_t)^{-1}$ with X_t is the above inverse GARCH model. Then Y_t is an GARCH model. This is the reason why we call X an inverse GARCH model.

Yamada-Watanabe SDEs I

- We now let $W = (W_t)_{t \in [0,T]}$ be a scalar Brownian motion.
- We consider an example introduced by Itô and Watanabe (1978):

$$X_t = \int_0^t 3X_s^{\frac{1}{3}} ds + \int_0^t 3X_s^{\frac{2}{3}} dW_s.$$

- Then $X_t = 0$ and $X_t = W_t^3$ are two different solutions, i.e., uniqueness does not hold.
- The coefficients $b(x) = 3x^{\frac{1}{3}}$ and $\sigma(x) = 3x^{\frac{2}{3}}$, although continuous in x, are *not* smooth at $x = 0$. They are not locally Lipschitz continuous.
- The continuously differentiable functions are locally Lipschitz continuous. However, $f(x) = |x|$ is locally Lipschitz but not continuously differentiable.

Yamada-Watanabe SDEs II

- Consider the Lipschitz continuous function $b:\mathbb{R}\to\mathbb{R}$ satisfying $b(0)\geq 0$.

Lemma (Well-posedness of Yamada-Watanabe SDEs)

Let $p\geq\frac{1}{2}$. For $x,\sigma>0$, the following one-dimensional SDE:

$$X_t^x = x+\int_0^t b(X_s)ds+\sigma\int_0^t (X_s)^p dW_s,\quad t\in[0,T]$$

admits a unique (nonnegative) strong solution.

- Proof. This is a corollary of Yamada and Watanabe (1971).

Yamada-Watanabe SDEs III

- Let $\rho(x) = \sigma x^p$ for $x \geq 0$ and $p \geq \frac{1}{2}$. Then, $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing function with $\rho(0) = 0$ and

$$\int_{(0,\epsilon)} \rho^{-2}(x)dx = +\infty, \quad \text{for all } \epsilon > 0 \tag{5}$$

- Verify the existence of weak solution of above SDE via martingale problems of Stroock and Varadhan (1969)
- Prove the pathwise uniqueness of the above SDE by introducing a sequence of auxiliary C^2-functions:
  - By (5), there exists a sequence of strictly decreasing $(a_k)_{k\geq 1} \subset (0,1]$ s.t. $a_0 = 1$, $a_\infty = 0$, and

$$\int_{a_k}^{a_{k-1}} \rho^{-2}(x)dx = k, \quad \forall\ k \geq 1.$$

Yamada-Watanabe SDEs IV

- For each $k \geq 1$, construct a $C(\mathbb{R})$-probability density function ρ_k with support (a_k, a_{k-1}) satisfying $0 \leq \rho_k(x) \leq \frac{2}{k\rho^2(x)}$ for all $x > 0$.
- Define a sequence of auxiliary C^2-functions by: for $k \geq 1$,

$$\psi_k(x) := \int_0^{|x|} \int_0^y \rho_k(s)dsdy, \quad x \in \mathbb{R}. \tag{6}$$

- Then, $|\psi_k'(x)| \leq 1$, $\lim_{k\to\infty} \psi_k(x) = |x|$ for $x \in \mathbb{R}$, and $(\psi_k)_{k\geq 1}$ is nondecreasing.
- We can verify the nonegativity of the solution by using the comparison theorem of SDE or the theory of linear continuous Markov processes.
- Question: Prove that the following estimate holds, for $x_1, x_2 > 0$,

$$E\left[|X_t^{x_1} - X_t^{x_2}|\right] \leq |x_1 - x_2| + L_b E\left[\int_0^t |X_s^{x_1} - X_s^{x_2}|\, ds\right], \tag{7}$$

where L_b is the Lipschitz coefficient of the drift $x \to b(x)$.

Linear Continuous Markov Processes I

- We here discuss a class of continuous and strong Markov processes $X = (X_t)_{t\geq 0}$ whose state space $I = (\ell, r)$ which is an open, closed or semi-open interval of $\mathbb{R}$.
- We assume that the death-time ζ is ∞, a.s., i.e., $P(\zeta < \infty) = 0$.
- Let the linear continuous Markov process (LCMP) X be regular, i.e., for any $x \in \mathrm{int}(I) = (\ell, r)$ and $y \in I$,

$$P_x(T_y < \infty) > 0,$$

where $T_y := \inf\{t > 0;\ X_t = y\}$.

Linear Continuous Markov Processes II

- In other words, for any regular LCMP X, starting with any interior point x, any point $y \in I$ can be reached by X with positive probability.

Lemma (Scale Function Formula)

For regular LCMP X, there exists a continuous, strictly increasing function S on I s.t. for all $a, b, x \in I$ with $\ell < a < x < b < r$,

$$P_x(T_b < T_a) = \frac{S(x) - S(a)}{S(b) - S(a)}.$$

In addition, if $\tilde{S}$ is another function with the same properties, then $\tilde{S}(x) = \alpha S(x) + \beta$ with $\alpha > 0$ and $\beta \in \mathbb{R}$.

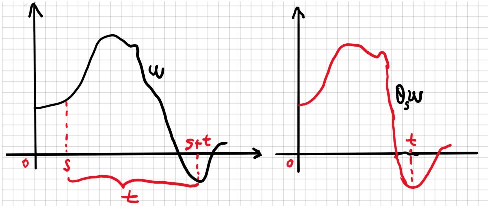
Linear Continuous Markov Processes III

- Proof. Let us first introduce shift operator. For any $s \geq 0$, one can construct $\mathcal{F}/\mathcal{F}$-measurable mapping $\theta_s : \Omega \to \Omega$ as:

$$X_{t+s}(\omega) = X_t(\theta_s\omega), \quad \forall\, \omega \in \Omega,\ s,t \geq 0. \tag{8}$$

- It is also convenient to use the canonical probability space $\Omega = C([0,\infty))$. Then, (9) is equivalent to

$$\theta_s\omega(t) = \omega(s+t), \quad \forall\, \omega \in \Omega,\ s,t \geq 0. \tag{9}$$



Linear Continuous Markov Processes IV

- Note that $T_a < T_\ell$ and $T_b < T_r$, however, we don't know the relationship of T_a, T_b and T_ℓ, T_r.
- Consider the event $\{T_r < T_\ell, T_a < T_b\}$.
- Clearly, on $\{T_a < T_b\}$, we have $T_a < T_\ell$ and $T_a < T_r$, then

$$T_\ell = T_a + T_\ell \circ \theta_{T_a}, \quad T_r = T_a + T_r \circ \theta_{T_a}.$$

- Using the strong Markov property, we get

$$\begin{aligned}
P_x(T_r < T_\ell, T_a < T_b) &= E_x\left[\mathbb{1}_{T_a<T_b}\mathbb{1}_{T_a+T_r\circ\theta_{T_a}<T_a+T_\ell\circ\theta_{T_a}}\right] \\
&= E_x\left[\mathbb{1}_{T_a<T_b}\mathbb{1}_{T_r<T_\ell}\circ\theta_{T_a}\right] \\
&= E_x\left\{E_x\left[\mathbb{1}_{T_a<T_b}\mathbb{1}_{T_r<T_\ell}\circ\theta_{T_a}|\mathcal{F}_{T_a}\right]\right\} \\
&= E_x\left\{\mathbb{1}_{T_a<T_b}E_x\left[\mathbb{1}_{T_r<T_\ell}\circ\theta_{T_a}|\mathcal{F}_{T_a}\right]\right\} \\
&= E_x\left\{\mathbb{1}_{T_a<T_b}E_{X_{T_a}}\left[\mathbb{1}_{T_r<T_\ell}\right]\right\} = P_x(T_a < T_b)P_a(T_r < T_\ell)
\end{aligned}$$

Linear Continuous Markov Processes V

- Similarly, $P_x(T_r < T_\ell, T_b < T_a) = P_x(T_b < T_a)P_b(T_r < T_\ell)$
- Then, it holds that

$$\begin{aligned} S(x) &:= P_x(T_r < T_\ell) \\ &= P_x(T_a < T_b)S(a) + P_x(T_b < T_a)S(b) \end{aligned}$$

- Note that $P_x(T_b < T_a) + P_x(T_a < T_b) = 1$.
- Then, it holds that

$$S(x) = (1 - P_x(T_b < T_a))S(a) + P_x(T_b < T_a)S(b)$$

- Solving $P_x(T_b < T_a)$ in terms of $S(x), S(a)$ and $S(b)$ to get the scale function formula
- Question: Prove that $I \ni x \to S(x)$ is strictly increasing and continuous.
- We call $x \to S(x)$ scale function of the regular LCMP X

Identification of Scale Function I

- If the scale function of a LCMP X can be taken to be $S(x) = x$, then we call this process is on its natural scale
- The following theorem can used to identify the scale function of some special LCMPs:

Theorem (Identification of Scale Function)

A locally bounded Borel function g is a scale function if and only if the stopped process $g(X)^{T_\ell \wedge T_r} = (g(X_{t\wedge T_\ell \wedge T_r}))_{t\geq 0}$ is a local martingale.

- Proof. $\Leftarrow$ Let $g(X)^{T_\ell \wedge T_r}$ is a local martingale.
- For $\ell < a < x < b < r$, since g is locally bounded and $X_t^{T_a \wedge T_b} \in [a, b]$, $g(X)^{T_a \wedge T_b}$ is a bounded martingale.

Identification of Scale Function II

- Using the optional stopping theorem, we have

$$g(x) = E_x[g(X_{T_a \wedge T_b})], \quad x \in (a, b).$$

- Note that

$$E_x[g(X_{T_a \wedge T_b})] = g(a)(1 - P_x(T_b < T_a)) + g(b)P_x(T_b < T_a)$$

- Then, it holds that

$$g(x) = g(a)(1 - P_x(T_b < T_a)) + g(b)P_x(T_b < T_a)$$

- This yields that

$$P_x(T_b < T_a) = \frac{g(x) - g(a)}{g(b) - g(a)}.$$

- By the definition of the scale function, g is a scale function.

Speed Measure of LCMP I

- For any open interval $J=(a,b)$ satisfying $[a,b]\subset I$, the exit time of J is defined by
$$\sigma_J := \inf\{t\geq 0;\ X_t\notin J\}.$$
- Then, P-a.s., $\sigma_J = T_a\wedge T_b$ for $x\in J$, and $\sigma_J=0$ for $x\notin J$.
- Define $m_J(x) := E_x[\sigma_J]$ for $x\in I$.
- Let $J_{c,d}=(c,d)\subset J$ (i.e., $a<c<d<b$). Then $\sigma_J\geq\sigma_{J_{c,d}}$.
- For $a<c<x<d<b$, we have
$$\begin{aligned} m_J(x) &= E_x[\sigma_J] = E_x[\sigma_{J_{c,d}} + \sigma_J\circ\theta_{\sigma_{J_{c,d}}}] \\ &= E_x[\sigma_{J_{c,d}}] + E_x[\sigma_J\circ\theta_{J_{c,d}}] = m_{J_{c,d}}(x) + E_x[\sigma_J\circ\theta_{J_{c,d}}] \end{aligned}$$
- Note that $x\in J_{c,d}$, then $m_{J_{c,d}}(x)>0$.

Speed Measure of LCMP II

- We also have from the strong Markov property that

$$
\begin{aligned}
E_x[\sigma_J \circ \theta_{J_{c,d}}] &= E_{X_{T_c}}[\sigma_J] P_x(T_c < T_d) + E_{X_{T_d}}[\sigma_J] P_x(T_d < T_c) \\
&= m_J(c) P_x(T_c < T_d) + m_J(d) P_x(T_d < T_c)
\end{aligned}
$$

- By Lemma 5, we have

$$
P_x(T_d < T_c) = \frac{S(x) - S(c)}{S(d) - S(c)}, \ P_x(T_c < T_d) = \frac{S(d) - S(x)}{S(d) - S(c)}
$$

- As a summay

$$
m_J(x) = m_{J_{c,d}}(x) + m_J(c) \frac{S(d) - S(x)}{S(d) - S(c)} + m_J(d) \frac{S(x) - S(c)}{S(d) - S(c)}
$$

- Since $m_{J_{c,d}}(x) > 0$, $x \to m_J(x)$ is a S-concave function.

Speed Measure of LCMP III

- Define a function on $I \times I$ as:

$$G_J(x,y) = \begin{cases} \dfrac{(S(x)-S(a))(S(b)-S(y))}{S(b)-S(a)}, & a \le x \le y \le b, \\[2ex] \dfrac{(S(y)-S(a))(S(b)-S(x))}{S(b)-S(a)}, & a \le y \le x \le b, \\[2ex] 0, & \text{otherwise.} \end{cases}$$

Speed Measure of LCMP IV

- Using the theory of S-concave function, we have

Theorem (Existence of Speed Measure)

There exists a unique Radon measure m defined on $\mathrm{int}(I)$ such that, for any $J = (a, b)$ satisfying $[a, b] \subset I$,

$$m_J(x) = \int_I G_J(x, y) m(dy), \quad x \in J.$$

- The measure m above is called speed measure of LCMP X.
- Question: For any open subset $J = (a, b)$, $x \in J$, and any nonnegative Borel function f, it holds that

$$E_x \left[\int_0^{\sigma_J} f(X_s) ds \right] = \int_I G_J(x, y) f(y) m(dy). \tag{10}$$

Speed Measure of LCMP V

- Hints: Pick c such that $a < c < b$. Define, for $x \in J$,

$$g_c(x) := E_x\left[\int_0^{\sigma_J} \mathbb{1}_{c<X_t<b} dt\right].$$

  Then g_c is a S-concave function on J and $g_c(a) = g_c(b) = 0$.

- The following theorem can be used to identify the speed measure of a LCMP:

Theorem (Identification of Speed Measure)

Let $\mathcal{A}$ be the infinitesimal generator of the regular LCMP X and its domain be $D(\mathcal{A})$. Let I be any sub-interval of $\mathbb{R}$. Define

$$\frac{d}{dS}f(x) := \lim_{y\to x} \frac{f(y)-f(x)}{S(y)-S(x)}, \text{ if exists.}$$

Then, for any bounded f on $D(\mathcal{A})$,

Speed Measure of LCMP VI

(i) $\frac{df}{dS}$ exists except possibly on the set $\{x;\ m(\{x\}) > 0\}$.

(ii) For $x_1, x_2 \in \mathrm{Int}(I)$ for which this S-derivative exists,

$$\frac{df}{dS}(x_2) - \frac{df}{dS}(x_1) = \int_{x_1}^{x_2} \mathcal{A}f(y)m(dy).$$

- Proof. For the bounded f on $D(\mathcal{A})$, Dykin's formula yields that, for $J = (a, b) \subset \mathrm{Int}(I)$ and $a < x < b$,

$$E_x\left[f(X_{T_a \wedge T_b})\right] - f(x) = E_x\left[\int_0^{T_a \wedge T_b} \mathcal{A}f(X_s)ds\right]$$

- Using (10) in Question, we have, for $\sigma_J = T_a \wedge T_b$,

$$E_x\left[\int_0^{\sigma_J} \mathcal{A}f(X_s)ds\right] = \int_I G_J(x, y)\mathcal{A}f(y)m(dy)$$

Speed Measure of LCMP VII

- This yields that

$$E_x\left[f(X_{T_a\wedge T_b})\right]-f(x)=\int_I G_J(x,y)\mathcal{A}f(y)m(dy),$$

- Note that

$$E_x\left[f(X_{T_a\wedge T_b})\right]=f(a)P_x(T_a<T_b)+f(b)P_x(T_b<T_a)$$

- By Lemma 5, we have

$$P_x(T_b<T_a)=\frac{S(x)-S(a)}{S(b)-S(a)},\ P_x(T_a<T_b)=\frac{S(b)-S(x)}{S(b)-S(a)}$$

Speed Measure of LCMP VIII

- Combine the above equalities, we have

$$\frac{f(b)-f(x)}{S(b)-S(x)}-\frac{f(x)-f(a)}{S(x)-S(a)}=\int_I H_J(x,y)\mathcal{A}f(y)m(dy)$$

where $H_J(x,y)$ is defined as

$$H_J(x,y)=\begin{cases}\dfrac{S(y)-S(a)}{S(x)-S(a)}\leq 1, & a<y\leq x,\\[2ex] \dfrac{S(b)-S(y)}{S(b)-S(x)}\leq 1, & x\leq y<b,\\[2ex] 0, & \text{otherwise.}\end{cases}$$

Feller's Boundary Classification I

- Consider the one-dimensional SDE:

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

- We impose the following assumptions:

(ND) $\sigma(x) > 0$ for $x \in I$;

(LI) for all $x \in I$, there exists $\epsilon > 0$ s.t.

$$\int_{x-\epsilon}^{x+\epsilon} \frac{1+|b(y)|}{\sigma^2(y)} dy < \infty.$$

- The generator of X is given by, for $f \in C^2(\mathbb{R})$,

$$\mathcal{A}f(x) = b(x)f'(x) + \frac{\sigma^2(x)}{2} f''(x), \quad x \in \mathbb{R}$$

Feller's Boundary Classification II

- In order to find a scale function S of X, by Theorem 6, we solve

$$\mathcal{A}S(x) = 0, \quad x \in \mathbb{R}. \tag{11}$$

- The solution of (11) is given by, for some $c \in \mathbb{R}$,

$$S(x) = \int_c^x \exp\left(-\int_c^y \frac{2b(z)}{\sigma^2(z)} dz\right) dy, \quad x \in \mathbb{R}. \tag{12}$$

- Note that $x \to S(x)$ is twice differential. Then $\frac{df}{dS}(x) = \frac{f'(x)}{S'(x)}$.
- From (12), we obtain

$$S'(x) = \exp\left(-\int_c^x \frac{2b(z)}{\sigma^2(z)} \mathrm{d}z\right), \quad S''(x) = -\frac{2b(x)}{\sigma^2(x)} S'(x).$$

Feller's Boundary Classification III

- Then, it holds that

$$\begin{aligned}\frac{df}{dS}(x_2)-\frac{df}{dS}(x_1)&=\int_{x_1}^{x_2}\left[\frac{f'(y)}{S'(y)}\right]'dy\\&=2\int_{x_1}^{x_2}\frac{\frac{1}{2}\sigma^2(y)f''(y)+f'(y)b(y)}{\sigma^2(y)S'(y)}dy\\&=2\int_{x_1}^{x_2}\frac{\mathcal{A}f(y)}{\sigma^2(y)S'(y)}dy=\int_{x_1}^{x_2}\mathcal{A}f(y)m(dy).\end{aligned}$$

- This implies that, the speed measure of X is given by

$$m(dx)=\frac{2}{\sigma^2(x)S'(x)}dx. \tag{13}$$

Feller's Boundary Classification IV

References.

1. Feller, W. (1952) The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics* **55,** 2nd Ser., 468–519.
2. Feller, W. (1954) The general diffusion operator and positivity preserving semi-groups in one dimension. *Annals of Mathematics* **60,** 2nd Ser., 417–436.
3. Feller, W. (1955) On second order differential operators. *Annals of Mathematics* **61,** 2nd Ser., 90–105.
4. Ito, K., and J. P. McKean, Jr (1965) Diffusion processes and their sample paths. Springer-Verlag.

- William Feller (1906-1970): Croatian-American mathematician specializing in probability theory. He obtained his Ph.D from University of Goettingen (supervisor: R. Courant who is the assistant of D. Hilbert)

Feller's Boundary Classification V

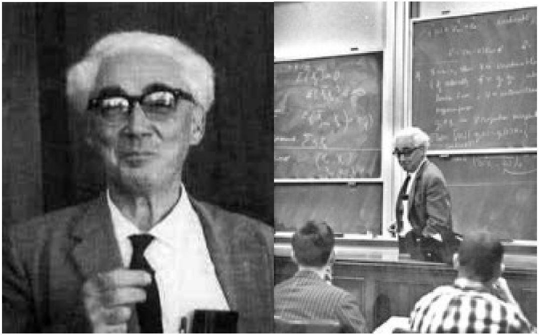


Figure: W. Feller (1906-1970): Croatian-American mathematician.

Feller's Boundary Classification VI

- Hereafter, define $s(x) = S'(x)$ and $m(x) = m(dx)/dx$.

Definition (Inaccessible, Absorbing, Reflecting Endpoints)

For the end-point b of the interval $I = (\ell, r)$,

(i) it is called inaccessible, if $b \in I^c$;

(ii) If $b \in I$, then b is called absorbing, if $P_b(T_y < \infty) = 0$ for all $y \in I \setminus \{b\}$;

(iii) it is called reflecting, if there exists $y \in I \setminus \{b\}$ s.t. $P_b(T_y < \infty) > 0$.

- Below, we only discuss the conditions under which the endpoints of I are inaccessible.

Feller's Test for Explosion I

- For the interval $I = (\ell, r)$, consider strictly monotone sequence $(\ell_k)_{k\geq 1}$ and $(r_k)_{k\geq 1}$ satisfying $\ell < \ell_k < r_k < r$, $\lim_{k\to\infty} \ell_k = \ell$, $\lim_{k\to\infty} r_k = r$ and

$$T_k := \inf\{t \geq 0;\ X_t \notin (\ell_k, r_k)\}, \quad k \geq 1.$$

- The explosion time is defined as

$$T := \inf\{t \geq 0;\ X_t \notin (\ell, r)\} = \lim_{k\to\infty} T_k.$$

- The related probability to T_k can computed by the scale function formula in Lemma 5.

Feller's Test for Explosion II

- Let us define the following quantities: for $y \in I$,

$$\begin{cases} \Sigma_r := \int_y^r \left(\int_y^v m(u)du \right) s(v)dv, \\ \Sigma_\ell := \int_\ell^y \left(\int_v^y m(u)du \right) s(v)dv, \\ N_r := \int_y^r \left(\int_y^v s(u)du \right) m(v)dv, \\ N_\ell := \int_\ell^y \left(\int_v^y s(u)du \right) m(v)dv. \end{cases} \tag{14}$$

Feller's Test for Explosion III

Theorem (Feller's Test for Explosion)

Recall (14). *Let* (ND), (LI) *hold. Then, it holds that*

(i) ℓ *and* r *are inaccessible (i.e.,* $P(T=\infty)=1$*) if and only if* $\Sigma_\ell=\Sigma_r=+\infty$

- (i1) r *is a natural boundary if* $\Sigma_r=N_r=+\infty$
- (i2) r *is an entrance boundary if* $\Sigma_r=+\infty$ *and* $N_r<+\infty$
- (i3) ℓ *is a natural boundary if* $\Sigma_\ell=N_\ell=+\infty$
- (i4) ℓ *is an entrance boundary if* $\Sigma_\ell=+\infty$ *and* $N_\ell<+\infty$

- For this theorem, please refer to Karlin and Taylor (1981), Table 6.2
- Question: For GBM,

$$X_t^x=x+\int_0^t\mu X_s^x ds+\int_0^t\sigma X_s^x dW_s,\quad x\in I:=(0,\infty),$$

where $\mu\in\mathbb{R}$ and $\sigma>0$. Prove that:

Feller's Test for Explosion IV

- the boundaries 0 and ∞ are all inaccessible, i.e., $P(X_t^x \in I,\ \forall\ t \geq 0) = 1$.

- Question: For Yamada-Wantanabe process,

$$X_t^x = x + \int_0^t (a + bX_s^x)ds + \int_0^t \sigma(X_s^x)^p dW_s, \quad x \in I := (0, \infty),$$

where $a, \sigma > 0$ and $b \in \mathbb{R}$. Prove that:

  - when $p = \frac{1}{2}$, the boundaries 0 and ∞ are inaccessible, i.e., $P(X_t^x \in I,\ \forall\ t \geq 0) = 1$ if and only if $2a \geq \sigma^2$;
  - when $p > \frac{1}{2}$, provide the conditions under which the boundaries 0 and ∞ are inaccessible.

Feller's Test for Explosion V

- Question: Consider the following so-called stepping-stone process given by

$$X_t^x = x + \int_0^t (a + bX_s^x)ds + \int_0^t \sigma\sqrt{(X_s^x - \ell)(r - X_s^x)}dW_s$$
$$x \in I := (\ell, r), \tag{15}$$

where $a, \sigma > 0$ and $b \in \mathbb{R}$. Do that

- provide the conditions under which the boundaries ℓ and r are inaccessible.

Course Outline

History of Feynman-Kac Formula I

- In the 1940s, R. Feynman discovered that the Schrödinger equation
  - the differential equation governing the time evolution of quantum states in quantum mechanics

  could be solved by a kind of averaging over paths, an observation which led him to a far-reaching reformulation of the quantum theory in terms of "path integrals"
- Upon learning of Feynman's ideas, M. Kac realized that a similar representation could be given for solutions of the heat equation with external cooling terms
  - a mathematician at Cornell University, where Feynman was, at the time, an Assistant Professor of Physics
- This representation is now known as Feynman-Kac formula
- Later it became evident that the expectation occurring in this representation is of the same type that occurs in derivative security pricing

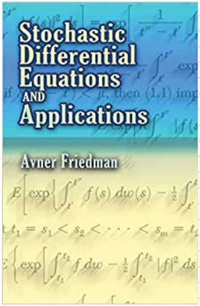
History of Feynman-Kac Formula II

- R. Feynman (1918-1988): 1965 Nobel Prize in Physics
- Mark Kac (1914-1984): probability, statistical physics, Feynman-Kac path integral


Figure: Left: R. Feynman; Right: M. Kac

Reference on Feynman-Kac Formula I

- I recommend you the book by Friedman (1975) on the theory of linear PDEs and their stochastic representation:



chapter 1 stochastic processes
2 Markov processes
3 Brownian Motions
4 The stochastic Integral
5 Stochastic Differential Equations
6 Elliptic & parabolic PDEs & Relations to SDEs
7 The Cameron-Martin-Girsanov Theorem
8 Asymptotic Estimates for solutions
9 Recurrent & Transient solutions

Figure: Friedman's Book, Volume 1

Reference on Feynman-Kac Formula II

- A. Friedman (1931-): Distinguished Professor of Math. & Phys. Sciences at Ohio State University; Areas of Expertise: PDEs, Mathematical Biology, SDEs, Control Theory and Free Boundary Problems.

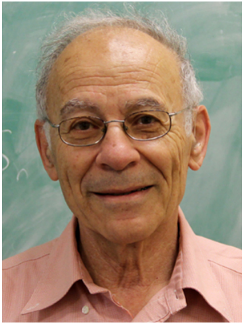


Figure: Avner Friedman (1931-)

Dirichlet Problem I

- Recall that $D \subseteq \mathbb{R}^n$ be a domain:
  - i.e., an open connected subset of $\mathbb{R}^n$
- Consider functions $b : D \to \mathbb{R}^{n\times 1}$, $\sigma : D \to \mathbb{R}^{n\times m}$ and $g : D \to \mathbb{R}$
- The second-order differential operator acted on $C^2(D)$ is defined as: for $f \in C^2(D)$,

$$\mathcal{A}f(x) := b(x)^\top \nabla_x f(x) + \frac{1}{2}\mathrm{tr}[a(x)\nabla_x^2 f(x)], \quad x \in D,$$

  where $a(x) := \sigma\sigma^\top(x)$, $\nabla_x = (\partial_{x_1}, \ldots, \partial_{x_n})^\top$ and ∇_x^2 is the corresponding Hessian matrix.
- We impose the following assumption $(\mathbf{A}_{FriD})$ introduced by Friedman (1975), page 144:

  (F1): The domain D is bounded and the boundary ∂D of D is in C^2
  - i.e., barriers exist at the all points of ∂D.

Dirichlet Problem II

(F2): The operator $\mathcal{A}$ is uniformly elliptic in D:

- there exists $C > 0$ s.t.

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq C|\xi|^2$$

for all $x \in D$ and $\xi \in \mathbb{R}^n$.

(F3): b, a are Lipschitz continuous in $\overline{D}$.

(F4): $g \leq 0$ and g is Hölder continuous in $\overline{D}$.

(F5): Given functions $f : \overline{D} \to \mathbb{R}$ and $\phi : \partial D \to \mathbb{R}$, they satisfy that

- f is Hölder continuous in $\overline{D}$;
- ϕ is continuous on ∂D.

Dirichlet Problem III

Theorem (Well-posedness of Dirichlet Problem)

Let $(\mathbf{A}_{FriD})$ *hold. Consider the Dirichlet problem given by*

$$(\mathcal{A}+g)u(x)-f(x)=0,\ \ in\ D;\ \ u(x)=\phi(x)\ on\ \partial D. \tag{16}$$

Then, there is a unique solution $u\in C^2(D)\cap C(\overline{D})$ *of Dirichlet problem* (16).

