FROM FOKKER-PLANCK TO SDES AND BACK VIA THE MARTINGALE PROBLEM

ANDRÉ SCHLICHTING

These notes are only for the use of the students in the title mentioned class at the University of Bonn, fall term 2015–2016.

Please send typos and corrections to schlichting@iam.uni-bonn.de

Contents

Date: February 9, 2016.

1. Introduction and outline

Let $X(\cdot)$ be a continuous Markov process with values in $\mathbb{R}^d$ and suppose that for $t \geq 0$ and test functions $\varphi \in C_c^\infty(\mathbb{R}^d)$ holds

$$E\left[\varphi(X(t+h)) - \varphi(X(t)) \mid X(t) = x\right] = hL_t\varphi(x) + o(h), \tag{1.1}$$

where $L_t : C_0^\infty(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is a linear operator. An immediate observation is, that if φ has maximum in x_0, then $E\left[\varphi(t+h) - \varphi(t) \mid x(t) = x_0\right] \leq 0$ and hence L_t must satisfy the weak maximum principle.

The linear operator L_t in the case of a continuous Markov process can be represented for some $a \in [0,T] \times \mathbb{R}^d \to S_d = \left\{a \in \mathbb{R}^{d\times d} : \langle\theta, a\theta\rangle \geq 0\right\}$ and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$L_t = \frac{1}{2}\sum_{i,j=1}^d a^{ij}(t,x)\partial_i\partial_j + \sum_{i=1}^d b^i(t,x)\partial_i.$$

Question: How can the Markov process $X(\cdot)$ be understood in terms of the generator L_t?

1.1. Forward equation: Fokker-Planck.

Let $P(s,x;t,\Gamma)$ denote the transition probability function, i.e.

$$P(s,x;t,\Gamma) = P(X(t) \in \Gamma \mid X(s) = x).$$

Since $x(\cdot)$ is a Markov process, the Chapman-Kolmogorov equations hold for $0 \leq s < t < u$

$$P(s,x;u,\Gamma) = \int P(t,y;u,\Gamma)\, P(s,x;t,dy).$$

Then, we can conclude from (1.1)

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi(y)P(s,x;t,dy) &= \lim_{h\to 0}\frac{1}{h}\int \varphi(y)\Big(\underbrace{P(s,x;t+h,dy)}_{\int P(t,z;t+h,dy)P(s,x;t,dz)} - P(s,x;t,dy)\Big) \\
&= \int P(s,x;t,dz)\lim_{h\to 0}\frac{1}{h}\underbrace{\left(\int P(t,z;t+h,dy)\varphi(y)-\varphi(z)\right)}_{=E[\varphi(x(t+h))-\varphi(x(t))|x(t)=z]} \\
&= \int L_t\varphi(z)\; P(s,x;t,dz) \qquad \forall \varphi\in C_c^\infty(\mathbb{R}^d).
\end{aligned}$$

Hence, we deduced the equation

$$\begin{aligned}
\partial_t P(s,x;t,\cdot) &= L_t^* P(s,x;t,\cdot), \quad t>s \\
\lim_{t\to s} P(s,x;t,\cdot) &= \delta_x(\cdot),
\end{aligned} \tag{1.2}$$

where L_t^* is the formal adjoint to L_t defined by

$$L_t^* g = \frac{1}{2}\sum_{i,j=1}^{d}\partial_i\partial_j\left(a^{ij}g\right) - \sum_{i=1}^{d}\partial_i\left(b^i g\right).$$

The solution $P(s,x;t,\cdot)$ is the parabolic Green function or fundamental solution for the Fokker-Planck equation.

- Need a density $P(s,x;t,dy)=p(s,x;t,y)\,dy$ to think of (1.2) as a equation for the density
- a and b need to be at least differentiable
- Works well if classic (strong) solution of the PDE exist

1.2. **Backward equation.** Define for $0\le s<t$ and $\varphi\in C_c^\infty(\mathbb{R}^d)$ the observable

$$u(s,x)=\int \varphi(y)P(s,x;t,dy) = E\left[\varphi(x(t))\mid x(s)=x\right].$$

Then

$$\begin{aligned}
-\partial_s u(s,x) &= -\partial_s\int \varphi(y)P(s,x;t,dy) \\
&= \lim_{h\to 0}\frac{1}{h}\int\Big(\underbrace{P(s-h,x;t,dy)}_{\int P(s,z;x,dy)P(s-h,x;s,dz)} - P(s,x;t,dy)\Big)\varphi(y) \\
&= \lim_{h\to 0}\frac{1}{h}\left(\int P(s-h,x;s,dz)u(s,z)-u(s,x)\right) \\
&= \lim_{h\to 0}\frac{1}{h}E\left(u(s,X(s))-u(s,X(s-h))\mid X(s-h)=x\right) = L_s u(s,x)
\end{aligned}$$

Hence, we found the formal adjoint to the Fokker-Planck equation

$$\partial_s P(s,x;t,\cdot) + L_s P(s,x;t,\cdot) = 0$$
$$\lim_{s\to t} P(s,x;t,\cdot) = \delta_x(\cdot).$$

- Existence holds under weaker assumptions on the coefficients (ellipticity and Hölder continuity)
- Corresponds to weak solution concepts for PDEs
- Lack of interpretation, due to its purely analytic content

1.3. **Itô.** We go back to a pathwise differential representation, where $X(t+h)-X(t)$ should look like Gaussian independent increments with drift $b(t,X(t))$ and covariance $a(t,X(t))$, which in infinitesimal form writes as

$$dX(t) = b(t,X(t))dt + \sigma(t,X(t))\,d\beta(t) \tag{1.3}$$

where β is a d-dim Brownian motion and σ a square root of a, i.e. $\sigma\sigma^T = a$.

- How to make sense of (1.3), since only $\beta \in C^{0,\alpha}([0,T])$ for $\alpha < \frac{1}{2}$
- How to use (1.3) for making calculations?
- Why does β as intermediate process pop up? Is it needed?
- Existence for σ, b Lipschitz is straight forward.

1.4. **Martingale solutions.** By the law of total expectation, holds for $0 \le t_1 < t_2$

$$E\left(\varphi(X(t_2)) \mid X(s), s \le t_1\right) = E\left(E\left(\varphi(X(t_2)) \mid X(s), s \le t_2\right) \mid X(s), s \le t_1\right).$$

Therewith, we can calculate

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t_2} E\left[\varphi(X(t_2)) \mid X(s), s \le t_1\right] \\
&= \lim_{h\to 0} \frac{1}{h} E\left[E\left[\varphi(X(_2+h)) - \varphi(X(t_2)) \mid X(s), s \le t_2\right] \mid X(s), s \le t_1\right] \\
&= E\left[L_{t_2}\varphi(X(t_2)) \mid X(s), s \le t_1\right].
\end{aligned}$$

Hence, integration leads to

$$E\left[\varphi(X(t_2)) - \varphi(X(t_1)) - \int_{t_1}^{t_2} L_t\varphi(X(t))\,\mathrm{d}t \mid X(s), s \le t_1\right] = 0.$$

By definition, this means that for any φ smooth enough

$$X_{\varphi(t)} = \varphi(X(t)) - \int_0^t L_s\varphi(X(s))\,\mathrm{d}s$$

is a martingale.

Definition 1.1 (Marginale problem)**.** Let L_t be given. A family of probability measures $\{P_x\}_{x\in\mathbb{R}^d}$ on $C([0,\infty],\mathbb{R}^d)$ is a solution to the martingale problem if

(i) $P(X(0)=x) = 1$ and $X_\varphi(\cdot)$ is a martingale for all test functions

(ii) There is only at most one P_x for every $x \in \mathbb{R}^d$.

- How does the Markov property come into play?
- What are necessary conditions of a well-defined solution to the martingale problem?
- What conclusion can be drawn from a well-defined solution?

1.5. **Interplay with measure valued solution of the PDE.** Having a well-defined solution to the martingale problem P_x, we can obtain measure valued solution to the Fokker-Planck equation by a superposition principle. Such a solution satisfies for all test functions $\varphi \in C_c^\infty(R^d)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi(x)\,\mathrm{d}\mu_t(x) = \int \left(\frac{1}{2}\sum_{i,j} a_{ij}(t,x)\partial_{ij}\varphi(x) + \sum_i b_i(t,x)\partial_i\varphi(x)\right)\mathrm{d}\mu_t(x)$$

and μ_t converges to $\bar{\mu}$ in the weak* topology as $t \to 0$ (initial value). The solution is defined for bounded a and b as long as $\int_0^T |\mu_t|(A)\,\mathrm{d}t < \infty$ for compact $A \subset \mathbb{R}^d$.

Let μ_0 be some initial measure with $|\mu_0|(\mathbb{R}^d) < \infty$. Then the measure μ_t^P defined by

$$\int \varphi\,\mathrm{d}\mu_t^P = \int_{\mathbb{R}^d} E^{P_x}[X(t)]\,\mathrm{d}\mu_0(x) \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^d)$$

is a measure-valued solution to the Fokker-Planck equation.

But, it holds also the converse, that is: Let μ_t be a measure-valued solution to the PDE, such that $\mu_t \in M_+(\mathbb{R}^d)$[1] for any $t \in [0,T]$ and $\mu_t(\mathbb{R}^d) \leq C$ for any $t \in [0,T]$. Then there exists a family of probability measure $\{P_x\}_{x\in\mathbb{R}^d}$ such that ν_x is a martingale solution and it holds the representation formula

$$\int_{\mathbb{R}^d} \varphi\,\mathrm{d}\mu_t = \int_{\mathbb{R}^d} E^{P_x}\left[x(t)\right]\mathrm{d}\mu_0(x).$$

2. Markov processes and parabolic partial differential equations

2.1. Probability spaces and random variables.

Definition 2.1 (Probability space)**.** Let Ω be a given set, then a σ-algebra $\mathcal{F}$ on Ω is a family $\mathcal{F}$ of subsets of Ω satisfying

(i) $\emptyset \in \mathcal{F}$;
(ii) $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$, where $F^c = \Omega \setminus F$;
(iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^\infty A_i \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a *measurable space.* A probability measure P is a function $P : \mathcal{F} \to [0,1]$ such that

[1]non-negative finite measures

(a) $P(\emptyset) = 0$, $P(\Omega) = 1$;
(b) if $A_1, A_2, \dots \in \mathcal{F}$ and $\{A_i\}_{i=1}^\infty$ is pairwise disjoint, then

$$P\left(\bigcup_{i=1}^\infty\right) = \sum_{i=1}^\infty P(A_i).$$

The triple $(\Omega, \mathcal{F}, P)$ is called a *probability space*. It is called a *complete* probability space if for any $G \subseteq F \in \mathcal{F}$ with $P(F) = 0$ follows $G \in \mathcal{F}$[2] and from now on the completeness assumption is implied without further reference.

Definition 2.2 (Measurability and random variables)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space, then $f : \Omega \to \mathbb{R}^d$ is $\mathcal{F}$*-measurable* if for all open $U \in \mathbb{R}^d$

$$f^{-1}(U) := \{\omega \in \Omega : f(\omega) \in U\} \in \mathcal{F}.$$

For any given $f : \Omega \to \mathbb{R}^d$, the σ*-algebra generated by* f is

$$\sigma(f) := \sigma\left(\left\{f^{-1}(U) : U \subset \mathbb{R}^d \text{ open}\right\}\right) = \sigma\left(\left\{f^{-1}(U) : U \in \mathcal{B}_{\mathbb{R}^d}\right\}\right).$$

A *random variable* f is an $\mathcal{F}$-measurable function $f : \Omega \to \mathbb{R}^d$. Each random variable f induces a probability measure on $\mathbb{R}^d$ by

$$\mu^f(A) = P(f^{-1}(A)), \qquad \text{for any } A \in \mathcal{F}.$$

A random variable f is *integrable*, denoted by $f \in L^1(\Omega, \mathcal{F}, P)$, if $\int_\Omega f(\omega)\, \mathrm{d}P(\omega) < \infty$. For any integrable random variable f

$$E[f] = \int_\Omega f(\omega)\, \mathrm{d}P(\omega) = \int_{\mathbb{R}^d} x \, \mathrm{d}\mu_f(x)$$

is called its *expectation*.

27.10.15

Definition 2.3 (Conditional expectation)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $f \in L^1(\Omega, \mathcal{F}, P)$ be an integrable random variable and let $\Sigma \subset \mathcal{F}$ be a σ-algebra. Then the *conditional expectation* of f given Σ, denoted by $E[f \mid \Sigma] : \Omega \to \mathbb{R}^d$, is defined as the a.s. unique function from Ω to $\mathbb{R}^n$ (i.e. a random variable)

(i) $E[f \mid \Sigma]$ is Σ measurable
(ii) $\int_S E[f \mid \Sigma](\omega)\, \mathrm{d}P(\omega) = \int_S f(\omega)\, \mathrm{d}P(\omega)$ for all $S \in \Sigma$.

Proof. For the existence note that the measure $\mu^f : \Sigma \ni S \mapsto \int_S f \, \mathrm{d}P$ is absolutely continuous wrt. P. Likewise, the restriction $\mu^f|_\Sigma$ is absolutely continuous wrt. the restriction $P|_\Sigma$. Hence, $E[f \mid \Sigma]$ is the Radon-Nikodym derivative of $\mu^f|_\Sigma$ wrt. $P|_\Sigma$.

Assume f^Σ and g^Σ are conditional expectation of f given Σ. Then for each $S \in \Sigma$ holds

$$\int_S f^\Sigma(\omega)\, \mathrm{d}P(\omega) = \int_S f(\omega)\, \mathrm{d}P(\omega) = \int_S g^\Sigma(\omega)\, \mathrm{d}P(\omega).$$

[2]Any probability space can be completed by adding to $\mathcal{F}$ all sets G above and extending P accordingly. Completeness is desirable because it maximizes the collection of measurable sets. In a complete probability space something less than negligible is also negligible.

Moreover, if the set $\left\{\omega \in \Omega : f^\Sigma(\omega) - g^\Sigma(\omega) \neq 0\right\}$ has positive P-measure, then w.l.o.g. $\left\{\omega \in \Omega : f^\Sigma(\omega) - g^\Sigma(\omega) \geq 0\right\}$ has positive measure. Hence, for some $\varepsilon > 0$, the set $S := \left\{\omega \in \Omega : f^\Sigma(\omega) - g^\Sigma(\omega) > \varepsilon\right\}$ has positive measure. By the measurability of f^Σ and g^Σ follows $D \in \Sigma$ and hence

$$0 = \int_D f^\Sigma(\omega) - g^\Sigma(\omega) \,\mathrm{d}P(\omega) \geq \varepsilon P(D) > 0$$

a contradiction. □

Proposition 2.4 (Properties of conditional expectation)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\Sigma \subset \mathcal{F}$ be a σ-algebra. Then*

(i) For any bounded random variable $f : \Omega \to \mathbb{R}^d$ holds for P-a.e. ω

$$\inf_{\omega\in\Omega} f(\omega) \leq E\left[f \mid \Sigma\right](\omega) \leq \sup_{\omega\in\Omega} f(\omega).$$

(ii) For any integrable random variable $f \in L^1(\Omega, \mathcal{F}, P)$

$$\left|E\left[f \mid \Sigma\right]\right| \leq E\left[|f| \mid \Sigma\right].$$

(iii) The conditional expectation continuous $E\left[\cdot \mid \Sigma\right] : L^1(\Omega, \mathcal{F}, P) \to L^1(\Omega, \Sigma, P)$.

Proof. Exercise: For (i) and (ii) define a set D where the inequality fails and show that D is Σ-measruable with P-measure zero. Then (iii) immediately follows from (ii). □

Let us turn to a more functional analytic point of view. Let (X, D) be a given Polish space (complete separable metric space). Then we call $(X, \mathcal{B}_X)$ the canonical measurable space, where $\mathcal{B}_X$ is its Borel σ-algebra[3]. In this case we denote by $\mathcal{P}(X)$ the set of all probability measures on $(X, \mathcal{B})$. In this case $\mathcal{P}(X)$ can be seen as a subset of the dual of $C_b(X)$ all bounded continuous functions on X with the inherited weak* topology. Therewith, $\mathcal{P}(X)$ will become a metric space.

The convergence in $\mathcal{P}(X)$ is governed by weak* (or narrow convergence):

Definition 2.5. A sequence of probability measures $(\mu_n) \subset \mathcal{P}(X)$ converges narrowly to $\mu \in \mathcal{P}(X)$ as $n \to \infty$ if

$$\lim_{n\to\infty} \int_X f(x) \,\mathrm{d}\mu_n(x) = \int f(x) \,\mathrm{d}\mu_n(x) \qquad \text{for all } f \in C_b(X).$$

Proposition 2.6. *The topology on $\mathcal{P}(X)$ induced by weak* convergence is metrizable.*

Proof (Sketch). The separability allows to find a dense countable subset $X_0 \subset X$. Therewith, we define the following class of countable D-Lipschitz functions

$$\tilde{\mathcal{C}}_1 = \left\{x \mapsto (q_1 + q_2 D(x, y)) \wedge k : q_1, q_2, k \in \mathbb{Q}, \quad q_2 \in (0, 1), \quad y \in X_0\right\}.$$

[3]the smalled σ algebra containing all open sets.

Then, let $\mathcal{C}_1$ be the collection obtained by taking the infimum of a finite numbers of functions. Hence each $h \in \mathcal{C}_1$ satisfies $\operatorname{Lip}(h, X) = \sup_{x,y\in X} \frac{|h(x)-h(y)|}{D(x,y)} < 1$. For an enumeration $\{f_n\}_{n\in\mathbb{N}} = \mathcal{C}_1$, define the distance[4] d on $\mathcal{P}(x)$ by

$$d(\mu,\nu) = \sum_{k=1}^{\infty} \frac{1}{2^k} \left| \int f_k \,\mathrm{d}\mu - \int f_k \,\mathrm{d}\nu \right| .$$

This distance d metrizables the weak* topology on $\mathcal{P}(X)$. Obviously, the topology induced by d is weaker than the weak* topology. Now, for any neighborhood N of $\mu \in \mathcal{P}(X)$ contains an open ball wrt. d around μ. Let $\{\mu_n\}_{n\in\mathbb{N}}$ be such that $d(\mu_n, \mu) \to 0$. Then, we conclude

$$\lim_{n\to\infty} \int f \,\mathrm{d}\mu_n = \int f \,\mathrm{d}\mu \qquad \text{for all } f \in \bar{\mathcal{C}}_1. \tag{2.1}$$

The above limit can be extended to λf with $f \in \mathcal{C}_1$ and $\lambda > 0$. Hence, (2.1) holds for f bounded uniformly continuous, implying the weak* convergence (Exercise or [2, Theorem 1.2]). □

Theorem 2.7 (Prokhorov's theorem)**.** *A family of probability measures* $\Pi \subset \mathcal{P}(X)$ *is called tight if for every* $\varepsilon > 0$ *exists a compact set* $K \subset X$ *such that*

$$\inf_{\mu\in\Pi} \mu(K) \geq 1 - \varepsilon,$$

Let X *be a Polish space. A family* $\Pi \subset \mathcal{P}(X)$ *is relative compact in* $\mathcal{P}(X)$[5] *if and only if* Π *is tight.*

Remark 2.8. The implication Π tight to Π relative compact does not need completeness of the space, i.e. holds in any separable metric space. For a proof see [2, Section 5].

On a Polish space each single measure $\mu \in \mathcal{P}(X)$ is tight (Ulam's theorem [2, Theorem 1.3]). Moreover, each measure satisfies the so called *inner regularity* assumption [2, Theorem 1.1]

$$\forall B \in \mathcal{B}_X \ \forall \varepsilon > 0 \ \exists K_\varepsilon \subset X \text{ compact} : \ \mu(B \setminus K_\varepsilon) \leq \varepsilon. \tag{2.2}$$

Let us extend the concept of conditional expectation towards regular conditional probability distributions r.c.p.d. We will put it into a functional analytic framework, where it corresponds to disintegration.

Theorem 2.9 (Disintegration)**.** *Let* X, Y *be Polish*[6] *and let* $\pi : X \to Y$ *be a Borel measurable map and let* $\nu = \pi_\sharp \mu \in \mathcal{P}(Y)$ *be its pushforward defined by*

$$\nu(A) = \pi_\sharp\mu(A) = \mu(\pi^{-1}(A)) \qquad \textit{for all Borel sets } A.$$

[4] induced by *Kantorovich-Rubinstein norm*

[5] wrt. the topology induced by weak* convergence

[6] It is sufficient to have a metric separable Radon space, where each $\mu \in \mathcal{P}(X)$ is inner regular (2.2), i.e. not necessarily complete.

Then, there exists ν-a.e. uniquely defined Borel family of probability measures (r.c.p.d.) $\{\mu_y\}_{y\in Y} \subset \mathcal{P}(X)$ such that

$$\mu_y(X \setminus \pi^{-1}(y)) = 0 \qquad \text{for } \nu\text{-a.e. } y \in X \tag{2.3}$$

and for every Borel map $f : X \to [0, \infty)$.

$$\int_X f(x)\,\mathrm{d}\mu(x) = \int_Y \int_{\pi^{-1}(y)} f(x)\,\mathrm{d}\mu_y(x)\ \mathrm{d}\nu(y).$$

Proof. Let us start by assuming that X is compact. Let $\lambda = (\mathrm{Id}, \pi)_\sharp \mu \in \mathcal{P}(X \times Y)$ with the product Borel algebra $\mathcal{B}_X \otimes \mathcal{B}_Y$. Let $\Pi = \{(x, y) \in X \times Y : \pi(x) = y\}$ be the graph of π. Then $\lambda(\Pi^c) = 0$. Indeed, since Y is separable, its Borel σ-algebra is countable generated by some $\mathcal{B}^0_Y$. Since Y is also a metric space, it contains the singletons, i.e. $\{y\} = \cap \left\{ B \in \tilde{\mathcal{B}}^0_Y : y \in B \right\}$. This implies that

$$\Pi^c = \{(x, y) \in X \times Y : \pi(x) = y\} = \bigcup_{B\in\mathcal{B}^0_Y} \{(x, y) \in X \times B : \pi(x) \notin B\}.$$

Since for fixed B holds

$$\lambda(\{(x, y) \in X \times B : \pi(x) \notin B\}) = \mu(\{x \in X : \pi(x) \notin B, \pi(x) \in B\}) = 0$$

the claim follows.

Next, since $C_b(X) = C(X)$ (due to compactness) is separable (by Stone-Weierstraß), we find a countable family $\mathcal{C}_0 \subset C(X)$ such that $1 \in \mathcal{C}_0$, $\mathcal{C}_0$ is closed wrt. the $\mathbb{Q}$-linear span and $\mathcal{C}_0$ is dense in $C(X)$[7]. Now, for any $f \in \mathcal{C}_0$ define the signed measure $\lambda^f(B) := \int f(x)\mathbb{1}_B(y)\,\mathrm{d}\lambda(x, y)$ for $B \in \mathcal{B}_Y$. Then

$$\left|\lambda^f(B)\right| \le \|f\|_{C(X)} \left| \int \mathbb{1}_B(y)\,\mathrm{d}\lambda(x, y) \right| = \|f\|_{C(X)}\, \nu(B).$$

Hence $\lambda^f \ll \nu$ and by Radon-Nikodym (for signed measure) let L^f_y for ν-a.e. y denote the derivative. Moreover, $L^f_y \ge 0$ if $f \ge 0$ (by the usual Radon-Nikodym). Since for any $g \in C(X)$ holds

$$\int_Y g(y)\,\mathrm{d}\lambda^f(y) = \int_{X\times Y} f(x)g(y)\,\mathrm{d}\lambda(x, y) = \int_Y L^f_y\,\mathrm{d}\nu(y),$$

we conclude that $f \mapsto L^f_y$ is linear for ν-a.e. $y \in Y$. Hence, we find a subset $Y_0 \subset Y$ of full ν-measure, such that the functional $\mathcal{C}_0 \ni f \mapsto L^f_y$ is $\mathbb{Q}$-linear and non-negative for all $y \in Y_0$.

It is left to show its boundedness. Therefore for fixed $f \neq 0$ let $q \in \mathbb{Q}$ such that $\|f\| \le q \le 2\|f\|$. Then

$$0 \le L^{q-f}_y \le L^q_y - L^f_y \le qL^1_y - L^f_y \le 2\|f\|\, L^1_y - L^f_y,$$

[7]take for instance the $\mathbb{Q}$-linear span of the class $\mathcal{C}_1 \cup 1$ from the proof of Proposition 2.6

and hence $\left|L_y^f\right| \leq 2\left\|f\right\|\left|L_y^1\right| \leq 2A\left\|f\right\|$ and so L_y^f is a continuous functional on the dense set $\mathcal{C}_0 \subset C(X)$ and can be extended to $C(X)$. Now, by the Ries-representation theorem exists a measure μ_y for $y \in Y_0$ such that $\int f(x)\,\mathrm{d}\mu_y(x) = L_y^f$ for $f \in C(X)$. Moreover, we have for all $y \in Y_0$

$$\mu_y(X) = L_y^1 = \frac{d\lambda^1}{d\nu} = \frac{d\nu}{d\nu} = 1$$

and hence $\mu_y(X)$ is a probability measure on X. Now, we set $\mu_y = 0$ for $y \in Y \setminus Y_0$ and claim that μ_y is the desired disintegration satisfying for $f : X \to [0,\infty)$ and $g : Y \to [0,\infty)$ both measurable

$$y \mapsto \int_X f(x)g(y)\,\mathrm{d}\mu_y(x) \quad \text{is measruable} \tag{2.4}$$

$$\int_Y \int_X f(x)g(y)\,\mathrm{d}\mu_y(x)\,\mathrm{d}\nu(y) = \int_{X\times Y} f(x)g(y)\,\mathrm{d}\lambda(x,y). \tag{2.5}$$

Both claims, are easily verified for $f \in \mathcal{C}_0$, since then $\int_X f(x)g(y)\,\mathrm{d}\mu_y(x) = g(y)L_y^f$ is a product of measurable functions in y and moreover

$$\begin{aligned}\int_Y \int_X f(x)g(y)\,\mathrm{d}\mu_y(x)\,\mathrm{d}\nu(y) &= \int_Y g(y)\frac{d\lambda^f}{d\nu}(y)\,\mathrm{d}\nu(y) = \int_Y g(y)\,\mathrm{d}\lambda^f(y)\\ &= \int_{X\times Y} f(x)g(y)\,\mathrm{d}\nu(x,y).\end{aligned}$$

In particular, (2.4) and (2.5) holds for $f \in C(X)$. Now, we show that the above also holds for $h : X \times Y \to [0,\infty)$. Since the product Borel σ-algebra $\mathcal{B}_X \otimes \mathcal{B}_Y$ is generated by $A \times B$ with $A \in \mathcal{B}_X$ and $B \in \mathcal{B}_Y$, we have to show (2.4) and (2.5) for functions $\mathbb{1}_{B(x,y),\varepsilon)}$[8]. We construct a sequence $h_n \in C(X \times Y)$ of product form such that such that $0 \leq h_n \leq 1$ and $h_n|_{B((x,y),\varepsilon)} = 1$ and $h_n = 0$ outside of $B((x,y),\varepsilon + 1/n)$. Take for instance $h_n(x,y) = h_n^1(x)h_n^2(y)$ with $h_n^1(x) = (1 - n\operatorname{dist}_X(x,A))_+$ and $h_n^2(y) = (1 - n\operatorname{dist}_X(x,B))_+$, where $(\cdot)_+ = \max\cdot, 0$. Then by dominated convergence we have that (2.4) and (2.5) hold for $\mathbb{1}_{B((x,y),\varepsilon})$ and hence for all measurable $h : X \times Y \to [0,\infty)$

$$\int_{X\times Y} h(x,y)\,\mathrm{d}\lambda(x,y) = \int_Y \int_X f(x,y)\,\mathrm{d}\mu_y(x)\,\mathrm{d}\nu(y).$$

The support statement (2.3) follows from

$$0 = \lambda(\Pi^c) = \int_Y \int_X \mathbb{1}_{\Gamma^c}\,\mathrm{d}\mu_y(x)\,\mathrm{d}\nu(y) = \int_y \mu_y\left(\{x \in X : \pi(x) \neq y\}\right)\mathrm{d}\nu(y)$$

and hence $\mu_y\left(\{x \in X : \pi(x) \neq y\}\right) = 0$ for ν-a.e. $y \in Y$.