- Proof. This follows from Theorem 6.2.4 of Friedman (1975), page 134.
- Question: Consider the SDE given by: for $(t,x)\in\mathbb{R}_+\times\mathbb{R}^n$,

$$X_s^{t,x}=x+\int_t^s b(X_r^{t,x})dr+\int_t^s\sigma(X_r^{t,x})dW_r,\ \ s\geq t. \tag{17}$$

Let σ_D be the exit time of D for X, i.e., $\sigma_D:=\inf\{t\geq 0;\ X_t\notin D\}$

Dirichlet Problem IV

- Note that even if $(t,x)\in\mathbb{R}_+\times D$, the Lipschitz continuity of b,σ on $\mathbb{R}^n$ does *not* yield that $X^{t,x}$ must be in D, P-a.s.
- If $(\mathbf{A}_{FriD})$ holds and $E_x[\sigma_D]<\infty$ for all $x\in D$, then the solution u of Dirichlet problem (16) admits the probabilitistic representation:

$$u(x)=E\left[\phi(X_{\sigma_D}^{0,x})\exp\left(\int_0^{\sigma_D}g(X_s^{0,x})ds\right)\right]$$
$$-E\left[\int_0^{\sigma_D}f(X_s^{0,x})\exp\left(\int_0^s g(X_r^{0,x})dr\right)ds\right].$$

- Question: Using Theorem 11 and then applying Itô formula to

$$u(X_t)\exp\left(\int_0^t g(X_s^{0,x})ds\right),\quad \text{on } t\in[0,\sigma_D].$$

Dirichlet Problem V

Example: Consider 1-dim SDE:

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \qquad \underline{\text{LCMP}}$$

$$X_t \in I = (l, r), \ -\infty \le l < r \le +\infty$$

Let $D \triangleq (a,b) \subset I$, i.e. $l < a < b < r$, and

$$Af(x) \triangleq \frac{\sigma^2(x)}{2} f''(x) + b(x) f'(x), \quad x \in I, f \in C^2(I)$$

We solve the follwing Dirilet problem:

$$\begin{cases} Au(x) = 0, & x \in D \\ u(x) = \phi(x), & x \in \partial D = \{a, b\} \end{cases}$$

Dirichlet Problem VI

Note that the exit time:

$$\sigma_D \triangleq \inf\{t \geq 0;\ X_t \notin D\} = T_a \wedge T_b$$

Then, for all $x \in D$

$$\mathbb{E}_x[\sigma_D] = \int_I G_D(x, y)\, m(dy),$$

where

$$\begin{cases} m(dx) = \dfrac{2}{\sigma^2(x) S'(x)} dx & \text{—— speed measure} \\ S(x) = \displaystyle\int_c^x \exp\left(-\int_c^y \frac{2b(z)}{\sigma^2(z)} dz\right) dy & \text{— scale function} \end{cases}$$

If $\mathbb{E}_x[\sigma_D] < +\infty$, then, for all $x \in D$,

Dirichlet Problem VII

$$
\begin{aligned}
u(x) &= \mathbb{E}_x\left[\phi\left(X_{T_a \wedge T_b}\right)\right] \\
&= \phi(a)\frac{S(b)-S(x)}{S(b)-S(a)} + \phi(b)\frac{S(x)-S(a)}{S(b)-S(a)}
\end{aligned}
$$

Initial-Boundary Problem I

- We incorporate time variable t into the Dirichlet problem and this results in the initial-boundary problem
- The initial-boundary problem is described as:

$$\begin{aligned} (\partial_t + \mathcal{A} + g)u(t,x) &= f(t,x), \quad \text{in } (t,x) \in [0,T) \times D, \\ u(T,x) &= \phi(x), \qquad \text{on } D, \\ u(t,x) &= h(t,x), \quad \text{on } [0,T) \times \partial D. \end{aligned} \tag{18}$$

- It is more reasonable to call (18) a terminal-boundary problem. However, we can change t to $T-t$ and then transfer it into an initial-boundary problem
- We impose the assumption $(\mathbf{A}_{FriIB})$:

  (FIB1): (F1)-(F4) in the assumption $(\mathbf{A}_{FriD})$ hold

Initial-Boundary Problem II

(FIB2): The functions $f : \overline{[0, T) \times D} \to \mathbb{R}$, $\phi : \overline{D} \to \mathbb{R}$ and $h : \overline{[0, T) \times \partial D}$ satisfy

- f is Hölder continuous in $\overline{[0, T) \times D}$;
- ϕ is continuous on $\overline{D}$;
- h is continuous on $\{T\} \times \overline{D} \cup [0, T] \times \partial D$ and $h(T, x) = \phi(x)$ for $x \in \partial D$.

- The following well-posedness of the initial-boundary problem (18) has been proved by Theorem 6.5.2 in Friedman (1975), page 147.

Initial-Boundary Problem III

Theorem (Well-posedness of Initial-Boundary Problem)

Let $(\mathbf{A}_{FriIB})$ *hold. Then, the initial-boundary problem* (18) *admits a unique solution* $u \in C^{1,2} := C^{1,2}([0,T) \times D) \cap C(\overline{[0,T) \times D})$ *such that*

$$\begin{aligned} u(t,x) = {} & E\left[h(\sigma_D^t, X_{\sigma_D^t}^{t,x}) \exp\left(\int_t^{\sigma_D^t} g(X_s^{t,x})ds\right) \mathbb{1}_{\sigma_T^t < T}\right] \\ & + E\left[\phi(X_T^{t,x}) \exp\left(\int_t^T g(X_s^{t,x})ds\right) \mathbb{1}_{\sigma_T^t = T}\right] \\ & - E\left[\int_t^{\sigma_D^t} f(s, X_s^{t,x}) \exp\left(\int_t^s g(X_r^{t,x})dr\right) ds\right], \end{aligned}$$

where $\sigma_T^t := \inf\{s \in [t,T);\ X_s^{t,x} \notin D\}$ *for* $x \in D$. *It is defined as* T *if the set is empty.*

Initial-Boundary Problem IV

- Let us make the following assumptions:

  ($\mathbf{A}_{b,\sigma}$) $b : D \to \mathbb{R}^{n\times 1}$ and $\sigma : D \to \mathbb{R}^{n\times m}$ are locally Lipschitiz continuous.

  ($\mathbf{A}_X$) For all $(t,x) \in [0,T] \times D$, the solution $X^{t,x}$ of SDE (17) neither explodes nor leaves D before T:

  - i.e., $P(\sup_{s\in[t,T]} |X_s^{t,x}| < \infty) = P(X_s^{t,x} \in D,\ \forall\ s \in [t,T]) = 1$.

- The assumption **($\mathbf{A}_{b,\sigma}$)** implies the pathwiseness uniqueness of SDE (17).

- The assumption **($\mathbf{A}_X$)** results in

$$\sigma_T^t := \inf\{s \in [t,T);\ X_s^{t,x} \notin D\} = \inf\emptyset = T$$

Initial-Boundary Problem V

- Then, Theorem 12 gives the stochastic representation of the following Cauchy problem:

$$\begin{aligned}(\partial_t + \mathcal{A} + g)u(t,x) &= f(t,x), \quad \text{in } (t,x) \in [0,T) \times D,\\ u(T,x) &= \phi(x), \qquad \text{on } D.\end{aligned} \tag{19}$$

- We next summarize the well-posedness of Cauchy problem (19) on the bounded domain D in the following theorem:

Initial-Boundary Problem VI

Theorem (Well-posedness of Cauchy Problem on Bounded Domain)

Let $(\mathbf{A}_{b,\sigma})$, $(\mathbf{A}_X)$ *and the following assumptions hold:*

$(\mathrm{C1})$: *The domain* $D \subseteq \mathbb{R}^n$ *is bounded.*

$(\mathrm{C2})$: *The operator* $\mathcal{A}$ *is uniformly elliptic in* D.

$(\mathrm{C3})$: f *is Hölder continuous in* $[0, T] \times \overline{D}$ *and* g *is Hölder continuous on* $\overline{D}$.

Then, Cauchy problem (19) *admits a unique solution* $u \in C^{1,2}$ *such that*

$$\begin{aligned} u(t,x) &= E\left[\phi(X_T^{t,x})\exp\left(\int_t^T g(X_s^{t,x})ds\right)\right] \\ &\quad - E\left[\int_t^T f(s, X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds\right]. \end{aligned} \tag{20}$$

Cauchy Problems I

- For $D = \mathbb{R}^n$, it is *not* bounded, and hence Theorem 13 fails.
- In order to study stochastic representation of the Cauchy problem on $\mathbb{R}^n$:

$$\begin{aligned} (\partial_t + \mathcal{A} + g)u(t,x) &= f(t,x), \quad \text{in } (t,x) \in [0,T) \times \mathbb{R}^n, \\ u(T,x) &= \phi(x), \qquad \text{on } \mathbb{R}^n, \end{aligned} \tag{21}$$

we have to impose the boundedness assumption for the coefficients of the equation.

Cauchy Problems II

- From Theorem 6.5.3. of Friedman (1975) the well-posedness of Cauchy problem (21) and its stochastic representation are given by:

Theorem (Well-posedness of Cauchy Problem on $\mathbb{R}^n$)

Let the following assumptions hold:

(CR1): *The operator $\mathcal{A}$ is uniformly elliptic, and b, σ are bounded, locally Lipschitz continuous on $\mathbb{R}^n$*

(CR2): *$g : \mathbb{R}^n \to \mathbb{R}$ is bounded and is locally Hölder continuous*

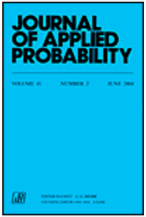
(CRf): *f is continuous in $[0, T] \times \mathbb{R}^n$, Hölder continuous in x uniformly w.r.t. $t \in [0, T]$, and $|f(t, x)| \leq C(1 + |x|^p)$*

(CRphi): *$\phi : \mathbb{R}^n \to \mathbb{R}$ is continuous and $|\phi(x)| \leq C(1 + |x|^p)$ for $C, p > 0$.*

Then, Cauchy problem (21) admits a unique solution $u \in C^{1,2}$ satisfying stochastic representation (20), and $|u(t, x)| \vee |\nabla_x u(t, x)| \leq C(1 + |x|^p)$.

Localization Idea I

- The stochastic representation of the above Cauchy problems is called Feynman-Kac Formula.
- However, many financial and physical applications do not satisfy the very restrictive assumptions imposed by these standard results.
- For instance, for the *unbounded* domain D, b, σ may be unbounded or grow faster than linearly or have unbounded derivatives, etc.
- Let us introduce a mixed argument of localization and probability proposed by Health and Schweizer (2000):



JOURNAL ARTICLE

Martingales versus PDEs in Finance: An Equivalence Result with Examples

David Heath and Martin Schweizer

Journal of Applied Probability

Vol. 37, No. 4 (Dec., 2000), pp. 947-957

Localization Idea II

- Hereafter, we assume that the domain $D \subseteq \mathbb{R}^n$ is *not* necessarily bounded
- Instead, we make the following assumption on the domain D:

  ($\mathbf{A}_D$) There exists a sequence $(D_k)_{k\geq 1}$ of bounded domains with C^2-boundary and $\overline{D}_k \subset D$ s.t. $\bigcup_{k=1}^{\infty} D_k = D$

- Under assumptions **($\mathbf{A}_{b,\sigma}$)** and **($\mathbf{A}_X$)**, we can have existence and uniqueness of D-valued strong solution of SDE: Consider the SDE given by: for $(t,x) \in [0,T] \times D$,

$$D \ni X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dW_r,\ s \in [t,T]$$

Localization Idea III

- For $(t,x) \in [0,T] \times D$, we define

$$u(t,x) := E\left[\phi(X_T^{t,x})\exp\left(\int_t^T g(X_s^{t,x})ds\right)\right] - E\left[\int_t^T f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds\right]. \qquad (22)$$

- We next study the continuity of $(t,x) \to u(t,x)$ under some additional assumptions.

Localization Idea IV

Lemma (Continuity of $(t,x)\to u(t,x)$)

Let $(\mathbf{A}_{b,\sigma})$, $(\mathbf{A}_X)$ and the following assumption hold:

(HSfgphi) $g: D\to\mathbb{R}$ *is continuous and is bounded from above,* $f:[0,T]\times D\to\mathbb{R}$ *and* $\phi: D\to\mathbb{R}$ *are continuous which satisfy* $|f(t,x)|\vee|\phi(x)|\leq C(1+|x|^p)$ *for* $C,p>0$;

(HSXmoment) $E[\sup_{s\in[t,T]}|X_s^{t,x}|^q]\leq C(1+|x|^q)$ *for all* $q\geq 1$.

Then, the function $u:[0,T]\times D\to\mathbb{R}$ *defined by* (22) *is continuous.*

- Proof. Denoted by $V_{t,x}$ the term in the expectation of (22). Define $H_{t,x}^{\epsilon}:=\{(s,y)\in[0,T]\times D;\ |s-t|+|y-x|<\epsilon\}$ for $\epsilon>0$. Then, for $k>1$ such that $pk\geq 1$, using (HSfgphi) and (HSXmoment),

$$\sup_{(s,y)\in H_{t,x}^{\epsilon}} E[|V_{s,y}|^k]\leq C_{k,T}\left\{1+E\left[\sup_{s\in[t,T]}|X_s^{t,x}|^{pk}\right]\right\}\leq C_{k,T}(1+|x|^{pk})$$

Localization Idea V

- This implies that $(V_{r,y})_{(r,y)\in H^{\epsilon}_{t,x}}$ is uniformly integrable
- The assumptions $(\mathbf{A}_{b,\sigma})$ and $(\mathbf{A}_X)$ yield that $(s,t,x)\to X_s^{t,x}$ is P-a.s. continuous. Hence $(t,x)\to\phi(X_T^{t,x})$ is P-a.s. continuous.
- $(s,t,x)\to g(X_s^{t,x})$ and $(s,t,x)\to f(s,X_s^{t,x})$ are P-a.s. uniformly continuous and bounded on any compact set of $[0,T]\times[0,T]\times D$.
- Then $(t,x)\to\int_t^T g(X_s^{t,x})ds$ and $(t,x)\to\int_t^T f(s,X_s^{t,x})ds$ are P-a.s. continuous.
- In summary, $(t,x)\to V_{t,x}$ is P-a.s. continuous. Therefore, the uniform integrability of $(V_{r,y})_{(r,y)\in H^{\epsilon}_{t,x}}$ implies that $(t,x)\to u(t,x)=E[V_{t,x}]$ is continuous.

If $p=0$ in the assumption (HSfgphi) of Lemma 15 (i.e., f,ϕ are all bounded), then Lemma 15 holds without the assumption (HSXmoment)

Localization Idea VI

- Based on the continuity of $(t, x) \to u(t, x)$, for each $k \geq 1$, consider the initial-boundary problem on $[0, T] \times D_k$:

$$\begin{aligned}(\partial_t + \mathcal{A} + g)v_k(t, x) &= f(t, x), \quad \text{in } (t, x) \in [0, T) \times D_k, \\ v_k(T, x) &= u(T, x) = \phi(x), \qquad \text{on } D_k, \\ v_k(t, x) &= u(t, x), \qquad \text{on } [0, T) \times \partial D_k.\end{aligned} \tag{23}$$

- As in Theorem 12 with the assumption $(\mathbf{A}_{FriIB})$, we then assume, for each $k \geq 1$,

  $(\mathbf{A}_{HS})$ (AHSba): The operator $\mathcal{A}$ is uniformly elliptic in D_k, i.e., there is a $l_k > 0$ s.t. $\xi^\top a(x)\xi \geq l_k|\xi|^2$ for all $x \in D_k$ and $\xi \in \mathbb{R}^n$;
  (AHSfg): g is Hölder continuous on $\overline{D}_k$ and f is uniformly Hölder continuous on $[0, T] \times \overline{D}_k$.

Localization Idea VII

Theorem (General Feynman-Kac Formula)

Let $(\mathbf{A}_{b,\sigma})$, $(\mathbf{A}_X)$, $(\mathbf{A}_D)$, $(\mathbf{A}_{HS})$, (HSfgphi), (HSXmoment) *hold. Then,* $u(t,x)$ *defined by the stochastic representation* (22) *is in* $C^{1,2}$ *and satisfies the Cauchy problem* (19), *i.e.,*

$$\begin{aligned}(\partial_t + \mathcal{A} + g)u(t,x) &= f(t,x), \quad \text{in } (t,x) \in [0,T) \times D,\\ u(T,x) &= \phi(x), \qquad \text{on } D.\end{aligned}$$

Moreover, there exists a unique classical solution of Cauchy problem (19).

- Proof. For each $k \geq 1$, the assumption $(\mathbf{A}_{b,\sigma})$ implies that b, a are Lipschitz continuous on the bounded $\overline{D}_k$.
- The assumptions (HSfgphi) and (HSXmoment) yield that $u(t,x)$ is continuous on $[0,T] \times \partial D_k \cup \{T\} \times \overline{D}_k$ by Lemma 15 ($\overline{D}_k \subset D$).

Localization Idea VIII

- Combine above claims and $(\mathbf{A}_{HS})$ to obtain that $(\mathbf{A}_{FriIB})$ is satisfied on $[0, T] \times D_k$
- Then, Theorem 12 yields that the initial-boundary problem (23) admits a unique classical solution $v_k(t, x)$.
- Now, for any $(t, x) \in [0, T] \times D$, the assumption $(\mathbf{A}_D)$ implies that one can find a $k \geq 1$ s.t. $x \in D_k$.
- Define σ_k as the exit time of $X^{t,x}$ from D_k from t before T, i.e., $\sigma_k := \inf\{s \in [t, T);\ X_s^{t,x} \notin D_k\}$ and $\inf \emptyset = T$.
- Since the path $s \to X_s^{t,x}$ is continuous, we have

$$(\sigma_k, X_{\sigma_k}^{t,x}) \in (0, T) \times \partial D_k \cup \{T\} \times D_k.$$

Localization Idea IX

- By Lemma 15, we obtain $u(\sigma_k, X_{\sigma_k}^{t,x}) < \infty$. Therefore, we can also apply the stochastic representation of $v_k(t,x)$ given in Theorem 12, one has

$$
\begin{aligned}
v_k(t,x) = {} & E\left[u(\sigma_k, X_{\sigma_k}^{t,x}) \exp\left(\int_t^{\sigma_k} g(X_s^{t,x})ds\right) \mathbb{1}_{\sigma_k<T}\right] \\
& + E\left[\phi(X_T^{t,x}) \exp\left(\int_t^{T} g(X_s^{t,x})ds\right) \mathbb{1}_{\sigma_k=T}\right] \\
& - E\left[\int_t^{\sigma_k} f(s, X_s^{t,x}) \exp\left(\int_t^{s} g(X_r^{t,x})dr\right) ds\right] \\
\overset{u(T,x)=\phi(x)}{=\!=\!=} {} & E\left[u(\sigma_k, X_{\sigma_k}^{t,x}) \exp\left(\int_t^{\sigma_k} g(X_s^{t,x})ds\right)\right] \\
& - E\left[\int_t^{\sigma_k} f(s, X_s^{t,x}) \exp\left(\int_t^{s} g(X_r^{t,x})dr\right) ds\right].
\end{aligned}
$$

Localization Idea X

- We next prove that $u(t,x)=v_k(t,x)$ if $(t,x)\in[0,T]\times D_k$.
- Since $T\geq\sigma_k\geq t$, we get

$$
\begin{aligned}
&\phi(X_T^{t,x})\exp\left(\int_t^T g(X_s^{t,x})ds\right)-\int_t^T f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds\\
&=\exp\left(\int_t^{\sigma_k} g(X_s^{t,x})ds\right)\phi(X_T^{t,x})\exp\left(\int_{\sigma_k}^T g(X_s^{t,x})ds\right)\\
&\quad-\exp\left(\int_t^{\sigma_k} g(X_r^{t,x})dr\right)\int_{\sigma_k}^T f(s,X_s^{t,x})\exp\left(\int_{\sigma_k}^s g(X_r^{t,x})dr\right)ds\\
&\quad-\int_t^{\sigma_k} f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds.
\end{aligned}
$$

Localization Idea XI

- This gives that

$$
\begin{aligned}
&E\left[\phi(X_T^{t,x})\exp\left(\int_t^T g(X_s^{t,x})ds\right)-\int_t^T f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds\middle|\mathcal{F}_{\sigma_k}\right]\\
&=\exp\left(\int_t^{\sigma_k} g(X_s^{t,x})ds\right)\\
&\times E\left[\phi(X_T^{t,x})\exp\left(\int_{\sigma_k}^T g(X_s^{t,x})ds\right)-\int_{\sigma_k}^T f(s,X_s^{t,x})\exp\left(\int_{\sigma_k}^s g(X_r^{t,x})dr\right)ds\middle|\mathcal{F}_{\sigma_k}\right]\\
&\quad-\int_t^{\sigma_k} f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds. \qquad (24)
\end{aligned}
$$

Localization Idea XII

- Using the strong Markov property of $X^{t,x}$, we have

$$E\left[\phi(X_T^{t,x})\exp\left(\int_{\sigma_k}^T g(X_s^{t,x})ds\right)-\int_{\sigma_k}^T f(s,X_s^{t,x})\exp\left(\int_{\sigma_k}^s g(X_r^{t,x})dr\right)ds\middle|\mathcal{F}_{\sigma_k}\right]$$
$$= u(\sigma_k, X_{\sigma_k}^{t,x}).$$

- Taking expectation on both sides of (24), for each $k \geq 1$ and $(t,x) \in [0,T] \times D_k$,

$$u(t,x) = E\left[u(\sigma_k, X_{\sigma_k}^{t,x})\exp\left(\int_t^{\sigma_k} g(X_s^{t,x})ds\right)\right]$$
$$- E\left[\int_t^{\sigma_k} f(s,X_s^{t,x})\exp\left(\int_t^s g(X_r^{t,x})dr\right)ds\right] = v_k(t,x).$$

- By $(\mathbf{A}_D)$ and (23), $u(t,x)$ satisfies Cauchy problem (19)

Localization Idea XIII

- The uniqueness follows from the stochastic representation (22) of $u(t,x)$ since $X^{t,x}$ is unique.
- Thus, we complete the proof of the theorem.
- Question: This problem is related to an option pricing problem under stochastic volatility model:

$$dS_t = rS_t + \sqrt{V_t}S_t dW_t,\ dV_t = \alpha(\beta - V_t)dt + \sigma\sqrt{V_t}dB_t,$$

  where B, W are two Brownian motions with $E[W_t B_t] = \rho t$ for $\rho \in (-1,1)$ and $2\alpha\beta \geq \sigma^2$.
- The P-price at time t of a European put on S with maturity T and strike K is then

$$u(t,S_t,V_t) = E\left[e^{-\int_t^T rds}\phi(S_T)\big|\mathcal{F}_t^{B,W}\right],\ \phi(x) = (K-x)^+$$

- Prove that $u \in C^{1,2}$ with $D = (0,\infty)^2$. How about the case $\phi(x) = (x-K)^+$?

Course Outline

History of FPK Equations I

- The Fokker-Planck equation is the equation governing the time evolution of the probability density of the Brownian particles.
- The Fokker-Planck equation is first established by Dutch physicist Adriaan Fokker and German physicist Max Planck.
- The Fokker-Planck equation is also known as the forward Kolmogorov equation, after Andrey Kolmogorov, who independently discovered the concept in 1931.
- The Fokker-Planck equation can be also derived from Chapman-Kolmogorov equation.
- Andrey Kolmogorov also finds backward Kolmogorov equation using Chapman-Kolmogorov equation.

History of FPK Equations II

Figure: Left: A. Fokker, Middle: M. Planck, Right: A. Kolmogorov.

- Adriaan Fokker (1887-1972): Dutch physicist and musician, he was the inventor of the Fokker organ.
- Max Planck (1858-1947): German physicist, 1918 Nobel Prize in Physics, he was the founder of Quantum Mechanics.
- A. Kolmogorov (1903-1987): Russian Mathematician, he was the founder of modern probability theory and one of the 20-th century's most eminent mathematicians.

Feller Semigroup I

- Let $b:\mathbb{R}^n\to\mathbb{R}^{n\times 1}$ and $\sigma:\mathbb{R}^n\to\mathbb{R}^{n\times m}$ satisfy $(\mathbf{A}_{lip})$
- Let us consider the following n-dimensional Itô SDE given by

$$X_t^x = x + \int_0^t b(X_s^x)ds + \int_0^t \sigma(X_s^x)dW_s, \quad (t,x)\in\mathbb{R}_+\times\mathbb{R}^n.$$

- Introduce the following spaces:

Let $\mathcal{P}(S)$ be the set of Borel probability measures on a topology space S

If (S,d) is a metric space, denoted by $\mathcal{P}_p(S)$ be the set of Borel probab. measures on S with finite p-order moments for $p\geq 1$

- Given the solution X^x of SDE, define the transition semigroup as: for any $f\in B(\mathbb{R}^n)$, $P_tf(x):=E[f(X_t^x)]$,
  - $B(S)$ is the set of bounded Borel functions on a topology space S.

Feller Semigroup II

Lemma (Feller Semigroup)

Let $(\mathbf{A}_{lip})$ *hold. Then* $(P_t)_{t\geq 0}$ *is a Feller semigroup.*

- Proof. For the semigroup property, note that X^x is a (strong) Markov process. Then for $s, t \geq 0$,

$$P_{s+t}f(x) = E[f(X_{s+t})|X_0 = x] = E\{E[f(X_{s+t})|X_0 = x, X_t]|X_0 = x\}$$
$$= E\{E_{X_t}[f(X_s)]|X_0 = x\} = E[P_s f(X_t)|X_0 = x] = P_t \circ P_s f(x)$$

  i.e., the semigroup property holds $P_{s+t} = P_t \circ P_s$.
- For Feller property, let $C_0(\mathbb{R}^n)$ be the set of continuous real-valued functions f on $\mathbb{R}^n$ satisfying $\lim_{|x|\to\infty} f(x) = 0$.
- Then, for any $f \in C_0(\mathbb{R}^n)$, since $x \to X_t^x$ is P-a.s. continuous and f is continuous, we have $x \to f(X_t^x)$ is P-a.s. continuous using CMT.

Feller Semigroup III

- Then, BCT yields the continuity of $x \to P_t f(x)$.
- Note that, for any $R > 0$,

$$
\begin{aligned}
|P_t f(x)| &\leq E[|f(X_t^x - x + x)|\mathbb{1}_{|X_t^x - x| \leq R}] \\
&\quad + E[|f(X_t^x - x + x)|\mathbb{1}_{|X_t^x - x| > R}] \\
&\leq \sup_{\{z;\ |z-x| \leq R\}} |f(z)| + \|f\|_\infty P(|X_t^x - x| > R) \\
&\leq \sup_{\{z;\ |z-x| \leq R\}} |f(z)| + \|f\|_\infty R^{-2} E[|X_t^x - x|^2] \\
&\leq \sup_{\{z;\ |z-x| \leq R\}} |f(z)| + \|f\|_\infty R^{-2} C_t
\end{aligned}
$$

- By letting x, then R, tend to ∞, we have $\lim_{|x| \to \infty} P_t f(x) = 0$.
- To this end, for any $x \in \mathbb{R}^n$, $t \to P_t f(x)$ is also continuous.
- It holds that $\mathcal{A}f = \lim_{t \downarrow 0} \frac{P_t f - f}{t}$.

Forward Kolmogorov Equation I

- Given a probability measure $\rho_0 \in \mathcal{P}(\mathbb{R}^n)$, define

$$\mu_t(dx) = \int_{\mathbb{R}^n} E\left[\delta_{X_t^{x_0}}(dx)\right] \rho_0(dx_0), \quad \text{on } \mathcal{B}(\mathbb{R}^n)$$

  where δ is the Dirac-delta measure and
  - $\mathcal{B}(\mathbb{R}^n)$ is the Borel-σ-algebra, i.e., the σ-algebra generated by open sets of $\mathbb{R}^n$;
  - $C_0^\infty(\mathbb{R}^n)$ is the set of functions $f \in C_0(\mathbb{R}^n)$ which are also infinitely differentiable.

- Obviously, $\mu_0 = \rho_0$, $\mu_t \in \mathcal{P}(\mathbb{R}^n)$ for all $t \geq 0$, and for any test function $f \in C_0^\infty(\mathbb{R}^n)$,

$$\langle \mu_t, f \rangle := \int_{\mathbb{R}^n} f(x)\mu_t(dx) = \int_{\mathbb{R}^n} P_t f(x_0)\rho_0(dx_0).$$

- Let $\mathcal{A}^*$ be the adjoint operator of the generator $\mathcal{A}$:

Forward Kolmogorov Equation II

- i.e., $\langle \mathcal{A}^* f, g\rangle_{L^2} = \langle \mathcal{A} g, f\rangle_{L^2}$, for $f, g \in C_0^2(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$;
- If b, σ are smooth, then

$$\mathcal{A}^* f(x) = \frac{1}{2} \sum_{i,j=1}^{n} \partial^2_{x_i x_j}((\sigma\sigma^\top)_{ij}(x) f(x)) - \sum_{i=1}^{n} \partial_{x_i}(b_i(x) f(x)).$$

- We next introduce Fokker-Planck-Kolmogorov (FPK) equation and the resulting forward Kolmogorov eqaution:
- Start with Itô formula for $f(X_t^{x_0})$, which yields that

$$f(X_t^{x_0}) = f(x_0) + \int_0^t \mathcal{A} f(X_s^{x_0}) ds + \int_0^t \nabla_x f(X_s^{x_0})^\top \sigma(X_s^{x_0}) dW_s.$$

Forward Kolmogorov Equation III

- Then, it holds that

$$\begin{aligned}\int_{\mathbb{R}^n} E[f(X_t^{x_0})]\rho_0(dx_0) = &\int_{\mathbb{R}^n} E[f(x_0)]\rho_0(dx_0)\\ &+\int_0^t\int_{\mathbb{R}^n} E\left[\mathcal{A}f(X_s^{x_0})\right]\rho_0(dx_0)ds\\ &+\int_0^t\int_{\mathbb{R}^n} \underbrace{E\left[\nabla_x f(X_s^{x_0})^\top\sigma(X_s^{x_0})dW_s\right]}_{=0}\rho_0(dx_0).\end{aligned}$$

- Therefore

$$\begin{aligned}\int_{\mathbb{R}^n} E[f(X_t^{x_0})]\rho_0(dx_0) = &\int_{\mathbb{R}^n} E[f(x_0)]\rho_0(dx_0)\\ &+\int_0^t\int_{\mathbb{R}^n} E\left[\mathcal{A}f(X_s^{x_0})\right]\rho_0(dx_0)ds.\end{aligned}$$

Forward Kolmogorov Equation IV

Lemma (Forward Kolmogorov Equation)

Let $(\mathbf{A}_{lip})$ *hold. Then, for all* $f \in C_0^\infty(\mathbb{R}^n)$,

$$\langle \mu_t, f\rangle = \langle \rho_0, f\rangle + \int_0^t \langle \mu_s, \mathcal{A}f\rangle ds, \quad t \geq 0. \tag{25}$$

If $\rho_0(dx) = u_0(x)dx$, *then* $\mu_t(dx) = p(t,x)dx$, *where* $p(t,x)$ *satisfies that*

$$\begin{aligned} \partial_t p(t,x) &= \mathcal{A}^* p(t,x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^n; \\ p(0,x) &= u_0(x), \quad x \in \mathbb{R}^n. \end{aligned} \tag{26}$$

- We call (25) a Fokker-Planck-Kolmogorov equation and Eq. (26) satisfied by the density function is said to be a forward Kolmogorov equation, which is first established by Fokker and Planck.

Forward Kolmogorov Equation V

- Warning: Note that we *don't* have the sufficient smoothness of $p(t,x)$ at the moment, the solution of the forward Kolmogorov equation (26) should be understood as in the distributional sense: for all $f \in C_0^\infty(\mathbb{R}^n)$,

$$\langle p(t,\cdot), f\rangle = \langle u_0, f\rangle + \int_0^t \langle p(s,\cdot), \mathcal{A}f\rangle ds, \quad t \geq 0.$$

- Since $\langle \mu_t, f\rangle = \int_{\mathbb{R}^n} P_t f(y)\rho_0(dy)$, if $\rho_0(dx) = \delta_{x_0}(x)dx$, then FPK equation (25) reads

$$\partial_t P_t f = P_t(\mathcal{A}f), \quad P_0 f = f.$$

- The word "forward" means that the above equation is obtained by perturbing the final position, i.e., $P_t(\mathcal{A}f)$ is the limit of $P_t(\frac{P_\epsilon f - f}{\epsilon})$ as $\epsilon \to 0$

Forward Kolmogorov Equation VI

- Question: Prove that $(P_t f)_{t\geq 0}$ satisfies the so-called backward equation $\partial_t P_t f = \mathcal{A}(P_t f)$ with $P_0 f = f$.
- Transition density function of X^{x_0}: Let $\rho_0(dx) = \delta_{x_0}(x)dx$ and use $p(t, x_0; x)$ to indicate the dependence of $p(t, x)$ on a given initial point $x_0 \in \mathbb{R}^n$.
- Then $P(X_t^{x_0} \in dx) = p(t, x_0; x)dx$. By Lemma 18, the transition density function $p(t, x_0; x)$ obeys that

$$\begin{aligned} \partial_t p(t, x_0; x) &= \mathcal{A}^* p(t, x_0; x), \quad (t, x) \in (0, \infty) \times \mathbb{R}^n; \\ p(0, x_0; x) &= \delta_{x_0}(x), \quad x \in \mathbb{R}^n. \end{aligned} \tag{27}$$

Forward Kolmogorov Equation VII

- If (b, σ) are smooth, we can expect to have a classical solution for Eq. (27). Typically, when $X^{x_0} = x_0 + W$ is an n-dimensional Brownain motion starting at $x_0 \in \mathbb{R}^n$ (i.e., $b = 0$ and $\sigma = I_{n\times n}$), then

$$\mathcal{A} = \mathcal{A}^* = \frac{1}{2}\Delta.$$

- Therefore, the forward equation (27) admits the classical solution (Einstein (1905)): for $(t, x) \in (0, \infty) \times \mathbb{R}^n$,

$$p(t, x_0; x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{t}} \exp\left(-\frac{|x - x_0|^2}{2t}\right),$$
$$p(0, x_0; x) = \delta_{x_0}(x), \quad x \in \mathbb{R}^n.$$

Forward Kolmogorov Equation VIII

- The above $p(t, x_0; x)$ is also called a fundamental solution and this yields that, Browanin Feller semigroup

$$P_t f(x_0) = \int_{\mathbb{R}^n} f(x) p(t, x_0; x) dx, \quad x_0 \in \mathbb{R}.$$

- Question: Let X^x be a OU process, i.e., for $(t, x) \in (0, \infty) \times \mathbb{R}$,

$$X_t^x = x + \int_0^t \alpha(\beta - X_s^x) ds + \sigma W_t,$$

where $\alpha, \sigma > 0$, $\beta \in \mathbb{R}$, and W is a scalar Brownian motion. Solve the forward Kolmogorov (26).

Non-Divergence Form of FPK Equation I

- Recall the adjoint operator given by, for $f \in C_0^\infty(\mathbb{R}^n)$,

$$\mathcal{A}^* f(x) = \frac{1}{2}\sum_{i,j=1}^n \partial^2_{x_i x_j}(a_{ij}(x)f(x)) - \sum_{i=1}^n \partial_{x_i}(b_i(x)f(x)). \tag{28}$$

- Let us assume that (b,σ) is sufficiently smooth. Then
- Subtract the coefficients of $\partial_{x_i} f$ and $\partial^2_{x_i,x_j} f$ from $\mathcal{A}^* f$, define

$$\begin{aligned}
\mathrm{b}_i(x) &:= \left(\sum_{j=1}^n \partial_{x_j} a_{ij}(x)\right) - b_i(x), \;\; i = 1,\dots,n;\\
g(x) &:= \frac{1}{2}\sum_{i,j=1}^n \partial^2_{x_i x_j} a_{ij}(x) - \sum_{i=1}^n \partial_{x_i} b_i(x).
\end{aligned}$$

Non-Divergence Form of FPK Equation II

- Therefore, it holds that

$$\mathcal{A}^* f(x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\partial^2_{x_i x_j} f(x) + \sum_{i=1}^{n} \mathrm{b}_i(x)\partial_{x_i} f(x) + g(x)f(x).$$

- Then, the forward Kolmogorov equation (26) can be written as in:

The Non-Divergence Form: for $(t,x) \in (0,\infty)\times\mathbb{R}^n$,

$$\partial_t p(t,x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\partial^2_{x_i x_j} p(t,x) + \sum_{i=1}^{n} \mathrm{b}_i(x)\partial_{x_i} p(t,x) + g(x)p(t,x),$$
$$p(0,x) = u_0(x), \quad x \in \mathbb{R}^n. \tag{29}$$

- The forward equation with non-divergence form becomes a Cauchy problem which has been discussed in detail in Section of Feynman-Kac Formula.

Non-Divergence Form of FPK Equation III

- In particular, we can refer to Chapter 4 of Friedman (1975) for a general result on the fundamental solution of the uniformly elliptic version of Cauchy problem (29).

Assumptions:

(A1) The adjoint operator $\mathcal{A}^*$ is uniformly elliptic on $\mathbb{R}^n$.

(A2) The coefficients $a, \mathrm{b}, \mathrm{g}$ are bounded and continuous functions in $\mathbb{R}^n$.

(A3) The coefficient a is Hölder continuous (exponent $\alpha \in (0,1]$) w.r.t. $x \in \mathbb{R}^n$, and b, g are Hölder continuous (exponent $\alpha \in (0,1]$) uniformly w.r.t. x in compact sets of $\mathbb{R}^n$.

Non-Divergence Form of FPK Equation IV

Lemma (Theorem 1.4.2 in Friedman (1964))

Let (A1)-(A3) hold. Then, there exists a fundamental solution $G(t_0, t; x_0, x)$ ($t_0 < t$) of Cauchy problem (29) satisfying, for all $f \in C_0(\mathbb{R}^n)$,

$$\partial_t G(t_0, t; x_0, x) = \mathcal{A}^* G(t_0, t; x_0, x), \quad \text{if } x \in \mathbb{R}^n,\ t_0 < t \leq T;$$

$$\int_{\mathbb{R}^n} G(t_0, t; x_0, x) f(x_0) dx_0 \to f(x), \quad t \downarrow t_0.$$

Moreover, for $k = (k_1, \ldots, k_n) \in \mathbb{N}^n$ with $0 \leq |k| \leq p$ and $p \geq 1$,

$$\left|\nabla_x^k G(t_0, t; x_0, x)\right| \leq C|t - t_0|^{-\frac{n+|k|}{2p}} \exp\left[-K\left(\frac{|x - x_0|^{2p}}{t - t_0}\right)^{\frac{1}{2p-1}}\right].$$

Non-Divergence Form of FPK Equation V

- However, we don't like the boundedness constraint of a, b on $\mathbb{R}^n$ in the assumption (A2).
- We can apply Theorem 16 with only local condition of coefficients to Cauchy problem (29). See the following example:

Example: Consider OU process, i.e., $m = n = 1$, $b(x) = \alpha(\beta - x)$, $\alpha > 0$, and $\sigma(x) = \sigma > 0$ for $x \in \mathbb{R}$. Then $\mathrm{b} = -\alpha\beta + \alpha x$ and $g(x) = -\alpha$. Thus, the forward equation with non-divergence form is given by

$$\begin{aligned} \partial_t p(t,x) &= \frac{\sigma^2}{2}\Delta_x p(t,x) + \alpha(x-\beta)\partial_x p(t,x) - \alpha p(t,x), \\ p_0(x) &= u_0(x). \end{aligned} \tag{30}$$

- Note that $D = \mathbb{R} = \cup_{k=1}^{\infty} D_k$ with $D_k := (-k, k)$ with smooth corners.

Non-Divergence Form of FPK Equation VI

- Then, the coefficients a, b, g satisfy the assumption in Theorem 16. By Theorem 16, if for some $p > 0$,

  (Aphi): $u_0 : \mathbb{R} \to \mathbb{R}$ is continuous and it satisfies

$$|u_0(x)| \leq C(1 + |x|^p), \quad x \in \mathbb{R},$$

- then Cauchy problem (30) has a unique classical solution.
- In fact, from our previous discussion, the fundamental solution of Cauchy problem (30) admits a closed-form representation.
- Question: Discuss the well-posedness of the forward equation with non-divergence form for Geometric Brownian Motion.

History of Gradient Flows I

- On May 3, 1941, Richard Courant gave an address to the AMS in which he proposed three methods for numerically solving variational PDEs.
  - Finite Element Method; Finite Difference Method; Gradient Method
- Richard Courant (1888-1972): German-born American mathematician and educator who made significant advances in the calculus of variations.
- Courant established one of America's most prestigious institutes of applied mathematics; upon his retirement the institute was named in his honour.
- Courant also wrote a two-volume elementary work on applied calculus, *Differential and Integral Calculus* (1934; originally published in German, 1927-29), and, with H. Robbins, a general work for the layperson, *What Is Mathematics?* (1941).

History of Gradient Flows II

Figure: Richard Courant (1888-1972)

History of Gradient Flows III

- The idea arose in the study of variational PDEs. Each of these equations has a function $V : \mathbb{R}^n \to \mathbb{R}$ s.t. a solution of the equation is a minimizer of V.
- The method of gradients starts with an initial point $x_0 \in \mathbb{R}^n$, and seeks to find a minimizer of V by following a curve X^{x_0} defined by ODE:

$$dX_t^{x_0} = -\nabla V(X_t^{x_0})dt, \quad X_0^{x_0} = x_0.$$

- It describes that the curve $X^{x_0} : [0, T] \to \mathbb{R}^n$ evolves in the direction of steepest decent of the energy V.

Example: The energy functional $V(x) = \frac{\alpha}{2}|x|^2$ for $\alpha > 0$ and $x \in \mathbb{R}^n$. Then, the gradient flow $X_t^{x_0} = x_0 e^{-\alpha t}$ is the unique solution of the gradient flow.

- The solution X^{x_0} is called an integral curve or gradient flow.

History of Gradient Flows IV

- Critical Point x^* of V: if $\nabla V(x^*) = 0$.

If the curve $X_t^{x_0}$ is *not* a critical point of V, then $X_t^{x_0}$ has the desirable property that V is always decreasing along it:

$$\frac{d}{dt} V(X_t^{x_0}) = \nabla V(X_t^{x_0}) \frac{dX_t^{x_0}}{dt} = -\left|\nabla V(X_t^{x_0})\right|^2 \leq 0.$$

- In addition, X^{x_0} has the desirable property that for a large class of functions it connects the initial point x_0 to a critical point of V:

If $V : \mathbb{R}^n \to \mathbb{R}$ satisfies Palais-Smale condition, and is smooth, Lipschitz continuously differentiable, bounded from below, and has isolated critical points. Then, $\lim_{t\to\infty} X_t^{x_0}$ exists and which is a critical point of V.

History of Gradient Flows V

- Palais-Smale condition satisfied by V: any subset $D \subset \mathbb{R}^n$ on which V is bounded and on which ∇V is *not* bounded away from zero contains in its closure a critical point of V.
- An example of a function that does not satisfy Palais-Smale condition: $V(x) = e^{-x}$ for $x \in \mathbb{R}$. On $D = [0, \infty)$, V is bounded and $|\nabla V(x)| = e^{-x}$ is *not* bounded away from zero, but V has no critical point on $\mathbb{R}$.
- In the aspect of the forward Kolmogorov equation, the “random version" of gradient flows result in a class of important FPK equations.

Gradient Flow I

- It is well-known that the generator $\mathcal{A} = \frac{1}{2}\Delta$ of Brownian motion is self-adjoint in $L^2(\mathbb{R}^n)$, since Brownian motion is a reversible Markov process.
- We expect to find the class of Itô diffusion process whose generator is self-adjoint in a right space.
- Let $V : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. Consider the following Itô diffusion process given by, for $x_0 \in \mathbb{R}^n$,

$$X_t^{x_0} = x_0 - \int_0^t \nabla V(X_s^{x_0})ds + \sqrt{2\sigma} W_t, \quad t \geq 0 \tag{31}$$

where $\sigma > 0$ and $\nabla = \nabla_x$ is the gradient operator.
  - If $\sigma = 0$, $dX_t^{x_0} = -\nabla V(X_t^{x_0})dt$ corresponds to the deterministic gradient flow for a curve $X^{x_0} : [0, T] \to \mathbb{R}^n$.

Gradient Flow II

- V is in general called a potential or an energy. Note that $V(x)=\frac{\alpha x^2}{2}$ for $x\in\mathbb{R}$ corresponds to Langevin equation.

- The generator of $X_t^{x_0}$ is given by

$$\mathcal{A}f(x)=-\nabla V(x)\cdot\nabla f(x)+\sigma\Delta f(x)$$

- The stochastic gradient flow (31) admits a unique invariant measure:

Lemma (Gibbs distribution)

Let $e^{-V(\cdot)/\sigma}\in L^1(\mathbb{R}^n)$. Then, the stochastic gradient flow (31) has a unique invariant density given by

$$\pi(x)=\frac{1}{\Gamma}e^{-\frac{V(x)}{\sigma}},\ x\in\mathbb{R}^n,\quad \Gamma:=\int_{\mathbb{R}^n}e^{-\frac{V(x)}{\sigma}}dx. \tag{32}$$

- The density function $\pi(x)$ given by (32) is called Gibbs distribution.

Gradient Flow III

- Proof. It is straightforward to verify $\int_{\mathbb{R}^n}(\mathcal{A}f(x))\pi(x)dx = 0$ for all $f \in \mathcal{D}(\mathcal{A})$ ($\mathcal{D}(\mathcal{A}) = C_0^2(\mathbb{R}^n)$). This yields that

$$\int_{\mathbb{R}^n}(-\nabla V(x)\nabla f(x) + \sigma\Delta f(x))\pi(x)dx = 0 \tag{33}$$

$$\Downarrow$$

$$\int_{\mathbb{R}^n}\{(-\nabla V(x)\nabla f(x))\pi(x) - \sigma\nabla f(x)\nabla\pi(x)\}dx = 0 \tag{34}$$

- For the uniqueness, the ergodic theory of Markov processes can be used.
- We next discuss the role of Gibbs distribution played in the study of forward Kolmogorov equations.

Gradient Flow IV

- By (28), the adjoint operator of $\mathcal{A}$ is given by

$$\begin{aligned}
\mathcal{A}^* f(x) &= \sum_{i=1}^n \partial_{x_i}(\partial_{x_i} V(x) f(x)) + \sigma \Delta f(x) \\
&= f(x) \sum_{i=1}^n \partial_{x_i}^2 V(x) + \sum_{i=1}^n \partial_{x_i} V(x) \partial_{x_i} f(x) + \sigma \Delta f(x) \\
&= f(x)\Delta V(x) + 2\nabla V(x) \cdot \nabla f(x) \\
&\quad \underbrace{-\nabla V(x) \cdot \nabla f(x) + \sigma \Delta f(x)}_{=\mathcal{A} f(x)} \text{-Non-Divergence} \\
&= \nabla \cdot ((\nabla V)(x) f(x)) + \sigma \Delta f(x). \qquad (35)
\end{aligned}$$

- The forward Kolmogorov equation becomes that, for $p := p(t,x)$,

$$\partial_t p = \nabla \cdot ((\nabla V) p) + \sigma \Delta p, \quad p_0 = u_0 \qquad (36)$$

Gradient Flow V

- We can implement a transform using Gibbs distribution for the solution of the forward Kolmogorov equation. After this transform, the resulting equation is the backward Kolmogorov equation:

Define $q(t,x) := \pi(x)^{-1}p(t,x)$. Then, q satisfies the following backward Kolmogorov equation given by, for $(t,x) \in (0,\infty) \times \mathbb{R}^n$,

$$\partial_t q(t,x) = \mathcal{A}q(t,x), \quad q(0,x) = \frac{u_0(x)}{\pi(x)}, \ x \in \mathbb{R}^n. \tag{37}$$

- We will leave the verification of the backward Kolmogorov equation (37) to a Question.
- Consequently, in order to study properties of solutions to the forward equation, it is sufficient to study the backward equation (37).

Gradient Flow VI

- Since $\mathcal{A}^* f = \mathcal{A} f + f\Delta V + 2\nabla f \nabla V$, $\mathcal{A}$ is *not* self-adjoint in $L^2(\mathbb{R}^n)$. Even if $\mathcal{A}$ is not self-adjoint in $L^2(\mathbb{R}^n)$, we can find a right space under which $\mathcal{A}$ is self-adjoint using Gibbs distribution.
- The right space is the following weight L^2-space as follows:

$$L_\pi^2(\mathbb{R}^n) := \left\{\phi : \mathbb{R}^n \to \mathbb{R};\ \int_{\mathbb{R}^n} |\phi(x)|^2 \pi(x) dx < \infty\right\}$$

- $L_\pi^2(\mathbb{R}^n)$ is a Hilbert space with inner product

$$\langle g, \phi\rangle_\pi := \int_{\mathbb{R}^n} g(x)\phi(x)\pi(x)dx.$$

Gradient Flow VII

- The following lemma proves that $\mathcal{A}$ is indeed self-adjoint in $L^2_\pi(\mathbb{R}^n)$.

Lemma (Self-Adjoint in $L^2_\pi(\mathbb{R}^n)$)

Let $e^{-V(\cdot)/\sigma} \in L^1(\mathbb{R}^n)$. Then, $\mathcal{A} = -\nabla V \cdot \nabla + \sigma\Delta$ is self-adjoint in $L^2_\pi(\mathbb{R}^n)$ and satisfies

(i) $\langle \mathcal{A}f, g\rangle_\pi = -\sigma\langle \nabla f, \nabla g\rangle_\pi$ *for all* $f, g \in C^2_0(\mathbb{R}^n)$;

(ii) $\mathcal{A}$ *is negative.*

- Proof. It is suffices to verify (i), which can be derived directly from $\mathcal{A} = -\nabla V \cdot \nabla + \sigma\Delta$. By (34), we have

$$\begin{aligned}\langle \mathcal{A}f, g\rangle_\pi &= \int_{\mathbb{R}^n} (-\nabla V \cdot \nabla f(x) + \sigma\Delta f(x))g(x)\pi(x)dx \\ &= \int_{\mathbb{R}^n} (-\nabla V \cdot \nabla f(x))\pi(x)g(x)dx + \sigma\int_{\mathbb{R}^n} \Delta f(x)g(x)\pi(x)dx\end{aligned}$$

Gradient Flow VIII

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} (-\nabla V \cdot \nabla f(x))\pi(x)g(x)dx - \sigma \int_{\mathbb{R}^n} \nabla f(x)\nabla(g(x)\pi(x))dx \\
&= \underbrace{\int_{\mathbb{R}^n} \{(-\nabla V \cdot \nabla f(x))\pi(x) - \sigma\nabla f(x)\nabla\pi(x)\}g(x)dx}_{=0 \ \ using \ \ (34)} - \sigma\langle\nabla f, \nabla g\rangle_\pi \\
&= -\sigma\langle\nabla f, \nabla g\rangle_\pi \Rightarrow \langle\mathcal{A}f, g\rangle_\pi = -\sigma\langle\nabla f, \nabla g\rangle_\pi = \langle\mathcal{A}g, f\rangle_\pi.
\end{aligned}
$$

- By (i), $\langle\mathcal{A}f, f\rangle_\pi = -\sigma\|\nabla f\|_\pi^2 \leq 0$ since $\sigma > 0$, which yields (ii).
- Thus, we complete the proof of the lemma.
- In the context of Markov processes, $\langle-\mathcal{A}f, f\rangle_\pi$ is called Dirichlet Form.
- We next show that the solution of the forward Kolmogorov equation converges to Gibbs distribution exponentially fast:

Gradient Flow IX

- The main tools are the following Poincare inequality and the backward Kolmogorov equation (37).

Assume that the potential V additionally satisfies ℓ-convexity condition: there exists $\ell > 0$ s.t. $\nabla^2 V(x) \geq \ell I$ for all $x \in \mathbb{R}^n$. If $g \in C^1(\mathbb{R}^n)$ satisfies $\langle g, 1\rangle_\pi = 0$, then $\|g\|_\pi^2 \leq \ell^{-1}\|\nabla g\|_\pi^2$.

- Now, assume that the initial density u_0 of the forward Kolmogorov equation satisfies $u_0/\pi \in L_\pi^2(\mathbb{R}^n)$.

Theorem (Large Time behavior Solution of Forward Kolmogorov Equation)

For any $t \geq 0$, it holds that

$$\|p(t,\cdot) - \pi(\cdot)\|_{\pi^{-1}} \leq e^{-\ell\sigma t}\,\|u_0(\cdot) - \pi(\cdot)\|_{\pi^{-1}}\,.$$

Gradient Flow X

- Proof. Recall $q(t,x)=\pi(x)^{-1}p(t,x)$ which satisfies the backward Kolmogorov equation (37).
- Then, using Lemma 21, we have

$$\|p(t,\cdot)-\pi(\cdot)\|_{\pi^{-1}}^2=\int_{\mathbb{R}^n}|q(t,x)-1|^2\pi(x)dx=\|q(t,\cdot)-1\|_{\pi}^2$$

- In order to apply the above Poincare inequality, we need to verify $\int_{\mathbb{R}^n}(q(t,x)-1)\pi(x)dx=0$ for all $t>0$.
- In fact, we get

$$\begin{aligned}\partial_t\left(\int_{\mathbb{R}^n}q(t,x)\pi(x)dx\right)&=\int_{\mathbb{R}^n}\partial_t q(t,x)\pi(x)dx\\&=\int_{\mathbb{R}^n}\mathcal{A}q(t,x)\pi(x)dx=\langle\mathcal{A}q(t,\cdot),1\rangle_{\pi}\\&\overset{(i)}{=}-\sigma\langle\nabla q(t,\cdot),\nabla 1\rangle_{\pi}=0.\end{aligned}$$

Gradient Flow XI

- Note that $\int_{\mathbb{R}^n} q(0,x)\pi(x)dx = \int_{\mathbb{R}^n} u_0(x)dx = 1$.
- Taking derivative on both sides of the above equality w.r.t. t, we have

$$\begin{aligned}\partial_t \|q(t,\cdot)-1\|_\pi^2 &= 2\left\langle \partial_t q(t,\cdot), q(t,\cdot)-1\right\rangle_\pi \\ &\overset{(37)}{=} 2\left\langle \mathcal{A}q(t,\cdot), q(t,\cdot)-1\right\rangle_\pi \\ &\overset{\mathcal{A}1=0}{=\!=} 2\left\langle \mathcal{A}(q(t,\cdot)-1), q(t,\cdot)-1\right\rangle_\pi \\ &\overset{(i)}{=\!=} -2\sigma\|\nabla(q(t,\cdot)-1)\|_\pi^2 \\ &\overset{\text{Poincare}}{\leq} -2\sigma\ell\|q(t,\cdot)-1\|_\pi^2.\end{aligned}$$

- As a summary

$$\partial_t \|p(t,\cdot)-\pi(\cdot)\|_{\pi^{-1}}^2 \leq -2\sigma\ell\|q(t,\cdot)-1\|_\pi^2.$$

Expansive Solution of FPK Equation I

- Let $V : \mathbb{R}^n \to \mathbb{R}$ be a smooth potential satisfying
  - $e^{-V(\cdot)/\sigma} \in L^1(\mathbb{R}^n)$;
  - $\nabla^2 V(x) \geq \ell I$ for all $x \in \mathbb{R}^n$.
- Then, by Lemma 21, for the operator $\mathcal{A} = -\nabla V \cdot \nabla + \sigma\Delta$, we have
  (i) $\mathcal{A}$ is a negative and self-adjoint operator on $L^2_\pi(\mathbb{R}^n)$;
  (ii) For $g \in C^1(\mathbb{R}^n)$ with $\langle g, 1\rangle_\pi = 0$, a spectral gap is given by

$$\|\langle \mathcal{A}g, g\rangle\|_\pi^2 = -\sigma\|\nabla g\|_\pi^2 \leq -\sigma\ell\|g\|_\pi^2.$$

- The spectral problem of $-\mathcal{A}$ is as follows:

$$-\mathcal{A}g_k = \lambda_k g_k, \quad k = 0, 1, \dots,$$

- The operator $-\mathcal{A}$ admits real, discrete spectrum satisfying $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$;
- The eigenfunctions $(g_k)_{k\geq 0}$ form an orthonormal basis in $L^2_\pi(\mathbb{R}^n)$.

Expansive Solution of FPK Equation II

- This implies that for any $\phi \in L^2_\pi(\mathbb{R}^n)$,

$$\phi = \sum_{k=0}^{\infty} \phi_k g_k, \quad \phi_k = \langle \phi, g_k \rangle_\pi. \tag{38}$$

- We seek the solution $q(t,x) = \pi(x)^{-1} p(t,x)$ of the backward Kolmogorov equation (37) in the following form:

$$q(t,x) = \sum_{k=0}^{\infty} q_k(t) g_k(x).$$

Expansive Solution of FPK Equation III

- Plugging the above $q(t,x)$ into the backward Kolmogorov equation (37), we get

$$\partial_t q(t,x) = \sum_{k=0}^{\infty} \partial_t q_k(t) g_k(x) = \sum_{k=0}^{\infty} q_k(t) \mathcal{A} g_k(x)$$
$$= -\sum_{k=0}^{\infty} q_k(t) \lambda_k g_k(x).$$

- Therefore, for $j \geq 0$,

$$\sum_{k=0}^{\infty} \partial_t q_k(t) \left(\int_{\mathbb{R}^n} g_k(x) g_j(x) \pi(x) dx \right)$$
$$= -\sum_{k=0}^{\infty} \lambda_k q_k(t) \left(\int_{\mathbb{R}^n} g_k(x) g_j(x) \pi(x) dx \right).$$

Expansive Solution of FPK Equation IV

- As a summary, for $j \geq 0$,

$$\sum_{k=0}^{\infty} \partial_t q_k(t) \langle g_k, g_j \rangle_\pi = -\sum_{k=0}^{\infty} \lambda_k q_k(t) \langle g_k, g_j \rangle_\pi .$$

- We then conclude the equations

$$\partial_t q_j(t) = -\lambda_j q_j(t), \quad j = 0, 1, \ldots .$$

- Assume that $q(0,x) = \frac{u_0(x)}{\pi(x)} \in L^2_\pi(\mathbb{R}^n)$.
- Then, by (38), we get

$$q(0,x) = \sum_{k=0}^{\infty} l_{0k} g_k(x), \quad l_{0k} = \langle q(0,\cdot), g_k \rangle_\pi . \tag{39}$$

Expansive Solution of FPK Equation V

- Thus, we obtain

$$\partial_t q_j(t) = -\lambda_j q_j(t), \quad q_j(0) = l_{0j}, \quad j = 0, 1, \ldots. \tag{40}$$

- The solution of (40) is given by, for $t \geq 0$,

$$q_0(t) = l_{00} = 1, \quad q_j(t) = l_{0j} e^{-\lambda_j t}, \ j = 1, 2, \ldots. \tag{41}$$

- Therefore, the solution of the backward Kolmogorov equation is as follows:

$$\pi(x)^{-1} p(t, x) = q(t, x) = 1 + \sum_{j=1}^{\infty} l_{0j} e^{-\lambda_j t} g_j(x)$$

Expansive Solution of FPK Equation VI

- Then, the forward Kolmogorov equation (36) admits the following expansive form:

$$p(t,x) = \pi(x) + \pi(x)\left(1 + \sum_{j=1}^{\infty} l_{0j} e^{-\lambda_j t} g_j(x)\right). \tag{42}$$

Course Outline

FPK Equations and Particle System I

- In the previous Chapter: FPK Equations, an important problem which is not discussed is the uniqueness of solutions of FPK equation. This issue can be studied by the approach of Martingale Problem.
- We here introduce the method of Propagation of Chaos on FPK equation. This in particularly implies the uniqueness of the FPK equation.
- Propagation of chaos is in fact establishing the convergence of the empirical measure of a particle system to the solution to a nonlinear equation, was first formulated by Kac (1956):
  - Kac (1956): Foundations of Kinetic Theory. *In Proceedings of the Third Berkeley Symposium on Mathematical Stats. and Probab.*, 1954-1955, vol. III, pages 171-197. University of California Press, Berkeley and Los Angeles.
- Kac (1956) studies the convergence of a toy particle system as a step to the rigorous derivation of the Boltzmann equation.