Let $\tilde{\mu}_y$ be another disintegration. Then, it holds for any $f \in \mathcal{C}_0$

$$\int_X f(x)\,\mathrm{d}\mu_y(x) = \int_X f(x)\,\mathrm{d}\tilde{\mu}_y(x) \qquad \text{for } \nu\text{-a.e. } y \in Y.$$

[8] wrt. the product metric

Since $\mathcal{C}_0$ is countable, we find that the above identity holds for a conull set $Y_0 \subset Y$. Since $\mathcal{C}_0$ is dense in $C(X)$, we get $\mu_y = \tilde{\mu}_y$ for all $y \in Y_0$.

Let us finish the proof and sketch how the case of X non-compact can be reduced to the compact case. By the inner regularity property (2.2) of μ and separability of X follows that there exists a countable family of compact sets $\{K_i\}_{i\in\mathbb{N}}$ and a μ-null set N such that

$$X = \bigcup_{i\in\mathbb{N}} K_i \cup N.$$

We obtain a unique disintegration of μ for each K_i and add them all together to get the unique one for X. □

03.11.15

Corollary 2.10. *Let $f \in L^1(X, \mathcal{B}_X, \mu)$ be an integrable random variable and let $(Y, \mathcal{B}_Y)$ be given. For $\pi : X \to Y$ a Borel map, let $\Sigma = \sigma(\pi) \subset \mathcal{B}_X$. Then, it holds for $\pi_\sharp\mu$-a.e. $y \in Y$*

$$\int_{\pi^{-1}(Y)} |f(x)| \,\mathrm{d}\mu_y(\mathrm{d}x) < \infty \quad \textit{and} \quad \int_{\pi^{-1}(Y)} f(x) \,\mathrm{d}\mu_y(\mathrm{d}x) = E\left[f \mid \Sigma\right](y).$$

The family $\{\mu_y\}_{y\in Y}$ is called regular conditional probability function*.*

2.2. **The path space $C([0,\infty);\mathbb{R}^d)$.** Let us make the setting more concrete. Let for $T > 0$, $\Gamma_T := C([0,T];\mathbb{R}^d)$ be the space of continuous trajectories and for $\gamma \in \Gamma$ let $e_t(\gamma) := \gamma(t)$ be the evaluation map. The distance is defined by $D_T : \Gamma_T \times \Gamma_T \to \mathbb{R}_+$

$$D_T(\gamma, \gamma') = \sup_{t\in[0,T]} |\gamma(t) - \gamma'(t)| \,.$$

We want to include the case $T = \infty$ and set $\Gamma := C([0,\infty), \mathbb{R}^d)$. The distance $D : \Gamma \times \Gamma \to [0,1]$ in this case is

$$D(\gamma, \gamma') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0\le t\le n} |e_t(\gamma) - e_t(\gamma')|}{1 + \sup_{0\le t\le n} |e_t(\gamma) - e_t(\gamma')|}.$$

Therewith (Γ_T, D_T) for all $T > 0$ including ∞ is a *Polish space* (complete separable metric space)[9].

Definition 2.11. The Borel σ-field of (Γ, D), denoted by $\mathcal{F}$ and called *canonical filtration*, is given by

$$\mathcal{F} = \sigma\left(\{e_t : t \ge 0\}\right) = \sigma\left(\bigcup_{t\ge 0} \bigcup_{U\subset\mathbb{R}^d \text{ Borel}} \{\gamma \in \Gamma : \gamma(t) \in U\}\right).$$

Proof. $\supseteq$ The map $\Gamma \ni \gamma \mapsto e_t(\gamma) \in \mathbb{R}^d$ is D-continuous and in particular measurable for $t \ge 0$.

[9] The convergence induced by D is uniform convergence on bounded time intervals, hence Cauchy-sequences wrt. D are continuous.

$\subseteq$ Fix $\gamma^0 \in \Gamma$, $t \geq 0$ and $\varepsilon > 0$, then by continuity follows

$$\left\{\gamma : \sup_{0\leq s\leq t} \left|e_s(\gamma) - e_s(\gamma^0)\right| < \varepsilon\right\} = \bigcup_{n=1}^{\infty} \left\{\gamma : \left|e_s(\gamma) - e_s(\gamma^0)\right| \leq \varepsilon \left(1 - \frac{1}{n}\right), \forall s \in [0,t] \cap \mathbb{Q}\right\}.$$

Hence the above set is element of $\sigma\left(\{e_t : t \geq 0\}\right)$. Moreover, these sets generate the topology of Γ.

□

Moreover, define for $t \geq 0$

$$\mathcal{F}_t = \sigma\left(\{e_s : 0 \leq s \leq t\}\right).$$

Clearly $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$ and for $t > 0$

$$\mathcal{F}_t = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right) \quad \text{as well as} \quad \mathcal{F} = \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right).$$

Prokhorov's theorem now allows to study relative compactness in $\mathcal{P}(\Gamma_T)$ by substituting tightness with an equivalent characterization due to the Arzela-Ascoli theorem (equiboundedness and equicontinuity).

Theorem 2.12 (Relative compactness)**.** *A family of probability measures* $\Pi \subset (\Gamma, \mathcal{F})$ *is relative compact if and only if*

$$\lim_{A\uparrow\infty} \inf_{P\in\Pi} P\left(|e_0| \leq A\right) = 1 \tag{2.6}$$

and for each $\rho > 0$ *and* $T < \infty$

$$\lim_{\delta\downarrow 0} \inf_{P\in\Gamma} P\left(\sup_{\substack{0\leq s\leq t\leq T\\ t-s\leq\delta}} |e_t - e_s| \leq \rho\right) = 1. \tag{2.7}$$

Proof. By the Arzela-Ascoli theorem (cf. [2, Theorem 7.2]) a set $K \subset \Gamma$ is relative compact if and only if

$$\sup_{\gamma\in K} |e_0(\gamma)| < \infty$$

and for each $T < \infty$

$$\lim_{\delta\to 0} \sup_{\gamma\in K} \sup_{\substack{0\leq s\leq t\leq T\\ t-s\leq\delta}} |e_s(\gamma) - e_t(\gamma)| = 0$$

Hence from the first part of Theorem 2.7 follows the necessity of (2.6) and (2.7).

For the sufficiency, let us consider a seqence $\rho = 1/n$ and $T = n$. Then we find for any $\varepsilon > 0$ by (2.7) $\delta_n(\varepsilon)$ such that

$$\inf_{P\in\mathcal{P}} P\left(\sup_{\substack{0\le s\le t\le n\\ t-s\le\delta_n(\varepsilon)}} |e_t - e_s| \le \frac{1}{n}\right) = 1 - \frac{\varepsilon}{2^{n+1}}$$

and from (2.6) we find $A(\varepsilon)$ such that

$$\inf_{P\in\mathcal{P}} (|e_0| \le A) \ge 1 - \frac{\varepsilon}{2}.$$

Then, we define the event

$$K_\varepsilon = \bigcap_{n\ge 1} \left\{\gamma : \sup_{\substack{0\le s\le t\le n\\ t-s\le\delta_n(\varepsilon)}} |e_t(\gamma) - e_s(\gamma)| \le \frac{1}{n}\right\} \cap \{\gamma : |e_0(\gamma)| \le A\}.$$

By construction, we obtain $P(K_\varepsilon) \ge 1-\varepsilon$ for any $P \in \Pi$. By another application of Arzela-Ascoli follows that K_ε is compact and the conclusion follows from the second part of (2.7). □

Proposition 2.13 (Disintegration)**.** *For a probability measure $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$, there are two canonical disintegrations*

i) For $s > 0$ let $\Gamma_T \ni \gamma \mapsto \gamma|_{[0,t]} \in \Gamma_t$ be the restriction map. Then, there exists a family $\{\boldsymbol{\eta}_{s,\gamma|_{[0,s]}}\} \subset \mathcal{P}(\Gamma_T)$ with $\boldsymbol{\eta}_{s,\gamma|_{[0,s]}}(\{\tilde\gamma : \tilde\gamma|_{[0,t]} = \gamma|_{[0,t]}\}) = 1$ such that for any $f \in C_b(\mathbb{R}^d)$ and $t \in [s,T]$ holds

$$E^{\boldsymbol{\eta}}\left[f(\gamma(t)) \mid \mathcal{F}_s\right](\gamma) = \int_{\Gamma_T} f(\gamma(t))\,\mathrm{d}\boldsymbol{\eta}_{s,\gamma|_{[0,s]}}(\gamma).$$

ii) Define for $t \in [0,T]$ the 1-marginals

$$\nu_t := (e_t)_\sharp \boldsymbol{\eta}.$$

There exists a family of measure $\{\boldsymbol{\eta}_{t,x}\}_{t\in[0,T];x\in\mathbb{R}^d} \subset \mathcal{P}(\Gamma_T)$ with $\boldsymbol{\eta}(\{\gamma : \gamma(t) = x\}) = 1$ such that for all $f \in C_b(\mathbb{R}^d)$ and any $s,t \in [0,T]$ holds

$$\int_{\Gamma_T} f(\gamma(s))\,\mathrm{d}\boldsymbol{\eta}(\gamma) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}\nu_s(x) = \iint_{\mathbb{R}^d\times\Gamma_T} f(\gamma(s))\,\mathrm{d}\boldsymbol{\eta}_{t,x}(\gamma)\,\mathrm{d}\nu_t(x).$$

Proof. Exercise: For (i) apply Corollary 2.10 with $X = \Gamma_T$, $Y = \Gamma_t$ and $\pi(\gamma) = \gamma|_{[0,t]}$. For (ii) apply Theorem 2.9 with $X = \Gamma_T$, $Y = \mathbb{R}^d$ and $\pi = e_t$. □

Remark 2.14. In general, there is no reason for $\boldsymbol{\eta}_{t,\gamma|_{[0,t]}} = \boldsymbol{\eta}_{t,\gamma(t)}$ (with $\boldsymbol{\eta}_{t,\gamma(t)}$ from Proposition 2.13). However, for Markov processes this identity holds in a certain sense.

2.3. **Extension and regularity of measures.**

Theorem 2.15 (Kolmogorov's extension theorem [12, Theorem 2.1.5])**.** *A family*

$$\left\{P_{t_1,\dots,t_n} \in \mathcal{P}\left((\mathbb{R}^d)^n\right) : n \geq 1, 0 \leq t_1 < \cdot < t_n\right\}$$

is called consistent *if*

i) for any permutation π of $\{1,\dots,n\}$ and sets $F_i \subset \mathbb{R}^d$, $i = 1,\dots n$ holds

$$P_{t_{\pi(1)},\dots,t_{\pi(n)}}\left(F_1 \times \cdots \times F_n\right) = P_{t_1,\dots,t_n}\left(F_{\pi(1)} \times \cdots \times F_{\pi(n)}\right);$$

ii) and for any $m \geq 1$

$$P_{t_1,\dots,t_n}\left(F_1 \times \cdots \times F_n\right) = P_{t_1,\dots,t_n,t_n+1,\dots,t_{n+m}}\left(F_1 \times \cdots \times F_n \times \mathbb{R}^d \times \cdots \times \mathbb{R}^d\right).$$

Given a consistent family, then there exists a unique probability measure $P \in \mathcal{P}\left((\mathbb{R}^d)^{[0,\infty]}\right)$[10] *such that its finite dimensional distributions are given by $\{P_{t_1,\dots,t_n}\}$, i.e.*

$$P_{t_1,\dots,t_n}\left(F_1 \times \cdots \times F_n\right) = P\left(\gamma(t_1) \in F_1, \dots, \gamma(t_n) \in F_n\right).$$

The above theorem is an easy and elegant construction for collections $\{\gamma(t)\}_{t\geq 0}$ of random variables, i.e. *stochastic processes.* It establishes an isomorphism between consistent families and measures on a function space. However, this function space is too large and the class of measurable subsets is too small, since the product Borel σ-algebra is the smallest one making all coordinate projections $t \mapsto \gamma(t)$ measurable.

Theorem 2.16 (Kolmogorov continuity theorem)**.** *Let $\{P_{t_1},\dots,P_{t_n}\}$ be a consistent family of finite dimensional distributions. If for each $T > 0$ there exists numbers $\alpha = \alpha_T > 0$, $r = r_T \geq 1 + \alpha_T$ and $C_T < \infty$ such that*

$$\int\int |x-y|^r \, P_{s,t}(\mathrm{d}x \times \mathrm{d}y) \leq C_T \, |t-s|^{1+\alpha}, \qquad 0 \leq s < t \leq T, \tag{2.8}$$

then there exists a unique $P \in \mathcal{P}(\Gamma)$ such that

$$P(\gamma(t_1) \in F_1, \dots, \gamma(t_n) \in F_n) = P_{t_1,\dots,t_n}(F_1 \times \cdots \times F_n)$$

for all $0 \leq t_1 < \cdots < t_n$ and $F_1,\dots,F_n \in \mathcal{B}_{\mathbb{R}^d}$. Moreover, it holds

$$\operatorname{supp} P \subseteq C^{0,\beta}_{\mathrm{loc}}([0,\infty)) \qquad \text{with } \beta \in \left[0, \tfrac{\alpha}{r}\right),$$

where $C^{0,\beta}_{\mathrm{loc}}([0,\infty))$ is the space of β-Hölder continuous functions on any compact $K \subset [0,\infty)$.

Proof (Sketch). The ingredients are the following results:

(1) "Restriction" Lemma [14, Lemmata 2.1.1 and 2.1.2]: A probability measure P on $\mathbb{R}^{[0,\infty]}$ can be restricted to a probability measure on Γ if for any $T > 0$ and any bounded countable set $S \subset [0,T)$ holds

$$P\left(\left\{\gamma \in (\mathbb{R}^d)^{[0,\infty)} : \gamma|_S \text{ is uniformly continuous}\right\}\right) = 1.$$

[10]Here, the product Borel σ-algebra is used.

(2) "Calculus" Lemma [14, Theorem 2.1.3]: Let $p, \Psi : C([0,\infty],\mathbb{R})$ be strictly increasing with $p(0) = \Psi(0) = 0$ and $\lim_{t\to\infty} \Psi(t) = \infty$. Given $T > 0$ and $\gamma \in \Gamma_T$, if

$$\int_0^T \int_0^T \Psi\left(\frac{|\gamma(t)-\gamma(s)|}{p(|t-s|)}\right) \mathrm{d}s\,\mathrm{d}t \leq B,$$

then for $0 \leq s < t \leq T$:

$$|\varphi(t) - \varphi(s)| \leq 8 \int_0^{t-s} \Psi^{-1}\left(\frac{4B}{u^2}\right) p(\mathrm{d}u).$$

For the proofs of the above statements see [14, Lemma 2.1.1, Lemma 2.1.2, Theorem 2.1.3].

Therewith, we can proof the theorem. We get a $P \in \mathcal{P}(\mathbb{R}^{[0,\infty)})$ from Kolmogorov's extension Theorem 2.15, whose pushforward $(e_s \times e_t)_\sharp P$ also satisfies (2.8). It is left to deduce with the above ingredients, that P is supported on Γ.

Therefore, let $\kappa \in (2, 2+\alpha)$, define $\beta = (\kappa-2)/r > 0$ and for $\lambda > 0$ the event

$$F_\lambda = \left\{\gamma \in \Gamma_T : \sup_{0\leq s\leq t\leq T} \frac{|\gamma(t)-\gamma(s)|}{|t-s|^\beta} \geq \frac{8\kappa}{\kappa-2}(4\lambda)^{\frac{1}{r}}\right\}.$$

Then, we obviously can estimate

$$\begin{aligned} &P\left(\left\{\gamma \in (\mathbb{R}^d)^{[0,\infty)} : \gamma|_S \text{ is uniformly continuous}\right\}\right) \\ &\geq P\left(\left\{\gamma \in \Gamma_T : \gamma \in C^{0,\beta}([0,T])\right\}\right) = 1 - \lim_{\lambda\to\infty} P(F_\lambda). \end{aligned} \tag{2.9}$$

From now on, we restrict P to Γ and denote it by $\tilde{P}$, which still satisfies (2.8) and after taking expectations

$$E^{\tilde{P}}\left[\int_0^T \int_0^T \left(\frac{|\gamma(t)-\gamma(s)|}{|t-s|^{\frac{\gamma}{r}}}\right)^r \mathrm{d}s\,\mathrm{d}t\right] \leq C_T \int_0^T \int_0^T |t-s|^{1+\alpha-\kappa}\,\mathrm{d}s\,\mathrm{d}t =: C_T A_T.$$

Observe, that $A_T < \infty$ due to $\kappa > 2$. Now, we introduce the event

$$\tilde{F}_\lambda := \left\{\gamma \in \Gamma_T : \int_0^T \int_0^T \left(\frac{|\gamma(t)-\gamma(s)|}{|t-s|^{\frac{\gamma}{r}}}\right)^r \mathrm{d}s\,\mathrm{d}t \geq \lambda\right\}.$$

Then, by Markov's inequality

$$\tilde{P}\left(\tilde{F}_\lambda\right) \leq \frac{C_T A_T}{\lambda}. \tag{2.10}$$

On the other side, for any $\gamma \notin \tilde{F}_\lambda$ follows by the calculus lemma the estimate

$$|\gamma(t)-\gamma(s)| < 8 \int_0^{|t-s|} \left(\frac{4\lambda}{u^2}\right)^{\frac{1}{r}} \mathrm{d}u^{\frac{\kappa}{r}} = \frac{8\kappa}{r} \int_0^{|t-s|} u^{\frac{\gamma}{r}-\frac{2}{r}-1} = \frac{8\kappa}{\kappa-2}(4\lambda)^{\frac{1}{r}} |t-s|^\beta.$$

we deduce that $\gamma \notin F_\lambda$, hence $F_\lambda \subseteq \tilde{F}_\lambda$. From here, we conclude from (2.10), that

$$P(F_\lambda) = \tilde{P}(F_\lambda) \leq \tilde{P}\left(\tilde{F}_\lambda\right) \leq \frac{C_T A_T}{\lambda}.$$

The estimate together with (2.9) shows that the assumption of the restriction lemma is satisfied and P has a restriction to Γ_T for any $T > 0$. The statement about the β-Hölder continuity follows from the second identity of (2.9). □

2.4. **Markov processes and transition probabilities.** For Markov processes the consistent families $\{P_{t_1,\dots,t_n}\}$ arise in the following way.

Definition 2.17 (Transition probability function)**.** A function $P(s,x;t,B)$ with $0 \le s < t$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}_{\mathbb{R}^d}$ satisfying

- (i) $P(s,x;t,\cdot) \in \mathcal{P}(\mathbb{R}^d)$ for all $0 \le s < t$ and $x \in \mathbb{R}^d$
- (ii) $P(s,\cdot;t,B)$ is $\mathcal{B}_{\mathbb{R}^d}$ measurable for all $0 \le s < t$ and $B \in \mathcal{B}_{\mathbb{R}^d}$
- (iii) if $0 \le s < t < u$ and $B \in \mathcal{B}_{\mathbb{R}^d}$, then
$$P(s,x;u,B) = \int P(t,y;u,B)\, P(s,x;t,\mathrm{d}y) \quad \textit{(Chapman-Kolmogorov equation)}.$$

is a *transition probability function TPF* on $\mathbb{R}^d$.

Definition 2.18 (Markov process)**.** Let $P(s,x;t,\cdot)$ be a TPF and $\mu \in \mathcal{P}(\mathbb{R}^d)$. A probability measure $\boldsymbol{\eta} \in \mathcal{P}((\mathbb{R}^d)^{(0,\infty)})$ is called the *Markov process with TPF $P(s,x;t,\cdot)$ and initial distribution μ*, denoted by $\mathrm{Markov}(P,\mu)$, if
$$\boldsymbol{\eta}(\gamma(0) \in B) = \mu(B), \quad B \in \mathcal{B}_{\mathbb{R}^d};$$
and for all $0 \le s < t$ and $B \in \mathcal{B}_{\mathbb{R}^d}$
$$\boldsymbol{\eta}\left(\gamma(t) \in B \mid \mathcal{F}_s\right) = P(s,\gamma(s);t,B). \tag{2.11}$$
If $P(s,x;t,\cdot) = P(t-s,x;\cdot)$, then the Markov process is called *time-homogeneous.*

Remark 2.19. The condition (2.11) can be rephrased (cf. Proposition 2.13 and Remark 2.14) in terms of disintegration as follows: For all $0 \le s < t$, $x \in \mathbb{R}^d$ and all $f \in C_b(\mathbb{R}^d)$ holds
$$\int_{\Gamma_T} f(\gamma(t))\, \mathrm{d}\boldsymbol{\eta}_{s,\gamma|_{[0,s]}}(\gamma) = \int_{\Gamma_T} f(\gamma(t))\, \mathrm{d}\boldsymbol{\eta}_{s,\gamma(s)}(\gamma)$$
and for all $B \in \mathcal{B}_{\mathbb{R}^d}$ holds
$$(e_t)_\sharp \boldsymbol{\eta}_{s,x}(B) = P(s,x;t,B).$$

09.11.15

Theorem 2.20. *Let $P(s,x;t,\cdot)$ be a TPF and $\mu \in \mathcal{P}(\mathbb{R}^d)$. Define $P_0(B) := \mu(B)$; $P_t(B) := \int P(0,y;t,B) P_0(\mathrm{d}y)$ and for $0 \le t_1 < \dots < t_{n+1}$ recursively with $\Delta \in \mathcal{B}_{(\mathbb{R}^d)^{n+1}}$*
$$P_{t_1,\dots,t_n,t_{n+1}}(\Delta) := \int \cdots \int_\Delta P(t_n,y_n;t_{n+1},dy_{n+1}) P_{t_1,\cdots,t_n}(\mathrm{d}y_1 \times \dots \times \mathrm{d}y_n). \tag{2.12}$$
Then $\{P_{t_1,\dots,t_n}\}$ is consistent.

Moreover any $\boldsymbol{\eta} \in \mathcal{P}((\mathbb{R}^d)^{(0,\infty)})$ is in $\mathrm{Markov}(P,\mu)$ *if and only if $\boldsymbol{\eta}$ has $\{P_{t_1,\dots,t_n}\}$ as finite dimensional distributions. In particular any TPF and initial distribution generates one and only one Markov process*

Proof. The proof of the consistency is straight-forward from the construction (2.12) by using the Chapman-Kolmogorov equation (2.12). Then, by Kolmogorov's extension Theorem 2.15 we obtain a measure $\boldsymbol{\eta} \in \mathcal{P}((\mathbb{R}^d)^{[0,\infty)})$ with finite dimensional distributions given by (2.12). Is is left to verify the Markov property (2.11). Let $t > 0$ and $B_0, B_1 \in \mathcal{B}_{\mathbb{R}^d}$, then

$$(e_0 \times e_t)_\sharp \boldsymbol{\eta} \left(B_0 \times B_1\right) = \int_{B_0} P(0, y; t, B_1)\ P_0(\mathrm{d}y) = E^{\boldsymbol{\eta}} \left[P(0, \gamma(0); t, B_1), \gamma(0) \in B_0\right].$$

Hence, we obtain $\boldsymbol{\eta} \left(\gamma(t) \in B_1 \mid \sigma(e_0)\right) = P(0, \gamma(0); t, B_1)$. Likewise, for $0 \leq s < t$ let $0 \leq t_1 < \cdots < t_n = s$ and $B, B_1, \ldots, B_n \in \mathcal{B}_{\mathbb{R}^d}$ be given. Then

$$\begin{aligned}
&(e_{t_1} \times \cdots \times e_{t_n} \times e_t)_\sharp \boldsymbol{\eta}(B_1 \times \cdots \times B_n \times B) \\
&= \int \cdots \int_{B_1 \times \cdots \times B_n} P(s, y_n; t, B)\ P_{t_1,\ldots,t_n}(\mathrm{d}y_1, \ldots, \mathrm{d}y_n) \\
&= E^{\boldsymbol{\eta}} \left[P(s, \gamma(s); t, B), \gamma(t_1) \in B_1, \ldots, \gamma(t_n) \in B_n\right].
\end{aligned}$$

Now, it is easy to verify that filtration $\mathcal{F}_s$ is generated by $\{e_t : t \in S \subset [0, s], S \text{ countable}\}$ (cp. Definition 2.11).