FPK Equations and Particle System II

- Let us start with a review of the FPK equation discussed in Chapter: FPK Equations, which is given by (25) in Lemma 18, i.e., for all $f \in C_0^\infty(\mathbb{R}^n)$,

$$\langle \mu_t, f \rangle = \langle \rho_0, f \rangle + \int_0^t \langle \mu_s, \mathcal{A} f \rangle ds, \quad t \in [0, T]$$

  where $\rho_0 \in \mathcal{P}(\mathbb{R}^n)$, and $T > 0$ is an arbitrary time horizon.

- The operator $\mathcal{A}$ is the second-order differential operator corresponding to the generator of the following Itô diffusion process given by, for $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s.$$

FPK Equations and Particle System III

- As it is proved in Lemma 18, the following probability measure valued function $(\mu_t)_{t\in[0,T]}$ is a solution of FPK equation (25):

$$\mu_t(dx) = \int_{\mathbb{R}^n} E\left[\delta_{X_t^{x_0}}(dx)\right] \rho_0(dx_0), \quad \text{on } \mathcal{B}(\mathbb{R}^n). \tag{43}$$

- If the uniqueness holds, then the solution of FPK equation (25) must be form (43).

Propagation of Chaos: Establish a particle system which includes N particles. For $i = 1, \ldots, N$, the state (e.g. position, velocity and so on) process of the i-th particle is given by a process $(X_t^i)_{t\in[0,T]}$. Define the empirical measure-valued process as $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$ for $t \in [0,T]$. For an arbitrary solution $\mu = (\mu_t)_{t\in[0,T]}$ of FPK equation (25), find an increasing function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ with $\alpha(0) = 0$ s.t. $d_T(\mu^N, \mu) \le \alpha(N^{-1})$, where $d_T(\cdot,\cdot)$ is a suitable metric.

FPK Equations and Particle System IV

- If the propagation of chaos is established, then for any two different solutions $\mu=(\mu_t)_{t\in[0,T]}$ and $\nu=(\nu_t)_{t\in[0,T]}$ of FPK equation (25), we have $\lim_{N\to\infty}\mu^N=\mu$ and $\lim_{N\to\infty}\mu^N=\nu$ w.r.t. d_T.
- Then $d(\mu,\nu)\leq d(\mu^N,\mu)+d(\mu^N,\nu)\to 0$ as $N\to\infty$, this yields that $\mu=\nu$, i.e., uniqueness holds.
- We next construct a (homogeneous) particle system required in the propagation of chaos.
- Let the number of particles be $N\geq 1$. For $i=1,\ldots,N$, the dynamics of the state of the i-th particle is given by

$$dX_t^i=b(X_t^i)dt+\sigma(X_t^i)dW_t^i,\quad X_0^i\in\mathbb{R}^n. \tag{44}$$

- Here $W^i=(W_t^i)_{t\in[0,T]}$, $i=1,\ldots,N$ and $W=(W_t)_{t\in[0,T]}$ are independent (m-dimensional) Brownian motions under the filtered probability space $(\Omega,\mathcal{F},\mathbb{F},P)$.

FPK Equations and Particle System V

- The initial values $(X_0^i)_{i\geq 1}$ is assumed to satisfy that

(A)$_{X_0^i}$ Let $q \geq 2$. The sequence of r.v.s $(X_0^i)_{i\geq 1}$ is i.i.d. according to the probability distribution $\rho_0 \in \mathcal{P}_p(\mathbb{R}^n)$ for some $p > q$.

- Equip $\mathcal{P}_p(\mathbb{R}^n)$ with Wasserstein distance $\mathcal{W}_p$: for $\mu, \nu \in \mathcal{P}_p$,

$$\mathcal{W}_p(\mu,\nu) = \begin{cases} \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^p \pi(dx,dy)\right)^{\frac{1}{p}}, & p \geq 1; \\ \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^p \pi(dx,dy), & 0 < p < 1, \end{cases}$$

where $\Pi(\mu,\nu)$ is the set of $\pi \in \mathcal{P}(\mathbb{R}^{2n})$ s.t. $\pi(A \times \mathbb{R}^n) = \mu(A)$ and $\pi(\mathbb{R}^n \times B) = \nu(B)$ for all $A, B \in \mathcal{B}(\mathbb{R}^n)$.

- Then, by Villani (2003), $(\mathcal{P}_p(\mathbb{R}^n), \mathcal{W}_p)$ is a Polish space since $(\mathbb{R}^n, |\cdot|)$ is Polish.

FPK Equations and Particle System VI

- As is well-known, under the assumption $\mathbf{(A)}_{X_0^i}$, by Glivenko-Cantelli's theorem, the empirical measure $\mu_0^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_0^i}$ tends weakly to ρ_0 as $N \to \infty$.
- Moreover, Theorem 1 in Fournier and Guillin (2015) yields that, there is a constant C depending only on n, p, q such that

$$E\left[\mathcal{W}_q(\mu_0^N, \rho_0)^q\right] \leq C\left(\int_D |x|^p \rho_0(dx)\right)^{\frac{q}{p}} \alpha(p, q, n, N). \tag{45}$$

Probab. Theory Relat. Fields (2015) 162:707–738
DOI 10.1007/s00440-014-0583-7

On the rate of convergence in Wasserstein distance of the empirical measure

Nicolas Fournier · Arnaud Guillin

Figure: A reference paper by Fournier and Guillin (2015)

FPK Equations and Particle System VII

- The main convergence rate $\alpha(p,q,n,N)$ is given by

$$\alpha(p,q,n,N) := \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{p-q}{p}}, & q > \frac{n}{2},\ p \neq 2q; \\ N^{-\frac{1}{2}} \ln(1+N) + N^{-\frac{p-q}{p}}, & q = \frac{n}{2},\ p \neq 2q; \\ N^{-\frac{q}{n}} + N^{-\frac{p-q}{p}}, & q < \frac{n}{2},\ p \neq \frac{n}{n-q}. \end{cases}$$

- We next establish the propagation of chaos by introducing a so-called propagator corresponding to FPK equation (25).

Propagator of FPK Equation I

- The propagator of FPK equation (25) is defined as: for $(t,x)\in[0,T]\times\mathbb{R}^n$,

$$P_{t,T}^{g}\phi(x) := E\left[\phi(X_T^{t,x})\exp\left(\int_t^T g(X_s^{t,x})ds\right)\right]. \qquad (46)$$

- The process $(X_s^{t,x})_{s\in[t,T]}$ satisfies: for $(t,x)\in[0,T]\times\mathbb{R}^n$,

$$X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dW_s, \quad s\in[t,T].$$

- Question: Verify that $(P_{0,T}^{g})_{T\geq 0}$ is a semigroup. Provide conditions under which $(P_{0,T}^{g})_{T\geq 0}$ is a Feller semigroup.
- The propagator (46) is the same to the representation of $u(t,x)$ in (22) with $f\equiv 0$.

Propagator of FPK Equation II

- Then, we can apply Theorem 16 [General Feynman-Kac Formula] to study the smoothness of the propagator $(t,x) \to P_{t,T}\phi(x)$.
- To this purpose, we review the assumptions imposed in Theorem 16 (with $D = \mathbb{R}^n$):

  (HSfgphi) $g : \mathbb{R}^n \to \mathbb{R}$ is continuous and is bounded from above, $\phi : \mathbb{R}^n \to \mathbb{R}$ are continuous which satisfying $|\phi(x)| \leq C(1+|x|^p)$ for $C, p > 0$;

  (HSXmoment) $E[\sup_{s\in[t,T]} |X_s^{t,x}|^q] \leq C|x|^q$ for all $q \geq 1$.
- If $p = 0$ in (HSfgphi), then we *don't* need (HSXmoment).
- Under the assumption $(\mathbf{A}_{lip})$, we also review

  $(\mathbf{A}_{HS})$ (AHSba): The operator $\mathcal{A}$ is uniformly elliptic in $D_k = (-k,k)^n$, i.e., there is a $l_k > 0$ s.t. $\xi^\top a(x)\xi \geq l_k|\xi|^2$ for all $x \in D_k$ and $\xi \in \mathbb{R}^n$;

  (AHSfg): g is Hölder continuous on $\overline{D}_k$.

Propagator of FPK Equation III

- Then, Theorem 16 gives the smoothness of the propagator of $P^g_{t,T}\phi$:

Lemma (Smoothness of Propagator)

Under the above assumptions, the propagator of $P^g_{\cdot,T}\phi(\cdot) \in C^{1,2}$*, and it also satisfies the following Cauchy problem:*

$$\begin{aligned} (\partial_t + \mathcal{A} + g)P^g_{t,T}\phi(x) &= 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n; \\ P^g_{T,T}\phi(x) &= \phi(x), \quad x \in \mathbb{R}^n. \end{aligned} \tag{47}$$

- Proof. The proof follows completely by verifying the assumptions imposed in Theorem 16.

Propagator of FPK Equation IV

- We can also have the estimate of the gradient $\nabla_x P_{t,T}\phi$ under the additional assumption on the Lipschitz-continuity of g:

Lemma (Gradient Estimate of Propagator)

Assume additionally that $g : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous. Then, there exists a positive constant $K_{b,\sigma,T}$ depending only on T and the Lipschitz coefficient b, σ s.t., for all $\phi \in Lip(\mathbb{R}^n)$ with $\|\phi\|_\infty := \sup_{x\in\mathbb{R}^n} |\phi(x)| \le C$,

$$\left\|\nabla_x P^g_{t,T}\phi\right\|_\infty \le K_{b,\sigma,T}, \quad \forall\ t \in [0,T]. \tag{48}$$

- Proof. Note that the Lipschitzian property of g, g is bounded from above and $\|\phi\|_\infty \le 1$.
- W.L.O.G. let $g(x) \le 0$ for all $x \in \mathbb{R}^n$.

Propagator of FPK Equation V

- Then, for all $x_1, x_2 \in \mathbb{R}^n$,

$$
\begin{aligned}
&\left|P^g_{t,T}\phi(x_1) - P^g_{t,T}\phi(x_2)\right|^2 \qquad (49)\\
&\leq 2E\left[\left|\phi(X_T^{t,x_1}) - \phi(X_T^{t,x_2})\right|^2 \exp\left(2\int_t^T g(X_s^{t,x})ds\right)\right]\\
&+ 2E\left[\left|\phi(X_T^{t,x_2})\right|^2 \left|\exp\left(\int_t^T g(X_s^{t,x_1})ds\right) - \exp\left(\int_t^T g(X_s^{t,x_2})ds\right)\right|^2\right]\\
&\leq 2\|\phi\|_{\mathrm{Lip}} E\left[\left|X_T^{t,x_1} - X_T^{t,x_2}\right|^2\right] + 8TC^2\|g\|^2_{\mathrm{Lip}} E\left[\int_t^T \left|X_s^{t,x_1} - X_s^{t,x_2}\right|^2 ds\right].
\end{aligned}
$$

Propagator of FPK Equation VI

- By the assumption $(\mathbf{A}_{lip})$ gives that

$$E\left[\sup_{s\in[t,T]}\left|X_s^{t,x_1}-X_s^{t,x_2}\right|^2\right]\le|x_1-x_2|^2+C_{b,\sigma,T}\int_t^T E\left[\left|X_s^{t,x_1}-X_s^{t,x_2}\right|^2\right]ds.$$

- Then, the Gronwall's lemma yields that

$$E\left[\sup_{s\in[t,T]}\left|X_s^{t,x_1}-X_s^{t,x_2}\right|^2\right]\le|x_1-x_2|^2e^{(T-t)C_{b,\sigma,T}}.\qquad(50)$$

Propagator of FPK Equation VII

- It follows from (49) and (50) that, for all $x_1, x_2 \in \mathbb{R}^n$,

$$\left|P^g_{t,T}\phi(x_1) - P^g_{t,T}\phi(x_2)\right|^2 \leq 2\left(\|\phi\|_{\rm Lip} + 4T^2C^2\|g\|^2_{\rm Lip}\right)e^{TC_{b,\sigma,T}}|x_1 - x_2|^2. \qquad (51)$$

- Since $P_{\cdot,T}\phi$ is the classical solution of Cauchy problem (47), we have $P_{\cdot,T}\phi \in C^{1,2}$.
- Then, the estimate (51) yields that the gradient estimate (48) by taking $K_{b,\sigma,T} := \sqrt{2(\|\phi\|_{\rm Lip} + 4T^2C^2\|g\|^2_{\rm Lip})e^{TC_{b,\sigma,T}}}$.
- Thus, we complete the proof of the lemma.
- What is the role of the propagator $P^g_{t,T}\phi$ in the construction of the propagation of chaos?

Propagator of FPK Equation VIII

- We in fact have the following important observation on the relation between μ_t and $P^0_{t,T}\phi$ (when $g \equiv 0$) given as follows:

An important observation (the proof is non-trivial): for all $\phi : \mathbb{R}^n \to \mathbb{R}$ satisfying (HSfgphi), it follows from (47) that

$$\begin{aligned} \partial_t \langle \mu_t, P^0_{t,T}\phi \rangle &= \langle \partial_t \mu_t, P^0_{t,T}\phi \rangle + \langle \mu_t, \partial_t P^0_{t,T}\phi \rangle \\ &= \langle \mu_t, \mathcal{A} P^0_{t,T}\phi \rangle + \langle \mu_t, \partial_t P^0_{t,T}\phi \rangle \\ &= \langle \mu_t, (\partial_t + \mathcal{A}) P^0_{t,T}\phi \rangle = 0. \end{aligned} \tag{52}$$

- This yields that

$$\langle \mu_t, P^0_{t,T}\phi \rangle = \langle \rho_0, P^0_{0,T}\phi \rangle, \quad \forall\ t \in [0, T]. \tag{53}$$

Propagator of FPK Equation IX

- The propagator $P^0_{t,T}\phi$ establishes the following relation satisfied by $\mu^N_T - \mu_T$ for any fixed $T > 0$:

Theorem (Decomposition of $\langle \mu^N_T - \mu_T, \phi \rangle$)

Let the above assumptions gold. Then, for any fixed $T > 0$, and all $\phi : \mathbb{R}^n \to \mathbb{R}$ satisfying (HSfgphi),

$$\langle \mu^N_T - \mu_T, \phi \rangle = \langle \mu^N_0 - \mu_0, P^0_{0,T}\phi \rangle + \frac{1}{N}\sum_{i=1}^N \int_0^T \nabla_x P_{s,T}\phi(X^i_s)^\top \sigma(X^i_s) dW^i_s. \tag{54}$$

- Proof. Recall the state process of the particle system $X^i = (X^i_t)_{t\in[0,T]}$ defined by (44) for $i \geq 1$.

Propagator of FPK Equation X

- Lemma 27 allows us to apply Itô's formula to $P^0_{t,T}\phi(X^i_t)$, and we have that, for $t \in [0, T]$,

$$P^0_{t,T}\phi(X^i_t) = P^0_{0,T}\phi(X^i_0) + \int_0^t (\partial_s + \mathcal{A})P^0_{s,T}\phi(X^i_s)ds + \int_0^t \nabla_x P^0_{s,T}\phi(X^i_s)^\top \sigma(X^i_s)dW^i_s.$$

- Note that $(\partial_t + \mathcal{A})P^0_{t,T}\phi = 0$, therefore

$$\frac{1}{N}\sum_{i=1}^N P^0_{t,T}\phi(X^i_t) = \frac{1}{N}\sum_{i=1}^N P^0_{0,T}\phi(X^i_0) + \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla_x P^0_{s,T}\phi(X^i_s)^\top \sigma(X^i_s)dW^i_s.$$

Propagator of FPK Equation XI

- Recall that $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$. Then, we have

$$\langle \mu_t^N, P_{t,T}^0\phi\rangle = \langle \mu_0^N, P_{0,T}^0\phi\rangle + \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla_x P_{s,T}^0\phi(X_s^i)^\top \sigma(X_s^i)dW_s^i.$$

- By (53), i.e., $\langle \mu_t, P_{t,T}^0\phi\rangle = \langle \rho_0, P_{0,T}^0\phi\rangle$ for all $t \in [0, T]$.

$$\begin{aligned}\langle \mu_t^N, P_{t,T}^0\phi\rangle - \langle \mu_t, P_{t,T}^0\phi\rangle &= \langle \mu_0^N, P_{0,T}^0\phi\rangle - \langle \rho_0, P_{0,T}^0\phi\rangle \\ &\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla_x P_{s,T}^0\phi(X_s^i)^\top \sigma(X_s^i)dW_s^i.\end{aligned}$$

- By taking $t = T$ and note that $P_{T,T}^0\phi = \phi$, we have (94).
- Thus we complete the proof of the theorem.

Metric between μ^N and μ I

- For $p \geq 1$, recall that $\mathcal{W}_p$ is Wasserstein metric with p-order on $\mathcal{P}_p(\mathbb{R}^n)$.
- For $q \geq 2$, define $\mathcal{S}_{p,q,T}$ as the set of $\mathcal{P}(\mathbb{R}^n)$-valued processes $\nu = (\nu_t)_{t\in[0,T]}$ such that

$$E\left[\sup_{t\in[0,T]}\left(\int_{\mathbb{R}^n}|x|^p\nu_t(dx)\right)^{\frac{q}{p}}\right] < \infty.$$

- For $p \geq 1$ and $q \geq 2$, introduce the metric on $\mathcal{S}_{p,q,T}$ as follows: for all $\nu^1, \nu^2 \in \mathcal{S}_{p,q,T}$,

$$d_{p,q,T}(\nu_1,\nu_2) := \left\{E\left[\sup_{t\in[0,T]}\mathcal{W}_p\left(\nu_t^1,\nu_t^2\right)^q\right]\right\}^{\frac{1}{q}}. \tag{55}$$

Metric between μ^N and μ II

- Consider $p = 1$ and using the Kantorovich-Rubinstein dual formula (see Villani (2003)),

$$d_{1,q,T}(\nu_1,\nu_2) = \left\{ E\left[\sup_{t\in[0,T]} \mathcal{W}_1\left(\nu_t^1,\nu_t^2\right)^q \right] \right\}^{\frac{1}{q}} \tag{56}$$

$$= \left\{ E\left[\sup_{t\in[0,T]} \left(\sup_{\phi\in\mathcal{R}_1} \int_{\mathbb{R}^n} \phi(x)\left(\nu_t^1(dx) - \nu_t^2(dx)\right)\right)^q \right] \right\}^{\frac{1}{q}}$$

where $\mathcal{R}_1$ is the set of Lipschitz functions $\phi : \mathbb{R}^n \to \mathbb{R}$ with the Lipschitz coefficient

$$\|\phi\|_{\rm Lip} := \sup_{x,y\in\mathbb{R}^n,\ x\neq y} \frac{|\phi(x)-\phi(y)|}{|x-y|} \leq 1.$$

Metric between μ^N and μ III

- In the metric (56), $\mathcal{R}_1$ can be replaced with $\mathcal{R}_1^b$ which is the set of bounded functions $\phi \in \mathcal{R}_1$, see Villani (2003).
- We can further reduces the set $\mathcal{R}_1$ in the metric $d_{1,q,T}$ to the set $\mathcal{R}_1^{b,1}$ which is the set of Lipschitz functions $\phi : \mathbb{R}^n \to \mathbb{R}$ with $\|\phi\|_{\rm Lip} + \|\phi\|_\infty \leq 1$. In other words, we have non-decreasing sets:

$$\mathcal{R}_1^{b,1} \subset \mathcal{R}_1^b \subset \mathcal{R}_1.$$

- The metric $d_{1,q,T}$ with $\mathcal{R}_1$ replaced by $\mathcal{R}_1^{b,1}$ becomes that

$$\begin{aligned} & d_{1,q,T}^{\rm BL}(\nu_1, \nu_2) \qquad (57) \\ & := \left\{ E\left[\sup_{t\in[0,T]} \left(\sup_{\phi\in\mathcal{R}_1^{b,1}} \int_{\mathbb{R}^n} \phi(x)\left(\nu_t^1(dx) - \nu_t^2(dx)\right)\right)^q \right] \right\}^{\frac{1}{q}} \end{aligned}$$

Metric between μ^N and μ IV

- The distance $d^{\mathrm{BL}}_{1,q,T}$ is called Fortet-Mourier distance (see Section 11.2 of Dudley (2004), page 390).
- Then $(\mathcal{S}_{1,q,T}, d^{\mathrm{BL}}_{1,q,T})$ is a complete metric space.
- We next introduce a weaker metric.
- As in Lucon and Stannat (2014) AAP, we establish a metric $d_{q,T}$ between $\mu=(\mu_t)_{t\in[0,T]}$ and $\mu^N=(\mu_t^N)_{t\in[0,T]}$ as

$$d_{q,T}(\mu,\nu^N):=\sup_{t\in[0,T]} d_{BL}(\mu_t,\mu_t^N). \tag{58}$$

- d_{BL} is a metric for $\mathcal{P}(\mathbb{R}^n)$-valued r.v.s, which is defined as

$$d_{BL}(\mu_t,\mu_t^N):=\sup_{\psi\in\mathcal{R}_1^{b,1}} E\left[\left|\int_{\mathbb{R}^n}\psi(x)(\mu_t-\mu_t^N)(dx)\right|^q\right]^{\frac{1}{q}}. \tag{59}$$

Metric between μ^N and μ V

- Here, we work with the construction of the propagation of chaos under the distance $d_{q,T}$ defined by (57).

Theorem (Propagation of Chaos)

Recall the assumption $(\mathbf{A})_{X_0^i}$ *with additional assumptions discussed in this chapter. Then, for any* $T>0$ *and* $N\geq 1$*, there exists a constant* $C>0$ *which is independent of* N *such that*

$$d_{q,T}(\mu,\mu^N)\leq C\left[\left(\int_D |x|^p\rho_0(dx)\right)^{\frac{q}{p}}\alpha(p,q,n,N)+\frac{1}{N^{q-1}}\right],\quad q\geq 2,$$

where the first convergence rate $\alpha(p,q,n,N)$ *is given by* (45).

Metric between μ^N and μ VI

- Proof. Using (94), it results in, for all $\phi \in \mathcal{R}_1^{b,1}$,

$$\langle \mu_T^N - \mu_T, \phi \rangle = \langle \mu_0^N - \mu_0, P_{0,T}^0 \phi \rangle + \frac{1}{N} \sum_{i=1}^N \int_0^T \nabla_x P_{s,T}^0 \phi(X_s^i)^\top \sigma(X_s^i) dW_s^i. \tag{60}$$

- Let $\phi_0 := K_{b,\sigma,T}^{-1} P_{0,T}^0 \phi$. Then, by applying the estimate (48) in Lemma 24, we obtain that $\|\phi_0\|_{\mathrm{Lip}} \leq 1$.
- Then, it follows from Kantorovich-Rubinstein dual formula that, for $q \geq 2$,

$$E\left[\left|\langle \mu_0^N - \mu_0, P_{0,T}^0 \phi \rangle\right|^q\right] = K_{b,\sigma,T} E\left[\left|\langle \mu_0^N - \mu_0, \phi_0 \rangle\right|^q\right] \leq K_{b,\sigma,T} E\left[\mathcal{W}_1(\mu_0^N, \rho_0)^q\right] \leq K_{b,\sigma,T} E\left[\mathcal{W}_q(\mu_0^N, \rho_0)^q\right].$$

Metric between μ^N and μ VII

- By the convergence rate estimate (45) under the assumption $\mathbf{(A)}_{X_0^i}$, it holds that

$$E\left[\left|\langle \mu_0^N-\mu_0, P_{0,T}^0\phi\rangle\right|^q\right]\leq K_{b,\sigma,T}E\left[\mathcal{W}_q(\mu_0^N,\rho_0)^q\right]$$
$$\leq K_{b,\sigma,T}C\left(\int_D|x|^p\rho_0(dx)\right)^{\frac{q}{p}}\alpha(p,q,n,N).$$

- Here, the constant $C>0$ which is independent of N is given in (45).
- We next estimate the 2nd term of the r.h.s. of the equality (60).

Metric between μ^N and μ VIII

- Using the estimate (48) in Lemma 24, the BDG inequality yields that, for $q \geq 2$,

$$
\begin{aligned}
&E\left[\sup_{t\in[0,T]}\left|\frac{1}{N}\sum_{i=1}^N\int_0^t \nabla_x P_{s,T}^0\phi(X_s^i)^\top\sigma(X_s^i)dW_s^i\right|^q\right]\\
&\quad\leq \frac{C_{q,T}}{N^q}\sum_{i=1}^N E\left[\left(\int_0^T\left|\nabla_x P_{s,T}^0\phi(X_s^i)\right|^2\left|\sigma(X_s^i)\right|^2ds\right)^{q/2}\right]\\
&\quad\leq \frac{C_{q,T}}{N^q}\sum_{i=1}^N E\left[\left(\int_0^T\left\|\nabla_x P_{s,T}^0\phi\right\|_\infty^2\left|\sigma(X_s^i)\right|^2ds\right)^{q/2}\right]\\
&\quad\leq \frac{C_{q,T}}{N^{q-1}}T^{\frac{q}{2}-1}K_{b,\sigma,T}^q\left(\int_0^T E\left[\frac{1}{N}\sum_{i=1}^N|\sigma(X_s^i)|^q\right]ds\right). \qquad (61)
\end{aligned}
$$

Metric between μ^N and μ IX

- Using the assumption $(\mathbf{A})_{lip}$ and $(\mathbf{A})_{X_0^i}$, we have

$$
\begin{aligned}
\sup_{t\in[0,T]} E\left[\frac{1}{N}\sum_{i=1}^{N}|\sigma(X_t^i)|^q\right] &\leq C_T\left\{1+\sup_{t\in[0,T]} E\left[\frac{1}{N}\sum_{i=1}^{N}|X_t^i|^q\right]\right\}\\
&\overset{X^i\ i.i.d.}{=\!=} C_T\left\{1+\sup_{t\in[0,T]} E\left[|X_t^1|^q\right]\right\}\\
&<\infty.
\end{aligned}
$$

- Thus, we complete the proof of the theorem.

Metric between μ^N and μ X

- Question: Let us assume that $\sigma(x) \equiv \sigma \in \mathbb{R}^{n\times m}$. At the moment, can you relax the assumption $(\mathbf{A}_{lip})$ on the coefficient b so that the propagation of chaos still holds? For instant, when b satisfies the so-called one-sided Lipschitz condition, i.e.,

$$(x_1 - x_2)^\top (b(x_1) - b(x_2)) \le L|x_1 - x_2|^2, \quad x_1, x_2 \in \mathbb{R}^n, \tag{62}$$

  where the coefficient $L \in \mathbb{R}$.

- Hints: One of main argument is to establish a sequence of functions $(b_n)_{n\ge 1}$ which have the same regularity to that of b such that $b_n \to b$ as $n \to \infty$ in some sense. The key point is to find a constant $C > 0$ independent of n s.t.

$$d_{q,T}(\mu^{N,n}, \mu^n) \le C\alpha(1/N).$$

McKean-Vlasov Equation I

- Let us introduce the following particle system with mean field: for $i = 1, \dots, N$,

$$dX_t^i = b(X_t^i, \overline{X}_t^\rho)dt + \sigma(X_t^i, \overline{X}_t^\rho)dW_t^i, \quad X_0^i \in \mathbb{R}^n. \tag{63}$$

- The random variables X_0^i, $i \geq 1$, are i.i.d. with common law $\rho_0 \in \mathcal{P}(\mathbb{R}^n)$.
- The mean-field term is defined as

$$\overline{X}_t^\rho := \frac{1}{N} \sum_{i=1}^N \rho(X_t^i), \tag{64}$$

where $\rho : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function.