For the uniqueness, we have to identify the finite dimensional distributions. Let $\boldsymbol{\eta} \in \mathrm{Markov}(P, \mu)$. Then, the initial distributions are equal, i.e. $(e_0)_\sharp \boldsymbol{\eta} = P_0 = \mu$. Moreover, by the Markov property follows $(e_t)_\sharp \boldsymbol{\eta}(B) = \int P(0, y; t, B) \mu(\mathrm{d}y) = P_t(B)$. Now, we can conclude by induction again using the Markov property with $s = t_n$

$$\begin{aligned}
&(e_{t_1} \times \cdots \times e_{t_n} \times e_t)_\sharp \boldsymbol{\eta}(B_1 \times \cdots \times B_n \times B) \\
&= \int \cdots \int_{B_1 \times \cdots \times B_n} P(t_n, y_n; t; B)\ (e_{t_1} \times \cdots \times e_{t_n})_\sharp \boldsymbol{\eta}(\mathrm{d}y_1, \ldots, \mathrm{d}y_n) \\
&= \int \cdots \int_{B_1 \times \cdots \times B_n} P(t_n, y_n; t; B)\ P_{t_1,\ldots,t_n}(\mathrm{d}y_1, \ldots, \mathrm{d}y_n) \\
&= P_{t_1,\ldots,t_n,t}(B_1 \times \cdots \times B_n \times B).
\end{aligned}$$

□

Now, we can combine the above Theorem with the characterization of Theorem 2.16, to give a condition under which $\boldsymbol{\eta}$ is supported on continuous trajectories, i.e. $\boldsymbol{\eta}$ is a *continuous Markov process.*

Corollary 2.21. *Let $P(s, x; t, \cdot)$ be a TPF such that for any $T > 0$ exists $\alpha = \alpha_T > 0$, $r = r_T \geq 1 + \alpha_T$ and $C = C_T$ such that for $0 \leq t_1 < t_2 \leq T$ holds*

$$\sup_{y_1 \in \mathbb{R}^d} \int |y - y_1|^r\, P(t_1, y_1; t_2, \mathrm{d}y) \leq C \left|t_2 - t_1\right|^{1+\alpha}.$$

Then for each $\mu \in \mathcal{P}(\mathbb{R}^d)$ exists a unique $\boldsymbol{\eta} \in \mathcal{P}(\Gamma)$ such that $\boldsymbol{\eta} \in \mathrm{Markov}(P, \mu)$. Moreover, it holds

$$\operatorname{supp} \boldsymbol{\eta} \subseteq C^{0,\beta}_{\mathrm{loc}}([0, \infty)) \qquad \textit{for } \beta \in \left(0, \tfrac{\alpha}{r}\right).$$

Proof. Combine Theorem 2.20 and 2.16. □

Example 2.22 (Wiener measure)**.** Define the following TPF

$$P_{\mathcal{W}}(s,x;t,B) := \int_B g_d(t-s,y-x)\,\mathrm{d}y \quad \text{with} \quad g_d(s,x) := \frac{1}{(2\pi s)^{d/2}} \exp\left(-\frac{|x|^2}{2s}\right).$$

Then the resulting measure from Corollary 2.21 is called *Wiener measure*. Especially by choosing as initial distribution $\mu = \delta_x$ for $x \in \mathbb{R}^d$, the resulting d-dimensional Wiener measure is denoted by $\mathcal{W}_x^{(d)}$ and satisfies for all $0 \le t_1 < t_2$

$$\mathcal{W}_x^{(d)}\left(\gamma(t_2) \in B \mid \mathcal{F}_{t_1}\right) = \int_B g_d(t_2 - t_1, y - \gamma(t_1))\,\mathrm{d}y.$$

Moreover $\mathcal{W}^{(d)} := \mathcal{W}_0^{(d)}$.

Exercise: Deduce that for any $r > 1$ exists $C_r < \infty$ such that

$$\int |y - y_1|^{2r} P_{\mathcal{W}}(t_1, y_1; t_2, \mathrm{d}y) \le C_r |t_2 - t_1|^r$$

and conclude that $\operatorname{supp} \mathcal{W}^{(d)} \subseteq C^{0,\beta}_{\mathrm{loc}}([0,\infty))$ for any $\beta \in (0, 1/2)$.

2.5. Maximum principle and C_0-semigroups.

Theorem 2.23 (Weak parabolic maximum principle)**.** *Let $a \in C_b([0,\infty) \times \mathbb{R}^d; S_d)$ and $b \in C_b([0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$ and set*

$$L_t = \tfrac{1}{2} a : \nabla^2 + b \cdot \nabla = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t,x) \partial_{x_i,x_j} + \sum_{i=1}^d b_i(t,x) \partial_i. \tag{2.13}$$

If $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$ bdd. from below satisfies:

$$s \in [0,T) : \partial_s f + L_s f \le 0 \qquad \text{and} \qquad \liminf_{s \to T} f(s,\cdot) \ge 0, \tag{2.14}$$

then f is non-negative. Moreover if f is in addition bounded and satisfies for some $c, g \in C_b([0,T])$ and all $s \in [0,T)$

$$\partial_s f + L_s f + c(s) f \ge -g(s), \tag{2.15}$$

then

$$f(s,x) \le \|f(T,\cdot)\| \exp\left(\int_s^T c(u)\,\mathrm{d}u\right) + \int_s^T g(t) \exp\left(\int_s^t c(u)\,\mathrm{d}u\right) \mathrm{d}t. \tag{2.16}$$

Proof. First assume that f satisfies

$$\partial_s f + L_s f < 0 \quad \text{for} \quad 0 \le s < T. \tag{2.17}$$

Then if $(s_0, x_0) \in [0,T) \times \mathbb{R}^d$ is a point such that

$$f(s_0,x_0) \le f(s,x) \quad \text{for all} \quad (s,x) \in [s_0,T) \times \mathbb{R}^d, \tag{2.18}$$

then it would hold

$$\partial_s f(s_0,x_0) \ge 0, \quad \nabla f(s_0,x_0) = 0 \quad \text{and} \quad \nabla^2 f(s_0,x_0) \ge 0.$$

Hence, in particular

$$(\partial_s + L_{s_0})f(s_0,x_0) = \partial_s f(s_0,x_0) + \tfrac{1}{2}a(s_0,x_0) : \nabla^2 f(s_0,x_0) + b(s_0,x_0)\cdot\nabla f(s_0,x_0) \geq 0.$$

Here, we used that for two matrices $A, B \geq 0$ also $A : B \geq 0$. This is a contradiction and hence there exists no point (s_0, x_0) satisfying (2.18).

Now, we define for $\delta > 0$ and $\varepsilon > 0$ the function

$$f_{\delta,\varepsilon}(s,x) = f(s,x) + \varepsilon(T-s) + \delta e^{-s}\,|x|^2\,.$$

Then, we get the estimate

$$(\partial_s + L_s)\, f_{\delta,\varepsilon} = -\varepsilon + \delta e^{-s}\left(\operatorname{tr}(a(s,x)) + 2x\cdot b(s,x) - |x|^2\right) \leq -\varepsilon + \delta\left(\|\operatorname{tr}(a)\| + \|b\|^2\right).$$

Hence, we can choose for each $\varepsilon > 0 : \delta(\varepsilon) = \varepsilon/2\left(\|\operatorname{tr}(a)\| + \|b\|^2\right)$ such that $(\partial_s + L_s)\, f_{\delta,\varepsilon} \leq -\varepsilon/2 < 0$. Now, suppose $f_{\delta,\varepsilon}(s,x) < 0$ for some $(s,x) \in [0,T)\times\mathbb{R}^d$. Since $\lim_{s\to T} f_{\delta,\varepsilon}(s,x) \geq 0$ and $f_{\delta,\varepsilon}(s,x) \to \infty$ as $|x| \to \infty$ for all $s \in [0,T]$, there has to exist a $(s_0,x) \in [0,T)\times\mathbb{R}^d$ such that (2.18) holds. This contradicts (2.17). Hence, for all $\varepsilon > 0$ and $0 \leq \delta \leq \delta(\varepsilon)$ holds $f_{\delta,\varepsilon} \geq 0$ and therefore also $f \geq 0$ in $[0,T)\times\mathbb{R}^d$.

Finally, f satisfies (2.15), then the following function

$$(s,x) \mapsto \|f(T,\cdot)\| - f(s,x)\exp\left(\int_s^T c(u)\,\mathrm{d}u\right) + \int_s^T g(t)\exp\left(-\int_t^T c(u)\,\mathrm{d}u\right)\mathrm{d}t$$

satisfies (2.14) and hence is non-negative, which translates to the estimate (2.16). □

Remark 2.24. The definition of L_t seems to be ad hock. However, one can proof the following theorem:

Let $L : C^\infty(\mathbb{R}^d) \to C(\mathbb{R}^d)$ linear satisfying

(i) L is local: If $\varphi(x) \equiv 0$ in $B_\varepsilon(x_0)$ for some $\varepsilon > 0$ then $L\varphi(x_0) = 0$.
(ii) L satisfies the weak maximum principle: If $\varphi \in C^\infty(\mathbb{R}^d)$ has a local maximum at x_0 then $L\varphi(x_0) \leq 0$.

Then there exists $a : \mathbb{R}^d \to S_d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ continuous such that $L = \frac{1}{2}a : \nabla^2 + b\cdot\nabla$.

Exercise: Proof the above theorem by the following steps:

1) For any $c \in \mathbb{R}$ follows $Lc \equiv 0$.
2) For any $\varphi \in C^2(\mathbb{R}^d)$ with $\varphi(x) = o(|x - x^0|)$ for some $x^0 \in \mathbb{R}^d$ follows $L\varphi(x^0) = 0$ (Consider the function $\varphi_\varepsilon(x) = \varphi(x) + \varepsilon\,|x - x^0|^2$).
3) Define $b_j(x) = (L\varphi_j)(x)$ and $a^{ij}(x) = (L\varphi_i\varphi_j)(x) - b_j(x)\varphi_i(x) - b_i(x)\varphi_j(x)$, where $\varphi_i(x) = x_i$. Show that $a^{ij}(x^0)$ is non-negative definite by testing with functions of the form $\sum_i \theta_i\cdot(x_i - x_i^0)^2$ with $\theta \in \mathbb{R}^d$.
4) Use a second order Taylor expansion of $\varphi \in C^\infty(\mathbb{R}^d)$ in x^0 and the definition of $b_i(x)$ and $a^{ij}(x)$ to verify the form of L.

Corollary 2.25 (C_0-semigroup)**.** *Let L_t be given by (2.13). Assume that there exists for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ an $f \in C^{1,2}([0,T] \times \mathbb{R}^d) \cap C_b([0,T] \times \mathbb{R}^d)$ such that for any $t > 0$*

$$s \in [0,t) : \partial_s f + L_s f = 0 \qquad \text{and} \qquad \lim_{s \to t} f(s,x) = \varphi(x). \tag{2.19}$$

Then, f is unique. The family of operators $T_{s,t}\varphi := f(s,\cdot)$ defines a strongly continuous semigroup*:*

(i) $T_{s,t} : C_0(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ *is a* linear non-negative contraction*: $\|T_{s,t}\varphi\| \le \|\varphi\|$;*
(ii) being strongly continuous $\lim_{s\to t} T_{s,t}\varphi(x) = \varphi(x)$ *for all $x \in \mathbb{R}^d$;*
(iii) and satisfying the semigroup *property $T_{t_1,t_3} = T_{t_1,t_2} T_{t_2,t_3}$ for $0 \le t_1 < t_2 < t_3$.*

Moreover, there exists a TPF P such that the semigroup $T_{s,t}$ has the representation

$$T_{s,t}\varphi(x) = \int \varphi(x) P(s,x;t,\mathrm{d}y)$$

and the TPF satisfies for some constant $A > 0$ only depending on a and b

$$\int |y-x|^4 \, P(s,x;t,\mathrm{d}y) \le A e^{t-s} (t-s)^2 , \qquad 0 \le s < t \text{ and } x \in \mathbb{R}^d. \tag{2.20}$$

Proof. The weak maximum principle of Theorem 2.23 implies the estimate

$$\min_{x \in \mathbb{R}^d} \varphi(x) \le f(s,\cdot) \le \max_{x \in \mathbb{R}^d} \varphi(x), \qquad 0 \le s < t,$$

which first of all implies uniqueness, since L_t is a linear operator. Moreover, it proves, that $T_{s,t}$ is linear non-negative contraction on $C_c^\infty(\mathbb{R}^d)$ and by density also on $C_0(\mathbb{R}^d)$. It is left to show that $T_{s,t}\varphi \in C_0(\mathbb{R}^d)$ for $\varphi \in C_0(\mathbb{R}^d)$. It is enough to show the claim for $\varphi \in C_c^\infty(\mathbb{R}^d)$. Therefore, construct for $M > 0$ the function $\varphi_M \in C_c^\infty(\mathbb{R}^d)$ such that

$$\mathbb{1}_{\bar{B}_M}(x) \le \varphi_M(x) \le \mathbb{1}_{\bar{B}_{M+1}}(x).$$

Here $B_R = B_R(0)$ is the euclidean ball around 0. Let $f_M \in (C^{1,2} \cap C_b)([0,T) \times \mathbb{R}^d)$ be the unique solution to $\partial_s f_M + L_s f_M = 0$ and $f_M(t,\cdot) = \varphi_M$. Now, we construct the following comparison function for $x_0 \in \mathbb{R}^d$ and $A > 0$

$$\psi_{x_0}(s,x) := e^{t-s} \left(A(t-s) + |x-x_0|^2 \right). \tag{2.21}$$

Then, it holds

$$\begin{aligned}
\partial_s \psi_{x_0}(s,x) + L_s \psi_{x_0}(s,x) &= -\psi_{x_0}(s,x) - A e^{t-s} + e^{t-s}\, a : \mathrm{Id} + e^{t-s}\, 2b \cdot (x - x_0) \\
&\le e^{t-s} \Big(-A - |x-x_0|^2 + \|\mathrm{tr}\, a\|_{L^\infty([0,T]\times\mathbb{R}^d)} \\
&\qquad + \|b\|^2_{L^\infty([0,T]\times\mathbb{R}^d)} + |x-x_0|^2 \Big) \\
&\le e^{t-s} \left(\|\mathrm{tr}\, a\|_{L^\infty([0,T]\times\mathbb{R}^d)} + \|b\|^2_{L^\infty([0,T]\times\mathbb{R}^d)} - A \right).
\end{aligned}$$

Hence, $\partial_s\psi_{x_0}(s,x) + L_s\psi_{x_0}(s,x) \le 0$ for $A := \|\operatorname{tr} a\|_{L^\infty([0,T]\times\mathbb{R}^d)} + \|b\|^2_{L^\infty([0,T]\times\mathbb{R}^d)}$. Now, by choosing x_0 such that $|x_0| \ge 2(M+1)$, it follows $\frac{|x_0|^2}{4}\varphi_M(\cdot) \le \psi_{x_0}(t,\cdot)$. Then, Theorem 2.23 implies for $0 \le s < t$

$$0 \le f_M(s,x) \le \frac{4}{|x_0|^2}\psi_{x_0}(s,x) \qquad \text{for all } x \in \mathbb{R}^d.$$

In particular, it follows by choosing $x = x_0$ and still $|x_0| \ge 2\,(M+1)$

$$0 \le f_M(s,x_0) \le \frac{A}{\varrho^2}e^{t-s}\,(t-s) \le \frac{4A}{|x_0|^2}e^{t-s}(t-s)$$

and therefore $f_M(s,\cdot) \in C_0(\mathbb{R}^d)$, finishing the proof of the first statement *(i)*. 24.11.15

The continuity statement *(ii)* follows by considering $\varphi \in C_c^\infty(\mathbb{R}^d)$. Let f be the solution to (2.19). Then the function $f_\varphi(s,x) := f(s,x) - \varphi(x)$ satisfies $\partial_s f_\varphi(s,x) + L_s f_\varphi(s,x) = L_s\varphi$. Hence, we get by (2.16)

$$\|f_\varphi(s,\cdot)\| \le \underbrace{\|f_\varphi(t,\cdot\|}_{=0} + \int_s^t \|L_u\varphi(\cdot)\|\,\mathrm{d}u \le C(\varphi,a,b)\;\;(t-s)\,.$$

By density of $C_c^\infty(\mathbb{R}^d)$ in $C_0(\mathbb{R}^d)$ we get *(ii)*.

For the semigroup property *(iii)*, let again $\varphi \in C_c^\infty(\mathbb{R}^d)$ be given and set $f(s,\cdot) = T_{s,t_3}\varphi$. Take a squence $\{\varphi_n\}_{n\in\mathbb{N}}$ such that $\varphi_n \to f(t_2,\cdot)$ uniformly and define $f_n(s,\cdot) = T_{s,t_2}\varphi_n$ for $0 \le s < t_2$. Then $f - f_n \in (C^{1,2}\cap C_b)\,([0,t_2)\times\mathbb{R}^d)$ and $\partial_s(f-f_n) + L_s(f-f_n) = 0$ for $0 \le s < t_2$. Moreover, by *(ii)* follows

$$\lim_{s\to t_2}\left(f(s,x) - f_n(s,x)\right) = f(t_2,x) - \varphi_n(x).$$

Thus an application of the weak maximum principle Theorem 2.23 leads to

$$\sup_{0\le s<t_2}\|f(s,\cdot) - f_n(s,\cdot)\|_\infty \le \|f(t_2,\cdot) - \varphi_n(\cdot)\|_\infty \to 0 \qquad \text{as } n\to\infty.$$

Resubstituting back all definitions leads to

$$T_{t_1,t_3}\varphi = f(t_1,\cdot) = \lim_{n\to\infty} f_n(t_1,\cdot) = \lim_{n\to\infty} T_{t_1,t_2}\varphi_n = T_{t_1,t_2}T_{t_2,t_3}\varphi$$

proofing *(iii)*.

Since for each $0 \le s < t \le T$ and each $x_0 \in \mathbb{R}^d$ the mapping $\varphi \mapsto T_{s,t}\varphi(x_0)$ is a linear non-negative contraction on $C_0(\mathbb{R}^d)$, there exists by the Riesz-representation theorem a regular countably additive non-negative measure $P(s,x_0;t,\cdot)$ on $\mathbb{R}^d$ such that

$$T_{s,t}\varphi(x_0) = \int \varphi(y)P(s,x_0,t,\mathrm{d}y) \qquad \text{and} \qquad P(s,x_0,t,\mathbb{R}^d) \le 1.$$

It is left to show, that $P(s,x_0,t,\mathbb{R}^d) = 1$. Therefore, let us construct for $M > 0$ the function $\varphi_M \in C_c^\infty(\mathbb{R}^d)$ such that $\mathbb{1}_{\bar{B}_M(x_0)}(x) \le \varphi_M(x) \le \mathbb{1}_{\bar{B}_{M+1}(x_0)}(x)$ and let f_M be the according solution to (2.19). Let ψ_{x_0} be as in the beginning of the proof (2.21). Then, it holds

$M^2(1-\varphi_M(\cdot)) \leq \psi_{x_0}(t,\cdot)$ and hence by Theorem 2.23 it follows $M^2(1-f_M(s,\cdot)) \leq \psi_{x_0}(s,\cdot)$ for all $0 \leq s < t$. This allows to estimate

$$P(s,x_0;t,B_{M+1}(x_0)) \geq T_{s,t}\varphi_M(x_0) = f_M(s,x_0) \geq 1 - \frac{\psi_{x_0}(s,x_0)}{M^2} \geq 1 - \frac{Ae^{t-s}(t-s)}{M^2}.$$

Hence, we get $P(s,x_0,t,\mathbb{R}^d) = 1$.

The last statement (2.20) follows by considering the test function

$$\omega_{x_0}(s,x) = Ae^{t-s}(t-s)^2 + Ae^{t-s}(t-s)\,|x-x_0|^2 + e^{t-s}\,|x-x_0|^4.$$

Exercise: *Show, there is a choice A large enough in terms of a, b such that*

$$\partial_s\omega_{x_0}(s,x) + L_s\omega_{x_0}(s,x) \leq 0 \qquad 0 \leq s < t.$$

First deduce the estimate[11] $L_s\,|x-x_0|^4 \leq 6\,\|\operatorname{tr} a\|_{L^\infty}\,|x-x_0|^2 + 4\,\|b\|_{L^\infty}^2\,|x-x_0|^2 + |x-x_0|^4$.

Therewith, we can conclude the proof. Let $\{\varphi_n\}_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^d)$ a sequence converging monotone to $|x-x_0|^4$, then we have by Theorem 2.23 that $T_{s,t}\varphi_n \leq \omega_{x_0}(s,\cdot)$ for all $0 \leq s < t$. Hence, we have

$$\int \varphi_n(y)P(s,x_0;t,\mathrm{d}y) \leq T_{s,t}\varphi(x_0) \leq \gamma(s,x_0) = Ae^{t-s}(t-s)^2.$$

□

2.6. **Existence for parabolic PDEs.** In the Example 2.22, we constructed the TPF $P_{\mathcal{W}}(s,x;t,B)$ from the density $p_{\mathcal{W}}(s,x;t,y) := g_d(t-s,y-x)$. It is easy to verify, that the function

$$u(s,x) = \int p_{\mathcal{W}}(s,x;t,y)u_T(y)\,\mathrm{d}y + \int_s^T \int p(s,x;t,y)c(t,y)\,\mathrm{d}y\,\mathrm{d}t$$

solves for $0 \leq s \leq T$

$$\partial_s u(s,x) + \Delta u(s,x) = -c(s,x) \qquad \text{and} \qquad u(T,x) = u_T(x).$$

For this reason, the density $p_{\mathcal{W}}(s,x;t,y)$ is called parabolic Green function of the heat equation or just *heat kernel.*

In this section, we will find a class of coefficients a, b such that the generator

$$L_t = \tfrac{1}{2}a : \nabla^2 + b\cdot\nabla = \frac{1}{2}\sum_{i,j=1}^d a^{ij}(t,x)\partial_{i,j} + \sum_{i=1}^d b_i(t,x)\partial_i.$$

has a parabolic Green function.

The first result is a existence theorem by [10], which we state and use without proof.

Theorem 2.26 (Existence for elliptic Hölder coefficients [10])**.** *Let $a : [0,\infty)\times\mathbb{R}^d \to S_d$ and $b : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ be bounded such that for $\lambda > 0$ and $0 < \gamma \leq 1$ exists $C < \infty$:*

[11] For any $n \in \mathbb{N}$ holds $\nabla^2\,|x|^n = n\,|x|^{n-2}\left(\operatorname{Id} + (n-2)\frac{x}{|x|}\otimes\frac{x}{|x|}\right)$

(i) ellipticity*:* $\langle \theta, a(s,x)\theta\rangle \geq \alpha\, |\theta|^2$ *for all* $\theta \in \mathbb{R}^d$

(ii) Hölder*:* $\|a(s,x) - a(t,y)\| + |b(s,x) - b(t,y)| \leq C\,(|x-y|^\gamma + |t-s|^\gamma)$*, where* $\|a\| := \sup_{\theta:|\theta|=1} |a\theta|$.

Then the generator

$$L_t = \tfrac{1}{2} a(t,\cdot) : \nabla^2 + b(t,\cdot)\cdot\nabla$$

posses a unique positive parabolic Green function $p(s,x;t,y)$*:*

- *p is jointly continuous for each $t > 0$ for $s \in [0,t)$ and $x, y \in \mathbb{R}^d$;*
- *and for any $u_T \in C_c^\infty(\mathbb{R}^d)$ and $c \in C_c^\infty([0,\infty)\times\mathbb{R}^d)$ the function*

$$u(s,x) = \int p(s,x;t,y) u_T(y)\,\mathrm{d}y + \int_s^T \int p(s,x;t,y) c(t,y)\,\mathrm{d}y\,\mathrm{d}t$$

satisfies for $s \in (0,T)$ and all $x \in \mathbb{R}^d$

$$\partial_s u(s,x) + L_s u(s,x) = -c(s,x) \qquad \textit{and} \qquad u(T,x) = u_T(x). \tag{2.22}$$

Corollary 2.27. *Let L_t satisfy the assumptions of Theorem 2.26 and let $P(s,x;t,\cdot)$ be the associated TPF (guaranteed by Corollary 2.25). Then for any $0 \leq s < t$ and $x \in \mathbb{R}^d$ holds*

$$P(s,x;t,B) = \int_B p(s,x;t,y)\,\mathrm{d}y, \qquad B \in \mathcal{B}_{\mathbb{R}^d}, \tag{2.23}$$

where $p(s,x;t,y)$ is the parabolic Green function of Theorem 2.26. Moreover, for $T > 0$ and any $u \in C_b^{1,2}([0,T)\times\mathbb{R}^d) \cap C_b([0,T]\times\mathbb{R}^d)$ holds

$$\int u(T,y) P(s,x;T,\mathrm{d}y) - u(s,x) = \int_s^T \int \left(\partial_t + L_t\right) u(t,y) P(s,x;t,\mathrm{d}y)\,\mathrm{d}t. \tag{2.24}$$

Proof. The representation (2.23) follows from (2.22) by choosing $c \equiv 0$. Then u solves $\partial_s u + L_s u = 0$ and $u(T,\cdot) = u_T(\cdot)$, hence $u(s,\cdot) = T_{s,T} u_T(\cdot)$, which is unique by Corollary 2.25. The formula (2.24) follows by considering

$$\psi(s,x) = \int u(T,y) P(s,x;T,\mathrm{d}y) - \int_s^T \int \left(\partial_t + L_t\right) u(t,y) P(s,x;t,\mathrm{d}y)\,\mathrm{d}t.$$

Then, $\psi \in C^{1,2}([0,T]\times\mathbb{R}^d)$ and it solves $\partial_s\psi + L_s\psi = \partial_s u + L_s u$ and since $\psi(T,\cdot) = u(T,\cdot)$, we can conclude by the weak maximum principle Theorem 2.23. □

Lemma 2.28 (Off-diagonal–diagonal matrix estimate)**.** *Let $a \in C^2(\mathbb{R}; S_d)$. Assume, that*

$$\Lambda_0 = \sup\left\{ \left|D^2 a^{ij}(x)\right| : 1 \leq i,j, \leq d \quad \textit{and} \quad x \in \mathbb{R} \right\} < \infty.$$

Then for all $1 \leq i,j, \leq d$ and $x \in \mathbb{R}$ holds

$$\left|(Da^{ij})(x)\right| \leq (2\Lambda_0)^{\frac{1}{2}} \left(a^{ii}(x) + a^{jj}(x)\right)^{\frac{1}{2}}.$$

Moreover, for any $s \in \mathbb{R}^{d\times d}$ symmetric holds

$$\left(\operatorname{tr}\left((Da)(x)\, s\right)\right)^2 \leq 4d^2\Lambda_0 \operatorname{tr}\left(s\, a(x)\, s\right). \tag{2.25}$$

Proof. If we have a non-negative function $\varphi \in C^2(\mathbb{R})$ with $\alpha = \sup_{x\in\mathbb{R}} \varphi''(x) < \infty$, then we can estimate by Taylor

$$0 \leq \varphi(x+y) \leq \varphi(x) + \varphi'(x)y + \frac{\alpha}{2}y^2.$$

Hence, the quadratic form $y \mapsto \alpha/2y^2 + \varphi'(x)y + \varphi(x)$ is non-negative for $y \in \mathbb{R}$. Therefore, its discriminant is strictly negative, i.e.

$$(\varphi'(x))^2 - 2\alpha\varphi(x) \leq 0 \qquad \text{hence} \qquad |\varphi'(x)| \leq (2\alpha)^{\frac{1}{2}} \, (\varphi(x))^{\frac{1}{2}} \, .$$

The above estimate can be applied to

$$\varphi_\pm(x) = a^{ii}(x) \pm 2a^{ij}(x) + a^{jj}(x) = \langle e_i \pm e_j, a(x)(e_i \pm e_j)\rangle$$

since $\left|\varphi''_\pm(x)\right| \leq 4\Lambda_0$. Hence, we get $\left|\varphi'_\pm\right| \leq (8\Lambda_0)^{\frac{1}{2}} \, (\varphi_\pm(x))^{\frac{1}{2}}$. Since, we also have that $a^{ij}(x) = \frac{1}{4}\left(\varphi_+(x) - \varphi_-(x)\right)$, it holds

$$\begin{aligned}\left|(Da^{ij})(x)\right| &\leq \tfrac{1}{4}\left(\left|\varphi'_+(x)\right| + \left|\varphi'_-(x)\right|\right) \leq (\Lambda_0/2)^{\frac{1}{2}} \left((\varphi_+(x))^{\frac{1}{2}} + (\varphi_-(x))^{\frac{1}{2}}\right) \\ &\leq (\Lambda_0)^{\frac{1}{2}} \, (\varphi_+(x) + \varphi_-(x))^{\frac{1}{2}} = (2\Lambda_0)^{\frac{1}{2}} \left(a^{ii}(x) + a^{jj}(x)\right)^{\frac{1}{2}} .\end{aligned}$$

To proof (2.25), we apply Cauchy-Schwarz to the rhs and then the just obtained estimate

$$\begin{aligned}(\operatorname{tr}(Da(x)s)) &\leq d^2 \sum_{i,j=1}^d \left((Da^{ij})(x)\right)^2 \left(s^{ij}\right)^2 \leq d^2 2\Lambda_0 \sum_{i,j=1}^d \left(a^{ii}(x) + a^{jj}(x)\right)^2 \left(s^{ij}\right)^2 \\ &= 4d^2\Lambda_0 \sum_{i,j=1}^d s^{ij}a^{jj}(x)s^{ji} = 4d^2\Lambda_0 \operatorname{tr}(s\, a(x)\, s) \, .\end{aligned}$$

□

Notation 2.29 (Derivatives)*.* We use the following notation $\varphi_{,i} = \partial_{x_i}\varphi$ as well as $\varphi_{,ij} = \partial_{x_i x_j}\varphi$. Moreover, $\operatorname{Hess}\varphi = (\varphi_{,ij})_{i,j=1}^d$ is the Hessian. For a multiindex $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, we set $|\alpha| = \alpha_1 + \dots + \alpha_d$ and define

$$\varphi^{(\alpha)} = D^\alpha\varphi = \frac{\partial^{|\alpha|}\varphi}{\partial_{x_1}^{\alpha_1}\cdots\partial_{x_d}^{\alpha_d}}.$$

Moreover, we define the $C^{0,n}(\mathbb{R}^d)$-norm by

$$\|\varphi(s,\cdot)\|^{(n)} = \sum_{|\alpha|\leq n} \left\|\varphi^{(\alpha)}(s,\cdot)\right\| .$$

Theorem 2.30 (Higher regularity)**.** *Let $a : [0,\infty) \times \mathbb{R}^d \to S_d$, $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $c : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ bounded continuous. Assume, that $a \in C_b^{0,m}([0,\infty) \times \mathbb{R}^d)$ for some $m \geq 2$ and that $b, c \in C^{0,n}([0,\infty) \times \mathbb{R}^d)$ for some $n \geq 1$. Given $T > 0$, $u_T \in C_b^n(\mathbb{R}^d)$, and $g \in C_b^{0,n}([0,T) \times \mathbb{R}^d)$, suppose that $u \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ satisfies*

$$\partial_s u + L_s u + c(s,\cdot)u = -g, \qquad 0 \leq s < T, \tag{2.26}$$

with $u(T,\cdot)=u_T$. Moreover, if for some $0\le l\le m\wedge n$,

$$u\in C_b^{0,l}([0,T]\times\mathbb{R}^d)\cap C^{0,l+2}([0,T)\times\mathbb{R}^d),$$

then $\partial_s u\in C^{0,l}([0,T)\times\mathbb{R}^d)$ and there exists A_l and B_l such that

$$\|u(s,\cdot)\|^{(l)}\le A_l\left(\|u_T\|^{(l)}+\sup_{s\le t<T}\|g(t,\cdot)\|^{(l)}\right)e^{B_l(T-s)},\qquad 0\le s<T.$$

The constants A_l and B_l depend only on l,d and the spatial derivatives of a upto order $l\vee 2$ and b,c upto order l.