McKean-Vlasov Equation II

- The empirical measure-valued process related to the particle system (63) is defined as:

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}, \quad \text{on } \mathcal{B}(\mathbb{R}^n).$$

- Then, the state process of particle system $(X^i)_{i\geq 1}$ can be rewritten as:

$$dX_t^i = b(X_t^i, \langle \mu_t^N, \rho\rangle)dt + \sigma(X_t^i, \langle \mu_t^N, \rho\rangle)dW_t^i, \quad X_0^i \in \mathbb{R}^n. \tag{65}$$

- Question: Let the coefficients (b, σ, ρ) satisfy the assumption

$(\mathbf{A}_{b,\sigma,\rho})$ $b : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $\sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n\times n}$ and $\rho : \mathbb{R}^n \to \mathbb{R}$ are Lipschitz continuous with linear growth, i.e., $|b(x, z)| \vee |\sigma(x, z)| \vee |\rho(x)| \leq C(1 + |x|)$ for all $(x, z) \in \mathbb{R}^n \times \mathbb{R}$.

Prove that the system (65) admits a unique strong solution.

McKean-Vlasov Equation III

- What is the limit of $\mu^N = (\mu_t^N)_{t\in[0,T]}$ as $N \to \infty$?
- In order to address this issue, we follow the similar argument to that used in the derivation of the forward Kolmogorov equation.
- To this purpose, define the operator as: for $(x, \mu) \in \mathbb{R}^n \times \mathcal{P}_1(\mathbb{R}^n)$, and $f \in C^2(\mathbb{R}^n)$,

$$\mathcal{A}^{\mu} f(x) := b(x, \langle \mu, \rho \rangle)^{\top} \nabla_x f(x) + \frac{1}{2} \mathrm{tr}[\sigma\sigma^{\top}(x, \langle \mu, \rho \rangle) \nabla_x^2 f(x)]. \quad (66)$$

- For any $f \in C_b^2(\mathbb{R}^n)$, we have from Itô formula that

$$\begin{aligned}\langle \mu_t^N, f \rangle &= \langle \mu_0^N, f \rangle + \int_0^t \langle \mu_s^N, \mathcal{A}^{\mu_s^N} f \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla_x f(X_s^i)^{\top} \sigma(X_s^i) dW_s^i\end{aligned}$$

McKean-Vlasov Equation IV

- Formally, using Martingale Convergence Theorem, a.s, $N \to \infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\int_0^t \nabla_x f(X_s^i)^\top \sigma(X_s^i)dW_s^i \to 0$$

- Then, we have as $N \to \infty$,

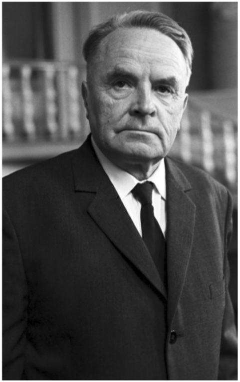
The McKean-Vlasov equation:

$$\langle \mu_t, f\rangle = \langle \mu_0, f\rangle + \int_0^t \langle \mu_s, \mathcal{A}^{\mu_s} f\rangle ds, \quad t \in [0, T]. \tag{67}$$

- A little bit of history:

McKean-Vlasov Equation V

(a) Mark Kac



(b) Anatoly Vlasov

- The story of these processes started with a stochastic toy model for the Vlasov equation of plasma proposed by Mark Kac in his paper "*Foundations of kinetic theory (1956)*".

McKean-Vlasov Equation VI

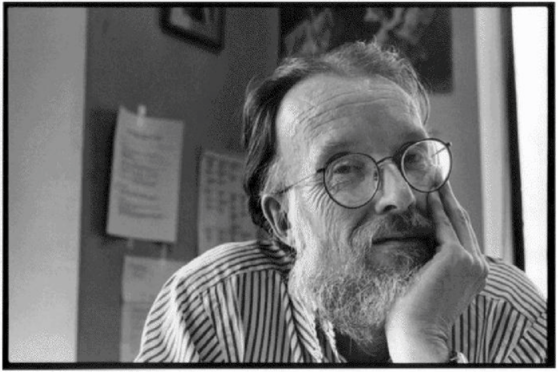


Figure: Henry P. McKean (1930-)

- In 1966, Henry P. McKean published his seminal paper "*A class of Markov processes associated with non-linear parabolic equations*".

McKean-Vlasov Equation VII

- We have two issues which should be addressed:
  - (Q1) Well-posedness of the McKean-Valsov equation (67).
  - (Q2) Propagation of Chaos of the McKean-Valsov equation (67).
- We can apply the similar argument to that in the study of the propagation of chaos of FPK equations.
- Recall (55), (56) and (57) in the previous section:
- For $p \geq 1$ and $q \geq 2$, introduce the metric on $\mathcal{S}_{p,q,T}$ as follows: for all $\nu^1, \nu^2 \in \mathcal{S}_{p,q,T}$,

$$d_{p,q,T}(\nu_1, \nu_2) := \left\{ E\left[\sup_{t\in[0,T]} \mathcal{W}_p\left(\nu_t^1, \nu_t^2\right)^q \right] \right\}^{\frac{1}{q}}.$$

McKean-Vlasov Equation VIII

- Consider $p=1$ and using the Kantorovich-Rubinstein dual formula (see Villani (2003)),

$$\begin{aligned} d_{1,q,T}(\nu_1,\nu_2) &= \left\{E\left[\sup_{t\in[0,T]}\mathcal{W}_1\left(\nu_t^1,\nu_t^2\right)^q\right]\right\}^{\frac{1}{q}} \\ &= \left\{E\left[\sup_{t\in[0,T]}\left(\sup_{\phi\in\mathcal{R}_1}\int_{\mathbb{R}^n}\phi(x)\left(\nu_t^1(dx)-\nu_t^2(dx)\right)\right)^q\right]\right\}^{\frac{1}{q}} \end{aligned}$$

where $\mathcal{R}_1$ is the set of Lipschitz functions $\phi:\mathbb{R}^n\to\mathbb{R}$ with the Lipschitz coefficient

$$\|\phi\|_{\rm Lip} := \sup_{x,y\in\mathbb{R}^n,\ x\neq y}\frac{|\phi(x)-\phi(y)|}{|x-y|}\leq 1.$$

McKean-Vlasov Equation IX

- In the metric (56), $\mathcal{R}_1$ can be replaced with $\mathcal{R}_1^b$ which is the set of bounded functions $\phi \in \mathcal{R}_1$, see Villani (2003).
- We can further reduces the set $\mathcal{R}_1$ in the metric $d_{1,q,T}$ to the set $\mathcal{R}_1^{b,1}$ which is the set of Lipschitz functions $\phi : \mathbb{R}^n \to \mathbb{R}$ with $\|\phi\|_{\mathrm{Lip}} + \|\phi\|_\infty \leq 1$. In other words, we have an non-decreasing sets:

$$\mathcal{R}_1^{b,1} \subset \mathcal{R}_1^b \subset \mathcal{R}_1.$$

- The metric $d_{1,q,T}$ with $\mathcal{R}_1$ replaced by $\mathcal{R}_1^{b,1}$ becomes that

$$d_{1,q,T}^{\mathrm{BL}}(\nu_1, \nu_2)$$
$$:= \left\{ E\left[\sup_{t\in[0,T]} \left(\sup_{\phi\in\mathcal{R}_1^{b,1}} \int_{\mathbb{R}^n} \phi(x) \left(\nu_t^1(dx) - \nu_t^2(dx)\right) \right)^q \right] \right\}^{\frac{1}{q}}$$

McKean-Vlasov Equation X

- The distance $d^{\mathrm{BL}}_{1,q,T}$ is called Fortet-Mourier distance (see Section 11.2 of Dudley (2004), page 390).
- Then $(\mathcal{S}_{1,q,T}, d^{\mathrm{BL}}_{1,q,T})$ is a complete metric space.
- Therefore, we introduce the following Itô SDE: for $x_0 \in \mathbb{R}^n$ and $\nu \in \mathcal{S}_{1,q,T}$,

$$X_t^{x_0,\nu} = x_0 + \int_0^t b(X_s^{x_0,\nu}, \langle \nu_s, \rho \rangle)ds + \int_0^t \sigma(X_s^{x_0,\nu}, \langle \nu_s, \rho \rangle)dW_s. \quad (68)$$

- Let us define that

$$L_t^{\nu}(dx) = \int_{\mathbb{R}^n} E\left[\delta_{X_t^{x_0,\nu}}(dx)\right] \rho_0(dx_0) \quad (69)$$

- Question: For any $\nu \in \mathcal{S}_{1,q,T}$, provide a mild condition on $\rho_0 \in \mathcal{P}(\mathbb{R}^n)$ under which $L^{\nu} = (L_t^{\nu})_{t\in[0,T]}$ belongs to $\mathcal{S}_{1,q,T}$.

McKean-Vlasov Equation XI

- Question: Under the above Question with the assumption $(\mathbf{A}_{b,\sigma,\rho})$, prove that $L^{\nu}:\mathcal{S}_{1,q,T}\to\mathcal{S}_{1,q,T}$ admits a fixed point μ. In other words, we have

$$L^{\mu}=\mu,\quad \text{under } (\mathcal{S}_{1,q,T},d^{\rm BL}_{1,q,T}). \tag{70}$$

- Based on the fixed point (70), for $f\in C_0^{\infty}(\mathbb{R}^n)$,

$$\langle\mu_t,f\rangle=\langle L_t^{\mu},f\rangle=\langle\rho_0,f\rangle+\int_0^t\langle L_s^{\mu},\mathcal{A}^{\mu_s}f\rangle ds=\langle\rho_0,f\rangle+\int_0^t\langle\mu_s,\mathcal{A}^{\mu_s}f\rangle ds.$$

- Then, we have

Existence of McKean-Valsov equation: The above fixed point $\mu\in\mathcal{S}_{1,q,T}$ given by (70) is a solution of McKean-Valsov equation (67) with $\mu_0=\rho_0$.

- Question: Establish the propagation of chaos for McKean-Valsov equation (67) with $\mu_0=\rho_0$.

Course Outline

Mathematical Model of Evolutionary Branching I

- Charles Robert Darwin (1809-1882): English naturalist whose scientific theory of evolution by natural selection became the foundation of modern evolutionary studies.

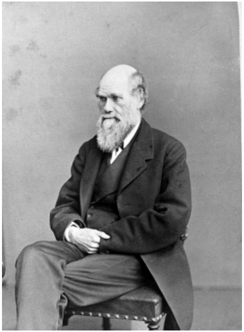
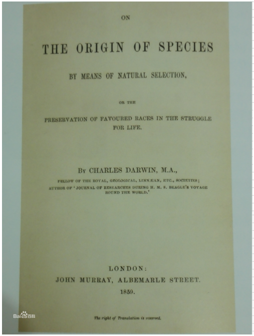


Figure: Charles Robert Darwin

Mathematical Model of Evolutionary Branching II

- In 1859, Darwin published "*On the Origin of Species by Means of Natural Selection*". The book immediately became controversial. http://darwin-online.org.uk/



ON

THE ORIGIN OF SPECIES

BY MEANS OF NATURAL SELECTION,

OR THE

PRESERVATION OF FAVOURED RACES IN THE STRUGGLE FOR LIFE.

BY CHARLES DARWIN, M.A.,

FELLOW OF THE ROYAL, GEOLOGICAL, LINNÆAN, ETC., SOCIETIES;
AUTHOR OF 'JOURNAL OF RESEARCHES DURING H. M. S. BEAGLE'S VOYAGE ROUND THE WORLD.'

LONDON:
JOHN MURRAY, ALBEMARLE STREET.
1859.

The right of Translation is reserved.

【目录】

Figure: Book: "Origin of Species"

Mathematical Model of Evolutionary Branching III

- Recently, the 2019 Novel Coronavirus, or 2019-nCoV outbreak in Wuhan.
- The recently emerged 2019-nCoV is not the same as the coronavirus that causes Middle East Respiratory Syndrome (MERS) or the coronavirus that causes Severe Acute Respiratory Syndrome (SARS).
- However, genetic analyses suggest this virus emerged from a virus related to SARS.
- There are ongoing investigations to learn more. This is a rapidly evolving situation.

总的来看，易感染新型冠状病毒的人群就是免疫力低下的人群。比如说，我们呼入了一个新型冠状病毒停留在肺部，然后这个病毒会分裂，免疫力好的人能阻止病毒的分裂，而免疫力差的人就不能阻挡病毒的分裂，从而最终确诊新型肺炎。

Mathematical Model of Evolutionary Branching IV

国家卫生健康委员会副主任李斌则称新型冠状病毒存在变异可能，病毒的变异情况并非罕见，有些病毒甚至会出现突变情况，比如结构改变导致某种功能突然强化，如果这种病毒变异成更容易传染和更难治愈的情况的话，那么疫情很可能会进一步扩大，不过中国科学院上海巴斯德研究所研究员郝沛等人分析认为目前状态下的武汉新型冠状病毒的传染性要低于2003年的“SARA”病毒，治疗难度也低于SARS病毒，所以我们应当尽快消灭这次疫情，因为时间越长，这种病毒的变异可能性就越高，应对的措施也就越复杂，抓紧时间控制住疫情发展状态，是当前的重中之重。

Figure: Replication and Mutation of 2019-nCoV.

Mathematical Model of Evolutionary Branching V

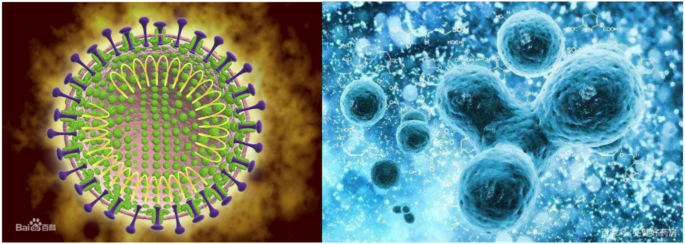


Figure: Coronavirus and its replication

- We next introduce a mathematical model for describing a molecular evolution.
- Consider first two species (e.g. two types of viruses) $\{1,2\}$:
- For $i=1,2$, let $X_t^i \geq 0$ be the population size of the species i at time $t \geq 0$.

Mathematical Model of Evolutionary Branching VI

- Let $r_i > 0$ be the fitness or replicating rate of the i-th species, i.e.,
  - every $\frac{1}{r_i}$ generations replicates once.
- Then, we have, for $i = 1, 2$,

$$\frac{dX_t^i}{dt} = r_i X_t^i, \quad t \geq 0. \tag{71}$$

- Assume that the scale of the population is finite and conservative.
- Make a normalization s.t. $X_t^1 + X_t^2 = 1$.
- We have the following system of equations as follows: for $i = 1, 2$,

$$\frac{dX_t^i}{dt} = X_t^i(r_i - g(X_t, r)), \quad t \geq 0. \tag{72}$$

Mathematical Model of Evolutionary Branching VII

- The term $g(x, r)$ is called

Average Fitness Function: for $(x_1, x_2) \in S_2$ and $r = (r_1, r_2) \in \mathbb{R}_+^2$,

$$g(x, r) := r_1 x_1 + r_2 x_2.$$

Here S_2 is a simplex, i.e., $S_2 := \{(x_1, x_2) \in \mathbb{R}_+^2;\ x_1 + x_2 = 1\}$.

- Taking the mutation into Replicator Equation (72).

Let q_{ij} be the probability of the type i mutating to the type j.

- Then, we have, for $i = 1, 2$,

$$\frac{dX_t^i}{dt} = \underbrace{r_1 X_t^1 q_{1i}}_{\text{no mutation}} + \underbrace{r_2 X_t^2 q_{2i}}_{\text{2 mut. to 1}} - X_t^i g(X_t, r) \tag{73}$$

Mathematical Model of Evolutionary Branching VIII

- Extending it to the general N types of species:

Replicator-Mutator (RM) Equation: For $i = 1, \ldots, N$, and $t \geq 0$,

$$\frac{dX_t^i}{dt} = \sum_{j=1}^{N} r_j X_t^j q_{ji} - X_t^i g(X_t, r). \tag{74}$$

Initial Behavior: $\sum_{i=1}^{N} X_0^i = 1$.
Average Fitness Function: $g(x, r) = \sum_{i=1}^{N} r_i x_i$ for $(x, r) \in S_N \times \mathbb{R}_+^N$.
Mutation Matrix: $Q = (q_{ij})$.

- Question: Solve RM equation (74). Make a transform $Y_t^{g,i} := X_t^i \exp(\int_0^t g(X_s, r)ds)$. Then $dY_t^{g,i} = \sum_{j=1}^{N} r_j Y_t^{g,j} q_{ji} dt$, i.e., $dY_t^g = (rI)_{N\times N} Q Y_t^g dt$.
- Question: How to estimate r_i and q_{ij} using RM equation (74)?

Mathematical Model of Evolutionary Branching IX

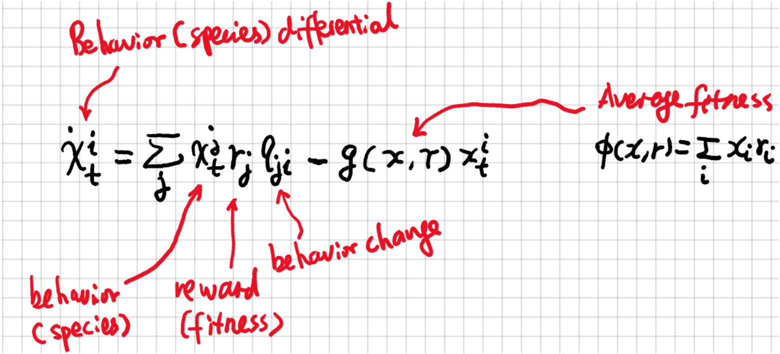


Figure: Summary of Replicator-Mutator Equations

Mean-Field Approach for RM Equation I

- We next introduce a Mean-Field approach for modifying RM equation in the study of the evolution of RNA virus populations in the following paper on PRL.

VOLUME 76, NUMBER 23 PHYSICAL REVIEW LETTERS 3 JUNE 1996

RNA Virus Evolution via a Fitness-Space Model

Lev S. Tsimring and Herbert Levine
Institute for Nonlinear Science, University of California, San Diego, La Jolla, California 92093-0402

David A. Kessler
Department of Physics, Bar-Ilan University, Ramat Gan 52900, Israel
(Received 16 January 1996)

Figure: Paper on PRL: A Mean-Field Approach for RM Dynamics

Mean-Field Approach for RM Equation II

An important observation: every sequence of RNA viruses can be characterized by its replication rate $r \in \mathbb{R}$. There may be different sequence which exhibit similar replication rates.

- Then, we can treat the all sequences which exhibits the similar replication rates as a population:
  - In other words, different r corresponds to different population.
- For any $t \geq 0$, let $r \to u(t, r)$ is a probability density function on $\mathbb{R}$, i.e., $\int_{\mathbb{R}} u(t, r)dr = 1$.
  - The solution X_t^i of RM equation (74) looks like a discrete version of $u(t, r)$:

$$\sum_{i=1}^{N} X_t^i = 1 \ \approx\!\approx\!\approx \ \sum_{i} \int_{r_i}^{r_{i+1}} u(t, r)dr = 1.$$

Mean-Field Approach for RM Equation III

- Without mutations, the discrete version (71) gives that, the replication dynamics is given by

$$\partial_t u(t,r) = r u(t,r).$$

- However, the solution $r \to u(t,r)$ of the above equation is *not* a probability density function:
  - A normalization should be made in the above equation as in the discrete version (72).
  - This yields that

$$\partial_t u(t,r) = \left(r - \int_{\mathbb{R}} r u(t,r) dr \right) u(t,r). \tag{75}$$

Mean-Field Approach for RM Equation IV

- Consider a simple mutation without the underlying genomic transition rate. Then Mean-Field version of RM equation is given by

RM Equation in Mean-Field Form: For the fitness space given by $\mathbb{R}$, for $(t,r) \in [0,T] \times \mathbb{R}$,

$$\partial_t u(t,r) = \underbrace{\frac{\sigma^2}{2}\Delta_r u(t,r)}_{\text{mutations}} + \underbrace{\left(r - \int_{\mathbb{R}} r u(t,r)dr\right) u(t,r)}_{\text{replication}}, \tag{76}$$

$$u(0,r) = u_0(r), \quad r \in \mathbb{R}, \tag{77}$$

where $u_0(r) \geq 0$ and $\int_{\mathbb{R}} u_0(r)dr = 1$.

- The solution of RM equation (76) and (77) admits a unique smooth solution which is studied by

Mean-Field Approach for RM Equation V

- Alfaro, M., and R. Carles (2014): Explicit solutions for replicator-mutator equations: extinction versus acceleration. *SIAM J. Appl. Math.* 74, 1919-1934.
- Question: Without referring to the paper by Alfaro and Carles (2014), establish the closed-form solution $u(t, r)$ of RM Equation (76) and (77) in Mean-Field form.

Extended RM Equations I

- Recall that $D \subseteq \mathbb{R}^n$ is domain, which is *not* necessarily bounded.
- Consider the following extended RM equation with fitness space D:

$$\partial_t u(t,r) = \underbrace{\mathcal{A}^* u(t,r)}_{\text{mutations}} + \underbrace{\left(g(r) - \int_D g(y)u(t,y)dy\right) u(t,r)}_{\text{replication with fitness function } g}, \tag{78}$$

$$u(0,r) = u_0(r), \quad r \in D,$$

where $u_0(r) \geq 0$ for $r \in D$ and $\int_D u_0(r)dr = 1$.

- Here, $\mathcal{A}^*$ is the adjoint operator of the operator $\mathcal{A}$ given by: for $f \in C^2(D)$,

$$\mathcal{A}f(x) = b(x)^\top \nabla_x f(x) + \frac{1}{2}\mathrm{tr}[\sigma\sigma^\top(x)\nabla_x^2 f(x)], \quad x \in D.$$

Extended RM Equations II

- Consider the Itô SDE given by: for $(t,x) \in [0,T] \times D$,

$$X_s^{t,x} = x + \int_t^s b(X_v^{t,x})dv + \int_t^s \sigma(X_v^{t,x})dW_v, \quad s \in [t,T].$$

- We assume $(\mathbf{A}_{b,\sigma})$ and $(\mathbf{A}_X)$ hold. Then, $X_s^{t,x} \in D$ for all $s \in [t,T]$, P-a.s.

Consider the weak solution of the extended RM equation (78), which is defined as: for all $f \in C_0^\infty(\mathbb{R}^n)$,

$$\begin{aligned}\langle u(t), f\rangle = \langle u_0, f\rangle &+ \int_0^t \langle u(s), (\mathcal{A}+g)f\rangle ds \\ &- \int_0^t \langle u(s), g\rangle \langle u(s), f\rangle ds,\end{aligned} \tag{79}$$

where the integral $\langle u(t), f\rangle := \int_D u(t,r)f(r)dr$.

Extended RM Equations III

- Relating the weak form (79) of the extended RM equation to the following FPK equation: for $f \in C_0^\infty(D)$,

$$\langle \mu_t, f\rangle = \langle \rho_0, f\rangle + \int_0^t \langle \mu_s, (\mathcal{A}+g)f\rangle ds - \int_0^t \langle \mu_s, f\rangle\langle \mu_s, g\rangle ds, \quad (80)$$

  where the initial datum $\rho_0 \in \mathcal{P}(D)$.
- Moreover, let the initial datum ρ_0 admit a density function given by $u_0(r)$ for $r \in D$, i.e., $\rho_0(dr) = u_0(r)dr$.
- Then, for $t \in [0, T]$, the solution $\mu_t(dr) = u(t, r)dr$, where $u(t, x)$ for $(t, x) \in [0, T] \times D$ solves the extended weak form of RM equation (79).
- We next discuss the well-posedness of FPK equation (80).
- Question: Establish a solution of FPK equation (80).
- Hints: Using the Itô diffusion process $X^{t,x}$.

Extended RM Equations IV

- Question: How to establish the propagation of chaos on FPK equation (80)?
- Hints: First, you should construct a particle system $X^i = (X_t^i)_{t\in[0,T]}$ related to FPK equation (80).
- We need to impose the assumption on the fitness function g:

($\mathbf{A}_{gl}$) (i) $g : D \to \mathbb{R}$ is continuous and bounded from above; (ii) there exists a polynomial $Q_g : \mathbb{R}_+ =\!\!\to \mathbb{R}_+$ such that

$$|g(x) - g(y)| \leq Q_g(|x| + |y|)|x - y|, \quad \forall\ x, y \in D.$$

($\mathbf{A}_{X_0}$) For $q \geq 2$, the sequence of initial states of particles $(X_0^i)_{i\geq 1}$ is i.i.d. according to $\rho_0 \in \mathcal{P}_{(\deg(Q_g)+1)q}(D)$.

($\mathbf{A}_{b,\sigma}$) i.e., **($\mathbf{A}_{lip}$)**

Extended RM Equations V

- For the particle system, we let the dynamics of the state process of i-th particle be

$$dX_t^i = b(X_t^i)dt + \sigma(X_t^i)dW_t^i, \quad X_0^i \in D.$$

- We then introduce the sequence of $\mathcal{P}(D)$-valued process

$$\mu_t^N = \frac{\frac{1}{N}\sum_{i=1}^N \delta_{X_t^i} \exp\left(\int_0^t g(X_s^i)ds\right)}{\frac{1}{N}\sum_{i=1}^N \exp\left(\int_0^t g(X_s^i)ds\right)}$$
$$\Updownarrow \tag{81}$$
$$\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i} \exp\left(\int_0^t (g(X_s^i) - \langle \mu_s^N, g\rangle)ds\right).$$

- Note that μ^N and μ are sub-probability measure.

Extended RM Equations VI

- Let $\mathcal{P}^s(D)$ be the set of sub-probability measures on $\mathcal{B}_D$, i.e., for any $\mu \in \mathcal{P}^s(D)$, μ is a finite measure on $\mathcal{B}_D$ such that $\mu(D) \leq 1$.
- We next introduce the (Alexandroff) one-point compactification.
- Add one point which is outside of D to D called "$\star$" and define $D_\star := D \cup \{\star\}$.
- Let D be topologized by a topology $\mathcal{T}$, and we then can define a topology $\mathcal{T}^\star$ for $D_\star$ as follows:
  - (i) each open subset of D is also in $\mathcal{T}^\star$, i.e., $\mathcal{T} \subset \mathcal{T}^\star$;
  - (ii) for each compact set $C \subset D$, define an element $U_C \in \mathcal{T}^\star$ by $U_C := (D \setminus C) \cup \{\star\}$. Let us define a bijection $\iota : \mathcal{P}^s(D) \to \mathcal{P}(D_\star)$ as:

$$(\iota\mu)(A) := \mu(A \cap D) + (1 - \mu(D))\delta_\star(A),$$

where $A \in \mathcal{B}(D_\star)$ and $\mu \in \mathcal{P}^s(D)$.

Extended RM Equations VII

- Then, the integral of $\mu \in \mathcal{P}^s(D)$ w.r.t. a measurable function $f : D_\star \to \mathbb{R}$ is defined as (if it is well-defined):

$$\begin{aligned}\int_{D_\star} f(x)(\iota\mu)(dx) &= \int_D f(x)\mu(dx) + f(\star)(1-\mu(D))\\ &= \langle \mu, f\rangle + f(\star)(1-\mu(D)).\end{aligned}$$

- Consider $\mu = (\mu_t)_{t\in[0,T]}$ as an arbitrary $\mathcal{P}(D)$-valued solution of extended EM equation, and we then define that

$$\mu_t^g := \exp\left(\int_0^t \langle \mu_s, g\rangle ds\right)\mu_t, \quad t\in[0,T]. \tag{82}$$

- This implies that, for all $f \in \mathcal{D}$,

$$\langle \mu_t^g, f\rangle = \langle \rho_0, f\rangle + \int_0^t \langle \mu_s^g, (\mathcal{A}+g)f\rangle ds. \tag{83}$$

Extended RM Equations VIII

- In view of (1.2) and (1.4) in Manita et al. (2015), the following equivalent representation holds, for all $\phi \in C_b^{1,2}([0,T] \times D)$,

$$\langle \mu_t^g, \phi(t,\cdot)\rangle = \langle \mu_0^g, \phi(0,\cdot)\rangle + \int_0^t \langle \mu_s^g, (\partial_t + \mathcal{A} + g)\phi(s,\cdot)\rangle \, ds.$$

- We also define that

$$\mu_t^{g,N} := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i} \exp\left(\int_0^t g(X_s^i) ds\right).$$

Extended RM Equations IX

- We introduce the following metrics:

$$\begin{aligned}\hat{d}_{q,T}(\mu,\mu^N) &:= \sup_{t\in[0,T]} d_{BL}(\iota\mu_t^{g,N},\iota\mu_t^g) \qquad (84)\\ &= \sup_{t\in[0,T]}\sup_{\psi\in\mathcal{R}_1} E\left[\left|\int_{D_\star}\psi(x)(\iota\mu_t^{g,N}-\iota\mu_t^g)(dx)\right|^q\right]^{\frac{1}{q}}.\end{aligned}$$

- Here, $\mathcal{R}_1$ is the set of (bounded) Lipschitz continuous functions $\psi: D_\star \to \mathbb{R}$ satisfying $\|\psi\|_\infty + \|\psi\|_{\rm Lip} \leq 1$ (where $\|\psi\|_\infty := \sup_{x\in D_\star}|\psi(x)|$ and $\|\psi\|_{\rm Lip}$ denotes the Lipschitz coefficient of ψ).
- Let $|\cdot|$ be the Euclidean norm. Then, we can define a metric $d_\star$ on $D_\star$ as in Mandelkern (1989):

Extended RM Equations X

- fix $x_0 \in D$ and define $l(x) := \frac{1}{1+|x-x_0|}$ for $x \in D$. For $x_1, x_2 \in D$, define $d_\star(x_1, x_2) := |x_1 - x_2| \wedge (l(x_1) + l(x_2))$, $d_\star(x, \star) := l(x)$ for $x \in D$, and $d_\star(\star, \star) = 0$. Then, the Lipschitz coefficient of ψ (as a seminorm) is given by

$$\|\psi\|_{\mathrm{Lip}} = \sup_{x_1 \neq x_2, x_1, x_2 \in D_\star} \frac{|\psi(x_1) - \psi(x_2)|}{d_\star(x_1, x_2)}.$$

- This implies that, for any $\psi \in \mathcal{R}_1$ and $x_1, x_2 \in D$,

$$|\psi(x_1) - \psi(x_2)| \leq |x_1 - x_2| \wedge (l(x_1) + l(x_2)) \leq |x_1 - x_2|. \tag{85}$$

- In other words, $\psi \in \mathcal{R}_1$ is also a (bounded) Lipschitz continuous function on D with the Lipschitz coefficient being less than one.