Proof. We want to conclude by finite induction over the spatial derivatives of u. In a first step, we apply D^α to (2.26) with $|\alpha|\le l$. We obtain after using several times the product rule and only keep the terms involving derivatives of order l or higher

$$\partial_s u^{(\alpha)}+L_s u^{(\alpha)}+\frac{1}{2}\sum_{k:\alpha_k\ge 1}\alpha_k a^{ij}_{,k}f^{(\hat\alpha_k)}_{,ij}+\sum_{\beta\le\alpha}c_{\alpha,\beta}u^{(\beta)}=-g^{(\alpha)},\tag{2.27}$$

where $\hat\alpha_k=\alpha-e_k$, with e_k the kth canonical basis vector of $\mathbb{R}^d$. Moreover, $c_{\alpha,\beta}$ only involves terms of a,b,c with spatial derivatives upto order l. In the next step, we define

$$w=\sum_{|\alpha|=l}\left(u^{(\alpha)}\right)^2.$$

Note that, since $\nabla u^{(\alpha)}=2u^{(\alpha)}\nabla u^{(\alpha)}$ and $D^2(u^{(\alpha)})=2u^{(\alpha)}D^2u^{(\alpha)}+2\nabla u^{(\alpha)}\otimes\nabla u^{(\alpha)}$, it holds

$$L_s w=2\sum_{|\alpha|=l}u^{(\alpha)}L_s u^{(\alpha)}+\sum_{|\alpha|=l}\left\langle\nabla u^{(\alpha)},a\nabla u^{(\alpha)}\right\rangle.$$

So, we can multiply (2.27) by $2u^{(\alpha)}$ and sum over $|\alpha|=l$, to arrive at

$$\begin{aligned}\partial_s w+L_s w&+\sum_{|\alpha|=l}\sum_{k:\alpha_k\ge 1}\alpha_k u^{(\alpha)}a^{ij}_{,k}u^{(\hat\alpha_k)}_{,ij}-\sum_{|\alpha|=l}\left\langle\nabla u^{(\alpha)},a\nabla u^{(\alpha)}\right\rangle\\&+2\sum_{|\alpha|=l}\sum_{\beta\le\alpha}c_{\alpha,\beta}u^{(\alpha)}u^{(\beta)}+2\sum_{|\alpha|=l}u^{(\alpha)}g^{(\alpha)}\\&=:\partial_s w+L_s w+I-II+III+IV=0\end{aligned}$$

We want to employ the bound (2.16) from the maximum principle Theorem 2.23. Therefore, we have to show that $I+II+III+IV\le C(s)w+G(s)$. We, apply the estimate (2.25) from Lemma 2.28 to

$$\left(a^{ij}_{,k}u^{(\hat\alpha_k)}_{,ij}\right)^2=\left(\operatorname{tr}\left(a_{,k}\,D^2u^{(\hat\alpha_k)}\right)\right)^2\le 4d^2\Lambda_k\operatorname{tr}\left(D^2u^{(\hat\alpha_k)}\,a\,D^2u^{(\hat\alpha_k)}\right),$$

with

$$\Lambda_k=\sup\left\{|\langle\theta,a_{,kk}\theta\rangle|\,/\,|\theta|^2:(s,x)\in[0,\infty)\times\mathbb{R}^d,\theta\in\mathbb{R}^d\setminus\{0\}\right\}.$$

Applying first Cauchy-Schwarz to I and then the above estimate, we deduce

$$\begin{aligned}\left(\sum_{k:\alpha_k\geq 1}\alpha_k u^{(\alpha)} a^{ij}_{,k} u^{(\hat\alpha_k)}_{,ij}\right)^2 &\leq \left(u^{(\alpha)}\right)^2\left(\sum_{k:\alpha_k\geq 1}\alpha_k^2\right)\left(\sum_{k:\alpha_k\geq 1}\left(a^{ij}_{,k}u^{(\hat\alpha_k)}_{,ij}\right)^2\right)\\ &\leq \left(u^{(\alpha)}\right)^2 4d^2l^2\Lambda_0\sum_{k:\alpha_k\geq 1}\operatorname{tr}\left(D^2u^{(\hat\alpha_k)}\,a\,D^2u^{(\hat\alpha_k)}\right),\end{aligned}$$

where $\Lambda_0 = \max_k \Lambda_k$. Since, $D^2u^{(\hat\alpha_k)}$ only involves derivatives also occuring in $\nabla u^{(\alpha)}$ for α running along $|\alpha| = l$, there exists a constant $C_1 = C_1(l, d, \|a\|^{(2)})$ such that[12]

$$\left(\sum_{|\alpha|=l}\sum_{k:\alpha_k\geq 1}\alpha_k u^{(\alpha)} a^{ij}_{,k} u^{(\hat\alpha_k)}_{,ij}\right)^2 \leq 4C_1 w\; II$$

Hence, we get using Young's inequality $2ab \leq a^2 + b^2$

$$I - II \leq 2\left(C_1 w\; II\right)^{\frac12} - II \leq C_1 w.$$

This leads to the estimate

$$\partial_s w + L_s w + C_1 w + III + VI \geq 0.$$

Let us continue with III

$$\begin{aligned}III &= 2\sum_{|\alpha|=l} c_{\alpha,\alpha}\left(u^{(\alpha)}\right)^2 + 2\sum_{|\alpha|=l}\sum_{\beta<\alpha} c_{\alpha,\beta}u^{(\alpha)}u^{(\beta)}\\ &\leq C_2 w + 2C_3\sqrt{w}\,\|u(s,\cdot)\|^{(l-1)} \leq \left(C_2 + C_3\|u(s,\cdot)\|^{(l-1)}\right)w + C_3\|u(s,\cdot)\|^{(l-1)}.\end{aligned}$$

Likewise, we can estimate IV

$$IV \leq 2\sqrt{w}\,\|g(s,\cdot)\|^{(l)} \leq \|g(s,\cdot)\|^{(l)} + \|g(s,\cdot)\|^{(l)} w.$$

A combination of all the estimates leads to

$$\begin{aligned}\partial_s w + L_s w + \left(C_1 + C_2 + C_3\|u(s,\cdot)\|^{(l-1)} + \|g(s,\cdot)\|^{(l)}\right)w&\\ \geq -\left(\|g(s,\cdot)\|^{(l)} + C_3\|u(s,\cdot)\|^{(l-1)}\right).&\end{aligned} \tag{2.28}$$

Now, the conclusion follows from (2.16) in Theorem 2.23 and a finite induction on l. □

01.12.15

For the final result, we need one more result by [10], which we again state and use without proof.

Theorem 2.31 (Analytic elliptic regularity [10])**.** *Let $a \in C_b^\infty([0,\infty)\times\mathbb{R}^d;\to S_d)$ and $b \in C_b^\infty([0,\infty)\times\mathbb{R}^d;\mathbb{R}^d)$ and assume a to be α-elliptic for some $\alpha > 0$, i.e.*

$$\langle\theta, a(s,x)\theta\rangle \geq \alpha\,|\theta|^2, \qquad (s,x)\in[0,\infty)\times\mathbb{R}^d \quad \text{and} \quad \theta\in\mathbb{R}^d.$$

[12]Note: for any symmetric matrix s holds $\operatorname{tr}(s\,a\,s) = s : as = \sum_{i=1}^d \langle s_i, as_i\rangle$, where s_i are the rows of s.

Then, for each $T > 0$ and $u_T \in C_c^\infty(\mathbb{R}^d)$, there exists an $u \in C_b^{1,2}([0,T] \times \mathbb{R}^d) \cap C^\infty([0,T) \times \mathbb{R}^d)$ such that $u(T,\cdot) = u_T$ and $\partial_s u + L_s u = 0$ for $0 \le s < T$.

Theorem 2.32 (Existence of Markov TPF)**.** *For diffusion matrix $a \in C_b^{0,2}([0,\infty) \times \mathbb{R}^d; S_d)$ and vectorfield $b \in C_b^{0,2}([0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, there exists a unique TPF $P(s,x;t,\cdot)$ s.t. $t \mapsto P(s,x;t,B)$ is Borel and for all $0 \le s < t$, $x \in \mathbb{R}^d$, $f \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d)$ holds*

$$\int f(t,y)P(s,x;t,\mathrm{d}y) - f(s,x) = \int_s^t \int (\partial_u + L_u) f(u,y) P(s,x;u,\mathrm{d}y) \tag{2.29}$$

Moreover, for each $(s,x) \in [0,\infty) \times \mathbb{R}^d$ exists $\boldsymbol{\eta}^{s,x} \in \mathcal{P}(\Gamma)$ s.t.

$$\boldsymbol{\eta}^{s,x}\left(\{\gamma : \gamma(t) = x, \forall t \in [0,s]\}\right) = 1 \tag{2.30}$$

and

$$\boldsymbol{\eta}^{s,x}\left(\gamma(t_2) \in B \mid \mathcal{F}_{t_1}\right) = P(t_1, \gamma(t_1); t_2, B) \qquad \boldsymbol{\eta}_{s,x}\text{-a.e.} \tag{2.31}$$

for $s \le t_1 < t_2$ and $B \in \mathcal{B}_{\mathbb{R}^d}$. In other words $\boldsymbol{\eta}^{s,x}$ is Markov(P,δ_x) *after time s.*

Proof. The proof uses an approximation of the coefficient by $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that

(a) $a_n, b_n \in C_b^\infty$.
(b) For all $n \ge 1$ exists $\alpha_n > 0$ s.t. $\langle \theta, a_n\theta\rangle \ge \alpha_n |\theta|^2$.
(c) For all $t > 0$

$$\sup_{s\in[0,t]} \sup_{x\in\mathbb{R}^d} \left(\|a_n(s,x) - a(s,x)\| + |b_n(s,x) - b(s,x)|\right) \to 0.$$

(d) There exists $C < \infty$ s.t.

$$\sup_{n\in\mathbb{N}} \sup_{s\in[0,\infty)} \left(\|a_n(s,\cdot)\|^{(2)} + \|b_n(s,\cdot)\|^{(2)}\right) \le C.$$

Exercise: Show that under the assumptions on a, b such coefficients exist. Use a mollification to satisfy (a), add some multiple of the identity to a_n for (b) and then check (c) and (d).

Now, let $\{L_t^n\}_{n\in\mathbb{N}}$ be the family of generators obtained from a_n, b_n and let $\left\{T_{s,t}^n : 0 \le s < t\right\}$ be the associated C_0-semigroups given from Corollary 2.25. Given $t > 0$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$, set $f_n(s,\cdot) = T_{s,t}^n\varphi$. Then by Theorem 2.31 holds $f_n \in C_b^{1,2}([0,t] \times \mathbb{R}^d) \cap C^\infty([0,\infty) \times \mathbb{R}^d)$, $\partial_s f_n + L_t^n f_n = 0$ and $f_n(t,\cdot) = \varphi$. By Theorem 2.30, we get $\sup_{0\le s<t} \|f_n(s,\cdot)\|^{(2)} \le \tilde{C} < \infty$ and hence we can pass to the limit in

$$\lim_{n\to\infty} \partial_s f_n + L_t f_n = 0 \qquad \text{uniformly in } [0,t) \times \mathbb{R}^d.$$

By density, we find that for any $\varphi \in C_0(\mathbb{R}^d)$: $T_{s,t}^n\varphi(x)$ converges uniformly to a limit for $(s,x) \in [0,t] \times \mathbb{R}^d$. If, we denote this limit by $T_{s,t}\varphi$, we find that $T_{s,t}^n \to T_{s,t}$ strongly as operators on C_0, especially the C_0-semigroup properties are preserved.

Now, we prove that if P satisfying (2.29) exists, then it has to be unique. Let $t > 0$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $f_n(s,\cdot) = T^n_{s,t}\varphi$. Then (2.29) holds with f replaced by f_n. Since $\partial_s f_n + L_s f_n \to 0$ uniformly on $[0,t) \times \mathbb{R}^d$ as $n \to \infty$, it holds

$$\int \varphi(y) P(s,x;t,\mathrm{d}y) - f_n(s,x) \to 0 \quad \text{as } n \to \infty.$$

Therefore, by the convergence of the semigroups

$$T_{s,t}\varphi = \int \varphi(y) P(s,x,;t,\mathrm{d}y). \tag{2.32}$$

By density, the identity (2.32) holds for all $\varphi \in C_0(\mathbb{R}^d)$ and by Corollary 2.25, we conclude by uniqueness of the semigroup.

The last step is to show the existence of the TPF P, which will be a compactness argument. Let $P_n(s,x;t,B)$ be the TPF of the semigroup $T^n_{s,t}$ from Corollary 2.25, which satisfies

$$\int |y-x|^4 P_n(s,x;t,\mathrm{d}y) \le A e^{B(t-s)}(t-s)^2, \qquad 0 \le s < t \quad \text{and} \quad x \in \mathbb{R}^d. \tag{2.33}$$

Especially, $\{P_n(s,x;t,\cdot)\}_{n\in\mathbb{N}}$ is tight and by Prokhorov relative compact for each $(s,x) \in [0,\infty) \times \mathbb{R}^d$ and $t > s$. Since, by Corollary 2.25, it also holds

$$\begin{array}{ccc} T^n_{s,t}\varphi(x) = & \int \varphi(y) P_n(s,x;t,\mathrm{d}y), & \varphi \in C_0(\mathbb{R}^d) \\ \downarrow & \downarrow & \\ T_{s,t}\varphi(x) = & \int \varphi(y) P(s,x;t,\mathrm{d}y), & \varphi \in C_0(\mathbb{R}^d), \end{array}$$

where $T^n_{s,t} \to T_{s,t}$ strongly, which then implies $P_n(s,x;t,\cdot)$ weakly* to $P_n(s,x;t,\cdot)$. Moreover, this representation implies that P is a TPF and jointly measurable in s,x and $t > s$. Moreover, P satisfies the same bound (2.33) as P_n. Now, to show that P satisfies (2.29), we use Corollary 2.27 applied with L^n to get P_n satisfying (2.29) with P_n and L^n replacing P and L. By construction we have $L^n_s f(x) \to L_s f(x)$ uniformly in $(s,x) \in [0,t) \times \mathbb{R}^d$ and we can conclude from the weak* convergence of P_n to P that (2.29) also holds for P and L.

Finally, by Corollary 2.21 exists for each $(s,x) \in [0,\infty) \times \mathbb{R}^d$ a $\boldsymbol{\eta}^{s,x} \in \mathcal{P}\left(C([s,\infty) \times \mathbb{R}^d)\right)$ being Markov(P,δ_x) and satisfying (2.31). By imposing $(e_u)_\sharp \boldsymbol{\eta}^{s,x} = \delta_x$ for all $u \in [0,s]$, we can uniquely extend $\boldsymbol{\eta}^{s,x} \in \mathcal{P}(\Gamma)$ satisfying (2.30). □

3. Martingale solutions

3.1. **Martingales.** As we only consider continuous stochastic processes, let us specialize the defintion of a Martingale to this setting, where we use the notation of Section 2.2.

Definition 3.1 (Martingale). Let $\boldsymbol{\eta} \in \mathcal{P}(\Gamma)$ and $s > 0$. A family $\{M_t : \Gamma \to \mathbb{R}\}_{t \ge s}$ is an *$\boldsymbol{\eta}$-martingale after time s* if

(i) For all $t \geq s$: M_t is adapted to $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$, i.e. M_t is $\mathcal{F}_t$ measruable;
(ii) For all $t \geq s$: $E^{\boldsymbol{\eta}}[|M_t|] < \infty$;
(iii) For all $s \leq t_1 < t_2$ holds

$$E^{\boldsymbol{\eta}}\left[M_{t_2}|\mathcal{F}_{t_1}\right] = M_{t_1} \qquad \boldsymbol{\eta}\text{-a.s.} \tag{3.1}$$

Likewise, it is called sub-martingale or super-martingale if "=" in (3.1) is replaced by "≥" or "≤", respectively.

Remark 3.2. We can rewrite (3.1) in disintegrated form as for each $\gamma|_{[s,t_1]} \in C([s,t_1];\mathbb{R}^d)$ holds

$$\int_\Gamma M_{t_2}(\gamma)\ \boldsymbol{\eta}_{t_1,\gamma|_{[s,t_1]}}(\mathrm{d}\gamma) = M_{t_1}(\gamma|_{[s,t_1]}).$$

Lemma 3.3. *Let $\{M_t\}_{t\geq s}$ be a $\boldsymbol{\eta}$-submartingale after time s. Let $g : \mathbb{R} \to \mathbb{R}_+$ be convex, non-decreasing. If $g \circ M_t$ is $\boldsymbol{\eta}$-integrable, then $g \circ M_t$ is $\boldsymbol{\eta}$-submartingale.*

Especially, if $\{M_t\}_{t\geq s}$ is a $\boldsymbol{\eta}$-martingale or non-negative $\boldsymbol{\eta}$-submartingale and for $r \geq 1$: $|M_t|^r$ is $\boldsymbol{\eta}$-integrable, then $|M_t|^r$ is a $\boldsymbol{\eta}$-submartingale.

Proof. We assume wlog. $s = 0$. By then Jensen inequality follows for $0 \leq t_1 < t_2$

$$E^{\boldsymbol{\eta}}\left[g(M(t_2)) \mid \mathcal{F}_{t_1}\right] \geq g\left(E^{\boldsymbol{\eta}}\left[M(t_2) \mid \mathcal{F}_{t_1}\right]\right) \geq g(M_{t_1}).$$

For the second assertion, it is left to show that if M_t is a martingale, then $|M_t|$ is a submartingale. This is immediate from

$$E^{\boldsymbol{\eta}}\left[|M_{t_2}| \mid \mathcal{F}_{t_1}\right] \geq \left|E^{\boldsymbol{\eta}}\left[M_{t_2} \mid \mathcal{F}_{t_1}\right]\right| = |M_{t_1}|\,.$$

□

Lemma 3.4. *Let $\{M_t\}_{t\geq s}$ be a $\boldsymbol{\eta}$-submartingale after time s, then for any $\lambda > 0$ and all $t > s$ holds*

$$\boldsymbol{\eta}\left(\sup_{s\leq u\leq t} M_t \geq \lambda\right) \leq \frac{1}{\lambda} E^{\boldsymbol{\eta}}\left[M_t, \sup_{s\leq u\leq t} M_t \geq \lambda\right] \tag{3.2}$$

In particular, if M_t is non-negative, then

$$\boldsymbol{\eta}\left(\sup_{s\leq u\leq t} M_t \geq \lambda\right) \leq \frac{1}{\lambda} E^{\boldsymbol{\eta}}\left[M_t\right] \tag{3.3}$$

and for all $r > 1$

$$E^{\boldsymbol{\eta}}\left[\left(\sup_{s\leq u\leq t} M_t\right)^r\right]^{1/r} \leq \frac{r}{r-1} E^{\boldsymbol{\eta}}\left[M_t^r\right]^{1/r}. \tag{3.4}$$

Proof. Again wlog. $s = 0$. If is enough to show for any $n \geq 1$ and $0 = t_0 < \cdots < t_n = t$

$$\boldsymbol{\eta}\left(\max_{0\leq k\leq n} M_{t_k} \geq \lambda\right) \leq \frac{1}{\lambda} E^{\boldsymbol{\eta}}\left(M_t, \max_{0\leq k\leq n} M_{t_k} \geq \lambda\right).$$

Let $A_0 := \{M_{t_0} \geq \lambda\}$ and for $1 \leq k \leq n$ let

$$A_k := \left\{ M_{t_k} \geq \lambda \text{ and } \max_{0\leq i<k} M_{t_i} < \lambda \right\}.$$

Clearly $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\{\max_{0\leq k\leq n} M_{t_k}\} = \bigcup_{k=0}^n A_k$, and $A_k \in \mathcal{F}_{t_k}$. Therefore,

$$\begin{aligned}\boldsymbol{\eta}\left(\max_{0\leq k\leq n} M_{t_k} \geq \lambda\right) &= \sum_{k=0}^n \boldsymbol{\eta}(A_k) \leq \frac{1}{\lambda}\sum_{k=0}^n E^{\boldsymbol{\eta}}\left[M_{t_k}, A_k\right] \\ &\leq \frac{1}{\lambda}\sum_{k=0}^n E^{\boldsymbol{\eta}}\left[M_t, A_k\right] = \frac{1}{\lambda}E^{\boldsymbol{\eta}}\left[M_T, \max_{0\leq k\leq n} M_{t_k} \geq \lambda\right]\end{aligned}$$

Now, the estimate (3.3) is an immediate consequence of (3.2), if M_t is non-negative. The estimate (3.4) follows from (3.3) as follows: Let $\overline{M}_t := \sup_{0\leq u\leq t} M_u$, then with Riemann-Stieltjes integration and integration by parts (see also [6, Theorem 3.4, Chapter 7])

$$\begin{aligned}E^{\boldsymbol{\eta}}\left[\overline{M}_t^r\right] &= -\int_0^\infty \lambda^r \,\mathrm{d}\boldsymbol{\eta}\left(\overline{M}_t \geq \lambda\right) = \int_0^\infty \boldsymbol{\eta}\left(\overline{M}_t \geq \lambda\right) \mathrm{d}\lambda^r \\ &\leq \int_0^\infty E^{\boldsymbol{\eta}}\left[M_t, \overline{M}_t \geq \lambda\right] \frac{\mathrm{d}\lambda^r}{\lambda} = rE^{\boldsymbol{\eta}}\left[M_t \int_0^{\overline{M}_t} \lambda^{r-2}\,\mathrm{d}r\right] = \frac{r}{r-1}E^{\boldsymbol{\eta}}\left[M_t\overline{M}_t^{r-1}\right].\end{aligned}$$

The conclusion follows from Hölder's inequality

$$E^{\boldsymbol{\eta}}\left[M_t\overline{M}_t^{r-1}\right] \leq \left(E^{\boldsymbol{\eta}}\left[M_t^r\right]\right)^{\frac{1}{r}} \left(E^{\boldsymbol{\eta}}\left[\overline{M}_t^r\right]\right)^{1-\frac{1}{r}}.$$

□

Lemma 3.5 (Markov martingales)**.** *Let a TPF P, $s > 0$ and $x \in \mathbb{R}^d$ be given. If $\boldsymbol{\eta}^{s,x} \in \mathcal{P}(\Gamma)$ is* Markov(P, δ_x) *after time s, then for any $f \in C_b^{1,2}(\mathbb{R}^d)$*

$$M_t^f := f_t \circ e_t - \int_s^t (\partial_u f_u + L_u f_u) \circ e_u \,\mathrm{d}u,$$

is an $\boldsymbol{\eta}^{s,x}$-martingale after time s, where $f_t(x) = f(t, x)$.

Proof. We can calculate using the relation between $\boldsymbol{\eta}^{s,x}$ and P

$$\boldsymbol{\eta}^{s,x}(\gamma(t_2) \in B \mid \mathcal{F}_{t_1}) = P(t_1, \gamma(t_1); t_2, B)$$

as follows

$$
\begin{aligned}
E^{\boldsymbol{\eta}^{s,x}}\left[M^f_{t_2} \mid \mathcal{F}_{t_1}\right] &= E^{\boldsymbol{\eta}^{s,x}}\left[f \circ e_{t_2} \mid \mathcal{F}_{t_1}\right] - \int_s^{t_1} (\partial_u f_u + L_u f_u) \circ e_u \,\mathrm{d}u \\
&\quad - \int_{t_1}^{t_2} E^{\boldsymbol{\eta}^{s,x}}\left[(\partial_u f_u + L_u f_u) \circ e_u \mid \mathcal{F}_{t_1}\right] \\
&= E^{\boldsymbol{\eta}^{s,x}}\left[M_{t_2} \mid \mathcal{F}_{t_1}\right] + \int f_{t_1}(y) P(t_1, \gamma(t_1); t_2, \mathrm{d}y) \qquad (3.5) \\
&\quad - f_{t_1}(\gamma(t_1)) - \int_{t_1}^{t_2} \int (\partial_u + L_u) f(u,y) P(t_1, \gamma(t_1); u, \mathrm{d}y) \,\mathrm{d}u.
\end{aligned}
$$

As a consequence of the weak maximum principle, we arrived at the following identity (2.24) for a generator L_t with bounded continuous coefficients a, b

$$
\int f(t,y) P(s,x;t,\mathrm{d}y) - f(s,x) = \int_s^t \int (\partial_u + L_u) f(u,y) P(s,x;u,\mathrm{d}y) \,\mathrm{d}u.
$$

Hence, the identity (3.5) is the martingale property $E^{\boldsymbol{\eta}^{s,x}}[M^f_{t_2} \mid \mathcal{F}_{t_1}] = M^f_{t_1}$. □

08.12.15

This observation motivates the definition:

Definition 3.6 (Martingale solution)**.** Given $a : \mathbb{R}_+ \times \mathbb{R}^d \to S_d$, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ measurable and $L_t := \frac{1}{2} a_t : \nabla^2 + b_t \cdot \nabla$. Let $T \in (0,\infty]$, a probability measure $\boldsymbol{\eta}^{s,x} \in \mathcal{P}(\Gamma_T)$ is called a *solution to the martingale problem for* L_t *starting from* (s,x), denoted by $\mathrm{MP}_s(L_t, \delta_x)$, if

(i) normalization:

$$
\boldsymbol{\eta}^{s,x}\left(\{\gamma \in \Gamma : \gamma(t) = x, \forall t \in [0,s]\}\right) = 1.
$$

(ii) boundedness:

$$
\int_s^T \int \left(\|a(t,y)\| + |b(t,y)|\right) \, (e_t)_\sharp \boldsymbol{\eta}^{s,x}(\mathrm{d}y) \,\mathrm{d}t < \infty. \qquad (3.6)
$$

(iii) Martingale property: For all $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$

$$
M^f_t := f_t \circ e_t - \int_s^t (\partial_u f_u + L_u f_u) \circ e_u \,\mathrm{d}u \qquad (3.7)
$$

is a $\boldsymbol{\eta}^{s,x}$-martingale after time s.

The martingale problem is *well-posed* if for any $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, there exists a unique $\boldsymbol{\eta}^{s,x}$ with the above properties.

If the *starting from* (s,x) is ommitted, then $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ is just called a solution to the martingale problem, denoted just by $\boldsymbol{\eta} \in \mathrm{MP}(L_t)$, if $\boldsymbol{\eta} \in \mathrm{MP}_0(L_t, (e_0)_\sharp \boldsymbol{\eta})$, i.e. the initial value is just implied by $\boldsymbol{\eta}$ itself.

Remark 3.7. The condition (3.6) immediately follows, if a, b are bounded.

Moreover, the boundedness assumption (3.6) implies that $E^{\boldsymbol{\eta}}\left[\left|M_t^f\right|\right] < \infty$, indeed

$$E^{\boldsymbol{\eta}}\left[\left|M_t^f\right|\right] \leq \|f\|_{C^{1,2}(\mathbb{R}_+\times\mathbb{R}^d)} \left(1 + \int_s^t \int \left(\|a(u,y)\| + |b(u,y)|\right) (e_u)_\sharp \boldsymbol{\eta}^{s,x} \,\mathrm{d}y \,\mathrm{d}u\right) < \infty \tag{3.8}$$

3.2. **Uniqueness.**

Lemma 3.8. *Let $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ satisfy (3.6) and $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$. Then, M_t^f defined in (3.7) is a $\boldsymbol{\eta}$-martingale after time s if and only if for all $g \in C_b\left((\mathbb{R}^d)^k\right)$ with $k \in \mathbb{N}$, all $0 \leq t_1 \leq \cdots \leq t_k \leq t$ with $t > s$ and all $h > 0$ holds*

$$E^{\boldsymbol{\eta}}\left[\left(f_{t+h}\circ e_{t+h} - f_t \circ e_t - \int_t^{t+h} (\partial_u f_u + L_u f_u)\circ e_u \,\mathrm{d}u\right) g \circ (e_{t_1} \times \cdots \times e_{t_k})\right] = 0. \tag{3.9}$$

Proof. Let M_t^f be a $\boldsymbol{\eta}$-marginal, then the lhs. of (3.9) satisfies

$$\begin{aligned} &E^{\boldsymbol{\eta}}\left[g\circ(e_{t_1}\times\cdots\times e_{t_k}) \cdot E^{\boldsymbol{\eta}}\left[f_{t+h}\circ e_{t+h} - f_t\circ e_t - \int_t^{t+h}(\partial_u f_u + L_u f_u)\circ e_u \,\mathrm{d}u \,\middle|\, \mathcal{F}_t\right]\right] \\ &= E^{\boldsymbol{\eta}}\left[g\circ(e_{t_1}\times\cdots\times e_{t_k}) \cdot \left(E^{\boldsymbol{\eta}}\left[M_{t+h}^f \mid \mathcal{F}_t\right] - M_t^f\right)\right] = 0 \end{aligned}$$

On the other hand, we get for by approximation of indicator functions with g

$$E^{\boldsymbol{\eta}}\left[M_{t+h}^f - M_t^f, \gamma(t_1)\in B_1,\dots,\gamma(t_k)\in B_k\right] = 0 \quad \text{with } B_i \in \mathcal{B}_{\mathbb{R}^d} \text{ for } i = 1,\dots,k.$$

By the boundedness (3.8) of $\left|M_t^f\right|$, we also get

$$\begin{aligned} &E^{\boldsymbol{\eta}}\left[\left(M_{t+h}^f - M_t^f\right)^+, \gamma(t_1)\in B_1,\dots,\gamma(t_k)\in B_k\right] \\ &= E^{\boldsymbol{\eta}}\left[\left(M_{t+h}^f - M_t^f\right)^-, \gamma(t_1)\in B_1,\dots,\gamma(t_k)\in B_k\right]. \end{aligned}$$

We recall that $\mathcal{F}_t$ is generated by $\{e_u : u \in S \subset [0,t] \text{ countable}\}$. Therewith, we can pass to the limit $k \to \infty$ and obtain for all $A \in \mathcal{F}_t$

$$E^{\boldsymbol{\eta}}\left[\left(M_{t+h}^f - M_t^f\right)^+, \gamma \in A\right] = E^{\boldsymbol{\eta}}\left[\left(M_{t+h}^f - M_t^f\right)^-, \gamma \in A\right]$$

and hence M_t^f a $\boldsymbol{\eta}$-martingale. ∎

Theorem 3.9 (Uniqueness and Markov property by first marginals)**.** *Let $T > 0$. If for any $\boldsymbol{\eta}, \tilde{\boldsymbol{\eta}} \in \mathcal{P}(\Gamma_T)$ solutions to* $\mathrm{MP}(L_t)$ *holds*

$$\forall t \in [0,T] : (e_t)_\sharp \boldsymbol{\eta} = (e_t)_\sharp \tilde{\boldsymbol{\eta}}, \tag{3.10}$$

then the martingale problem has a unique solution and this solution satisfies the Markov property, i.e. for any $t > s > 0$ and $B \in \mathcal{B}_{\mathbb{R}^d}$ holds

$$\boldsymbol{\eta}\left(\gamma(t) \in B \mid \mathcal{F}_s\right) = \boldsymbol{\eta}\left(\gamma(t) \in B \mid e_s\right). \tag{3.11}$$

Proof. It is enough to identify the finite dimensional distributions. We conclude by induction and the induction basis is just (3.10). Hence, we assume now, that for some $n \geq 1$ and all $0 \leq s_1 < \cdots < s_n \leq T$ holds $(e_{s_1} \times \cdots \times e_{s_n})_\sharp \boldsymbol{\eta} = (e_{s_1} \times \cdots \times e_{s_n})_\sharp \tilde{\boldsymbol{\eta}}$. Let $B_1, \ldots, B_n \in \mathcal{B}_{\mathbb{R}^d}$ be given. We define the conditional measures

$$\boldsymbol{\eta}_{s_1,\ldots,s_n}(\bullet) := \boldsymbol{\eta}\left(\bullet \mid \gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n\right) := \frac{\boldsymbol{\eta}\left(\bullet, \gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n\right)}{\boldsymbol{\eta}(\gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n}$$

and likewise for $\tilde{\boldsymbol{\eta}}$. Moreover, we define the shifted evaluation $e_t^{s_n} := e_{t+s_n}$ and $N_t^f = M_{t+s_n}^f$. We will check that N_t^f is a $\boldsymbol{\eta}_{s_1,\ldots,s_n}$-martingale. Therefore, we apply Lemma 3.8 and we observe that for any $0 \leq t_1 < \cdots \leq t_k \leq t$, any $g \in C_b\left((\mathbb{R}^d)^k\right)$ and $h > 0$ with ξ defined by

$$\begin{aligned}\xi = &\left(f_{t+s_n+h} \circ e_{t+h}^{s_n} - f_{t+s_n} \circ e_t^{s_n} - \int_t^{t+h} (\partial_u f_u + L_u f_u) \circ e_u^{s_n} \,\mathrm{d}u\right) \times \\ &\times g \circ (e_{t_1}^{s_n} \times \cdots \times e_{t_k}^{s_n})\end{aligned}$$

holds

$$E^{\boldsymbol{\eta}_{t_1,\ldots,t_n}}[\xi] = \frac{1}{\boldsymbol{\eta}\left(\gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n\right)} E^{\boldsymbol{\eta}}\left[\xi, \gamma(s_1) \in B_1, \ldots, \gamma(s_n) \in B_n\right]. \tag{3.12}$$

Therefore, by an application of Lemma 3.8 with $\tilde{g} \in C_b((\mathbb{R}^d)^{k+n})$, where

$$\bar{g}(x_1, \ldots, x_k, y_1, \ldots, y_n) = g(x_1, \ldots, x_k) \prod_{i=1}^n \mathbb{1}_{B_i}(y_i),$$

follows that the rhs. of (3.12) is zero, since M_t^f is a $\boldsymbol{\eta}$-martingale. But, then also N_t^f is a $\boldsymbol{\eta}_{s_1,\ldots,s_n}$-martingale. In the same way, we define $\tilde{\boldsymbol{\eta}}_{s_1,\ldots,s_n}$ and then N_t^f will be a $\tilde{\boldsymbol{\eta}}_{s_1,\ldots,s_n}$-martingale by an analog argument. Now, both solutions have the same initial value, since by construction

$$\begin{aligned}\boldsymbol{\eta}_{s_1,\ldots,s_n}(e_0^{s_n} \in B) &= \frac{\boldsymbol{\eta}\left(\gamma(s_1) \in B_1, \ldots, \gamma(s_{n-1}) \in B_{n-1}, \gamma(s_n) \in B_n \cap B\right)}{\boldsymbol{\eta}(\gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n} \\ &= \frac{\tilde{\boldsymbol{\eta}}\left(\gamma(s_1) \in B_1, \ldots, \gamma(s_{n-1}) \in B_{n-1}, \gamma(s_n) \in B_n \cap B\right)}{\tilde{\boldsymbol{\eta}}(\gamma_{s_1} \in B_1, \ldots, \gamma_{s_n} \in B_n} \\ &= \tilde{\boldsymbol{\eta}}_{s_1,\ldots,s_n}(\gamma \circ e_0^{s_n} \in B).\end{aligned}$$

But, then also by the assumption (3.10) for all later time $u \geq 0$: $\tilde{\boldsymbol{\eta}}_{s_1,\ldots,s_n}(e_u^{s_n} \in B) = \tilde{\boldsymbol{\eta}}_{s_1,\ldots,s_n}(e_u^{s_n} \in B)$. We can especially choose $u = s_{n+1} - s_n$ and conclude the induction step

$$\boldsymbol{\eta}\left(\gamma(s_1) \in B_1, \ldots, \gamma(s_n) \in B_n, \gamma(s_{n+1}) \in B\right) = \tilde{\boldsymbol{\eta}}\left(\gamma(s_1) \in B_1, \ldots, \gamma(s_n) \in B_n, \gamma(s_{n+1}) \in B\right).$$

For the Markov property, define for some $F \in \mathcal{F}_s$ such that $\boldsymbol{\eta}(F) > 0$

$$\boldsymbol{\eta}_{[0,s]}(\bullet) := \frac{E^{\boldsymbol{\eta}}\left[\mathbb{1}_F\ \boldsymbol{\eta}(\bullet \mid \mathcal{F}_s)\right]}{\boldsymbol{\eta}(F)} \quad \text{and} \quad \boldsymbol{\eta}_s(\bullet) := \frac{E^{\boldsymbol{\eta}}\left[\mathbb{1}_F\ \boldsymbol{\eta}(\bullet, F \mid e_s)\right]}{\boldsymbol{\eta}(F)}$$

If we now define for $N_t^f := M_{s+t}^f$ with $t \geq 0$, we can show by an analog argument as before and by an application of Lemma 3.8 that N_t^f is a martingale wrt. $\boldsymbol{\eta}_{[0,s]}$ as well as $\boldsymbol{\eta}_s$. Moreover, by the martingale property $E^{\boldsymbol{\eta}}\left[M_s^f \mid \mathcal{F}_s\right] = M_s^f = E^{\boldsymbol{\eta}}\left[M_s^f \mid e_s\right]$ and

$$\boldsymbol{\eta}_{[0,s]}\left(e_0^s \in B\right) = \boldsymbol{\eta}\left(e_0^s \in B | F\right) = \boldsymbol{\eta}_s\left(e_0^s \in B\right),$$

where $e_t^s = e_{s+t}$ is again the shifted evaluation map. Hence, both $\boldsymbol{\eta}_{[0,s]}$ and $\boldsymbol{\eta}_s$ are solutions to the martingale problem after time s with the same initial value. Then, by the hypothesis (3.10) follows for any $t \geq 0$ and any $f \in \mathbb{C}_b(\mathbb{R}^d)$

$$\boldsymbol{\eta}_{[0,s]}\left(e_t^s \in B\right) = \boldsymbol{\eta}_s\left(e_t^s \in B\right),$$

which translates to

$$\begin{aligned}\boldsymbol{\eta}\left(e_{t+s} \in B \mid \mathcal{F}_s\right) &= E^{\boldsymbol{\eta}}\left[\mathbb{1}_F\ \boldsymbol{\eta}\left(e_t^s \in B \mid \mathcal{F}_s\right)\right] \\ &= E^{\boldsymbol{\eta}}\left[\mathbb{1}_F\ \boldsymbol{\eta}\left(e_t^s \in B \mid e_s\right)\right] = \boldsymbol{\eta}\left(e_{t+s} \in B \mid e_s\right)\end{aligned}$$

for all $t \geq 0$ and all $f \in C_b(R_d)$, which proves (3.11). □

3.3. **Stopping times.** Before starting with defining the quadratic variation, we define stopping times.

Definition 3.10 (Stopping time)**.** A map $\tau : \Gamma \to [0, \infty]$ is called a $\mathcal{F}_t$-*stopping time*, if for all $t \in [0, \infty]$

$$\{\tau \leq t\} := \{\gamma \in \Gamma : T(\gamma) \leq t\} \in \mathcal{F}_t$$

For a stopping time τ, the *pre-τ-σ-algebra* $\mathcal{F}_\tau$ is defined by

$$\mathcal{F}_\tau := \{F \in \mathcal{F} : F \cap \{\tau \leq t\} \in \mathcal{F}_t\ \forall t \in [0, \infty]\}.$$

Lemma 3.11 (Basic properties of stopping times)**.** *Let τ, $\tilde{\tau}$ be stopping times. Then:*

(i) If $\tau \leq \tilde{\tau}$, then $\mathcal{F}_\tau \leq \mathcal{F}_{\tilde{\tau}}$.
(ii) $\mathcal{F}_{\tau \wedge \tilde{\tau}} = \mathcal{F}_\tau \cap \mathcal{F}_{\tilde{\tau}}$.
(iii) If $f \in \mathcal{F}_{\tau \vee \tilde{\tau}}$, then $F \cap \{\tau \leq \tilde{\tau}\} \in \mathcal{F}_{\tilde{\tau}}$.
(iv) $\mathcal{F}_{\tau \vee \tilde{\tau}} = \sigma\left(\mathcal{F}_\tau, \mathcal{F}_{\tilde{\tau}}\right)$.

Proof. Exercise. □

Example 3.12 (First entrance times)**.** Let $A \subset \mathbb{R}^d$ be a closed set. Then the first entrance time τ_A defined for continuous stochastic process $X_t : \Gamma \to \mathbb{R}^d$ by

$$\tau_A = \inf\{t \geq 0 : X_t \in A\}$$

is a stopping time.

Lemma 3.13. *Let $X_t : \Gamma \to \mathbb{R}^d$ be a right-continuous stochastic process and τ a $\mathcal{F}_t$-stopping time, then X_τ is $\mathcal{F}_\tau$-measurable.*

Proof. Exercise. □

Theorem 3.14 (Optional stopping)**.** *Let $\{M_t\}_{t\geq 0}$ be a submartingale and let $\tau, \tilde{\tau}$ be $\mathcal{F}_t$-stopping times. Then for each $A < \infty$*

$$E\left[M_{\tau\wedge A} \mid \mathcal{F}_{\tilde{\tau}}\right] \geq M_{\tau\wedge\tilde{\tau}\wedge A}, \qquad a.s..$$

If in addition

(i) τ *is finite a.s.,*
(ii) $E\left[|M_\tau|\right] < \infty$, *and*
(iii) $\lim_{A\to\infty} E\left[X_A, \tau > A\right] = 0$,

then

$$E\left[M_\tau \mid \mathcal{F}_{\tilde{\tau}}\right] \geq M_{\tau\wedge\tilde{\tau}}, \qquad a.s..$$

Proof. See [14, Theorem 1.2.5] or [3, Theorem 1.9.27]. □

15.12.15

3.4. **Stochastic integrals and quadratic variation.** In this section, we fix a generic probability space $(\Gamma, \mathcal{F}, P)$ with filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

The following definition will naturally pop up in the proofs of the next results

Definition 3.15 (Local martingale)**.** A stochastic process $\{M_t\}_{t\geq 0}$ is called *local martingale* if there exists a sequence of stopping times $\tau_n \leq \tau_{n+1}$ with $\tau_n \to \infty$ as $n \to \infty$ such that $M_t^{\tau_n} := M_{t\wedge\tau_n}$ is a martingale for each n.

Definition 3.16 (Variations)**.** For $n \in \mathbb{N}$ and $T > 0$, the subdivision $\Delta_n[0,T]$ is defined as

$$\Delta_n[0,T] := \left\{(t_0, \ldots, t_n) : 0 \leq t_0 < \cdots < t_n \leq T\right\}.$$

Moreover, we set

$$|\Delta_n[0,T]| = \max_{0\leq k\leq n-1} |t_{k+1} - t_k|$$

Then, for a continuous process X_t, its variation is defined by

$$|X|_{\mathrm{var},T} := \sup_{n\in\mathbb{N}:\Delta_n[0,T]} \sum_{k=0}^{n-1} \left|X_{t_{k+1}} - X_{t_k}\right|.$$

Likewise, its quadratic variation is defined by

$$\langle X\rangle_T := \sup_{n\in\mathbb{N}:\Delta_n[0,T]} \sum_{k=0}^{n-1} \left(X_{t_{k+1}} - X_{t_k}\right)^{\otimes 2},$$

where $X^{\otimes 2} := X \otimes X \in \mathbb{R}^{d\times d}$.

In the following, we will not make all vector-vector or matrix-vector products explicit.

Proposition 3.17 (Stochastic integral wrt. continuous functions)**.** *Let* $\{M_t\}_{t\geq 0}$ *be a a continuous (local) martingale and let* $\{V_t\}_{t\geq 0}$ *be a continuous* $\mathcal{F}_t$*-measurable process of bounded variation. Then*

$$W(t) = \int_0^t V_u \,\mathrm{d}M_u = V_t M_t - V_0 M_0 - \int_0^t M_u \,\mathrm{d}V_u$$

is a local martingale.

Proof. Let the stopping time γ_N be defined such that $|M^{\gamma_N}|$ and the total variation $|V|_{\mathrm{var},t}$ are bounded by N. Let us recall the definition of the Riemann-Stieltjes integral, which leads to

$$\int_0^t V_u \,\mathrm{d}M_u = \lim_{|\Delta_n[0,t]|\to 0} \sum_{k=0}^{n-1} V_{t_k^n} \left(M^{\gamma_N}_{t^n_{k+1}} - M^{\gamma_N}_{t^n_k} \right).$$

Then, for any finite n, the rhs above is a martingale by the law of total expectation on the conditionals $t_1, \dots, t_n$. Moreover, we can rearrange the rhs by a simple discrete integration by parts as

$$\sum_{k=0}^{n-1} V_{t_k^n} \left(M^{\gamma_N}_{t^n_{k+1}} - M^{\gamma_N}_{t^n_k} \right) = V_t M_t^{\gamma_N} - V_0 M_0^{\gamma_N} - \sum_{k=0}^{m-1} M^{\gamma_N}_{t^n_{k+1}} \left(V_{t^n_{k+1}} - V_{t^n_k} \right).$$

Since M^{γ_N} is bounded and V is of finite variation, we can pass to the limit in the above formulation and the last term converges to $\int_0^t M_s^{\gamma_N} \,\mathrm{d}V_s$ in L^1 as $n \to \infty$. The result follows by also letting $N \to \infty$. □

Lemma 3.18. *Let* $T > 0$ *and* $\{M_t\}_{0\leq t\leq T}$ *be a (local) continuous martingale of finite variation, then* $\{M_t\}_{0\leq t\leq T}$ *is constant, i.e. it is deterministic.*

Proof. We may assume $M_0 = 0$. For $N \geq 0$, let us consider the stopping time

$$\tau_N = \inf \left\{ s \in [0,T] : |M_s| \geq N, |M|_{\mathrm{var},s} \geq N \right\} \wedge T.$$

The stopped process $(M_{t\wedge T_N})_{0\leq t\leq T}$ is a martingale by the optional stopping Theorem 3.14 and therefore for $s \leq t$,

$$E\left[|M_{t\wedge\tau_N} - M_{s\wedge\tau_N}|^2 \right] = E\left[|M_{t\wedge\tau_N}|^2 \right] - E\left[|M_{s\wedge\tau_N}|^2 \right].$$

Consider now a sequence of subdivisions $\Delta_n[0,T]$ whose mesh tends to 0. By summing up the above inequality on the subdivision, we obtain

$$\begin{aligned}
E\left[|M_{\tau_N}|^2\right] &= E\left[|M_0|^2\right] + \sum_{k=0}^{n-1}\left(M_{t_{k+1}\wedge\tau_N} - M_{t_k\wedge\tau_N}\right)^2 \\
&\leq \sup_k \left|M_{t_k^n\wedge\tau_N} - M_{t_{k-1}^n\wedge\tau_N}\right| E\left(\sum_{k=0}^{n-1}\left|M_{t_{k+1}\wedge\tau_N} - M_{t_k\wedge\tau_N}\right|\right) \\
&\leq \underbrace{\sup_k \left|M_{t_k^n\wedge\tau_N} - M_{t_{k-1}^n\wedge\tau_N}\right|}_{\to 0 \text{ as } n\to\infty} \; |M|_{\mathrm{var},T}
\end{aligned}$$

Hence, by letting $n \to +\infty$, we get $\mathbb{E}(M_{\tau_N}^2) = 0$. This implies $M_{\tau_N} = 0$. Letting now $N \to \infty$, we conclude $M_T = 0$. □

Example 3.19. Let $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ Lipschitz be given. Then define its flow for $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ as the solution to

$$\forall t > s : \dot{X}^{s,x}(t) = b_t \circ X^{s,x}(t) \quad \text{and} \quad \forall t \leq s : X^{s,x}(t) := x.$$

Let $\boldsymbol{\eta}^{s,x} = \delta_{X^{s,x}}$. Then $\boldsymbol{\eta}^{s,x}$ is the only solution to the martingale problem $\mathrm{MP}_s(L_t^0, \delta_x)$ with $L_t^0 := b_t \cdot \nabla$. In particular its variation is 0.

Theorem 3.20. *Let $\{M_t\}_{t\geq 0}$ be a (local) continuous square integrable martingale such that $M_0 = 0$. There is a unique continuous and increasing process denoted $\{\langle M\rangle_t\}_{t\geq 0}$ that satisfies the following properties:*

(i) $\langle M\rangle_0 = 0$;
(ii) The process $\left\{M_t^{\otimes 2} - \langle M\rangle_t\right\}_{t\geq 0}$ *is a martingale.*

Moreover, for every $t \geq 0$ and for every sequence of subdivisions $\Delta_n[0,t]$ such that $|\Delta_n[0,t]| \to 0$ as $n \to \infty$, the following convergence takes place in probability:

$$\lim_{n\to+\infty}\sum_{k=1}^{n}\left(M_{t_k^n} - M_{t_{k-1}^n}\right)^{\otimes 2} = \langle M\rangle_t.$$

The process $\{\langle M\rangle_t\}_{t\geq 0}$ is called the quadratic variation *process of $\{M_t\}_{t\geq 0}$.*

Proof. We can assume that the martingale $(M_t)_{t\geq 0}$ is bounded, because else consider the stopped one exceeding some N and let $N \to \infty$ afterwards. We prove that if $\Delta_n[0,t]$ is a sequence of subdivisions of the interval $[0,t]$ such that $|\Delta_n[0,t]| \to 0$ as $n \to \infty$ then the limit

$$\lim_{n\to+\infty}\sum_{k=1}^{n}\left(M_{t_k^n} - M_{t_{k-1}^n}\right)^{\otimes 2}$$

exists in L^2 and thus in probability.

Toward this goal, we introduce some notations. If $|\Delta[0,T]|$ is a subdivision of the time interval $[0,T]$ and if $(X_t)_{t\geq 0}$ is a stochastic process, then we denote

$$S_t^{\Delta[0,T]}(X) = \sum_{i=0}^{k-1} \left(X_{t_{i+1}} - X_{t_i}\right)^{\otimes 2} + (X_t - X_{t_k})^{\otimes 2},$$

where k is such that $t_k \leq t \leq t_{k+1}$.