Extended RM Equations XI

- We have that, for all $\psi \in \mathcal{R}_1$,

$$\int_{D_\star} \psi(x)(\iota\mu_t^{g,N} - \iota\mu_t^g)(dx) = \int_D \psi_\star(x)(\mu_t^{g,N} - \mu_t^g)(dx)$$
$$= \langle \mu_t^{g,N} - \mu_t^g, \psi_\star \rangle, \tag{86}$$

- Here $\psi_\star(x) := \psi(x) - \psi(\star)$ for $x \in D$. Then $\psi_\star$ is a (bounded) Lipschitz continuous function on D with the Lipschitz coefficient being less than one.

Extended RM Equations XII

Lemma

Let assumptions $(\mathbf{A}_{loc})$, $(\mathbf{A}_X)$ *and* $(\mathbf{A}_{gl})$-(i) *hold. Suppose also that*

- $(\mathbf{A}_{D,\sigma})$ *there exists a sequence* $(D_k)_{k\in\mathbb{N}}$ *of bounded domains with* $\overline{D}_k \subset D$ *such that* $\bigcup_{k=1}^{\infty} D_k = D$, *each* D_k *has a* C^2*-boundary; and for each* $k \geq 1$, $\sigma\sigma^\top(x)$ *is uniform elliptic on* $\mathbb{R}^n$ *for* $(t,x) \in [0,T) \times D_k$.

Then, the propagator $P^g_{t,s}f$ *satisfies that, for* $(s,x) \in [t,T] \times D$,

$$\partial_t P^g_{t,s}f(x) + (\mathcal{A}+g)P^g_{t,s}f(x) = 0, \quad P^g_{s,s}f(x) = f(x). \tag{87}$$

Moreover $P^g_{\cdot,s}f \in C^{1,2}((0,s]\times D) \cap C([0,s]\times D)$, *and there exists a unique classical solution of the Cauchy problem* (87).

Extended RM Equations XIII

- We next show the uniqueness of solutions to extended RM equation using the propagator $P^g_{t,s}$.

Lemma

Let assumptions $(\mathbf{A}_{X_0})$, $(\mathbf{A}_{b,\sigma})$, $(\mathbf{A}_{D,\sigma})$ *and* $(\mathbf{A}_{gl})$ *hold. Let* $\mu=(\mu_t)_{t\in[0,T]}$ *be a solution of extended RM equation satisfying* $\mu_0=\rho_0$ *and the integrability condition:*

$$\sup_{t\in[0,T]}\int_D |x|^{\deg(Q_g)+1}\mu_t(dx)<+\infty. \tag{88}$$

Then μ *is unique.*

- Proof. Let $\hat{\mu}=(\hat{\mu}_t)_{t\in[0,T]}$ be another solution with $\hat{\mu}_0=\rho_0$.
- We then define $\hat{\mu}^g_t:=\exp(\int_0^t\langle\hat{\mu}_r,g\rangle dr)\hat{\mu}_t$ for $t\in[0,T]$.

Extended RM Equations XIV

- We first show that

$$\langle \mu_t^g, \psi \rangle = \langle \hat{\mu}_t^g, \psi \rangle, \quad t \in [0, T], \ \psi \in \mathcal{R}_1. \tag{89}$$

- Let $s \in [0, T]$ be an arbitrary (fixed) time and define $h(t, x) := P_{t,s}^g \psi(x)$ with $(t, x) \in [0, s] \times D$ and $\psi \in \mathcal{R}_1$.
- However, h is not in $C_b^{1,2}([0, s] \times D)$.
- We hence introduce the following cut-off function $\xi \in C_0^\infty(\mathbb{R}^n)$ and it satisfies that $\xi(x) \in [0, 1]$ for all $x \in \mathbb{R}^n$, and $\xi(x) = 1$ for $|x| \leq 1$; $\xi(x) = 0$ for $|x| > 2$. Moreover, for $N \geq 1$, define $\xi_N(x) := \xi(x/N)$ for $x \in \mathbb{R}^n$.
- We have $h_N(t, x) := \xi_N(x) h(t, x)$ is in $C_b^{1,2}([0, s] \times D)$. Therefore, for all $t \in [0, s]$,

$$\langle \mu_t^g, h_N(t, \cdot) \rangle = \langle \mu_0^g, h_N(0, \cdot) \rangle + \int_0^t \langle \mu_r^g, (\partial_t + \mathcal{A} + g) h_N(r, \cdot) \rangle dr.$$

Extended RM Equations XV

- In terms of the definition of h_N, it holds that

$$(\partial_t + \mathcal{A} + g)h_N(t,x) = h(t,x)\mathcal{A}\xi_N(x) + \nabla_x \xi_N(x)^\top \sigma\sigma^\top \nabla_x h(t,x).$$

- Note that both $\nabla_x^2 \xi_N$ and $\nabla_x \xi_N$ are bounded (the boundedness is independent of N) and supported on $\{x;\ N \leq |x| \leq 2N\} \cap D$.
- Moreover, for $N \leq |x| \leq 2N$, there exists a constant C (which is independent of N) such that

$$|\nabla_x \xi_N(x)| = \frac{1}{N}\left|\nabla_x \xi(x/N)\right| \leq \frac{C}{|x|}.$$

Extended RM Equations XVI

- Let $\alpha := \deg(Q_g) + 1$. Then, there exits a constant C (which is independent of N) such that

$$\left|\int_D \left\{h(t,x)\mathcal{A}\xi_N(x) + \nabla_x \xi_N(x)^\top \sigma\sigma^\top \nabla_x h(t,x)\right\} \mu_t^g(dx)\right| \\ \leq C \int_{D_N} \{|x|^\alpha + 1\}\, \mu_t^g(dx), \tag{90}$$

  where $D_N := \{x : N \leq |x| \leq 2N\} \cap D$.

- Then, by the assumption (88), i.e., $\sup_{t\in[0,T]} \int_D |x|^\alpha \mu_t(dx) < \infty$ and the assumption $(\mathbf{A}_{gl})$-(i), there exists a constant C (which is independent of N) such that

$$\sup_{t\in[0,T]} \int_{D_N} \{|x|^\alpha + 1\}\, \mu_t^g(dx) \leq C \sup_{t\in[0,T]} \int_D \{|x|^\alpha + 1\}\, \mu_t(dx) < +\infty,$$

Extended RM Equations XVII

- This gives that the r.h.s. of (90) tends to 0 as $N \to \infty$. Hence

$$\lim_{N\to\infty}\int_D \left\{h(t,x)\mathcal{A}\xi_N(x) + \nabla_x\xi_N(x)^\top\sigma\sigma^\top\nabla_x h(t,x)\right\}\mu_t^g(dx) = 0.$$

- This yields that, for all $t \in [0,s]$,

$$\begin{aligned}\langle \mu_t^g, h(t,\cdot)\rangle &= \lim_{N\to+\infty}\langle \mu_t^g, h_N(t,\cdot)\rangle = \lim_{N\to+\infty}\langle \mu_0^g, h_N(0,\cdot)\rangle \\ &= \langle \mu_0^g, h(0,\cdot)\rangle. \end{aligned} \tag{91}$$

- Note that $\langle \mu_s^g, h(s,\cdot)\rangle = \langle \mu_s^g, \psi\rangle$ and hence (91) gives that

$$\langle \mu_s^g, \psi\rangle = \langle \mu_0^g, h(0,\cdot)\rangle = \langle \rho_0, h(0,\cdot)\rangle. \tag{92}$$

- The same reasoning yields that

$$\langle \hat{\mu}_s^g, \psi\rangle = \langle \hat{\mu}_0^g, h(0,\cdot)\rangle = \langle \rho_0, h(0,\cdot)\rangle. \tag{93}$$

Extended RM Equations XVIII

- Then, the equality (89) follows from (92) and (93) with the arbitrariness of $s \in [0, T]$. By choosing $\psi \equiv 1$, it follows from (89) that

$$\exp\left(\int_0^t \langle \mu_s, g\rangle ds\right) = \langle \mu_t^g, 1\rangle = \langle \hat{\mu}_t^g, 1\rangle = \exp\left(\int_0^t \langle \hat{\mu}_s, g\rangle ds\right)$$

- Therefore

$$\langle \mu_t, \psi\rangle = \langle \hat{\mu}_t, \psi\rangle, \quad t \in [0, T], \ \psi \in \mathcal{R}_1.$$

Extended RM Equations XIX

Lemma

Let the conditions of Lemma 28 hold. Then, for any fixed $T>0$, it holds that

$$\langle \mu_T^{g,N} - \mu_T^g, f\rangle = \langle \mu_0^{g,N} - \mu_0^g, P_{0,T}^g f\rangle + \frac{1}{N}\sum_{i=1}^N \int_0^T \exp\left(\int_0^s g(X_r^i)dr\right) \nabla_x P_{s,T}^g f(X_s^i)^\top \sigma(X_s^i) dW_s^i, \quad (94)$$

where $P_{t,T}^g f$ for $t\in[0,T]$ is the propagator.

- Proof. The key claim is

$$\partial_t \langle \mu_t^g, P_{t,T}^g f\rangle = 0, \quad \text{for all } t\in[0,T]. \quad (95)$$

Extended RM Equations XX

- In fact, for all $(t, y) \in [0, T] \times D$,

$$
\begin{aligned}
&\exp\left(\int_0^t g(X_s^y)ds\right) P_{t,T}^g f(X_t^y) \\
&= P_{0,T}^g f(y) + \int_0^t \exp\left(\int_0^s g(X_r^y)dr\right) (\partial_t + \mathcal{A} + g)\, P_{s,T}^g f(X_s^y)ds \\
&\quad + \int_0^t \exp\left(\int_0^s g(X_r^y)dr\right) \nabla_x P_{s,T}^g f(X_s^y)^\top \sigma(X_s^y)dW_s \\
&= P_{0,T}^g f(y) + \int_0^t \exp\left(\int_0^s g(X_r^y)dr\right) \nabla_x P_{s,T}^g f(X_s^y)^\top \sigma(X_s^y)dW_s.
\end{aligned}
$$

- It holds again that

$$
\mathbb{E}\left[\exp\left(\int_0^t g(X_s^y)ds\right) P_{t,T}^g f(X_t^y)\right] = P_{0,T}^g f(y).
$$

Extended RM Equations XXI

- Then, we have

$$\langle \mu_t, P^g_{t,T} f\rangle = \frac{\int_D E\left[\exp\left(\int_0^t g(X_s^y)ds\right) P^g_{t,T} f(X_t^y)\right]\rho_0(dy)}{\int_D E\left[\exp\left(\int_0^t g(X_s^y)ds\right)\right]\rho_0(dy)}.$$

- This yields from (82) that

$$\begin{aligned}\langle \mu_t^g, P^g_{t,T} f\rangle &= \int_D \mathbb{E}\left[\exp\left(\int_0^t g(X_s^y)ds\right) P^g_{t,T} f(X_t^y)\right]\rho_0(dy)\\ &= \int_D P^g_{0,T} f(y)\rho_0(dy).\end{aligned}$$

Extended RM Equations XXII

Theorem

Let previous assumptions hold. Let $\mu = (\mu_t)_{t\in[0,T]}$ be the $\mathcal{P}(D)$-valued solution of extended RM equation and $\mu^N = (\mu_t^N)_{t\in[0,T]}$ be the sequence of $\mathcal{P}(D)$-valued processes. Then, for any $T > 0$ and $N \geq 1$, there exists a constant $C > 0$ which is independent of N such that, for any $p \geq 1$,

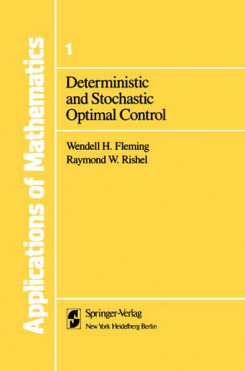
$$\hat{d}_{q,T}(\mu, \mu^N) \leq C\left(\alpha(p,q,n,N) + \frac{1}{N^{q-1}}\right), \quad q \geq 2, \tag{96}$$

$$\alpha(p,q,n,N) := \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{p-q}{p}}, & q > \frac{n}{2},\ p \neq 2q; \\ N^{-\frac{1}{2}} \ln(1+N) + N^{-\frac{p-q}{p}}, & q = \frac{n}{2},\ p \neq 2q; \\ N^{-\frac{q}{n}} + N^{-\frac{p-q}{p}}, & q < \frac{n}{2},\ p \neq \frac{n}{n-q}. \end{cases} \tag{97}$$

Course Outline

Deterministic Control Problem I

- I recommend you the book "*Deterministic and Stochastic Optimal Control*" by Fleming and Rishel (1975):



Chapter I: The simplest problem in calculus of Variations
2: The optimal Control problem
3: Existence & Continuity properties of optimal Controls
4: Dynamic programming
5: SDE & Markov Diffusion processes
6: Optimal control of Markov Diffusion processes

Figure: Book by Fleming, W.H. and R.W. Rishel.

Deterministic Control Problem II

- Consider a deterministic control problem described as follows:
  - Time horizon: T; time variable: $t \in [0, T]$;
  - State function of a controlled system: $X_t^{x,u} \in \mathbb{R}^n$; control function u_t;
- The state dynamics of the controlled system is described as: for $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$X_s^{t,x,u} = x + \int_t^s b(X_r^{t,x,u}, u_r)dr, \quad s \in [t, T]. \tag{98}$$

- $b(x, u) : \mathbb{R}^n \times U \to \mathbb{R}^n$ satisfies a uniform Lipschitz condition in U.
- The function u_s for $s \in [t, T]$ is called a control or strategy, which is assumed to take values in a compact subset U of $\mathbb{R}^m$.

Set of Admissible Controls $\mathcal{U}_t^T$: it is defined as:

$$\mathcal{U}_t^T := \{u_s : [t, T] \to U;\ s \to u_s \text{ is measurable}\}.$$

Deterministic Control Problem III

- A deterministic control problem can be described as follows:
  - Terminal Cost Function $g : \mathbb{R}^n \to \mathbb{R}$;
  - Running Cost Function $f : \mathbb{R}^n \times U \to \mathbb{R}$.
- The optimization problem is to find an optimal control $u^* \in \mathcal{U}_t^T$ that minimizes the following objective functional:

$$\begin{aligned} J(t,x;u^*) &= \inf_{u \in \mathcal{U}_t^T} J(t,x;u) \\ &:= \inf_{u \in \mathcal{U}_t^T} \left[g(X_T^{t,x,u}) + \int_t^T f(X_s^{t,x,u}, u_s) ds \right]. \end{aligned} \tag{99}$$

- Define the value function corresponding to the control problem (99):

$$V(t,x) := \inf_{u \in \mathcal{U}_t^T} \left[g(X_T^{t,x,u}) + \int_t^T f(X_s^{t,x,u}, u_s) ds \right]. \tag{100}$$

Deterministic Control Problem IV

- Obviously, the value function satisfies a terminal condition $V(T,x) = g(x)$.

Dynamic Programming Principle (DPP): For the value function $V(t,x)$ defined by (110), and $t < s \leq T$, it holds that

$$V(t,x) = \inf_{u \in \mathcal{U}_t^T} \left[V(s, X_s^{t,x,u}) + \int_t^s f(X_r^{t,x,u}, u_r) dr \right] \tag{101}$$

which is documented in Bellman (1957): "*Dynamic Programming*".

- It is saying that if one knows the value function at time $s > t$, one may determine the value function at time t by optimizing from time t to time s and using $V(\cdot, s)$ as the terminal cost.
- DPP means that $V(t,x)$ satisfies a semigroup property, but, which is running backwards.

Deterministic Control Problem V

- Proof of DPP: The proof is based on the following observation on the admissible control set:

For $t < s \leq T$, we have $\mathcal{U}_t^T = \mathcal{U}_t^s \oplus \mathcal{U}_s^T$. Here $\oplus$ means that if $u^1 : [t, s] \to U \in \mathcal{U}_t^s$ and $u^2 : [s, T] \to U \in \mathcal{U}_s^T$, then $u^t \oplus u^s$ is defined as: for $r \in [t, T]$,

$$u^t \oplus u^s := \begin{cases} u_r^1, & r \in [t, s]; \\ u_r^2, & r \in [s, T]. \end{cases}$$

- Then, $u := u^1 \oplus u^2 \in \mathcal{U}_t^T$ if $u^1 \in \mathcal{U}_t^s$ and $u^2 \in \mathcal{U}_s^T$.
- On the other hand, if $u \in \mathcal{U}_t^T$, then by restricting the domain of u to $[t, s]$, we obtain an admissible control in $\mathcal{U}_t^s$;

Deterministic Control Problem VI

- Similarly, by restricting the domain of u to $[s, T]$, we obtain an admissible control in $\mathcal{U}_s^T$;
- Therefore, we proved that $\mathcal{U}_t^T = \mathcal{U}_t^s \oplus \mathcal{U}_s^T$.
- By the definition (110) of the value function:

$$
\begin{aligned}
V(t,x) &:= \inf_{u\in\mathcal{U}_t^T}\left[g(X_T^{t,x,u}) + \int_t^T f(X_s^{t,x,u}, u_s)ds\right] \\
&= \inf_{u\in\mathcal{U}_t^T}\left[g(X_T^{t,x,u}) + \int_t^s f(X_r^{t,x,u}, u_r)dr + \int_s^T f(X_r^{t,x,u}, u_r)dr\right] \\
&= \inf_{u=u^1\oplus u^2;\ u^1\in\mathcal{U}_t^s,\ u^2\in\mathcal{U}_s^T} \\
&\qquad \left[g(X_T^{t,x,u}) + \int_t^s f(X_r^{t,x,u}, u_r)dr + \int_s^T f(X_r^{t,x,u}, u_r)dr\right]
\end{aligned}
$$

Deterministic Control Problem VII

- Meanwhile, we decompose the state process $X^{t,x,u}$ into the ones in the time interval $[t,s]$ and $[s,T]$, i.e., $X = X^1 \oplus X^2$:
  - $dX_r^{1,u^1} = b(X_r^{1,u^1}, u_r^1)dr,\ r \in (t,s];\ X_t^{1,u^1} = x;$
  - $dX_r^{2,u^2} = b(X_r^{2,u^2}, u_r^2)dr,\ r \in (s,T];\ X_s^{2,u^2} = X_s^{1,u^1} = X_s^{t,x,u};$
- Therefore

$$V(t,x) = \inf_{u^1 \in \mathcal{U}_t^s} \inf_{u=u^1\oplus u^2;\ u^2 \in \mathcal{U}_s^T,\ X_s^{2,u^2}=X_s^{1,u^1}} \left[g(X_T^{2,u^2}) + \int_t^s f(X_r^{1,u^1}, u_r^1)dr + \int_s^T f(X_r^{2,u^2}, u_r^2)dr \right].$$

Deterministic Control Problem VIII

- Note that X^{1,u^1} depends only on x and u^1, not on X^{2,u^2} or u^2. Since the first integral depends only on X^{1,u^1} and u^1, this may be rearranged as:

$$
\begin{aligned}
V(t,x) = \inf_{u^1\in\mathcal{U}_t^s} \Bigg\{ & \int_t^{s} f(X_r^{1,u^1}, u_r^1)dr \\
& + \inf_{u^2\in\mathcal{U}_s^T,\ X_s^{2,u^2}=X_s^{1,u^1}} \left[g(X_T^{2,u^2}) + \int_s^T f(X_r^{2,u^2}, u_r^2)dr \right] \Bigg\} \\
= \inf_{u^1\in\mathcal{U}_t^s} & \left[\int_t^{s} f(X_r^{1,u^1}, u_r^1)dr + V(X_s^{1,u^1}, s) \right] \\
= \inf_{u\in\mathcal{U}_t^T} & \left[\int_t^{s} f(X_r^{t,x,u}, u_r)dr + V(X_s^{t,x,u}, s) \right].
\end{aligned}
$$

Deterministic Control Problem IX

- If the value function $V(t,x)$ is $C^{1,1}$, then we have the following

Hamilton-Jacobi-Bellman (HJB) equation: The value function satisfies that

$$\begin{aligned} \partial_t V(t,x) + H(\nabla_x V(t,x), x) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^n; \\ V(T,x) = g(x), \quad x \in \mathbb{R}^n, \end{aligned} \tag{102}$$

where $H(p,x)$ for $(p,x) \in \mathbb{R}^n \times \mathbb{R}^n$ is called Hamiltonian, which is defined as:

$$H(p,x) := \inf_{u \in U} \left[b(x,u)^\top p + f(x,u) \right].$$

Deterministic Control Problem X

- Proof. Using DPP, for a sufficient small $h > 0$,

$$V(t,x) = \inf_{u \in \mathcal{U}_t^{t+h}} \left[V(t+h, X_{t+h}^{t,x,u}) + \int_t^{t+h} f(X_r^{t,x,u}, u_r) dr \right].$$

- However, it holds that

$$V(t+h, X_{t+h}^{t,x,u}) = V(t,x) + \int_t^{t+h} b(X_r^{t,x,u}, u_r)^\top \nabla_x V(r, X_r^{t,x,u}) dr.$$

- When $T \to \infty$, the finite control problem becomes an infinite horizon control problem, which is formulated as:

$$\begin{aligned} J(x; u^*) &= \inf_{u \in \mathcal{U}_t^\infty} J(x; u) \\ &:= \inf_{u \in \mathcal{U}_t^\infty} \left[\int_t^\infty e^{-\lambda s} f(X_s^{t,x,u}, u_s) ds \right]. \end{aligned} \tag{103}$$

Deterministic Control Problem XI

- Then, we have that

Value Function: For $x \in \mathbb{R}^n$,

$$V(x) := \inf_{u \in \mathcal{U}_0^\infty} \left[\int_0^\infty e^{-\lambda s} f(X_s^{0,x,u}, u_s) ds \right]. \tag{104}$$

DPP: For $t < s$, it holds that

$$V(x) = \inf_{u \in \mathcal{U}_0^\infty} \left[e^{-\lambda s} V(X_s^{0,x,u}) + \int_0^s e^{-\lambda r} f(X_r^{0,x,u}, u_r) dr \right] \tag{105}$$

HJB Equation: For $x \in \mathbb{R}^n$,

$$H(\nabla_x V(x), x) - \lambda V(x) = 0 \tag{106}$$

Deterministic Control Problem XII

- We next illustrate the HJB equation approach in terms of calculus of variations rather than optimal control:

Let $L(q,x)$ be the Lagrangian, which is a sufficiently smooth function in $q,x\in\mathbb{R}^n$. Fix two points $x,y\in\mathbb{R}^n$, and consider the class of admissible trajectories connecting these points: for $t>0$,

$$\mathcal{U}_t^{x,y} := \left\{\phi\in C^1([0,t];\mathbb{R}^n);\ \phi(0)=x,\ \phi(t)=y\right\}.$$

The basic problem of the calculus of variations is to find the optimal curve $\phi^*\in\mathcal{U}_t^{x,y}$ s.t.

$$J(\phi^*) = \inf_{\phi\in\mathcal{U}_t^{x,y}} J(\phi) := \inf_{\phi\in\mathcal{U}_t^{x,y}} \int_0^t L(\dot\phi(s),\phi(s))ds$$

Deterministic Control Problem XIII

- Now assume that ϕ^* exists and we want to see what is the property satisfied by ϕ^*.

Euler-Lagrange Equation: The optimal curve $\phi^* \in \mathcal{U}_t^{x,y}$ satisfies that, for $r \in [0, t]$,

$$\frac{d}{dr}[\nabla_q L(\dot{\phi}^*(r), \phi^*(r))] = \nabla_x L(\dot{\phi}^*(r), \phi^*(r)). \tag{107}$$

- Proof. Let $\psi \in C^1([0, t]; \mathbb{R}^n)$ with $\psi(0) = \psi(t) = 0$, i.e., $\psi \in \mathcal{U}_t^{0,0}$.
- Define $\phi^s(r) := \phi^*(r) + s\psi(r)$ for $r \in [0, t]$ and $s \in \mathbb{R}$. Hence $\phi^s \in \mathcal{U}_t^{x,y}$.
- Define also that $\Phi(s) := J(\phi^s)$. We next compute $\Phi'(s)$ for $s \in \mathbb{R}$:

Deterministic Control Problem XIV

- Recall that

$$\Phi(s) = \int_0^t L(\dot{\phi}^*(r) + s\dot{\psi}(r), \phi^*(r) + s\psi(r))dr.$$

- Therefore, by setting $r_s := \dot{\phi}^*(r) + s\dot{\psi}(r)$,

$$\Phi'(s) = \int_0^t [\nabla_q L(r_s)^\top \dot{\psi}(r) + \nabla_x L(r_s)^\top \psi(r)]dr.$$

- Using integration by parts, we have

$$\begin{aligned}\int_0^t \nabla_q L(r_s)^\top \dot{\psi}(r)dr &= \nabla_q L(r_s)^\top \psi(r)\Big|_{r=0}^{r=t} - \int_0^t \frac{d}{dr}[\nabla_q L(r_s)]^\top \psi(r)dr \\ &= -\int_0^t \frac{d}{dr}[\nabla_q L(r)]^\top \psi(r)dr.\end{aligned}$$

Deterministic Control Problem XV

- Thus, it holds that

$$\Phi'(s) = \int_0^t \left\{ \nabla_x L(r_s)^\top - \frac{d}{dr}[\nabla_q L(r_s)]^\top \right\} \psi(r) dr.$$

- Since $\phi^* \in \mathcal{U}_t^{x,y}$ is a minimizer of $J(\phi)$ over $\phi \in \mathcal{U}_t^{x,y}$, we have

$$\Phi'(0) = 0.$$

- This yields Euler-Lagrange equation (107).
- We next connect the Euler-Lagrange equation to the so-called Hamilton equation.

Let $p \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$ be given. Assume that the equation $\nabla_q L(q, x) = p$ in the unknown q has a unique smooth solution $q(p, x) \in \mathbb{R}^n$.

Deterministic Control Problem XVI

- Define the Hamiltonian as $H(p,x) := pq(p,x) - L(q(p,x),x)$ for $(p,x) \in \mathbb{R}^n \times \mathbb{R}^n$.
- Let $\phi^*(r)$ for $r \in [0,t]$ be the solution of the Euler-Lagrange equation (107).
- Question: Prove that ϕ^* satisfies the Hamilton equation: for $r \in [0,t]$,

$$\begin{cases} \dot{\phi}^*(r) = \nabla_p H(p(r), \phi^*(r)), \\ \dot{p}(r) = -\nabla_x H(p(r), \phi^*(r)). \end{cases} \tag{108}$$

Stochastic Control Problem I

- Consider the following controlled diffusion process described as: for $(t,x) \in [0,T] \times \mathbb{R}^n$,

$$X_s^{t,x} = x + \int_t^s b(X_r^{t,x}, u_r)dr + \int_t^s \sigma(X_r^{t,x}, u_r)dW_r, \ \ s \in [t,T]. \tag{109}$$

- Here, $b(x,u) : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma(x,u) : \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$.
- The control $(u_t)_{t\in[0,T]}$ is a progressively measurable process, valued in $U \subset \mathbb{R}^m$.
- We impose the assumption on the coefficients (b,σ):

($\mathbf{A}_c$) b,σ satisfy a uniform Lipschitz condition in U: for any $x,y \in \mathbb{R}^n$ and $u \in U$,

$$|b(x,u) - b(y,u)| + |\sigma(x,u) - \sigma(y,u)| \leq L|x-y|,$$

for some $K > 0$.