An easy computation on conditional expectations shows that if $(X_t)_{t\geq 0}$ is a martingale, then the process $X_t^{\otimes 2} - S_t^{\Delta[0,T]}(X)$, is also a martingale for $t \leq T$. Also, if $\Delta[0,T]$ and $\Delta'[0,T]$ are two subdivisions of the time interval $[0,T]$, we will denote by $\Delta \vee \Delta'[0,T]$ the subdivision obtained by putting together the points $\Delta[0,T]$ and the points of $\Delta'[0,T]$. Let now $\Delta_n[0,T]$ be a sequence of subdivisions of $[0,T]$ such that $|\Delta_n[0,T]| \to 0$ as $n \to \infty$.

Let us show that the sequence $S_T^{\Delta_n[0,T]}(M)$ is a Cauchy sequence in L^2.

Exercise: Show using the fact that the process $S^{\Delta_n[0,T]}(M) - S^{\Delta_p[0,T]}(M)$ is a martingale (as a difference of two martingales) the estimate

$$\begin{aligned} &E\left[\left|S_T^{\Delta_n[0,T]}(M) - S_T^{\Delta_p[0,T]}(M)\right|^2\right] \\ &\leq E\left[\sup_k \left|(M_{s_{k+1}} + M_{s_k} - 2M_{t_l})^{\otimes 2}\right|^2\right]^{1/2} E\left[\left|S_T^{\Delta_n \vee \Delta_p[0,T]}(M)\right|^2\right]^{1/2}, \end{aligned}$$

where s_k are the points of the subdivision $\Delta_n \vee \Delta_p[0,T]$ and for fixed s_k, we denote by t_l the point of $\Delta_n[0,T]$ which is the closest to s_k and such that $t_l \leq s_k \leq t_{l+1}$.

Since the martingale M is assumed to be continuous, when $n, p \to \infty$,

$$E\left[\sup_k \left|(M_{s_{k+1}} + M_{s_k} - 2M_{t_l})^{\otimes 2}\right|^2\right] \to 0.$$

Since M is bounded also $E\left[\left|S_T^{\Delta_n \vee \Delta_p[0,T]}(M)\right|^2\right]$ is bounded and therefore, in the L^2 sense the following convergence holds

$$\langle M \rangle_t = \lim_{n\to\infty} \sum_{k=1}^{n} \left(M_{t_k^n} - M_{t_{k1}^n}\right)^{\otimes 2}.$$

The process $\{M_t^2 - \langle M \rangle_t\}_{t\geq 0}$ is seen to be a martingale because for every n and $T \geq 0$, the process $M_t^2 - S_t^{\Delta_n[0,T]}(M)$ is a martingale for $t \leq T$. Let us now show that the obtained process $\langle M \rangle$ is a continuous process. From Doob's inequality, for $n, p \geq 0$ and $\varepsilon > 0$,

$$P\left(\sup_{0\leq t\leq T} \left|S_t^{\Delta_n[0,T]}(M) - S_t^{\Delta_p[0,T]}(M)\right|^2 \geq \varepsilon\right) \leq \frac{E\left[\left|S_T^{\Delta_n[0,T]}(M) - S_T^{\Delta_p[0,T]}(M)\right|^2\right]}{\varepsilon^2}.$$

From Borel-Cantelli lemma, there exists therefore a sequence n_k such that the sequence of continuous stochastic processes $\left(S_t^{\Delta_{n_k}[0,T]}(M)\right)_{0\leq t\leq T}$ almost surely uniformly converges to the process $(\langle M\rangle_t)_{0\leq t\leq T}$. This proves the existence of a continuous version for $\langle M\rangle$. Finally, to prove that $\langle M\rangle$ is increasing, it is enough to consider a an increasing sequence of subdivisions whose mesh tends to 0. Let us now prove that $\langle M\rangle$ is the unique process such that $M^{\otimes 2}-\langle M\rangle$ is a martingale. Let A and A' be two continuous and increasing stochastic processes such that $A_0=A'_0=0$ and such that $(M_t^{\otimes 2}-A_t)_{t\geq 0}$ and $(M_t^{\otimes 2}-A'_t)_{t\geq 0}$ are martingales. The process $(N_t)_{t\geq 0}=(A_t-A'_t)_{t\geq 0}$ is then seen to be a martingale that has a bounded variation. From Lemma 3.18, this implies that $(N_t)_{t\geq 0}$ is constant and therefore equal to 0 due to its initial condition.

Finally, let $\Delta_n[0,t]$ be a sequence of subdivisions whose mesh tends to 0. We have for every $\varepsilon>0$,

$$P\left(\left|A_t-\sum_{k=1}^n\left(M_{t_k^n}-M_{t_{k-1}^n}\right)^{\otimes 2}\right|\geq\varepsilon\right)\leq P\left(T_N\leq t\right) + P\left(\left|A_t^N-\sum_{k=1}^n\left(M_{t_k^n\wedge T_N}-M_{t_{k-1}^n\wedge T_N}\right)^{\otimes 2}\right|\geq\varepsilon\right).$$

This easily implies the announced convergence in probability of the quadratic variations to A_t. ☐

Theorem 3.21 (L^2-integrals wrt. martingales)**.** *Let $\{M_t\}_{t\geq 0}$ be a continuous square integrable martingale and $X\in L^2(\mathrm{d}\langle M\rangle;\mathbb{R}^d)$, then there exists a unique continuous square integrable local martingale $t\mapsto\int_0^t X_s\,\mathrm{d}M_s$ such that whenever a sequence of left-continuous step functions X^n satisfies*

$$\sum_{n\in\mathbb{N}}\left|E\left[\int_0^n\left(X_s^n-X_s\right)^{\otimes 2}:\mathrm{d}\langle M\rangle_s\right]\right|<\infty, \tag{3.13}$$

then

$$\lim_{n\to\infty}\sup_{0\leq t\leq T}\left|\int_0^T\left(X_s^n-X_s\right)\cdot\mathrm{d}M_s\right|=0, \tag{3.14}$$

almost surely and in $L^2(\mathrm{d}\,\langle M\rangle)$. Moreover,

$$\left\langle t\mapsto\int_0^t X_s\cdot\mathrm{d}M_s\right\rangle_t=\int_0^t X_s^{\otimes 2}:\mathrm{d}\langle M\rangle_s, \tag{3.15}$$

Note, that if $\sigma\in L^2(\mathrm{d}\,\langle M\rangle\,;\mathbb{R}^{d\times d}$, then (3.13) *and* (3.14) *hold with $X^{\otimes 2}$ replaced by $\sigma\sigma^T$ and* (3.15) *becomes*

$$\left\langle t\mapsto\int_0^t\sigma_s\,\mathrm{d}M_s\right\rangle_t=\int_0^t\sigma_s\sigma_s^T\,\mathrm{d}\langle M\rangle_s, \tag{3.16}$$

Theorem 3.22 (Itô formula)**.** *Let a stochastic process be given as $X_t = X_0 + V_t + M_t$, with $V_0 = M_0 = 0$, V_t a continuous adapted process of bounded variation and M_t a local martingale, then for any $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$*

$$\mathrm{d}f(t,X_t) = \partial_t f(t,X_t)\,\mathrm{d}t + \nabla f(t,X_t)\cdot \mathrm{d}X_t + \tfrac{1}{2}\nabla^2 f(t,X_t) : \mathrm{d}\langle X\rangle_t$$

Sketch of Proof. The one-dimensional case is straightforward. First, for some $n \in \mathbb{N}$ take a partition $\Delta_n[0,t]$ and do a Taylor expansion of f to first order in time and second order in space. Then, after rewriting of the second order term in space, one can show the convergence towards $\mathrm{d}\left\langle M\right\rangle_t = \mathrm{d}\left\langle X\right\rangle_t$. □

3.5. **Weak solutions to SDEs and the martingale problem.**

Definition 3.23 (Weak solutions to SDEs)**.** A stochastic differential equation with $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to S_d$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ measurable

$$\mathrm{d}X_t = b(t,X_t)\,\mathrm{d}t + \sigma(t,X_t)\,\mathrm{d}B_t \tag{3.17}$$

has a *weak solution* with initial distribution $\mu \in \mathcal{P}(\mathbb{R}^d)$, it there exists continuous d-dimensional martingales $\{X_t\}_{t\geq 0}$ and $\{B_t\}_{t\geq 0}$ such that

(1) $\{B_t\}$ satisfies $\left\langle B\right\rangle_t = t\,\mathrm{Id}$, i.e. is a Brownian motion[13]
(2) X_0 has law μ
(3) boundedness: $\int_0^t \left(|\sigma(s,X_s)|^2 + |b(s,X_s)|\right)\mathrm{d}s < \infty$, a.s. for all t
(4) (3.17) holds.

Lemma 3.24. *Assume that $\{X_t\}_{t\geq 0}$ together $\{B_t\}_{t\geq 0}$ is a weak solution to (3.17) and $\sigma \in C_b(\mathbb{R}_+ \times \mathbb{R}^d; S_d)$ and $b \in C_b(\mathbb{R}_+ \times R^d; \mathbb{R}^d)$. Moreover, define the operator*

$$L_t := \frac{1}{2}\sigma\sigma^T(t,x) : \nabla^2 + b(t,x)\cdot\nabla,$$

then $\boldsymbol{\eta} = \mathrm{law}\, X \in \mathcal{P}(\Gamma)$ is a solution to $\mathrm{MP}(L_t)$.

Proof. By Itô's formula follows

$$f(t,X_t) - f(0,X_0) = \int_0^t \partial_u f(u,X_u)\,\mathrm{d}u + \int_0^t \nabla f(u,X_u)\,\mathrm{d}X_u + \frac{1}{2}\int_0^t D^2 f(u,X_u) : \mathrm{d}\left\langle X\right\rangle_u .$$

Now, since

$$\mathrm{d}X_s = b(s,X_s)\,\mathrm{d}s + \sigma(s,X_s)\cdot \mathrm{d}B_s,$$

we obtain setting $f(t,x) = x$

$$M_t := X_t - X_0 - \int_0^t b(s,X_s)\,\mathrm{d}s = \int_0^t \sigma(u,X_u)\,\mathrm{d}B_u$$

[13] by Lévy's theorem.

and hence obtain the splitting $X_t = X_0 + V_t + M_t$ with V_t of finite variation. Moreover, M_t is a martingale by Proposition 3.17 and its quadratic variation calculates by (3.16) with $\langle B\rangle_u = s\,\mathrm{Id}$ as

$$\langle M\rangle_t = \int_0^t \sigma\sigma^T(u, X_u)\,\mathrm{d}u.$$

Therewith and by rearranging the first identity of the proof, we obtain

$$f_t \circ e_t - f_0 \circ e_0 - \int_0^T (\partial_u f_u + L_u f_u) \circ e_u \,\mathrm{d}u = -\int_0^t \nabla f(s, X_s) \cdot \mathrm{d}M_s,$$

where the right hand side is a martingale again by Proposition 3.17. Finally the boundedness assumption in $\mathrm{MP}(L_t)$ is satisfied because of the pathwise boundedness assumption (3) in Definition 3.23. □

Remark 3.25. Note, that the Brownian motion $\{B_t\}_{t\geq 0}$, which seems to be essential in defining the weak solution, does not play a role for the solution of $\mathrm{MP}(L_t)$.

Before coming to the reverse statement, let us calculate the quadratic variation of the martingales associated to $\mathrm{MP}(L_t)$.

Proposition 3.26 (Quadratic variation of martingale solution)**.** *Let $\boldsymbol{\eta} \in \mathcal{P}(\Gamma)$ be a solution to $\mathrm{MP}(L_t)$. Then for $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ the quadratic variation of the $\boldsymbol{\eta}$-martingale $\left\{M^f\right\}_{t\geq s}$ after time s as defined in (3.7) is given by*

$$\langle M^f\rangle_t = \int_s^t \langle \nabla f_u, a_u \nabla f_u\rangle \circ e_u \,\mathrm{d}u. \tag{3.18}$$

Proof. For the proof, we omit $\circ e_t$ in the notation and we set $s = 0$ and omit after time s. Now, since for all $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d) : M_t^f$ is a $\boldsymbol{\eta}$-martingale, then also $M_t^{f^2}$ is a $\boldsymbol{\eta}$-martingale. Moreover, we set $\mathcal{L}_u = \partial_u + L_u$. Then, we observe

$$\begin{aligned}
\left(M_t^f\right)^2 &= f_t^2 - 2f_t \int_0^t \mathcal{L}_u f_u \,\mathrm{d}u + \left(\int_0^t \mathcal{L}_u f_u \,\mathrm{d}u\right)^2 \\
&= f_t^2 + 2\int_0^t \mathcal{L}_u f_u \left(\int_u^t \mathcal{L}_r f_r \,\mathrm{d}r - f_t\right) \mathrm{d}u \\
&= M_t^{f^2} + \int_0^t \mathcal{L}_u f_u^2 \,\mathrm{d}u + 2\int_0^t \mathcal{L}_u f_u \left(\int_u^t \mathcal{L}_r f_r \,\mathrm{d}r - f_t\right) \mathrm{d}u.
\end{aligned}$$

Now, we use the identity

$$\mathcal{L}_u f_u^2 = 2 f_u \mathcal{L}_u f_u + \langle \nabla f_u, a_u \nabla f_u\rangle,$$

which shows, that

$$M_t^{f^2} = (M_t^f)^2 - \langle M^f\rangle_t + 2\int_0^t \mathcal{L}_u f_u \left(f_t - f_u - \int_u^t \mathcal{L}_r f_r \,\mathrm{d}r\right) \mathrm{d}u.$$

It is left to show, that

$$\int_0^t \mathcal{L}_u f_u \left(f_t - f_u - \int_u^t \mathcal{L}_r f_r \,\mathrm{d}r \right) \mathrm{d}u = \int_0^t \mathcal{L}_u f_u \left(M_t^f - M_u^f \right) \mathrm{d}u$$

is a martingale. Therefore, we consider for $r \in (0,t)$

$$E\left[\int_0^t \mathcal{L}_u f_u \left(M_t^f - M_u^f \right) \mathrm{d}u \mid \mathcal{F}_r \right] = \left(\int_0^r + \int_r^t \right) \mathcal{L}_u f_u E\left[M_t^f - M_u^f \mid \mathcal{F}_r \right] \mathrm{d}u$$
$$= \int_0^r \mathcal{L}_u f_u \left(M_r^f - M_u^f \right) \mathrm{d}u.$$

The conclusion follows from Theorem 3.20. □

Theorem 3.27 (Well-posedness via the SDE)**.** *Assume for simplicity* $\sigma \in C_b(\mathbb{R}_+ \times \mathbb{R}^d; S_d)$*,* $\sigma^{-1} \in C_b(\mathbb{R}_+ \times \mathbb{R}^d; S_d)$ $b \in C_b(\mathbb{R}_+ \times \mathbb{R}^d; \mathbb{R}^d)$*. Assume that* $\boldsymbol{\eta} \in \mathcal{P}(\Gamma)$ *is solution to* $\mathrm{MP}(L_t)$ *with* $L_t = \frac{1}{2}\sigma_t \sigma_t^T : D^2 + b_t \cdot \nabla$*. The there exists a weak solution* $\{X_t\}_{t\geq 0}$ *and* $\{B_t\}_{t \geq 0}$ *to* (3.17) *in the sense of Definition 3.23 such that* law $X = \boldsymbol{\eta}$*.*

Proof. By the definition of solutions to $\mathrm{MP}(L_t)$ follows that

$$f_t \circ e_t - f_0 \circ e_0 - \int_0^t (\partial_u + L_u) f_u \circ e_u \,\mathrm{d}u$$

is a $\boldsymbol{\eta}$-martingale. In particular, by setting $f(x) = x$, we get that

$$M_t := e_t - e_0 - \int_0^t b_u \circ e_u \,\mathrm{d}u$$

is a martingale. We can use (3.18) to find its quadratic variation, i.e. set $f_i(x) = x_i$ for $i = 1, \ldots, d$ and hence

$$\langle M \rangle_t = \int_0^t \sigma_u \sigma_u^T \circ e_u \,\mathrm{d}u.$$

Now, we check that B_t defined by

$$B_t := \int_0^t \sigma_u^{-T} \circ e_u \,\mathrm{d}M_u$$

is a Brownian motion. This is indeed that case, because of (3.16) follows

$$\langle B \rangle_t = \int_0^t \sigma_u^{-T} \sigma_u^{-1} \circ e_u \,\mathrm{d} \langle M_u \rangle = t \,\mathrm{Id}\,.$$

The boundedness (3) in Definition 3.23 is consequence of the boundedness assumption on σ and b. □

22.15.15

4. From the Fokker-Planck equation to MP via superposition

This section is mainly based on [9, 15].

4.1. **Definition and basic properties.** Let us recall the form of the diffusion operator L_t and its formal adjoint for $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$.

$$L_t f := \tfrac{1}{2} a_t : \nabla^2 f + b_t \cdot \nabla f \qquad \text{and} \qquad L_t^* f := \tfrac{1}{2} \nabla^2 (a_t f_t) - \nabla \cdot (b_t f_t).$$

Definition 4.1 (Weak solutions to FPE)**.** A Borel curve $(\nu_t)_{t\in[0,T]} \in \mathcal{M}(\mathbb{R}^d)$ is a weak solution of the Fokker-Planck equation

$$\partial_t \nu_t = L_t^* \nu_t, \qquad \text{on } (0,T) \times \mathbb{R}^d \tag{4.1}$$

if

$$\int_0^T \int (|a_t| + |b_t|) \, \mathrm{d}|\nu_t| \, \mathrm{d}t < \infty. \tag{4.2}$$

and for every $f \in C_c^{1,2}\big((0,T) \times \mathbb{R}^d\big)$ holds

$$\int_0^T \int (\partial_t f(t,x) + L_t f(t,x)) \, \nu_t(\mathrm{d}x) \, \mathrm{d}t = 0 \tag{4.3}$$

In the following by a solution to FPE(L_t), a weak solution in the above sense is understood.

To be able to speak about initial values, we need to show some continuity properties of the curve $t \mapsto \nu_t$.

Lemma 4.2 (Weak*-continuity)**.** *Let $\{\nu_t\}_{t\in[0,T]}$ be a solution of (4.1), then there exists a unique weakly*-continuous representative $\{\tilde{\nu}_t\}_{t\in[0,T]}$ with $\tilde{\nu}_t = \nu_t$ for a.e. $t \in [0,T]$, i.e. it holds for all $t \in [0,T]$ and all $f \in C_b([0,T] \times \mathbb{R}^d)$*

$$\lim_{s\to t} \int f_s \, \mathrm{d}\nu_s = \int f_t \, \mathrm{d}\tilde{\nu}_t.$$

Especially, it holds for $f \in C_b^{1,2}\big([0,T] \times \mathbb{R}^d\big)$ and $0 \le t_1 < t_2 \le T$

$$\int f_{t_2} \, \mathrm{d}\nu_{t_2} - \int f_{t_1} \, \mathrm{d}\nu_{t_1} = \int_{t_1}^{t_2} \int (\partial_t f(t,x) + L_t f(t,x)) \, \nu_t(\mathrm{d}x) \, \mathrm{d}t. \tag{4.4}$$

Proof. Let us take $f(t,x) = g(t)h(x)$ for some $g(t) \in C_c^1([0,T])$ and $h \in C_c^2(\mathbb{R}^d)$, then (4.3) becomes

$$-\int_0^T g'(t) \int_{\mathbb{R}^d} h \, \mathrm{d}\nu_t \, \mathrm{d}t = \int_0^T g(t) \int_{\mathbb{R}^d} L_t h \, \mathrm{d}\nu_t \, \mathrm{d}t.$$

Therefore, the map $t \mapsto \nu_t(h) := \int h \, \mathrm{d}\nu_t$ weakly differentiable with distributional derivative

$$\dot{\nu}_t(h) = \int_{\mathbb{R}^d} L_t h \, \mathrm{d}\nu_t$$

satisfying

$$|\dot{\nu}_t(h)| \le \|h\|_{C^2(\mathbb{R}^d)} V(t) \qquad \text{with} \qquad V(t) := \int \left(\tfrac{1}{2} |a_t| + |b_t|\right) \mathrm{d}\nu_t. \tag{4.5}$$

Since ν_t is a solution to the FPE we get $V \in L^1([0,T])$ by (4.2). Now, let L_h be the set of the Lebesgue points of $\nu_t(h)$, then $[0,T]\setminus L_h$ is a null-set. Let $\mathcal{C} \subset C_c^2(\mathbb{R}^d)$ be countable dense, wrt. the usual C^2-norm, subset. Then, we can define $L_{\mathcal{C}} := \bigcap_{h\in\mathcal{C}} L_h$, which is still a co-null set in $[0,T]$. By (4.5), we find for $s \le t$ with $s,t \in L_{\mathcal{C}}$

$$|\nu_t(h) - \nu_s(h)| \le \|h\|_{C^2} \int_s^t V(u)\,\mathrm{d}u.$$

Hence, by Arzela-Ascoli the curve $t \mapsto \nu_t$ extends uniquely to a continuous curve $\{\tilde\nu_t\}_{t\in[0,T]}$ in $C_c^2(\mathbb{R}^d)'$. We still have to show, that $\{\mu_t\}_{t\in[0,T]}$ is also tight, such that the extension is indeed a weakly*-continuous curve in $\mathcal{P}(\mathbb{R}^d)$. For this, we introduce for $R \ge 1$ a smooth cut-off function $\chi_R : \mathbb{R}^d \to [0,1]$ such that

$$\chi_R(x) = \begin{cases} 1 & ,|x| \le R \\ 0 & ,|x| \ge R+1 \end{cases} \quad \text{such that} \quad |\nabla \chi_R| \le 4 \quad \text{and} \quad \left|\nabla^2 \chi_R\right| \le 4.$$

Since $L_t\chi_R(x) = 0$ for $|x| \le R$, we get from (4.5) the estimate

$$|\nu_t(\chi_R) - \nu_s(\chi_R)| \le A_R := 4\int_0^T \int_{R<|x|\le R+1} \left(\tfrac{1}{2}\,|a_u| + |b_u|\right) \mathrm{d}\nu_u\,\mathrm{d}u$$

and

$$\sum_{k=1}^\infty A_k = \int_0^T V(t)\,\mathrm{d}t < \infty.$$

Since for any $s \in L_{\mathcal{C}}$, the measure ν_s is tight, we find for $\varepsilon > 0$ a $k \in \mathbb{N}$ such that $\nu_s(\chi_k) > 1 - \varepsilon/2$ and $A_k < \varepsilon/2$. This shows for all $t \in L_{\mathcal{C}}$

$$\nu_t\left(\overline{B_{k+1}(0)}\right) \ge \nu_t(\chi_k) \ge 1 - \varepsilon.$$

The representation (4.4) follows in two steps. First, let us consider for a test function $f \in C_b^{1,2}([0,T]\times\mathbb{R}^d)$ with $f_t \in C_c^2(\mathbb{R}^d)$ for all $t \in [0,T]$ the function $f_\varepsilon(t,x) := \eta_\varepsilon(t) f(t,x)$ with $\eta_\varepsilon \in C_c^\infty((t_1,t_2))$ such that

$$0 \le \eta_\varepsilon(t) \le 1, \qquad \lim_{\varepsilon\to0} \eta_\varepsilon(t) = \mathbb{1}_{(t_1,t_2)}(t), \qquad \lim_{\varepsilon\to0} \eta_\varepsilon' = \delta_{t_1} - \delta_{t_2} \quad \text{in } C([0,T])'.$$

Then, we obtain from (4.3)

$$\begin{aligned} 0 &= \int_0^T \int \left(\partial_t\left(\eta_\varepsilon f\right) + L_t\left(\eta_\varepsilon f\right)\right) \mathrm{d}\nu_t\,\mathrm{d}t \\ &= \int_0^T \eta_\varepsilon(t) \int \left(\partial_t f_t + L_t f_t\right) \mathrm{d}\nu_t\,\mathrm{d}t + \int_0^T \eta_\varepsilon'(t) \int f_t\,\mathrm{d}\tilde\nu_t\,\mathrm{d}t. \end{aligned}$$

We can pass to the limit in the last integral due to the weak* continuity of $\tilde\nu_t$.

Now, we conclude by density as follows. For $R > 0$, define the cutoff-function $\chi_R : \mathbb{R}^d \to [0,1]$

$$\chi_R(x) = \begin{cases} 1 & , |x| \le R \\ 0 & , |x| \ge 2R \end{cases} \quad \text{such that} \quad |\nabla \chi_R| \le 4R^{-1} \quad \text{and} \quad \left|\nabla^2 \chi_R\right| \le 4R^{-2}. \tag{4.6}$$

Then for $f \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$, we set $f_t^R := \chi_R f_t \in C_c^2(\mathbb{R}^d)$ for all t and we can apply (4.4) by the previous step. Moreover, we get

$$L_t f_t^R = (L_t f_t)\chi_R + f_t L_t \chi_R + \langle \nabla f_t, a_t \nabla \chi_R \rangle$$

and hence

$$\left|f_t f_t^R\right| \le |L_t f_t| + |f_t| \, |L_t \chi_R| + |a_t| \, |\nabla f_t| \, |\nabla \chi_R| \le C \, \|f_t\|_{C^2(\mathbb{R}^d)} \left(|a_t| + |b_t|\right).$$

By dominated convergence, (4.4) also holds for $f \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$. ☐

Remark 4.3. The Lemma 4.2 allows to speak about initial values, i.e. we say the $t \mapsto \nu_t$ solves the FPE with initial distribution $\bar{\nu}_0 \in \mathcal{P}(\mathbb{R}^d)$, if for all $\varphi \in C_c(\mathbb{R}^d)$ holds

$$\lim_{t \to 0} \int \varphi \, \mathrm{d}\nu_t = \int \varphi \, \mathrm{d}\bar{\nu}_0$$

Moreover, it is a straightforward application of definitions to show that solutions to the martingale problems give rise to solutions of the FPE.

Lemma 4.4 (From MP to FPE)**.** *Let $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ be a solution to* $\mathrm{MP}(L_t)$*. Then $\nu_t := (e_t)_\sharp \boldsymbol{\eta}$ is a solution to the Fokker-Planck equation* (4.1)*.*

Proof. Let $f \in C_c^{1,2}([0,T] \times \mathbb{R}^d)$, then for any $0 \le s \le t \le T$ holds by the definition of the martingale problem (3.7)

$$E^{\boldsymbol{\eta}}\left[\varphi \circ e_t\right] - E^{\boldsymbol{\eta}}\left[\varphi \circ e_s\right] = E^{\boldsymbol{\eta}}\left[\int_s^t L_u \varphi \circ e_u \, \mathrm{d}u\right].$$

By the definition of the pushforward and setting $s = 0$ and $t = T$, we get (4.1). Moreover, the boundedness assumption (3.6) directly translates to (4.2). ☐

We want to show the converse statement, i.e. every solution of the FPE gives rise to a solution to MP. We will do this in several steps with weaker and weaker assumptions on the coefficients.

4.2. **Superposition for bounded smooth coefficients.** The base case are bounded smooth coefficients. In detail, smooth means that the a and b have two spatial derivative, which are bounded uniformly on $\mathbb{R}^d$. In time, we only need a L^1 type bound, i.e.

$$\int_0^T \|a_t\|_{C^2(\mathbb{R}^d)} + \|b_t\|_{C^2(\mathbb{R}^d)} \, \mathrm{d}t < \infty. \tag{4.7}$$

The a priori estimates, which we obtained from the maximum principle Theorem 2.23 are valid under this slightly weaker assumption on the coefficient. Indeed, by revising the proof of Theorem 2.30, one can show

Corollary 4.5 (Quantitative a priori estimate). *Let $f \in C_b^{1,2}([0,T]\times\mathbb{R}^d)$, $\bar f \in C_b^2(\mathbb{R}^d)$ and $g \in C_c^\infty([0,T]\times\mathbb{R}^d)$ such that*

$$\partial_s f(s,x) + L_s f(s,x) = g(s,x) \quad \textit{for } (s,x) \in (0,T)\times\mathbb{R}^d \quad \textit{and} \quad \lim_{t\to T} f(s,x) = \bar f(x),$$

then there exists a constant C only depending on d such that

$$\sup_{t\in[0,T]} \|f_t\|^{(2)} \le C\left(\left\|\bar f\right\|^{(2)} + T\sup_{0\le t\le T}\|g_t\|^{(2)}\right)\exp\left(\int_0^T\left(\|a_t\|^{(2)} + \|b_t\|^{(2)}\right)\mathrm{d}t + T\sup_{0\le t\le T}\|g_t\|^{(2)}\right), \tag{4.8}$$

where $\|f\|^{(2)} := \sup_{x\in\mathbb{R}^d}\max\{|f(x)|, |\nabla f(x)|, |\nabla^2 f(x)|\}$.

Proof. Tracking back all the constants occurring in the final estimate (2.28) of the proof of Theorem 2.30, shows that C_1, C_2 and C_3 can be estimated by $C\left(1 + \|a_s\|^{(l)} + \|b_s\|^{(l)}\right)$ with C a combinatorial factor depending only on d and l. Then, a finite induction with $l = 1, 2$ and an application the bound (2.16) from the weak maximum principle Theorem 2.23 shows (4.8). □

Theorem 4.6 (Uniqueness). *Let $a : [0,T]\times\mathbb{R}^d \to S_d$ and $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be Borel maps satisfying*

$$\int_0^T \|a_t\|_{C^2(B)} + \|b_t\|_{C^2(B)}\,\mathrm{d}t < \infty, \quad \textit{for every bounded open } B \subset \mathbb{R}^d. \tag{4.9}$$

Let $\{\nu_t\}_{t\in[0,T]} \subset \mathcal{M}(\mathbb{R}^d)$ is a weakly-continuous solution to FPE (4.1). If $\nu_0 \le 0$ then $\nu_t \le 0$ for all $t \in [0,T]$. Thus, for any $\bar\nu \in \mathcal{M}(\mathbb{R}^d)$ exists at most one weakly*-continuous solution $\{\nu_t\}_{t\in[0,T]}$ with $\nu_0 = \bar\nu$.*

Proof. Let $g \in C_c^\infty([0,T]\times\mathbb{R}^d)$ with $g \ge 0$. It is enough to show that $\int_0^T\int g_t\,\mathrm{d}\nu_t\,\mathrm{d}t \le 0$. Fix $R \ge 1$ large enough such that $\operatorname{supp} g \subset B_R(0)$. Then, let χ_R be the cut-off function from (4.6). Therewith, set $a^R = a\chi_R$ and $b^R = b\chi_R$, then (4.9) becomes (4.7). Since L_t is a local operator, we have that $L_t f = L_t^R f$ on $(0,T)\times B_R(0)$ for any $f \in C_b^{0,2}((0,T)\times\mathbb{R}^d)$. For $\varepsilon > 0$ let $a^{R,\varepsilon}$ and $b^{R,\varepsilon}$ be a space-time mollification of a^R and b^R, respectively. Let $L^{R,\varepsilon}$ be the associated diffusion operator, which now has smooth coefficients satisfying (4.7) uniformly in ε. Let $f^{R,\varepsilon}$ be the solution to

$$\partial_t f_t^{R,\varepsilon} + L_t f_t^{R,\varepsilon} = g, \qquad f_T^{R,\varepsilon} = 0,$$

which by the weak maximum principle is non-negative since g is non-negative. Moreover, we can take $f^{R,\varepsilon}\chi_R$ as a test function in the weak formulation (4.3). Therewith and by noting that $f^\varepsilon \leq 0$ as well as $\nu_0 \leq 0$, we can calculate

$$
\begin{aligned}
0 &\geq -\int f_0^{R,\varepsilon}\chi_R \,\mathrm{d}\nu_0 = \int_0^T \int \left(\chi_R \partial_t f_t^{R,\varepsilon} + L_t\left(f_t^{R,\varepsilon}\chi_R\right)\right) \mathrm{d}\nu_t \,\mathrm{d}t \\
&= \int_0^T \int \left(\chi_R g - \chi_R L_t^{R,\varepsilon} f_t^{R,\varepsilon} + L_t\left(f_t^{R,\varepsilon}\chi_R\right)\right) \mathrm{d}\nu_t \,\mathrm{d}t \\
&= \int_0^T \int \left(\chi_R\left(g - L_t^{R,\varepsilon} f_t^{R,\varepsilon} + L_t f_t^{R,\varepsilon}\right) + f_t^{R,\varepsilon} L_t \chi_R + \left\langle \nabla f_t^{R,\varepsilon}, a\nabla\chi_R\right\rangle\right) \mathrm{d}\nu_t \,\mathrm{d}t \\
&\geq \int_0^T \int g \,\mathrm{d}\nu_t \,\mathrm{d}t \\
&\quad - \sup_{t\in[0,T]} \left\|f_t^{R,\varepsilon}\right\|_{C^2(\mathbb{R}^d)} \int_0^T \int \left(\chi_R\left(\left|a^{R,\varepsilon} - a\right| + \left|b^{R,\varepsilon} - b\right|\right) + |L_t\chi_R| + |a|\,|\nabla\chi_R|\right) \mathrm{d}|\nu_t| \,\mathrm{d}t.
\end{aligned}
$$

Letting $\varepsilon \to 0$: Since $a^R = a$ and $b^R = b$ on $B_R(0)$ and $\sup_{t\in[0,T]} \left\|f_t^{R,\varepsilon}\right\|_{C^2(\mathbb{R}^d)}$ is uniformly bounded in ε by (4.8), the second integral converges to

$$\int_0^T \int \left(|L_t\chi_R| + |a|\,|\nabla\chi_R|\right) \mathrm{d}|\nu_t| \,\mathrm{d}t \leq \frac{4}{R}\int_0^T \int \left(|a_t| + |b_t|\right) \mathrm{d}\nu_t \,\mathrm{d}t.$$

The conclusion follows from the boundedness assumption by letting $R \to \infty$. □

Theorem 4.7 (Well-posedness via the FPE)**.** *Let $a : [0,T]\times\mathbb{R}^d \to S_d$ and $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be Borel maps satisfying (4.7). Then, for every $\bar\nu \in \mathcal{P}(\mathbb{R}^d)$, there exists $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d)$ solving $\mathrm{MP}(L_t)$ with coefficients a, b such that $(e_0)_\sharp\boldsymbol{\eta} = \bar\nu$.*

Proof. The proof follows along the very same lines as the proof of Theorem 2.32. We, first consider a^ε and b^ε as space-time mollifications of the coefficients a, b, respectively such that

(a) $a^\varepsilon, b^\varepsilon \in C_b^\infty$
(b) $\langle\theta, a^\varepsilon\theta\rangle \geq \alpha^\varepsilon\,|\theta|^2$
(c) $\int_0^T \sup_{x\in\mathbb{R}^d}\left(|a^\varepsilon(t,x) - a(t,x)| + |b^\varepsilon(t,x) - b(t,x)|\right)\mathrm{d}t \to 0$ as $\varepsilon \to 0$
(d)
$$\int_0^T \|a_t^\varepsilon\|_{C^2(\mathbb{R}^d)} + \|b_t^\varepsilon\|_{C^2(\mathbb{R}^d)}\,\mathrm{d}t < \infty \qquad \text{uniformly in } \varepsilon.$$

Now, using the improved a priori estimate (4.8), instead of the one from Theorem 2.30, we can redo the proof of of Theorem 2.20 to obtain a TPF P, which by Lemma 3.5 gives rise to a solution to the martingale problem. □

12.01.16

Theorem 4.8 (Superposition for bounded smooth coefficients)**.** *Let $a : [0,T]\times\mathbb{R}^d \to S_d$ and $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be Borel maps satisfying (4.9). Then, the superposition principle holds for every solution $\{\nu_t\}_{t\in[0,t]} \subset \mathcal{P}(\mathbb{R}^d)$ of $\mathrm{FPE}(L_t)$, i.e. there exists $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ solving $\mathrm{MP}(L_t)$ such that $\nu_t = (e_t)_\sharp\boldsymbol{\eta}$.*

Proof. By Lemma 4.2, we find a unique weakly*-continuous representative $\tilde\nu_t$, which provides a initial value $\tilde\nu_0$. Then by Theorem 4.7, we get a solution $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d)$ to $\mathrm{MP}(L_t)$ starting from $\tilde\nu_0$. Moreover, by Lemma 4.4 follows that $(e_t)_\sharp\boldsymbol{\eta}$ is a solution to $\mathrm{FPE}(L_t)$ starting from $\tilde\nu_0$. The conclusion follows from the uniqueness Theorem 4.6. □

4.3. **Approximation methods for coefficients.**

Lemma 4.9 (Push forward of generator)**.** *Let $L_t := \frac{1}{2}a_t : \nabla^2 + b_t\cdot\nabla$, let $\pi \in C^2(\mathbb{R}^d;\mathbb{R}^d)$ with uniformly bounded first and second derivatives and let $(\nu_t)_{t\in[0,T]}$ be a solution to* $\mathrm{FPE}(L_t)$*. Define the push-forward coefficients by*

$$\pi(a)_t^{i,j}(x) := E^{\nu_t}\left[\left\langle \nabla\pi^i, a_t\nabla\pi^j\right\rangle \mid \pi = x\right] = \frac{d\pi_\sharp\left(\langle\nabla\pi^i, a_t\nabla\pi^j,\nu\rangle_t\right)}{d\pi_\sharp\nu_t}(x)$$

and

$$\pi(b)_t^{i}(x) := E^{\nu_t}\left[L_t(\pi^i) \mid \pi = x\right] = \frac{d\pi_\sharp\left(L_t(\pi^i)\nu_t\right)}{d\pi_\sharp\nu_t}(x).$$

Then $\pi(L)_t := \frac{1}{2}\pi(a)_t : \nabla^2 + \pi(b)_t\cdot\nabla$ is a diffusion operator on $\mathbb{R}^d$ and $\pi_\sharp\nu := (\pi_\sharp\nu_t)_{t\in[0,T]}$ is a solution of $\mathrm{FPE}(\pi(L)_t)$*, i.e. $\partial_t\pi_\sharp\nu_t = \pi(L)_t^*\pi_\sharp\nu_t$.*

Proof. Let's consider the weak formulation of $\mathrm{FPE}(L_t)$ for any $\varphi \in C_c^{1,2}([0,T]\times\mathbb{R}^d)$ we have that $\varphi_t\circ\pi$ is again an admissible testfunction and we have

$$L_t\left(\varphi_t\circ\pi\right) = \frac{1}{2}\sum_i \partial_i\varphi_t\circ\pi\,\left(a : D^2\pi^i + b\cdot\nabla\pi^i\right) + \frac{1}{2}\sum_{i,j}\partial_{i,j}\varphi_t\circ\pi\,\left\langle\nabla\pi^i, a\nabla\pi^j\right\rangle$$

Therewith, we get

$$0 = \int_0^T\int\left(\partial_t\varphi_t\circ\pi + L_t(\varphi_t\circ\pi)\right)\mathrm{d}\nu_t\,\mathrm{d}t = \int_0^T\int\left(\partial_t\varphi_t + \pi(L)_t\varphi_t\right)\mathrm{d}\pi_\sharp\nu_t\,\mathrm{d}t.$$

The boundedness follows similar by using the boundedness of φ in $C^2(\mathbb{R}^d;\mathbb{R}^d)$. □

Lemma 4.10 (Mollification by convolutions)**.** *Let $L_t := \frac{1}{2}a_t : \nabla^2 + b_t\cdot\nabla$, let $\rho\in\mathcal{P}(\mathbb{R}^d)$ be smooth and let $(\nu_t)_{t\in[0,T]}$ be a solution to* $\mathrm{FPE}(L_t)$*. Therewith, define the mollified coefficients a^ρ and b^ρ as*

$$(a^\rho)_t := \frac{\mathrm{d}(a\nu_t)*\rho}{\mathrm{d}\nu_t*\rho} \quad \textit{and} \quad (b^\rho)_t := \frac{\mathrm{d}\,(b\nu_t)*\rho}{\mathrm{d}\nu_t*\rho}. \tag{4.10}$$

Then, $(\nu_t*\rho)_{t\in[0,T]}$ is a solution of $\mathrm{FPE}(L_t^\rho)$ *with $L_t^\rho := \frac{1}{2}a_t^\rho : \nabla^2 + b_t^\rho\cdot\nabla$.*

Proof. Let $f\in C_b^{1,2}([0,T]\times\mathbb{R}^d)$ be a test function, then since gradients and convolutions commute, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int f\,\mathrm{d}\left(\nu*\rho\right) = \int\partial_t f*\rho\,\mathrm{d}\nu = -\int L\left(f*\rho\right)\mathrm{d}\nu = -\int L^\rho f\,\mathrm{d}\left(\nu*\rho\right).$$

□

Lemma 4.11 (Regularity of mollified measures). *Let ρ be a smooth probability density on $\mathbb{R}^d$ with $\rho > 0$ and $|\nabla^i \rho| \le C\rho$ for $i \in \{1, \dots, k\}$ and some $C > 0$. Let $\mu, \nu \in \mathcal{M}^+(\mathbb{R}^d)$ be a non-negative Borel measure on $\mathbb{R}^d$ with $\mu \ll \nu$. Then $\mu * \rho \ll \nu * \rho$ and*

$$\mathbb{R}^d \ni x \mapsto \frac{\mathrm{d}\mu * \rho}{\mathrm{d}\nu * \rho}(x) = \frac{\int \rho(x-y)\, \mu(\mathrm{d}y)}{\int \rho(x-y)\, \nu(\mathrm{d}x)} \in C^k(\mathbb{R}^d).$$

Moreover, for any convex lsc. $\Theta : \mathbb{R} \to [0, \infty]$ holds

$$\int \Theta\left(\frac{\mathrm{d}\mu * \rho}{\mathrm{d}\nu * \rho}\right) \mathrm{d}\nu * \rho \le \int \Theta\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\nu. \tag{4.11}$$

In particular for $(\mu_t)_{t\in[0,T]} \subset \mathcal{M}^+(\mathbb{R}^d)$ a Borel curve and $(\nu_t)_{t\in[0,T]} \subset \mathcal{P}(\mathbb{R}^d)$ a weak continuous curve such that $\mu_t \ll \nu_t$ for every $t \in [0,T]$ it holds in addition*

$$\sup_{t\in[0,T]} \left\| \frac{\mathrm{d}\mu * \rho}{\mathrm{d}\nu * \rho} \right\|_{C^k(B)} < \infty \quad \text{for any bounded } B \subset \mathbb{R}^d. \tag{4.12}$$

Proof. The regularity properties are straightforward to proof and left as Exercise. For the proof of (4.11), we note that $\Psi : \mathbb{R}_+^2 \to \mathbb{R}_+$ defined by $\Psi(a,b) = \Theta(a/b)b$ is jointly convex and 1-homogeneous. Now, for a 1-homogeneous convex function, the following form of Jensen's inequality is true for any measure $m \in \mathcal{M}_+(\mathbb{R}^d)$

$$\Psi\left(\int f \,\mathrm{d}m, \int g \,\mathrm{d}m\right) \le \int \Psi(f,g) \,\mathrm{d}m.$$

Therewith, we can estimate

$$\begin{aligned} \Theta\left(\frac{\mathrm{d}\mu * \rho}{\mathrm{d}\nu * \rho}(x)\right) \nu * \rho(x) &= \Psi\left(\int \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(y)\rho(x-y)\nu(\mathrm{d}y), \int \rho(x-y)\nu(\mathrm{d}y)\right) \\ &\le \int \Psi\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(y), 1\right) \rho(x-y)\nu(\mathrm{d}y) \\ &= \int \Theta\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(y)\right) \rho(x-y)\nu(\mathrm{d}y). \end{aligned}$$

An integration in x gives the desired result. □

4.4. **Tightness.** We will now show that the statement of Kolmogorov's continuity Theorem 2.16 about Hölder continuity of the canonical process can translated to martingales. The crucial ingredient are Burkholder-Davis-Gundy inequalities.

Theorem 4.12 (Burkholder-Davis-Gundy inequality [13, Theorem 42.1]). *Let $\Theta : \mathbb{R}_+ \to [0, \infty]$ be of moderate growth, i.e. there exists $\alpha > 1$ such that*

$$\sup_{x>0} \frac{\Theta(\alpha x)}{\Theta(x)} < \infty.$$

Then, there exist universal constants c_Θ, C_Θ, such that for any continuous local martingale holds

$$c_\Theta E\left[\Theta\left(\sqrt{\langle M\rangle_T}\right)\right] \leq E\left[\Theta\left(M_T^*\right)\right] \leq C_\Theta E\left[\Theta\left(\sqrt{\langle M\rangle_T}\right)\right],$$

where $M_T^* = \sup_{0\leq t\leq T}|M_t|$.

The basic function of moderate growth are power function $x \mapsto x^p$ for any $p > 0$.

Theorem 4.13. *Let $\theta, \Theta_1, \Theta_2 : \mathbb{R}_+ \to \mathbb{R}_+$ be function with Θ_1, Θ_2 monotone increasing, convex, lsc.,*

$$\lim_{x\to\infty}\theta(x) = \lim_{x\to\infty}\frac{\Theta_1(x)}{x} = \lim_{x\to\infty}\frac{\Theta_2(x)}{x} = \infty \tag{4.13}$$

and Θ_2 is of moderate growth .Then, there exists some coercive[14] *function $\Psi : C([0,T];\mathbb{R}) \to [0,+\infty]$ such that for any progressively measurable processes φ, α, β on a filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ such that*

$$[0,T] \ni t \mapsto M_t := \varphi_t - \int_0^t \beta_s \,\mathrm{d}s, \quad \textit{and} \quad [0,T] \ni t \mapsto M_t^2 - \int_0^t \alpha_s \,\mathrm{d}s$$

are $\mathbb{P}$-a.s. continuous local martingales, it holds

$$E\left[\Psi(\varphi)\right] \leq E\left[\theta(\varphi_0) + \int_0^T \left(\Theta_1\left(|\beta_s|\right) + \Theta_2\left(|\alpha_s|\right)\right)\mathrm{d}s\right]. \tag{4.14}$$

Proof. For a paths $\gamma^1, \gamma^2 \in C([0,T];\mathbb{R})$ we denote by $\gamma^1 + \gamma^2$ the path obtained by a pointwise sum. We will prove, that there exists Ψ_1 and Ψ_2 depending on Θ_1 and Θ_2, respectively, such that Ψ defined by

$$\Psi(\gamma) := \theta\left(|\gamma_0|\right) + \inf_{\gamma^1+\gamma^2=\gamma}\left(\Psi_1(\gamma^1) + \Psi_2(\gamma^2)\right) \tag{4.15}$$

satisfies the statement of the theorem.

For $\varepsilon > 0$, let $K_i = K_i(\varepsilon) \in \mathbb{N}$, $i = 1, 2$ be the smallest integer such that for some $\sigma > 0$

$$\frac{\Theta_1\left(\varepsilon K_1/T\right)}{K_1/T} \geq \frac{1}{\varepsilon^\sigma} \quad \text{and} \quad \frac{\Theta_2\left(\varepsilon^2 K_2/T\right)}{K_2/T} \geq \frac{1}{\varepsilon^\sigma}, \tag{4.16}$$

which existence is ensured by (4.13). For the sage of notation, we also introduce $\delta_i := T/K_i$ for $i = 1, 2$. Therewith, we can define the events

$$A_i(\varepsilon) := \left\{\gamma \in C([0,T];\mathbb{R}) : \sup_{k=1,\dots,K_i}\ \sup_{s\in[\delta_i(k-1),\delta_i k]}\left|\gamma_s - \gamma_{\delta_i(k-1)}\right| \leq \varepsilon\right\}.$$

Moreover, we set for some $C_{\Psi_i} > 0$, $i = 1, 2$ to be determined

$$\Psi_i(\gamma) := C_{\Psi_i}\sum_{m\geq 0}(m+1)\mathbb{1}_{A_i(2^{-m})^c}(\gamma).$$

[14]the sub-level sets $\{\Psi < t\}$ are pre-compact

Then, whenever $\Psi_i(\gamma) \le m$ it holds $\gamma \in A_i\left(2^{-n}\right)$ for every $n \ge m$. Therewith, we define Ψ via (4.15). Now, for the coercivity of Ψ let $\gamma \in \{\Psi \le m\}$, then we can decompose $\gamma = \gamma^1 + \gamma^2$ such that γ^i has the following modulus of continuity

$$\omega_{i,m}(h) := \sup_{\substack{x,y\in[0,T] \\ |x-y|\le h}} |\gamma_x - \gamma_y| \le 2 \inf_{n\ge m} \inf_{k=1,\dots,K_i(2^{-n})} \sup_{s\in[\delta_i(k-1),\min\{\delta_i k,\delta_i(k-1)+h\}]} \left|\gamma_s - \gamma_{\delta_i(k-1)}\right|$$
$$\le \begin{cases} 2^{-(n-1)} & , h \in \left[\delta_i\left(2^{-(n+1)}\right), \delta_i\left(2^{-n}\right)\right) \text{ with } n \ge m, \\ K_i\left(2^{-m}\right) 2^{-(m-1)} & , h \ge \delta_i\left(2^{-m}\right). \end{cases} \tag{4.17}$$

Therewith, we can conclude by the Arzela-Ascoli Theorem, that $\{\Psi \le m\}$ is pre-compact in $C([0,T];\mathbb{R})$ and hence Ψ coercive.

Now, we want to show the estimate (4.14). Therefore, we observe that since M_t is a martingale by assumption, we can conclude from Theorem 3.20 that its quadratic variation process is $t \mapsto \int_0^t \alpha_s \,\mathrm{d}s$. Let us define $\gamma_t^1 := \int_0^t \beta_s \,\mathrm{d}s$ and $\gamma_t^2 := M_t$, then we can estimate $E[\Psi(\varphi)]$ by

$$E[\Psi(\varphi)] \le E\left[\theta(\varphi_0) + \Psi_1(\gamma^1) + \Psi_2(\gamma^2)\right]$$
$$\le E\left[\theta(\varphi_0)\right] + \sum_{m\ge 0}(m+1)\left(C_{\Psi_1}\mathbb{P}\left(\gamma^1 \notin A_1\left(2^{-m}\right)\right) + C_{\Psi_2}\mathbb{P}\left(\gamma^1 \notin A_1\left(2^{-m}\right)\right)\right).$$

Now, we have for $i = 1,2$ and writing $\varepsilon = 2^{-m}$

$$\mathbb{P}\left(\gamma^1 \notin A_i(\varepsilon)\right) = \mathbb{P}\left(\sup_{k=1,\dots,K_i} \left(\gamma^i\right)_k^* > \varepsilon\right) \le \sum_{k=1}^{K_i} \mathbb{P}\left(\left(\gamma^i\right)_k^* > \varepsilon\right),$$

where $\left(\gamma^i\right)_k^* := \sup_{s\in[\delta_i(k-1),\delta_i k]} \left|\gamma_s^i - \gamma_{\delta_i(k-1)}^i\right|$.

In the case $i = 1$, we have $|\gamma_s^1 - \gamma_t^1| \le \int_s^t |\beta_r| \,\mathrm{d}r$ and hence by Markov's

$$\mathbb{P}\left(\left(\gamma^1\right)_k^* > \varepsilon\right) \le \mathbb{P}\left(\frac{K_1}{T}\int_{T(k-1)/K_1}^{Tk/K_1} |\beta_r| \,\mathrm{d}r > \frac{K_1\varepsilon}{T}\right)$$
$$\le \frac{1}{\Theta_1\left(\frac{K_1\varepsilon}{T}\right)} E\left[\Theta_1\left(\frac{K_1}{T}\int_{T(k-1)/K_1}^{Tk/K_1} |\beta_r| \,\mathrm{d}r\right)\right]$$
$$\le \varepsilon^\sigma E\left[\int_{T(k-1)/K_1}^{Tk/K_1} \Theta_1\left(|\beta_r|\right) \mathrm{d}r\right], \tag{4.18}$$

where the last estimate follows from Jensen's inequality and the choice of K_1.