Stochastic Control Problem II

- Denote by $\mathcal{U}$ the set of control processes $(u_t)_{t\in[0,T]}$ such that

$$E\left[\int_0^T (|b(0,u_t)|^2 + |\sigma(0,u_t)|^2)dt\right] < \infty.$$

- We next introduce the objective functional:

Terminal Payoff Function: $g:\mathbb{R}^n \to \mathbb{R}$ is measurable and satisfies a quadratic growth condition:

$(\mathbf{A}_g)$ i.e., $|g(x)| \leq K(1+|x|^2)$ for all $x \in \mathbb{R}^n$.

Running Payoff Function: $f:\mathbb{R}^n \times U \to \mathbb{R}$ is measurable and satisfies a quadratic growth condition in x:

$(\mathbf{A}_f)$ i.e., $|f(x,u)| \leq K(1+|x|^2) + l(u)$ for all $(x,u) \in \mathbb{R}^n \times U$,

where $l: U \to \mathbb{R}_+$ is a positive function.

Stochastic Control Problem III

- We next introduce the admissible control sets:

Admissible Control Set: For $(t,x) \in [0,T] \times \mathbb{R}^n$, denote by $\mathcal{U}_{t,x}$ the subset of controls $(u_t)_{t\in[0,T]} \in \mathcal{U}$ such that

$$E\left[\int_t^T |f(s, X_s^{t,x}, u_s)|ds\right] < \infty.$$

- **Question**: Under the assumption $(\mathbf{A}_f)$, prove $\mathcal{U}_{t,x} \neq \emptyset$.

Value Function: For $(t,x) \in [0,T] \times \mathbb{R}^n$,

$$\begin{aligned} V(t,x) &:= \sup_{u\in\mathcal{U}_{t,x}} J(t,x,u) \\ &:= \sup_{u\in\mathcal{U}_{t,x}} E\left[g(X_T^{t,x}) + \int_t^T f(s, X_s^{t,x}, u_s)ds\right]. \end{aligned} \tag{110}$$

Stochastic Control Problem IV

- For any $(t,x) \in [0,T) \times \mathbb{R}^n$, if there exists $u^* \in \mathcal{U}_{t,x}$ such that $V(t,x) = J(t,x,u^*)$, then u^* is called an optimal control.

Markovian Controls: For any $u \in \mathcal{U}_{t,x}$, if there exists a measurable function $\mathrm{u} : [0,T] \times \mathbb{R}^n \to U$ such that $u_s = \mathrm{u}(s, X_s^{t,x})$ for $s \in [t,T]$.

- Optimal Markovian Control: the optimal control $u_s^* = \mathrm{u}^*(s, X_s^{t,x,*})$. Here $X^{t,x,*}$ satisfies SDE (99) with u replaced by u^*.
- We next introduce the dynamic programming principle (DPP), which is a fundamental principle in the theory of stochastic control.

Stochastic Control Problem V

- To this purpose, let $\tau \in \mathcal{T}_{t,T}$ be the set of stopping times taking values on $[t, T]$.

Theorem (Stochastic Version of DPP)

For $(t,x) \in [0, T] \times \mathbb{R}^n$, it holds that

$$V(t,x) = \sup_{u \in \mathcal{U}_{t,x}} \sup_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s) ds \right],$$

$$V(t,x) = \sup_{u \in \mathcal{U}_{t,x}} \inf_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s) ds \right].$$

- Proof. For any $u \in \mathcal{U}_{t,x}$, i.e., u is an admissible control, using the pathwise uniqueness of SDE (99),

$$X_s^{t,x} = X_s^{\tau, X_\tau^{t,x}}, \qquad s \geq \tau \in \mathcal{T}_{t,T}. \tag{111}$$

Stochastic Control Problem VI

- Then, for any $\tau \in \mathcal{T}_{t,T}$, using (111),

$$
\begin{aligned}
J(t,x,u) &= E\left[g(X_T^{t,x}) + \int_\tau^T f(s,X_s^{t,x},u_s)ds + \int_t^\tau f(s,X_s^{t,x},u_s)ds\right]\\
&= E\left\{E\left[g(X_T^{t,x}) + \int_\tau^T f(s,X_s^{t,x},u_s)ds + \int_t^\tau f(s,X_s^{t,x},u_s)ds \Big| \mathcal{F}_\tau\right]\right\}\\
&= E\left\{E\left[g(X_T^{t,x}) + \int_\tau^T f(s,X_s^{t,x},u_s)ds \Big| \mathcal{F}_\tau\right] + \int_t^\tau f(s,X_s^{t,x},u_s)ds\right\}\\
&= E\left[J(\tau,X_\tau^{t,x},u) + \int_t^\tau f(s,X_s^{t,x},u_s)ds\right].
\end{aligned}
$$

- Since $J(t,x,u) \leq V(t,x)$, it holds that

$$
J(t,x,u) \leq E\left[V(\tau,X_\tau^{t,x}) + \int_t^\tau f(s,X_s^{t,x},u_s)ds\right], \ \forall\, \tau \in \mathcal{T}_{t,T}.
$$

Stochastic Control Problem VII

- Therefore, for all $u \in \mathcal{U}_{t,x}$,

$$J(t,x,u) \leq \inf_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right]$$
$$\leq \sup_{u \in \mathcal{U}_{t,x}} \inf_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right].$$

- This yields that

$$V(t,x) \leq \sup_{u \in \mathcal{U}_{t,x}} \inf_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right].$$

- On the other hand, fix some arbitrary admissible control $u \in \mathcal{U}_{t,x}$ and $\tau \in \mathcal{T}_{t,T}$.

Stochastic Control Problem VIII

- By the definition of the value function, for any $\epsilon > 0$ and $\omega \in \Omega$-a.s., there exists $u^{\epsilon}(\omega) \in \mathcal{U}_{\tau(\omega), X^{t,x}_{\tau(\omega)}(\omega)}$, which is an ϵ-optimal control for $V(\tau(\omega), X^{t,x}_{\tau(\omega)}(\omega))$, i.e.,

$$V(\tau(\omega), X^{t,x}_{\tau(\omega)}(\omega)) - \epsilon \leq J(\tau(\omega), X^{t,x}_{\tau(\omega)}(\omega), u^{\epsilon}(\omega)). \tag{112}$$

- Let us define that

$$\tilde{u}_t(\omega) := \begin{cases} u_t(\omega), & t \in [0, \tau(\omega)]; \\ u^{\epsilon}_t(\omega), & t \in [\tau(\omega), T]. \end{cases} \tag{113}$$

- Warning: Measurability issue on $\tilde{u} \in \mathcal{U}_{t,x}$: but it can be shown by the measurable selection theorem.

Stochastic Control Problem IX

- Then, by (112), we have

$$
\begin{aligned}
V(t,x) \geq J(t,x,\tilde{u}) &= E\left[J(\tau, X_\tau^{t,x}, u^\epsilon) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right] \\
&\geq E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right] - \epsilon.
\end{aligned}
$$

- By the arbitrariness of $u \in \mathcal{U}_{t,x}$, $\tau \in \mathcal{T}_{t,T}$ and $\epsilon > 0$, we get

$$
V(t,x) \geq \sup_{u \in \mathcal{U}_{t,x}} \sup_{\tau \in \mathcal{T}_{t,T}} E\left[V(\tau, X_\tau^{t,x}) + \int_t^\tau f(s, X_s^{t,x}, u_s)ds\right].
$$

- Thus, we complete the proof of DPP.

Stochastic Control Problem X

- DPP can yield the Hamilton-Jacobi-Bellman (HJB) equation:

HJB Equation: If the value function $V \in C^{1,2}$, then V satisfies the HJB equation given by: for $(t,x) \in [0,T) \times \mathbb{R}^n$,

$$\partial_t V(t,x) + H(x, \nabla_x V(t,x), \nabla_x^2 V(t,x)) = 0, \qquad V(T,x) = g(x), \quad x \in \mathbb{R}^n, \tag{114}$$

where the Hamiltonian $H(x,p,M)$ for $(x,p,M) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times n}$ is defined as:

$$H(x,p,M) := \sup_{u\in U} \left[b(x,u)^\top p + \frac{1}{2}\mathrm{tr}[\sigma\sigma^\top(x,u)M] + f(x,u) \right].$$

Stochastic Control Problem XI

- Theorem VI.6.2 in Fleming and Rishel (1975), page 169 proves the well-posedness of HJB equation (114) when $\sigma(x,u)=\sigma(x)$, i.e., the volatility of the controlled process is independent of the control u:
- Let $Q=(0,T)\times\mathbb{R}^n$;

(a) The policy space $U\subset\mathbb{R}^m$ is compact.

(b) $b(x,u)=\tilde{b}(x)+\sigma(x)\theta(x,u)$.

(c) $\tilde{b},\sigma\in C^2(\mathbb{R}^n)$; σ,σ^{-1} and $\sigma_x,\tilde{b}_x$ are bounded in $\mathbb{R}^n$; $\theta\in C^1(\mathbb{R}^n\times U)$, θ,θ_x are bounded.

(d) $f\in C^1(\mathbb{R}^n\times U)$, f,f_x satisfy the polynomial growth condition.

(e) $g\in C^2(\mathbb{R}^n)$, g,g_x satisfy the polynomial growth condition.

Theorem (Well-posedness of Smooth Solution of HJB Equation)

Under assumptions (a)-(e), the HJB equation (114) *admits a classical solution with polynomial growth.*

Backward Stochastic Differential Equations I

- We first introduce the framework of backward stochastic differential equation (BSDE).
- For this purpose, define the following space for stochastic processes:

Space S_T^p: the set of $\mathbb{R}$-valued progressively measurable processes $Y = (Y_t)_{t\in[0,T]}$ such that

$$E\left[\sup_{t\in[0,T]} |Y_t|^p\right] < \infty, \qquad p \geq 1.$$

Backward Stochastic Differential Equations II

Space $\mathrm{H}^p_{m,T}$: the set of $\mathbb{R}^m$-valued progressively measurable processes $Z = (Z_t)_{t\in[0,T]}$ such that

$$E\left[\int_0^T |Z_t|^p dt\right] < \infty, \qquad p \geq 1.$$

- Given a real-valued r.v. ξ and a random mapping $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, assume that

  ($\mathbf{A}_{\mathbf{BSDE0}}$) $\xi \in L^2(\Omega; \mathbb{R})$;

  ($\mathbf{A}_{\mathbf{BSDE}f}$) Write $f(t, y, z) = f(\cdot, t, y, z)$, it is progressively measurable for all $(y, z) \in \mathbb{R} \times \mathbb{R}^m$; $f(t, 0, 0) \in \mathrm{H}^2_{1,T}$; $f(t, \cdot, \cdot)$ is Lipschitz in (y, z) uniformly w.r.t. (t, ω), i.e. there exists a constant K s.t., $dt \otimes P$-a.s.,

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq K[|y_1 - y_2| + |z_1 - z_2|]$$

Backward Stochastic Differential Equations III

- The standard BSDE is given by

One-dimensional BSDE: For a terminal horizon $T > 0$,

$$\begin{aligned} dY_t &= -f(t, Y_t, Z_t)dt + Z_t^\top dW_t, \\ Y_T &= \xi. \end{aligned} \tag{115}$$

Theorem (Well-posedness of BSDE)

Given (ξ, f) *satisfying* $(\mathbf{A}_{\mathbf{BSDE}0})$ *and* $(\mathbf{A}_{\mathbf{BSDE}f})$, *then BSDE* (115) *admits a unique solution* $(Y, Z) \in \mathrm{S}_T^2 \times \mathrm{H}_{m,T}^2$.

- Proof. Use a fixed point argument.

Backward Stochastic Differential Equations IV

- Let $(U, V) \in \mathcal{X} := \mathrm{S}^2_T \times \mathrm{H}^2_{m,T}$. Then, we define a square-integrable martingale by assumptions $(\mathbf{A_{BSDE0}})$ and $(\mathbf{A_{BSDE}}_f)$ as:

$$M_t := E\left[\xi + \int_0^T f(s, U_s, V_s)ds \Big| \mathcal{F}_t\right]$$

- The martingale representation theorem yields that, there exists $Z \in \mathrm{H}^2_{m,T}$ such that

$$M_t = M_0 + \int_0^t Z_s^\top W_s, \quad t \in [0, T].$$

- Given $Z \in \mathrm{H}^2_{m,T}$ above, we then define

$$Y_t := E\left[\xi + \int_t^T f(s, U_s, V_s)ds \Big| \mathcal{F}_t\right] = M_t - \int_0^t f(s, U_s, V_s)ds.$$

Backward Stochastic Differential Equations V

- Obviously, we have $Y_T = \xi$. Then, we can summarize that

$$Y_t = \xi + \int_0^t Z_s^\top W_s - \int_0^t f(s, U_s, V_s) ds. \tag{116}$$

- Accordingly, we define a mapping Φ on $\mathcal{X}$ as follows:

$$(Y, Z) = \Phi(U, V).$$

- By BDG inequality, we obtain $(Y, Z) \in \mathcal{X}$, i.e., $\Phi : \mathcal{X} \to \mathcal{X}$.
- Therefore, (Y, Z) is a solution of BSDE (115) if and only if (Y, Z) is a fixed point of Φ.

Backward Stochastic Differential Equations VI

- We next prove that $\Phi : \mathcal{X} \to \mathcal{X}$ is a contraction mapping on Banach space $\mathcal{X}$ with norm given by

$$\|(Y,Z)\|_\lambda := \left\{ E\left[\int_0^T e^{\lambda s}(|Y_s|^2 + |Z_s|^2)ds \right] \right\}^{\frac{1}{2}}$$

by taking a suitable parameter $\lambda \in \mathbb{R}$.

- Question: Prove $\Phi : \mathcal{X} \to \mathcal{X}$ is a contraction mapping.
- By (116), $\xi = \xi + \int_0^T Z_s^\top dW_s - \int_0^T f(s, Y_s, Z_s)ds$. Then, making difference between it and (116) to get that (Y, Z) is the solution of BSDE.
- Thus, we complete the proof of the theorem.

Backward Stochastic Differential Equations VII

- We next introduce comparison principles of solutions of BSDE (115).

Theorem (Comparison Theorem)

Let (ξ^i, f^i) *satisfies assumptions* $(\mathbf{A}_{\mathbf{BSDE0}})$ *and* $(\mathbf{A}_{\mathbf{BSDE}f})$ *for* $i = 1, 2$. *Let* (Y^i, Z^i) *be the solution of BSDE* (115) *with* (ξ^i, f^i). *Assume that*

(i) $\xi^1 \leq \xi^2$, *a.s.;*

(ii) $f^1(t, Y_t^1, Z_t^1) \leq f^2(t, Y_t^1, Z_t^1)$, $dt \otimes dP$*-a.s.*

(iii) $f^2(t, Y_t^1, Z_t^1) \in \mathrm{H}^2_{m,T}$.

Then, $Y_t^1 \leq Y_t^2$, *for all* $t \in [0, T]$, P*-a.s.*

- Question: Prove Theorem 34.

Backward Stochastic Differential Equations VIII

- Question: Consider the following linear BSDE:

$$\begin{aligned} dY_t &= -(A_t Y_t + Z_t^\top B_t + C_t)dt + Z_t^\top dW_t, \\ Y_T &= \xi. \end{aligned} \tag{117}$$

where $A = (A_t)_{t\in[0,T]}$, $B = (B_t)_{t\in[0,T]}$ are bounded progressively measurable processes and $C = (C_t)_{t\in[0,T]} \in \mathrm{H}^2_{1,T}$. Then

$$Y_t = X_t^{-1} E\left[X_T \xi + \int_t^T X_s C_s ds \Big| \mathcal{F}_t \right],$$

where the process $X = (X_t)_{t\in[0,T]}$ satisfies the following forward linear SDE:

$$dX_t = X_t(A_t dt + B_t^\top dW_t), \quad X_0 = 1,$$

i.e., $X_t = \mathcal{E}_t(\int_0^\cdot A_s ds + \int_0^\cdot B_s^\top dW_s)$.

Nonlinear Feynamn-Kac Formula I

- Consider the following assumptions:

  ($\mathbf{A_{NFYf}}$) The deterministic function $f(t,x,y,z):[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m$ is continuous, it satisfies a linear growth condition in (x,y,z) and a Lipschitz condition in (y,z) uniformly w.r.t. (t,x);

  ($\mathbf{A_{NFYphi}}$) The function $\phi:\mathbb{R}^n\to\mathbb{R}$ is continuous and it satisfies a linear growth condition.

- Consider the semilinear Cauchy problem: for $(t,x)\in[0,T)\times\mathbb{R}^n$,

$$\begin{aligned}\partial_t u+\mathcal{A}u+f(t,x,u,\sigma^\top\nabla_x u)&=0, \\ u(T,x)&=\phi(x),\quad x\in\mathbb{R}^n.\end{aligned}\tag{118}$$

- We next introduce the following forward-backward SDE as follows:

$$dY_s=-f(s,X_s,Y_s,Z_s)ds+Z_s^\top dW_s,\ s\in[t,T],\ Y_T=\phi(X_T).$$

Nonlinear Feynamn-Kac Formula II

- The forward SDE is given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad s \in [t, T],$$

where the generator of X is $\mathcal{A}$.

- For $(t, x) \in [0, T] \times \mathbb{R}^n$, consider

$$X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dW_r, \quad s \in [t, T].$$

- Let $(Y_s^{t,x}, Z_s^{t,x})_{s\in[t,T]}$ be the above BSDE with $X_s = X_s^{t,x}$ for $s \in [t, T]$.
- Then, $u(t, x) := Y_t^{t,x}$ is a deterministic function on $[0, T] \times \mathbb{R}^n$.
- Note that $u(T, x) = Y_T^{T,x} = \phi(X_T^{T,x}) = \phi(x)$ for $x \in \mathbb{R}^n$.

Nonlinear Feynamn-Kac Formula III

- Using Markov property of X and the uniqueness of the solution to BSDE, we have that

$$Y_t = u(t, X_t), \quad t \in [0, T]. \tag{119}$$

- Then, the function $u(t, x)$ in (119) is in fact related to the solution of the semilinear Cauchy problem (118):

Viscosity Solution of Semilinear Cauchy problem (118): $u(t, x) := Y_t^{t,x}$ is continuous on $[0, T] \times \mathbb{R}^n$ and it is a viscosity solution.

- Now, assume that $u(t, x)$ is a classical solution of the semilinear Cauchy problem (118), and it also satisfies a linear growth condition. Moreover, $|\nabla_x u(t, x)| \leq K(1 + |x|^p)$ for $K, p > 0$.

Nonlinear Feynamn-Kac Formula IV

- Using Itô formula to $u(t, X_t)$, we have

Let $t \in [0, T]$ and define

$$Y_t = u(t, X_t), \quad Z_t = \sigma(X_t)^\top \nabla_x u(t, X_t) \tag{120}$$

is the solution of BSDE.

- Question: Prove (120).

Stochastic Maximum Principle I

- Recall the controlled diffusion process given by: $X_0 = x \in \mathbb{R}^n$, and

$$dX_t = b(X_t, u_t)dt + \sigma(X_t, u_t)dW_t, \ \ t \in [0, T].$$

- One wants to maximize the following objective functional given by

Objective Functional: For $x \in \mathbb{R}^n$ and $u \in \mathcal{U}$,

$$J(u) := E\left[g(X_T) + \int_0^T f(t, X_t, u_t)dt\right].$$

- We impose the following assumptions:

($\mathbf{A_{SMPfg}}$) The terminal payoff function $g : \mathbb{R}^n \to \mathbb{R}$ is a concave C^1-function; The running payoff function $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$ is continuous in (t, x) for all $u \in U$; f, g satisfy a quadratic growth in x.

Stochastic Maximum Principle II

- We next introduce the so-called generalized Hamiltonian:

Generalized Hamiltonian: $\Pi : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times m} \times U \to \mathbb{R}$, which is defined as:

$$\Pi(t, x, y, z, u) := b(x, u)^\top y + \mathrm{tr}[\sigma(x, u)^\top z] + f(t, x, u). \tag{121}$$

- Then, the controlled diffusion process can be rewritten as:

$$dX_t = \nabla_y \Pi(t, X_t, Y_t, Z_t, u_t)dt + \sigma(X_t, u_t)dW_t, \ \ t \in [0, T].$$

- We further assume that the gradient $\nabla_x \Pi(t, x, y, z, u)$ exists.

The Adjoint Equation: it is the following BSDE given by, for $u \in \mathcal{U}$,

$$\begin{aligned} dY_t &= -\nabla_x \Pi(t, X_t, Y_t, Z_t, u_t)dt + Z_t dW_t, \quad t \in [0, T]; \\ Y_T &= \nabla_x g(X_T). \end{aligned} \tag{122}$$

Stochastic Maximum Principle III

- Then, the stochastic maximum principle is given by

Theorem (Stochastic Maximum Principle)

Let $u^* \in \mathcal{U}$ and X^* be the controlled diffusion process with control u^*. Assume that there is a solution (Y^*, Z^*) to the adjoint equation (122) such that, for $t \in [0, T]$, P-a.s.,

$$\Pi(t, X_t^*, Y_t^*, Z_t^*, u_t^*) = \sup_{u \in \mathcal{U}} \Pi(t, X_t^*, Y_t^*, Z_t^*, u),$$

and $(x, u) \to \Pi(t, x, Y_t^*, Z_t^*, u)$ is a concave function for all $t \in [0, T]$. Then u^* is an optimal control, i.e., $J(u^*) = \sup_{u \in \mathcal{U}} J(u)$.

Stochastic Maximum Principle IV

- Proof. For any $u \in \mathcal{U}$, we have

$$
\begin{aligned}
& J(u^*) - J(u) \\
& \quad = E\left[g(X_T^*) - g(X_T) + \int_0^T (f(s, X_s^*, u_s^*) - f(s, X_s, u_s))ds\right].
\end{aligned}
$$

- It follows from the concavity of g that

$$
\begin{aligned}
E\left[g(X_T^*) - g(X_T)\right] &\geq E\left[(X_T^* - X_T)^\top \nabla_x g(X_T^*)\right] \\
&= E\left[(X_T^* - X_T)^\top Y_T^*\right].
\end{aligned}
$$

Stochastic Maximum Principle V

- By Itô formula, we obtain

$$
\begin{aligned}
E\left[(X_T^* - X_T)^\top Y_T^*\right] &= E\left[\int_0^T (X_s^* - X_s)^\top dY_s^*\right] \\
&+ E\left[\int_0^T Y_s^{*,\top} d(X_s^* - X_s)\right] + E\left[\int_0^T \mathrm{tr}[(\sigma(X_s^*, u_s^*) - \sigma(X_s, u_s))^\top Z_s^*] ds\right] \\
&= -E\left[\int_0^T (X_s^* - X_s)^\top \nabla_x \Pi(s, X_s^*, Y_s^*, Z_s^*, u_s^*) ds\right] \\
&\quad + E\left[\int_0^T (b(X_s^*, u_s^*) - b(X_s, u_s))^\top Y_s^* ds\right] \\
&\quad + E\left[\int_0^T \mathrm{tr}[(\sigma(X_s^*, u_s^*) - \sigma(X_s, u_s))^\top Z_s^*] ds\right].
\end{aligned}
$$

Stochastic Maximum Principle VI

- Using the definition of Π, one has

$$
\begin{aligned}
&E\left[\int_0^T (f(s,X_s^*,u_s^*)-f(s,X_s,u_s))ds\right]\\
&=E\left[\int_0^T (\Pi(s,X_s^*,Y_s^*,Z_s^*,u_s^*)-\Pi(s,X_s,Y_s^*,Z_s^*,u_s))ds\right]\\
&\quad -E\left[\int_0^T (b(X_s^*,u_s^*)-b(X_s,u_s))^\top Y_s^* ds\right]\\
&\quad -E\left[\int_0^T \mathrm{tr}[(\sigma(X_s^*,u_s^*)-\sigma(X_s,u_s))^\top Z_s^*]ds\right]
\end{aligned}
$$

Stochastic Maximum Principle VII

- Combing the above equalities, we have

$$J(u^*) - J(u) \geq -E\left[\int_0^T (X_s^* - X_s)^\top \nabla_x \Pi(s, X_s^*, Y_s^*, Z_s^*, u_s^*)ds\right]$$
$$+ E\left[\int_0^T (\Pi(s, X_s^*, Y_s^*, Z_s^*, u_s^*) - \Pi(s, X_s, Y_s^*, Z_s^*, u_s))ds\right].$$

- Using the assumption $\Pi(t, X_t^*, Y_t^*, Z_t^*, u_t^*) = \sup_{u\in\mathcal{U}} \Pi(t, X_t^*, Y_t^*, Z_t^*, u)$, and $(x, u) \to \Pi(t, x, Y_t^*, Z_t^*, u)$ is a concave function, we thus proves the theorem.
- Thus, we complete the proof of the theorem.
- We next establish the relationship between HJB equation and the stochastic maximum principle.

Stochastic Maximum Principle VIII

- First, recall the controlled diffusion process described as (109), i.e., for $(t,x)\in[0,T]\times\mathbb{R}^n$,

$$X_s^{t,x}=x+\int_t^s b(X_r^{t,x},u_r)dr+\int_t^s \sigma(X_r^{t,x},u_r)dW_r,\ \ s\in[t,T].$$

- Recall the value function defined by (110), i.e.,

Value Function: For $(t,x)\in[0,T]\times\mathbb{R}^n$,

$$\begin{aligned} V(t,x) &:= \sup_{u\in\mathcal{U}_{t,x}} J(t,x,u) \\ &:= \sup_{u\in\mathcal{U}_{t,x}} E\left[g(X_T^{t,x})+\int_t^T f(s,X_s^{t,x},u_s)ds\right]. \end{aligned}$$

Stochastic Maximum Principle IX

- Recall the HJB equation given by (114), i.e.,

HJB Equation: If the value function $V \in C^{1,2}$, then V satisfies the HJB equation given by: for $(t,x) \in [0,T) \times \mathbb{R}^n$,

$$\partial_t V(t,x) + \sup_{u\in\mathcal{U}} \tilde{H}(t,x,u,\nabla_x V(t,x),\nabla_x^2 V(t,x)) = 0,$$

$$V(T,x) = g(x), \quad x \in \mathbb{R}^n,$$

where $\tilde{H}(t,x,u,p,M)$ for $(t,x,u,p,M) \in [0,T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^{n\times n}$ is defined as:

$$\tilde{H}(t,x,u,p,M) := b(x,u)^\top p + \frac{1}{2}\mathrm{tr}[\sigma\sigma^\top(x,u)M] + f(t,x,a).$$

Stochastic Maximum Principle X

- Then, we have the following connection:

Theorem (Connection between HJB Equation and Maximum Principle)

Let HJB equation has a solution $V \in C^{1,3}([0,T) \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$. Assume that there is an optimal control $u^* \in \mathcal{U}$ for the value function $V(t,x)$ and X^* is the controlled diffusion process with u^*. Then

$$\begin{aligned}&\tilde{H}(t, X_t^*, u_t^*, \nabla_x V(t, X_t^*), \nabla_x^2 V(t, X_t^*))\\&\quad = \sup_{u \in \mathcal{U}} \tilde{H}(t, X_t^*, u, \nabla_x V(t, X_t^*), \nabla_x^2 V(t, X_t^*)),\end{aligned}$$

and the pair $(Y_t^*, Z_t^*) = (\nabla_x V(t, X_t^*), \nabla_x^2 V(t, X_t^*)\sigma(X_t^*, u_t^*))$ is a solution of the adjoint equation (BSDE) (122).

- Question: Prove this theorem and discuss that why here we need $V \in C^{1,3}$.

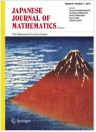
Definition of MFG I

- Consider a large system of agents (players) which behave similarly.
- The interactions among agents (players) are negligible but each agent's actions affect the mean of the population.
- Every agent (player) acts according to his/her control problem by taking into account other agents (players)' decisions.

Mean field differential game (MFG) studies the existence of a representative agent (player) such that the large system of agents (players) similar to this representative agent when the number of agents (players) goes to infinity.

- The seminal work on MFG are:
- Academics: Lions and Lasry (2007): Mean field games. *Jpn. J. Math.* 2, 229-260.

Definition of MFG II



Japanese Journal of Mathematics
March 2007, Volume 2, Issue 1, pp 229–260 | Cite as

Mean field games

Authors | Authors and affiliations

Jean-Michel Lasry [1]
Pierre-Louis Lions [2] [3]
Email author

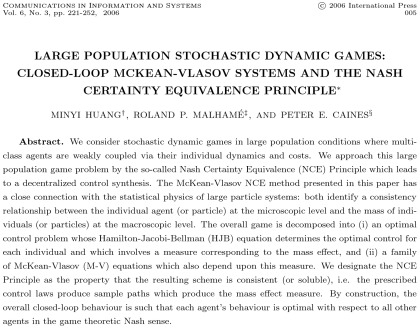
1. Institut de Finance, Université Paris-Dauphine, Paris Cedex 16, France
2. Collège de France, Paris, France
3. Ceremade-UMR CNRS 7549, Université Paris-Dauphine, Paris Cedex 16, France

Special Feature: The 1st Takagi Lecture
First Online: 28 March 2007

Figure: Seminal work on MFG

Definition of MFG III

- Industry: Huang, Malhame and Caines (2006): Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* no. 3, 221-251.



COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 6, No. 3, pp. 221-252, 2006

© 2006 International Press
005

LARGE POPULATION STOCHASTIC DYNAMIC GAMES: CLOSED-LOOP MCKEAN-VLASOV SYSTEMS AND THE NASH CERTAINTY EQUIVALENCE PRINCIPLE*

MINYI HUANG†, ROLAND P. MALHAMÉ‡, AND PETER E. CAINES§

Abstract. We consider stochastic dynamic games in large population conditions where multi-class agents are weakly coupled via their individual dynamics and costs. We approach this large population game problem by the so-called Nash Certainty Equivalence (NCE) Principle which leads to a decentralized control synthesis. The McKean-Vlasov NCE method presented in this paper has a close connection with the statistical physics of large particle systems: both identify a consistency relationship between the individual agent (or particle) at the microscopic level and the mass of individuals (or particles) at the macroscopic level. The overall game is decomposed into (i) an optimal control problem whose Hamilton-Jacobi-Bellman (HJB) equation determines the optimal control for each individual and which involves a measure corresponding to the mass effect, and (ii) a family of McKean-Vlasov (M-V) equations which also depend upon this measure. We designate the NCE Principle as the property that the resulting scheme is consistent (or soluble), i.e. the prescribed control laws produce sample paths which produce the mass effect measure. By construction, the overall closed-loop behaviour is such that each agent's behaviour is optimal with respect to all other agents in the game theoretic Nash sense.

Figure: Seminal work on MFG

Definition of MFG IV

- Two Approaches:

(i): The coupled Hamilton-Jacobi-Bellman with Focker-Planck which comes from dynamic programming in control theory;
(ii): PDEs and Forward-Backward SDE (FBSDEs) of McKean-Vlasov type which comes from stochastic analysis.

- We next provide an example for illustrating a static MFG:

Meeting Game Example: There are N professors who will attend an important meeting. This meeting will start at time t_0 (which is known).

- There N professors start from different locations to attend but are symmetric in a sense that they share the same characteristics (for instant they take the same distance to the venue or they go with the same speed, etc).

Definition of MFG V

But, we have to consider the following factors:

(i) someones are always late, the department chair decided to actually start the meeting only when the 75% of them gather to the venue.

(ii) Each professor given his/her preferences has a target time t_i of arrival, for $i = 1, \dots, N$.

(iii) Because of unpredictable circumstances (weather conditions, traffic etc), they actually arrive at the venue at time X_i.

$$X_i = t_i + \sigma_i \xi_i, \quad \sigma_i > 0,\ \xi_i \text{ i.i.d.} \sim N(0,1). \tag{123}$$

(iv) where, we explain that

- t_i is the desired arrival time which is the control for professor i.
- $\sigma_i \xi_i$ models unpredictable events.

Definition of MFG VI

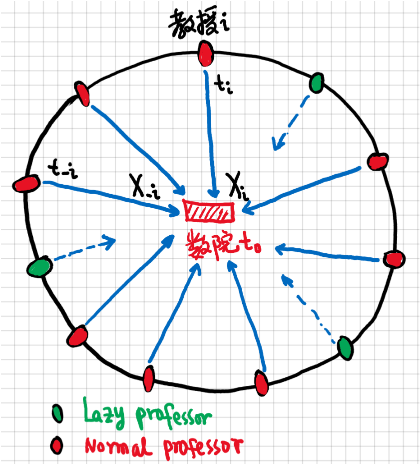


Figure: MFG on meeting

Definition of MFG VII

- By (i), the actual time T the meeting starts, is a function of the empirical distribution μ_X^N of the arrival times $X = (X_1, ..., X_N)$,

$$\mu_X^N(dx) := \frac{1}{N}\sum_{i=1}^N \delta_{X_i}(dx), \quad \text{on } \mathcal{B}(\mathbb{R}). \tag{124}$$

- In other words,
  $T = f(\mu_X^N) = \inf\{t \in (-\infty, t_0];\ \mu_X^N((-\infty, t]) = 0.75\}$.
- The expected cost of professor i is given by: for control $(t_1, \ldots, t_N) \in \mathbb{R}_+^N$,

$$J_i(t_1, \ldots, t_i, \ldots, t_N)$$
$$= E\left[\underbrace{A(X_i - t_0)^+}_{\text{reputation cost}} + \underbrace{B(X_i - T)^+}_{\text{overdue cost}} + \underbrace{C(T - X_i)^+}_{\text{cost of early arrival}}\right], \tag{125}$$

where, we note that T depends on the control $(t_1, \ldots, t_N)$.

Definition of MFG VIII

- We next give the game formulation:

Game Formulation:

- $\mathbb{I}$: the set of agents (players); $|\mathbb{I}| = N$, the number of players;
- U_i: the set of actions (controls) for agent (player) i; $U = U_1 \times \cdots \times U_N$;
- $u = (u_1, \ldots, u_N) \in U$; $u_{-i} := (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N)$;
- $\mathbb{J}$: the set of payoff functions $J : U \to \mathbb{R}$.

- Then, we have the following definitions:

Definition (Nash Equilibrium)

Let $J_i \in \mathbb{J}$ be the payoff function of player i. An action (control) $u^* \in U$ is called a Nash equilibrium of the game $\mathbb{I}$ if and only if for every player $i = 1, \ldots, |\mathbb{I}|$, $J_i(u^*) \geq J_i(v, u^*_{-i})$ for all $v \in U_i$.

Definition of MFG IX

- We introduce the concept of the best response function:

Definition (Best Response Function (BRF))

For player $i = 1, \ldots, |\mathbb{I}|$, a function $B_i : U \to U_i$ is said to be a best response of player i to the actions of the other players if for all $u \in U$,

$$B_i(u) := \{u_i \in U_i;\ J_i(u_i, u_{-i}) \geq J_i(v, u_{-i}),\ \forall\ v \in U_i\}.$$

- The fixed point of $B := B_1 \times \cdots \times B_N : U \to U$:

Let $u^* \in U$ be the fixed point of $B : U \to U$ if and only if u^* is a Nash equilibrium of the game $\mathbb{I}$:

- $u^* = B(u^*) = B_1(u^*) \times \cdots \times B_N(u^*)$, i.e., $J_i(u^*) \geq J_i(v, u^*_{-i})$ for all $v \in U_i$.

- Example: Prisoners' Dilemma:

Definition of MFG X

"囚徒困境"是1950年美国兰德公司的梅里尔·弗勒德（Merrill Flood）和梅尔文·德雷希尔（Melvin Dresher）拟定出相关困境的理论，后来由顾问艾伯特·塔克（Albert Tucker）以囚徒方式阐述，并命名为"囚徒困境"。两个共谋犯罪的人被关入监狱，不能互相沟通情况。如果两个人都不揭发对方，则由于证据不确定，每个人都坐牢1年；若一人揭发，而另一人沉默，则揭发者因为立功而立即获释，沉默者因不合作而入狱20年；若互相揭发，则因证据确凿，二者都判刑5年。由于囚徒无法信任对方，因此倾向于互相揭发，而不是同守沉默。最终导致纳什均衡仅落在非合作点上的博弈。

Definition of MFG XI

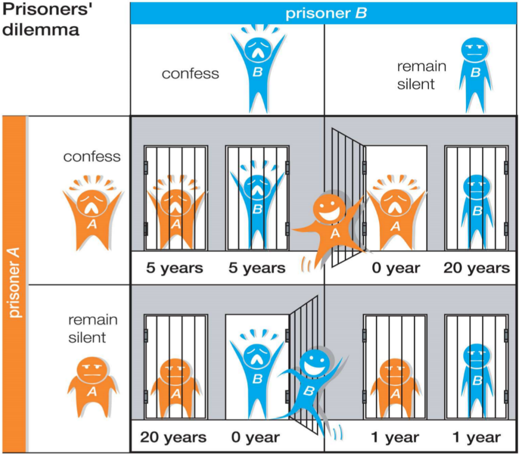


Figure: Game and Nash Equilibrium

Definition of MFG XII

- We next give an example on a game which has no Nash equilibrium:

Matching Pennies: There are $|\mathbb{I}| = N = 2$ players who show each other simultaneously the face of a coin:

- if the faces they show are the same, then player 2 pays 1 dollar to player 1;
- if the faces they show are different, then player 1 pays 1 dollar to player 2.

- Formulation of this game: $U_1 = U_2 = \{\text{Head}, \text{Tail}\}$; The payoff of the game is

		play 2	
		Head	Tail
play 1	Head	1, -1	-1, 1
	Tail	-1, 1	1, -1

- Question: Check that this game has no Nash equilibrium.

Definition of MFG XIII

- In order to guarantee the existence of Nash equilibrium, we introduce the idea of Nash equilibrium in mixed (relaxed) strategies.

Mixed (Relaxed) Strategies: A mixed (relaxed) strategy for player i in a strategic game is a probability measure $\mu_i \in \mathcal{P}(U_i)$ for his/her actions (control or strategy) given the actions of the other players.

Definition (Nash Equilibrium in Mixed (Relaxed) Strategies)

We call $\mu^* \in \mathcal{P}(U) = \mathcal{P}(\Pi_{i=1}^N U_i)$ a Nash equilibrium in mixed (relaxed) strategies if $J_i(\mu^*) \geq J_i(\mu)$ for all $\mu \in \mathcal{P}(U)$.

- If $u^* \in U$ is a Nash equilibrium, then $\mu^* = \delta_{u^*} \in \mathcal{P}(U)$ is a Nash equilibrium in degenerate mixed strategies.
- Question: Nash Theorem: Prove that "every strategic game with a finite action (control) set, has a Nash equilibrium in mixed strategies".

Definition of MFG XIV

- Nash Theorem stresses the importance of the finiteness of the plays. We next discuss how to deal with the game with the large number N of players.
- As we introduced at the beginning of this part, we expect to get a representative player for is game as $N \to \infty$.

A Representative Agent (Player) of Meeting Game: Glivenko-Cantelli Lemma yields that, if $X_1, \ldots, X_N, \ldots$ are i.i.d., then there exists a probability measure $\mu \in \mathcal{P}(\mathbb{R})$ s.t. $\mu_X^N \Rightarrow \mu$. Moreover, as $N \to \infty$,

$$\sup_{x \in \mathbb{R}} \left| \mu_X^N((-\infty, x]) - \mu((-\infty, x]) \right| \to 0, \quad P\text{-}a.s.$$

Consider a simplification: For $i = 1, \ldots, N$, $X_i = X$, $t_i = t$, $\epsilon_i \to \epsilon \sim N(0,1)$ and $\sigma_i \to \sigma > 0$. Then

$$X = t + \sigma\epsilon \text{ and } J_i(t_1, \ldots, t_N) \to J(t, f(\mu))$$

Definition of MFG XV

- We compute the objective functional $J(t, f(\mu))$ of the representative agent by defining $T^* = f(\mu)$:

$$
\begin{aligned}
J(t, T^*) &= E\left[A(X-t_0)^+ + B(X-T^*)^+ + C(T^*-X)^+\right] \\
&= AE\left[(t-t_0+\sigma\epsilon)^+\right] + BE\left[(X-T^*)\mathbb{1}_{X>T^*}\right] \\
&\quad + CE\left[(T^*-X)\mathbb{1}_{X\leq T^*}\right] \\
&= AE\left[(t-t_0+\sigma\epsilon)^+\right] + BE\left[X-T^*\right] \\
&\quad + (B+C)E\left[(T^*-X)\mathbb{1}_{X\leq T^*}\right] \\
&= A\int_{\mathbb{R}}(t-t_0+\sigma x)^+\varphi(x)dx + B(t-T^*) \\
&\quad + (B+C)\int_{-\infty}^{\frac{T^*-t}{\sigma}}(T^*-t-\sigma x)\varphi(x)dx
\end{aligned}
$$

- Question: Find a minimizer $t^* = t^*(T^*)$ of $t \to J(t, T^*)$. Prove that $t^*(\cdot)$ has a fixed point.

Interacting Controlled Diffusion Processes I

- We here give an abstract model for illustrating the construction of Nash equilibrium of the mean-field differential game by using two approaches mentioned in previous sections.
- For agent (player) i, the state process with his/her control u^i is given by the following interacting diffusion process with mean field:

$$dX_t^i = \alpha(\overline{X}_t - X_t^i)dt + u_t^i dt + \sigma\left(\sqrt{1-\rho^2}dW_t^i + \rho dW_t^0\right), \quad (126)$$

- The parameters in (126) satisfy that

The Mean Field Term: $\overline{X}_t = \frac{1}{N}\sum_{i=1}^N X_t^i$; the control u^i is an $\mathbb{R}$-valued progressively measurable process (the set of thus control is given by $\mathcal{U}^i$); Brownian Motions: W^i, $i = 0, 1, \ldots, N$ are independent (scalar) Brownian motions; $\sigma > 0$, $\rho \in [-1, 1]$.

Interacting Controlled Diffusion Processes II

- The objective (cost) functional of agent i is defined as: for $(u^1, \ldots, u^N) \in \mathcal{U} := \mathcal{U}^1 \times \cdots \times \mathcal{U}^N$,

$$J_i(u^1, \ldots, u^N) := E\left[g(X_T^i, \overline{X}_T) + \int_0^T f(X_t^i, \overline{X}_t, u_t^i)dt\right]. \tag{127}$$

- The terminal cost function and running cost function are given by

The Terminal Cost Function: for $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and $\overline{x} : \frac{1}{N}\sum_{i=1}^N$,

$$g_i(x) := g(x_i, \overline{x}) := \frac{C}{2}|\overline{x} - x_i|^2.$$

The Running Cost Function: For $u^i \in \mathbb{R}$,

$$f_i(x, u^i) := f(x_i, \overline{x}, u^i) := \frac{|u^i|^2}{2} - qu^i(\overline{x} - x_i) + \frac{\epsilon}{2}|\overline{x} - x_i|^2.$$

Interacting Controlled Diffusion Processes III

- The function $f_i(x, u^i)$ is convex in (x, u^i) if $q^2 \leq \epsilon$.
- We next apply the HJB equation approach to find the Nash equilibrium with finite N:
- To this purpose, let $X_t = (X_t^1, \ldots, X_t^N)$ for $t \geq 0$, and define the value function of agent i as: for $(t, x) \in [0, T] \times \mathbb{R}^N$ and $u^{-i} = (u_t^{-i})_{t\in[0,T]}$ being fixed

$$V_i(t, x) := \inf_{u^i \in \mathcal{U}^i} E\left[g(X_T^i, \overline{X}_T) + \int_t^T f(X_t^i, \overline{X}_t, u_t^i) dt \middle| X_t = x \right]. \tag{128}$$

Interacting Controlled Diffusion Processes IV

- Then, the value function $V_i(t,x)$ satisfies the HJB equation:

$$
\begin{aligned}
0 = \partial_t V_i(t,x) + \inf_{u^i\in\mathbb{R}} \Bigg\{ & \sum_{j=1}^{N} [\alpha(\overline{x}-x_j)+u^j]\partial_{x_j} V_i(t,x) \\
& + \frac{\sigma^2}{2}\sum_{j,k=1}^{N}(\rho^2+\delta_{jk}(1-\rho^2))\partial^2_{x_j,x_k} V_i(t,x) \\
& + \frac{|u^i|^2}{2} - qu^i(\overline{x}-x_i) + \frac{\epsilon}{2}|\overline{x}-x_i|^2 \Bigg\}.
\end{aligned}
$$

- The terminal condition $V_i(T,x)=\frac{c}{2}|\overline{x}-x_i|^2$ for $x\in\mathbb{R}^N$.
- Using the first-order condition w.r.t. u^i, it follows that

$$
u^{*,i}(t,x) = q(\overline{x}-x_i) - \partial_{x_i} V_i(t,x), \quad i=1,\dots,N.
$$

Interacting Controlled Diffusion Processes V

- In terms of the terminal condition $V_i(T, x) = \frac{C}{2}|\overline{x} - x_i|^2$, we guess the value function V_i in the following form:

$$V_i(t, x) = \frac{\eta_t}{2}(\overline{x} - x_i)^2 + \mu_t, \quad t \in [0, T].$$

- Here $t \to \eta_t$ and $t \to \mu_t$ are deterministic functions with $\eta_T = C$ and $\mu_T = 0$.
- Plugging $u^{*,i}$ and the expression of $V_i(t, x)$ into the HJB equation, we obtain

$$\begin{cases} \partial_t \eta_t = 2(\alpha + q)\eta_t + \left(1 - \frac{1}{N^2}\right)\eta_t^2 - (\epsilon - q^2), \\ \\ \partial_t \mu_t = -\frac{\sigma^2}{2}(1 - \rho^2)\left(1 - \frac{1}{N}\right)\eta_t. \end{cases}$$

Interacting Controlled Diffusion Processes VI

- The HJB equation approach gives Closed-Loop Equilibria.
- Question: Apply Stochastic Maximum Principle to find the Nash Equilibrium, which is an Open-Loop Equilibria.
- To illustrate the application of Stochastic Maximum Principle to find the Nash Equilibrium, we discuss a Linear-Quadratic (LQ) model coming from R. Carmona, Jean-Pierre Fouque and Li-Hsien Sun (2013) below.

Approximate Nash Equilibria I

- The procedure of finding Approximate Nash Equilibrium as $N \to \infty$:

Step 1: Fix an $(\mathcal{F}_t^{W_0})_{t\in[0,T]}$-adapted process $(m_t)_{t\in[0,T]}$ (being thought of as a candidate for the limit of $\overline{X}_t$ as $N \to \infty$);

Step 2: Consider the following control problem of a representative agent (player) given by:

$$\inf_{u\in\mathcal{U}} E\left[g(X_T, m_T) + \int_0^T f(X_t, m_t, u_t)dt\right],$$

where $dX_t = \alpha(m_t - X_t)dt + u_t dt + \sigma(\sqrt{1-\rho^2}dW_t + \rho dW_t^0)$, and W^i for $i \geq 1$, W^0 and W are independent Brownian motions.

Approximate Nash Equilibria II

Step 3: Solving the fixed point problem given by

$$m_t = E\left[X_t|\mathcal{F}_t^{W_0}\right], \quad t \in [0, T].$$

- Note that in Step 2, $m = (m_t)_{t\in[0,T]}$ is a process. Then, it is convenient to solve the control problem of the representative agent using Stochastic Maximum Principle (see Theorem 35):
- Therefore, the Hamiltonian is given by: for $(t, x, y) \in [0, T] \times \mathbb{R} \times \mathbb{R}$ and $(z, u) \in \mathbb{R}^2 \times \mathbb{R}$,

$$\Pi(\omega, t, x, y, z, u) := \alpha(m_t(\omega) - x)y + uy + \left[\sigma\sqrt{1-\rho^2}, \sigma\rho\right] z \\ + f(x, m_t(\omega), u).$$

- It is strictly convex in (x, u) under the condition $q^2 \leq \epsilon$.

Approximate Nash Equilibria III

- Then, FOC gives that, the optimal control u^* satisfies that

$$\frac{\partial \Pi}{\partial u} = 0 \Longleftrightarrow u^* = q(m_t - x) - y.$$

- The corresponding adjoint forward-backward equations are given by:

$$\begin{aligned} dX_t^* &= \Pi_y(t, X_t^*, Y_t^*, Z_t^*, u_t^*)dt + \sigma(\sqrt{1-\rho^2}dW_t + \rho dW_t^0); \\ dY_t^* &= -\Pi_x(t, X_t^*, Y_t^*, Z_t^*, u_t^*)dt + Z_t^{*,0}dW_t^0 + Z_t^{*,1}dW_t, \\ Y_T^* &= \nabla_x g(X_T^*), \end{aligned}$$

where $Z_t^* = (Z_t^{*,0}, Z_t^{*,1})$.

Approximate Nash Equilibria IV

- Therefore

$$
\begin{aligned}
dX_t^* &= [(\alpha+q)(m_t - X_t^*) - Y_t^*]dt + \sigma(\sqrt{1-\rho^2}dW_t + \rho dW_t^0);\\
dY_t^* &= [(\alpha+q)Y_t^* + (\epsilon - q^2)(m_t - X_t^*)]dt + Z_t^{*,0}dW_t^0 + Z_t^{*,1}dW_t,\\
Y_T^* &= C(X_T^* - m_T).
\end{aligned}
$$

- Hence, for $m_t^X := E[X_t^*|\mathcal{F}_t^{W_0}]$ and $m_t^Y := E[Y_t^*|\mathcal{F}_t^{W_0}]$, we obtain that

$$
m_t^X = m_0^X + \int_0^t [(\alpha+q)(m_s - m_s^X) - m_s^Y]ds + \sigma\rho W_t^0; \tag{129}
$$

$$
m_t^Y = m_T^Y - \int_t^T [(\alpha+q)m_s^Y + (\epsilon - q^2)(m_s^Y - m_s^X)]ds + Z_t^{*,0}dW_t^0; \tag{130}
$$

$$
m_T^Y = C(m_T^X - m_T). \tag{131}
$$

Approximate Nash Equilibria V

- Question: Prove (129), (130) and (131).
- By Step 3, we have $m_t = m_t^X$ by the fixed point, and hence $m_T^Y = 0$. Thus, we obtain

$$m_t^Y = -\int_t^T e^{(\alpha+q)(s-t)} Z_s^{0,*} dW_s^0, \quad t \in [0, T].$$

- Then, $dm_t = dm_t^X = -m_t^Y dt + \rho\sigma dW_t^0$.
- In order to obtain $m = (m_t)_{t\in[0,T]}$, we have to find $Z^{*,0}$, and this implies that we need to find the solution (Y^*, Z^*) of BSDE (130).
- Now, we assume that the first solution component Y_t^* of BSDE (130) has the form given by:

$$Y_t^* = -\eta_t(m_t - X_t^*),$$

where $t \to \eta_t$ is a C^1-deterministic function.

Approximate Nash Equilibria VI

- Plugging it into (130) to have that $Z^{0,*} = 0$ and $Z^{1,*} = \eta_t \sigma \sqrt{1-\rho^2}$ with

$$\partial_t \eta_t = 2(\alpha + q)\eta_t + \eta_t^2 - (\epsilon - q^2), \quad \eta_T = C. \tag{132}$$

- Therefore $m_t^Y \equiv 0$ and hence $m_t = m_t^X = E[X_0^*] + \sigma\rho W_t^0$.

MFG Strategy with Finite Players: $u_t^{*,i} = (\eta_t + q)(\overline{X}_t^* - X_t^{*,i})$;
MFG Strategy with Infinite Players: $u_t^* = q(m_t - X_t^*) - Y_t^*$.

- Question: Does it hold that $J_i(u_t^{*,1}, \ldots, u_t^{*,N}) \to J(u^*)$ as $N \to \infty$?
- We next apply HJB equation approach to find an approximating Nash Equilibrium:
- In order to apply HJB equation approach, we assume that Nash Equilibrium has a Markovian feedback form, i.e. $u_t = u(t, X_t)$.

Approximate Nash Equilibria VII

- Then, the limiting state process is given by

$$dX_t = \alpha(m_t - X_t)dt + u(t, X_t)dt + \sigma\left(\sqrt{1-\rho^2}dW_t + \rho dW_t^0\right).$$

- Here $m_t = E[X_t|\mathcal{F}_t^{W_0}]$ by the fixed point in Step 3.
- For $\rho_0 \in \mathcal{P}(\mathbb{R})$, define the measure-valued process as:

$$\mu_t(dx) = \int_{\mathbb{R}} E\left[\delta_{X_t^{x_0,\mu}}(dx)\middle|\mathcal{F}_t^{W_0}\right]\rho_0(dx_0), \quad \text{on } \mathcal{B}(\mathbb{R}). \tag{133}$$

- Here $X^{x_0,\nu}$ for $\nu_t \in \mathcal{P}_1(\mathbb{R})$ satisfies that

$$\begin{aligned} X_t^{x_0,\nu} &= x_0 + \int_0^t \alpha(\langle \nu_s, I\rangle - X_s^{x_0,\nu})dt + u(s, X_s^{x_0,\nu})ds \\ &\quad + \sigma\left(\sqrt{1-\rho^2}W_t + \rho W_t^0\right). \end{aligned} \tag{134}$$

Approximate Nash Equilibria VIII

- Then, $\mu=(\mu_t)_{t\in[0,T]}$ defined by (133) is an $(\mathcal{F}_t^{W_0})_{t\in[0,T]}$-adapted $\mathcal{P}(\mathbb{R})$-valued process with $\mu_0=\rho_0$.
- Hence, $\langle\mu_t,I\rangle=\int_{\mathbb{R}}E[X_t^{x_0,\mu}|\mathcal{F}_t^{W_0}]\rho_0(dx_0)$ with $I(x):=x$.
- The process $\mu=(\mu_t)_{t\in[0,T]}$ defined by (133) satisfies the following stochastic FPK equation given by

Stochastic FPK equation: for all $f\in C_0^{\infty}(\mathbb{R})$,

$$\langle\mu_t,f\rangle=\langle\rho_0,f\rangle+\int_0^t\langle\mu_s,\mathcal{A}^{\mu_s}f\rangle ds+\int_0^t\langle\mu_s,\mathcal{L}f\rangle dW_s^0, \tag{135}$$

where the operators are given by: for $\nu\in\mathcal{P}_1(\mathbb{R})$,

$$\mathcal{A}^{\nu}f(x):=[\alpha(\langle\nu,I\rangle-x)+u(t,x)]f'(x)+\frac{\sigma^2}{2}f''(x);$$
$$\mathcal{L}f(x):=\sigma\rho f'(x),\quad x\in\mathbb{R}. \tag{136}$$

Approximate Nash Equilibria IX

- Question: Similarly to (70), prove the existence of solutions to stochastic FPK equation (135) using the fixed point argument.
- Let $\rho_0(dx) = u_0(x)dx$, i.e., u_0 is the initial density function of μ.
- Then, $\mu_t(dx) = p(t,x)dx$ where $p(t,x)$ satisfies the following stochastic forward Kolmogorov equation:

$$\partial_t p(t,x) + \mathcal{A}^* p(t,x) + \mathcal{L}^* p(t,x) dW_t^0 = 0, \quad (t,x) \in (0,T] \times \mathbb{R}; \\ p(0,x) = u_0(x), \quad x \in \mathbb{R}, \tag{137}$$

where $\mathcal{A}^*$ (resp. $\mathcal{L}^*$) are the adjoint operator of $\mathcal{A}$ (resp. $\mathcal{L}$).
- Question: Write the expression of the adjoint operators $\mathcal{A}^*$ and $\mathcal{L}^*$. Prove that $m_t = \langle \mu_t, I \rangle = \int_{\mathbb{R}} xp(t,x)dx$ satisfies that

$$dm_t = \rho\sigma dW_t^0, \quad m_0 = \int_{\mathbb{R}} xu_0(x)dx. \tag{138}$$

Approximate Nash Equilibria X

- Then, by Step 2, we conclude our limiting MFG model for a representative agent:

The MFG Value Function: for $(t, x, m) \in [0, T] \times \mathbb{R} \times \mathbb{R}$,

$$V(t, x, m) := \inf_{u \in \mathcal{U}} E\left[g(X_T, m_T) + \int_t^T f(X_t, m_t, u_t) dt \middle| X_t = x, m_t = m \right].$$

The State Process $(X_t, m_t)_{t \in [0,T]}$:

$$dX_t = \alpha(m_t - X_t)dt + u(t, X_t)dt + \sigma\left(\sqrt{1-\rho^2}dW_t + \rho dW_t^0\right);$$
$$dm_t = \rho\sigma dW_t^0. \tag{139}$$

- Question: Derive the HJB equation of the above control problem of a representative agent and solve this control problem.

Approximate Nash Equilibria XI

- If $\rho = 0$, i.e., there is no common noise in the model, then the above equation reduces to MFG equation proposed in P.L. Lions's paper:
- Pierre-Louis Lions (1956-): French mathematician. His research interest is Nonlinear PDE, he is the recipient of the 1994 Fields Medal.

Figure: Pierre-Louis Lions (1956-)

References on This Course

1 Fleming, W., and R. Rishel (1975): *Deterministic and Stochastic Optimal Control*. Springer-Verlag, New York.

2 Friedman, A. (1965): *Partial Differential Equations of Parabolic Type*. Dover Publications.

3 Friedman, A. (1975): *Stochastic Differential Equations and Applications*. Dover Publications.

4 Karatzas, I., and S.E. Shreve (1991): *Borwonian Motion and Stochastic Calculus*. Springer-Verlag, New York.

5 Revuz, D., and M. Yor (2005): *Continuous Mrtingales and Brownian Motion*. Springer-Verlag, New York.

6 Villani, C. (2003): *Topics in Optimal Transportation*. American Mathematical Soc.

7 Villani, C. (2008): *Optimal Transport: Old and New*. Springer Science & Business Media, New York.

End

THANK YOU FOR YOUR ATTENTION!