Now, the case $i = 2$. The first estimate is again by Markov's inequality

$$\mathbb{P}\left(M_k^* > \varepsilon\right) = \mathbb{P}\left(K_2/T\left(M_k^*\right)^2 > \varepsilon^2 K_2/T\right) \le \frac{1}{\Theta_2\left(\varepsilon^2 K_2/T\right)} E\left[\Theta_2\left(\left(M_k^*\right)^2 K_2/T\right)\right].$$

The rhs. is in a form to which we can apply the Burkholder-Davis-Gundy inequality. Indeed, we can use $[0,T/K_2] \ni s \mapsto \sqrt{K_2/T} M_{s+(k-1)T/K_i}$ as martingale and observe that $x \mapsto \Theta_2(x^2)$ is also of moderate growth. Therewith, we can estimate

$$\begin{aligned} \mathbb{P}\left(M_k^* > \varepsilon\right) &\leq \frac{C_{\Theta_2}}{\Theta_2\left(\varepsilon^2 K_2/T\right)} E\left[\Theta_2\left(\frac{K_2}{T}\int_{T(k-1)/K_2}^{Tk/K_2} |\alpha_r| \,\mathrm{d}r\right)\right] \\ &\leq C_{\Theta_2}\varepsilon^{\sigma} E\left[\int_{T(k-1)/K_2}^{Tk/K_2} \Theta_2\left(|\alpha_r|\right)\mathrm{d}r\right]. \end{aligned} \tag{4.19}$$

Now, we can conclude

$$\begin{aligned} E\left[\Psi\left(\varphi\right)\right] \leq E\left[\theta\left(\varphi_0\right)\right] + \sum_{m\geq 0}(m+1)2^{-\sigma m}\Bigg(& C_{\Psi_1} E\left[\int_0^T \left|\Theta_1\left(\beta_r\right)\right|\mathrm{d}r\right] \\ &+ C_{\Theta_2} C_{\Psi_2} E\left[\int_0^T \left|\Theta_2\left(\alpha_r\right)\right|\mathrm{d}r\right]\Bigg) \end{aligned}$$

from where the thesis follows by suitable choice of C_{Ψ_1} and C_{Ψ_2}. □

Corollary 4.14 (Hölder continuity)**.** *In the setting of Theorem 4.13, let* $\Theta_1(x) = |x|^{p_1}$ *and* $\Theta_2(x) = |x|^{p_2}$*, for* $p_1, p_2 \in (1,\infty)$*. Then for any* $r > 0$ *with* $r < r(p_1,p_2) := \min\left\{1-\frac{1}{p_1}, \frac{1}{2}\left(1-\frac{1}{p_2}\right)\right\}$*, then it holds*

$$\mathbb{P}\left(\limsup_{h\to 0}\sup_{|t-s|\leq h}\frac{|\varphi_t-\varphi_s|}{|t-s|^r} = 0\right) = 1.$$

Hence $\varphi \in C^{0,r}([0,T];\mathbb{R})$ *for* $\mathbb{P}$*-a.e.* φ*.*

Proof. The necessary conditions (4.16) from the proof of Theorem 4.13 become by abbreviating $\delta_i = T/K_i$ for any $\sigma > 0$

$$\left(\varepsilon/\delta_1\right)^{p_1}\delta_1 \overset{!}{\geq} \frac{1}{\varepsilon^\sigma} \quad\text{and}\quad \left(\varepsilon^2/\delta_2\right)^{p_1}\delta_2 \overset{!}{\geq} \frac{1}{\varepsilon^\sigma}$$

This leads to the conditions

$$\delta_1 \leq \varepsilon^{\frac{p_1+\sigma}{p_1-1}} \quad\text{and}\quad \delta_2 \leq \varepsilon^{\frac{2p_2+\sigma}{p_2-1}}$$

for which $\delta_1 = \delta_2 = \varepsilon^{\frac{1}{r}}$ with $r < r(p_1,p_2)$ is sufficient for σ small enough.

Now, from the estimates (4.18) and (4.19), we deduce

$$\begin{aligned} &\sum_{m=0}^{\infty}\mathbb{P}(A_1(2^{-m})^c) \leq \sum_{m=0}^{\infty} 2^{-\sigma m} E\left[\int |\beta_s|^{p_1}\,\mathrm{d}s\right] < \infty \\ \text{and}\quad &\sum_{m=0}^{\infty}\mathbb{P}(A_2(2^{-m})^c) \leq \sum_{m=0}^{\infty} 2^{-\sigma m} E\left[\int |\alpha_s|^{p_2}\,\mathrm{d}s\right] < \infty, \end{aligned}$$

By the Borel-Cantelli-Lemma there exists $\mathbb{P}$-a.s. an $m \geq 0$ such that the curve $(\varphi_t)_{t\in[0,T]}$ can be decomposed into γ^1 and γ^2 having the modulus of continuity $w_{i,m}$ defined in (4.17). Hence, we deduced a Hölder estimate with finite norm. But since, we can choose r slightly larger, we deduce the conclusion. □

19.01.16

Corollary 4.15 (Bound for $\mathrm{MP}(L_t)$)**.** *Let $L_t := \frac{1}{2}a_t : \nabla^2 + b_t \cdot \nabla$ with a, b Borel maps and let $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ be a solution of $\mathrm{MP}(L_t)$. Then in the setting of Theorem 4.13 holds for any $f \in C_b^{1,2}\left([0,T]\times\mathbb{R}^d\right)$*

$$\int \Psi\left(f_t\right) \mathrm{d}\eta_t \leq \int \theta\left(|f_0|\right) \mathrm{d}\eta_0 + \int_0^T \int \left(\Theta_1\left(|L_t f_t|\right) + \Theta_2\left(\langle \nabla f_t, a_t \nabla f_t\rangle\right)\right) \mathrm{d}\eta_t \,\mathrm{d}t,$$

where $\eta_t := (e_t)_\sharp \boldsymbol{\eta}$.

Proof. Since $\boldsymbol{\eta}$ is a solution to $\mathrm{MP}(L_t)$, we have that $M_t := f_t\circ e_t - \int_0^T L_s f_s \,\mathrm{d}s$ is $\boldsymbol{\eta}$-martingale by definition. Moreover $M_t^2 - \int_0^T \langle \nabla f_t, a_t \nabla f_t\rangle \,\mathrm{d}t$ is a $\boldsymbol{\eta}$-martingale by Proposition 3.26. Then, the conclusion follows from Theorem 4.13 by setting $\beta_s := |L_s f_s|$ and $\alpha_s := \langle \nabla f_s, a_s \nabla f_s\rangle$. □

4.5. **Limit.** For this subsection, we assume that we already have constructed a family of probability measures $\boldsymbol{\eta}^n \in \mathcal{P}(\Gamma_T)$ obtained from the superposition principle for a suitable approximating sequence ν^n solving $\mathrm{FPE}(L_t^n)$ for some generator L^n with coefficients converging to the one of L. Moreover, we assume that this family $\{\boldsymbol{\eta}^n\}$ has some limit point $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$.

The question remains how to pass to the limit in the martingale formulation for $\boldsymbol{\eta}^n$, which is not direct, since the limit is taken in the weak formulation and involves the coefficients not necessarily continuous. Therefore, let us recall: $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ is a solution to $\mathrm{MP}(L_t)$ if and only if for all $s,t \in [0,T]$ with $s \leq t$ and all $f \in C_c^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ with $\|f\|_{C^{1,2}} \leq 1$

$$\int \left(f_t \circ e_t - f_s \circ e_s - \int_s^t \left(\partial_r + L_r\right) f_r \circ e_r \,\mathrm{d}r\right) \mathrm{d}\boldsymbol{\eta} = 0.$$

The same identity holds for $\boldsymbol{\eta}^n$ and L_t^n. Hence, we have to show the limit

$$\int\int_s^t L_r^n f_r \circ e_r \,\mathrm{d}r \,\mathrm{d}\boldsymbol{\eta}^n - \int\int_s^t L_r f_r \circ e_r \,\mathrm{d}r \,\mathrm{d}\boldsymbol{\eta} \quad \to \quad 0 \qquad \text{as} \qquad n \to \infty. \tag{4.20}$$

Lemma 4.16 (Limits for push-forward approximation)**.** *Let ν be a solution to $\mathrm{FPE}(L_t)$ with coefficient a_t and b_t. Let $\pi^n \in C_b^2(\mathbb{R}^d;\mathbb{R}^d)$ such that $\pi^n(x) \to x$ uniformly in x such that for all $x \in \mathbb{R}^d$ locally uniformly in n*

$$D\pi^n(x) \to \mathrm{Id}; \qquad D^2\pi^n(x) \to 0 \quad \textit{and} \quad \left|D^i\pi^n(x)\right| \leq C \textit{ for some } C \geq 0 \textit{ and } i \in \{1,2\}.$$

Let $\nu^n = \pi^n_\sharp \nu^n$, and assume $\boldsymbol{\eta}^n$ to be a superposition solution for ν^n with generator $L_t^n := \pi^n(L)_t$ from Lemma 4.9. Then any weak limit point $\boldsymbol{\eta}$ is a superposition of ν solving $\mathrm{MP}(L_t)$.*

Proof. Let $\bar a$ and $\bar b$ some continuous compactly supported coefficients to be chosen later and denote by $\bar L_t$ the according generator. Then by (4.20), it is enough to show

$$\limsup_{n\to\infty} \int \left|L_t^n f - \bar L f\right| \mathrm{d}\nu^n + \int \left|L f - \bar L f\right| \mathrm{d}\pi_\sharp \nu. \tag{4.21}$$

Here, we used that since $\boldsymbol{\eta}$ is weak* limit point of $\boldsymbol{\eta}^n$, the pushforward measure $\nu_t^n = (e_t)_\sharp \boldsymbol{\eta}^n$ also weakly* converge to $\nu = (e_t)_\sharp \boldsymbol{\eta}$. Moreover, since $\bar L f \in C_c([0,T]\times \mathbb{R}^d)$, we also have $\int \bar L f \,\mathrm{d}\nu^n \to \int \bar L f \,\mathrm{d}\nu$.

By the definition of $\pi^n(L)$ from Lemma 4.9 and applying the push-forward, we get

$$\int \left|\pi^n(L) f - \bar L f\right| \mathrm{d}\pi^n_\sharp \nu = \int \left| E^\nu\left[L(f\circ\pi) \mid \pi\right] - (\bar L f)\circ \pi\right| \mathrm{d}\nu.$$

Since $\left(\bar L\right)\circ \pi^n$ is a sole function of π^n alone, we can rewrite and estimate

$$\begin{aligned}\int \left| E^\nu\left[L(f\circ\pi^n) \mid \pi^n\right] - (\bar L f)\circ \pi^n\right| \mathrm{d}\nu &= \int \left| E^\nu\left[L(f\circ\pi^n) - \left(\bar L f\right)\circ \pi \mid \pi^n\right]\right| \mathrm{d}\nu \\ &\le \int \left| L\left(f\circ\pi^n\right) - \left(\bar L f\right)\circ \pi^n\right| \mathrm{d}\nu,\end{aligned}$$

where we used that conditional expectation is reducing the $L^1(\nu)$-norm. Now, by explicitly writing down the difference and neglect for a moment the n for the sake of notation

$$\begin{aligned} L\left(f\circ \pi\right) - \left(\bar L f\right)\circ \pi = \frac12 \sum_{i,j=1}^d \left(\left\langle \nabla \pi^i, a \nabla \pi^j\right\rangle - \bar a^{i,j}\circ \pi\right)\left(\partial_{i,j} f\right)\circ \pi \\ + \sum_{i=1}^d \left(L(\pi^i) - \bar b^i \circ \pi\right)\left(\partial_i f\right)\circ \pi. \end{aligned}$$

Therewith, we deduce the following estimate by using the convergence assumptions on π

$$\limsup_{n\to\infty} \int \left|\pi^n(L) f - \bar L f\right| \mathrm{d}\pi^n_\sharp f \le \|f\|_{C^{1,2}} \int \left(\frac12 \left|a - \bar a\right| + \left|b - \bar b\right|\right) \mathrm{d}\nu.$$

It is clear that the same bound applies to the second term in (4.21) and hence, we can conclude the statement by density of continuous compactly supported function in $L^1(\nu)$. ☐

Lemma 4.17 (Limits for mollifications)**.** *Let ν be a solution to* $\mathrm{FPE}(L_t)$ *with coefficient a_t and b_t. Let $\rho^n \in \mathcal{P}(\mathbb{R}^d)$ be with smooth density and $\rho^n > 0$ such that $\rho^n(x)\,\mathrm{d}x \xrightharpoonup{*} \delta_0$ as $n\to\infty$. Let $\nu^n = \nu * \rho^n$ and L^n be the operator with coefficients as in (4.10) from Lemma 4.10. If $\boldsymbol{\eta}^n$ is a superposition for ν^n solving* $\mathrm{MP}(L_t^n)$*, then any weak* limit point $\boldsymbol{\eta}$ is a superposition of ν solving* $\mathrm{MP}(L_t)$*.*

Proof. We use the same strategy as in the proof of Lemma 4.16. We introduce some continuous compactly supported coefficients $\bar a$ and $\bar b$ having modulus function $\bar\omega$, i.e. satisfying

$$\max\left|\bar a(x) - \bar a(y)\right|, \left|\bar b(x) - \bar b(y)\right| \le \bar\omega\left(x-y\right).$$

Moreover, we introduce their mollifications

$$\bar{a}^n := \frac{\mathrm{d}\left(\bar{a}\nu * \rho^n\right)}{\mathrm{d}\nu * \bar{\rho}^n} \quad \text{and} \quad \frac{\mathrm{d}\left(\bar{b}\nu * \rho^n\right)}{\mathrm{d}\nu * \bar{\rho}^n}$$

and denote the according generators by $\bar{L}$ and $\bar{L}^n$, respectively. Therewith, we can estimate

$$\begin{aligned}\int \left|\bar{L}^n f - \bar{L} f\right| \mathrm{d}\nu^n &\leq \|f\|_{C^{1,2}} \left(\int |\bar{a}^n - \bar{a}|\,\mathrm{d}\nu^n + \int \left|\bar{b}^n - \bar{b}\right| \mathrm{d}\nu^n\right) \\ &= \|f\|_{C^{1,2}} \int \left|\left(\bar{\nu} * \rho^n\right)(x) - \bar{a}(x)\left(\nu * \rho^n\right)(x)\right| \mathrm{d}x \\ &\quad + \|f\|_{C^{1,2}} \int \left|\left(\bar{b}\nu * \rho^n\right)(x) - \bar{b}(x)\left(\nu * \rho^n\right)(x)\right| \mathrm{d}x \\ &\leq 2\|f\|_{C^{1,2}} \int \left(\int \bar{\omega}\left(|y-x|\right) \rho^n(y-x)\,\mathrm{d}x\right) \nu\left(\mathrm{d}y\right) \\ &= 2\|f\|_{C^{1,2}} \int \bar{\omega}\,\mathrm{d}\rho^n \to 2\|f\|_{C^{1,2}}\,\bar{\omega}(0) = 0.\end{aligned}$$

Therefore, we can conclude

$$\begin{aligned}\limsup_{n\to\infty} \int \left|L^n f - \bar{L} f\right| \mathrm{d}\nu^n &= \limsup_{n\to\infty} \int \left|L^n - \bar{L}^n f\right| \mathrm{d}\nu^n \\ &\leq \limsup_{n\to\infty} \int \left(|a^n - \bar{a}^n| + \left|b^n - \bar{b}^n\right|\right) \mathrm{d}\nu^n \\ &\leq \int \left(|a - \bar{a}| + \left|b - \bar{b}\right|\right) \mathrm{d}\nu.\end{aligned}$$

Hence, we can again conclude by density for continuous compactly supported functions in $L^1(\nu)$. □

4.6. **Superposition principle.**

Lemma 4.18 (de la Valleé-Poussin [4]). *Let μ be a positive Borel measure on $(0,\infty)$ and $f \in L^1(\mu)$ non-negative. Then, there exists $\phi : [0,\infty) \to [0,\infty)$ increasing with $\phi(y) \to \infty$ as $y \to \infty$ and*

$$\int_0^\infty \phi(x)\, f(x)\, \mu(\mathrm{d}x) < \infty.$$

Moreover, there exists $\Phi : [0,\infty) \to [0,\infty)$ increasing with $\lim_{y\to\infty} \Phi(y)/y = \infty$ such that

$$\int_0^\infty \Phi\left(f(x)\right) \mu(\mathrm{d}x) < \infty.$$

The function Φ can be chosen such that $\Phi(0) = 0$, $\Phi \in C^\infty$ and strictly convex.

Theorem 4.19 (Superposition principle)**.** *Let $\nu = \bar{\nu}_{t t\in[0,T]} \subset \mathcal{P}(\mathbb{R}^d)$ be a weakly*-continuous solution*[15] *to* FPE(L_t). *Then, there exists $\boldsymbol{\eta} \in \mathcal{P}(\Gamma_T)$ solving* MP(L_t) *such that for every $t \in [0,T]$, it holds $(e_t)_\sharp \boldsymbol{\eta} = \nu_t$.*

Proof. The starting point is the superposition principle as proven for smooth coefficients in Theorem 4.8. From there, we will generalize it to bounded, locally bounded and general coefficients. Each case, will take the three steps we prepared: Approximation of coefficients providing existence, tightness of martingales providing limits points and finally taking the limit in the martingale problem.

Case I: Bounded coefficients. Let us assume, that a, b are bounded in $L^1_t L^\infty_x$, i.e.

$$\int_0^T \sup_{x\in\mathbb{R}^d} \left(|a(t,x)| + |b(t,x)|\right) \mathrm{d}t < \infty. \tag{4.22}$$

We use as a convolution kernel $\rho(x) = a \exp\left(-\sqrt{1+|x|^2}\right)$ with a a normalization constant. For $\varepsilon \in (0,1)$, we set $\rho^\varepsilon(x) := \varepsilon^n \rho(x/\varepsilon)$ and find the bounds $|D^i \rho^e ps| \leq C\varepsilon^{-2}\rho^\varepsilon$ for $i \in \{1,2\}$ and C is an absolute constant. The mollified density $\nu^\varepsilon = \nu * \rho^\varepsilon$ is by Lemma 4.10 a solution to FPE(L^ε_t) with L^ε_t having coefficients $a^\varepsilon, b^\varepsilon$ from (4.10). The coefficients $a^\varepsilon, b^\varepsilon$ satisfy (4.7) by (4.12) in Lemma 4.11. Hence, we obtain existence of a superposition solution $\boldsymbol{\eta}^\varepsilon \in \mathcal{P}(\Gamma_T)$ by Theorem 4.8.

By construction we have that ν^ε_0 converges weakly* to ν_0 and is in particular uniformly tight. Therefore, we find by the de la Valleé-Poussin Lemma 4.18 an increasing function $\theta : \mathbb{R} \to \mathbb{R}$ with $\lim_{z\to\infty} \theta(z) = \infty$ such that $\sup_{\varepsilon>0} \int \theta\left(|x|\right) \mathrm{d}\nu^\varepsilon_0 \leq 1$.

For $R \geq 1$, let $\chi_R : \mathbb{R}^d \to [0,1]$ the cut-off from (4.6) and denote with $x^i_R(x) := x_i \chi_R(x) \in C^2_b(\mathbb{R}^d)$ for $i = 1, \dots, d$. Then for any fixed $p \in (1,\infty)$ let $\Theta_1(x) = \Theta_2(x) = |x|^p$. Then, we can apply Corollary 4.15 and obtain a coercive functional $\Psi : C\left([0,T];\mathbb{R}\right) \to [0,\infty]$ such that

$$\int \Psi(x^i_R \circ \gamma) \boldsymbol{\eta}^\varepsilon(\mathrm{d}\gamma) \leq \int \theta\left(\left|x^i_R\right|\right) \mathrm{d}(e_0)_\sharp \boldsymbol{\eta}^\varepsilon + \int_0^T \left(\left|L^\varepsilon_t x^i_R\right|^p + \left|\left\langle \nabla x^i_R, a^\varepsilon_t \nabla x^i_R\right\rangle\right|^p\right) \mathrm{d}(e_t)_\sharp \boldsymbol{\eta}^\varepsilon \,\mathrm{d}t.$$

Now, by construction $\|\nabla x^i_R\|_\infty \leq C$ uniformly and $\|\nabla^2 x^i_R\|_\infty \leq C/R$. Therewith, we can let $R \to \infty$ and obtain using the lower semicontinuity of Ψ, Fatou's lemma and dominated convergence, the bound

$$\int \Psi(\gamma^i) \boldsymbol{\eta}^\varepsilon(\mathrm{d}\gamma) \leq \int \theta\left(\left|x^i\right|\right) \nu^\varepsilon_0(\mathrm{d}x) + \int_0^T \left(\left|(b^\varepsilon)^i_t\right|^p + \left|(a^\varepsilon)^{i,i}_t\right|^p\right) \nu^\varepsilon_t(\mathrm{d}x)\,\mathrm{d}t.$$

Now, by the smoothing inequality (4.11) follows

$$\int \Psi(\gamma^i) \boldsymbol{\eta}^\varepsilon(\mathrm{d}\gamma) \leq 1 + \int \left(\left|b^i_t\right|^p + \left|a^{i,i}_t\right|^p\right) \nu_t(\mathrm{d}x)\,\mathrm{d}t < \infty.$$

[15] this is no restriction by Lemma 4.2

So tightness follows since $\gamma \mapsto \sum_{i=1}^d \Psi(\gamma^i)$ is coercive on Γ_T. Hence, in particular, we find weak* limit points of $\boldsymbol{\eta}^\varepsilon$ in $\mathcal{P}(\Gamma_T)$.

Now, the case of bounded coefficients satisfying (4.22) is settled by an application of Lemma 4.17.

Case II: Locally bounded coefficients. We generalize the coefficient to only be locally bounded

$$\int_0^T \sup_{x\in B} \left[|a(t,x)| + |b(t,x)|\right] \mathrm{d}t < \infty, \qquad \text{for every bounded Borell } B \subset \mathbb{R}^d. \tag{4.23}$$

For the approximation step, we use the following push-forward. Let $M \geq 1$, χ_M again the cut-off function from (4.6). Therewith, define $\pi_M(x) = x\chi_M(x)$. Then, by Lemma 4.9, we have

$$\pi(b)_t^i(x) = E^{\nu_t}\left[L_t(\pi_M^i)|\pi_M = x\right] \leq \|\pi_M\|_{C^2} \sup_{|z|\leq 2M} (|a(z)| + |b(z)|) < \infty$$
$$\pi(a)_t^{i,j}(x) = E^{\nu_t}\left[\left\langle \nabla\pi_M^i, a_t \nabla \pi_M^j\right\rangle \mid \pi_M = x\right] \leq \|\pi_M\|_{C^1} \sup_{|z|\leq 2M} |a(z)| < \infty.$$

Hence, the push-forwarded coefficients are bounded and we find a superposition $\boldsymbol{\eta}^M$.

For the tightness, we can again use θ as before, since ν_0^M is uniformly tight. For the construction of Θ_1, Θ_2, we have to rely on the de la Valleé-Poussin Lemma 4.18 which improves the integral bound (4.2) to

$$\int_0^T \int \left(\Theta\left(|b_t|\right) + \Theta\left(|a_t|\right)\right) \mathrm{d}\nu_t \, \mathrm{d}t < \infty,$$

for some convex Θ having superlinear growth in the sense of (4.13) from Theorem 4.13. The moderate growth assumption of Θ can be easily satisfied by a slightly worsen function. Then, by the same reasoning as before but using $\Theta_1 = \Theta_2 = \Theta$ instead of the pth power, we conclude tightness of $\boldsymbol{\eta}^M$.

The limit in the martingale problem follows now from Lemma 4.16.

Case III: General coefficients. The only observation, which we need to conclude this step, is that after a convolution of a solution ν of FPE(L_t) with ρ^ε from the Case I, we obtain some ν^ε with coefficients of the form (4.10), which satisfies by (4.12) from Lemma 4.11 the locally bounded estimate (4.23) of Case II and moreover preserve the bound (4.2) by (4.11)

$$\int_0^T \int \left(|a_t^\varepsilon| + |b_t^\varepsilon|\right) \mathrm{d}\nu_t^\varepsilon \, \mathrm{d}t \leq \int_0^T \int \left(|a_t| + |b_t|\right) \mathrm{d}\nu_t \, \mathrm{d}t < \infty.$$

The tightness follows as in Case II and the limit is taken again by Lemma 4.17. □

26.01.16

4.7. **Energy estimate and renormalized solutions for** FPE.

4.7.1. *Heuristics and idea.* In this section we prepare the well-posedness for solutions $(\nu_t)_{t\in[0,T]}$ of $\mathrm{FPE}(L_t)$ with coefficient in some suitable Sobolev space. The main ingredient will be estimates for the Lebesgue density $u_t(x)\,\mathrm{d}x = \nu_t(\mathrm{d}x)$. Therefore, for a function $\beta \in C^2(\mathbb{R};\mathbb{R})$ the equation satisfied by $t \mapsto \int \beta(u_t)\,\mathrm{d}x$ is derived. The goal is to derive suitable Gronwall estimates, i.e. for $x \mapsto |x|^r$ for some $r > 1$. This will lead to uniqueness results in suitable Lebesgue spaces. This idea goes back to [5]. Let us start with the smooth case, where this calculation becomes transparent.

Lemma 4.20. *Assume $a, b \in C_b^{1,2}([0,T]\times\mathbb{R}^d)$. Then, for any $\beta \in C^2(\mathbb{R};\mathbb{R})$ and any $f, u \in C_b^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ holds*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int f_t \beta(u_t(x))\,\mathrm{d}x = \int \left(\partial_t f_t + L_t f_t\right)\beta(u_t(x))\,\mathrm{d}x - \int f_t \frac{\beta''(u_t)}{2}\left\langle \nabla u_t, a_t \nabla u_t\right\rangle \mathrm{d}x \\ - \int f_t\left(\beta'(u_t)u_t - \beta(u_t)\right)(\operatorname{div} L_t)\,\mathrm{d}x,\end{aligned} \tag{4.24}$$

where

$$\operatorname{div} L := \nabla\cdot b - \tfrac{1}{2}\nabla^2 : a.$$

In particular if $\beta(0) = 0$ and β is convex, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \beta(u_t(x))\,\mathrm{d}x \le \int \left(\beta'(u_t)u_t - \beta(u_t)\right)(\operatorname{div} L_t)_-\,\mathrm{d}x, \tag{4.25}$$

where $(a)_- := \max\{a, 0\}$ is the negative part.

Proof. For the proof we neglect the subscript t. From the chain rule, we get $\frac{\mathrm{d}}{\mathrm{d}t}\beta(u) = \beta'(u)L^* u$. Let us investigate first investigate only the drift part

$$\beta'(u)\nabla\cdot(bu) = \beta'(u)u\;\nabla\cdot b + \beta'(u)\;b\cdot\nabla u = \nabla\cdot(b\beta(u)) + \left(\beta'(u)u - \beta(u)\right)\nabla\cdot b.$$

For the diffusion part, we use the identities

$$\nabla^2 : (au) = \nabla\cdot(a\nabla u) + \nabla\cdot(u\nabla\cdot a) \quad \text{with} \quad (\nabla\cdot a)_i = \sum_{j=1}^d \partial_j a_{i,j}.$$

Therewith, we can calculate

$$\begin{aligned}\beta'(u)\nabla^2 : (au) &= \beta'(u)\left(\nabla\cdot(a\nabla u) + \nabla\cdot(u\nabla\cdot a)\right)\\ &= \beta'(u)\nabla\cdot(a\nabla u) + \nabla\cdot(\beta(u)\nabla\cdot a) + \left(\beta'(u)u - \beta(u)\right)\nabla^2 : a\\ &= \nabla\cdot(a\nabla\beta(u)) - \beta''(u)\left\langle\nabla u, a\nabla u\right\rangle + \nabla\cdot(\beta(u)\nabla\cdot a) + \left(\beta'(u)u - \beta(u)\right)\nabla^2 : a\\ &= \nabla^2 : (a\beta(u)) - \beta''(u)\left\langle\nabla u, a\nabla u\right\rangle + \left(\beta'(u)u - \beta(u)\right)\nabla^2 : a.\end{aligned}$$

So in total, we arrive at the identity

$$\partial_t\beta(u) = L^*\beta(u) - \tfrac{1}{2}\beta''(u)\left\langle\nabla u, a\nabla u\right\rangle - \left(\beta'(u)u - \beta(u)\right)\operatorname{div} L.$$

Multiplying by f and integrating by parts leads to (4.24). The second claim, follows by noting that with β convex also $\beta'(u)u - \beta(u) \geq 0$ and choosing $f \equiv 1$. □

Definition 4.21 (Renormalized solution). Let $a : [0,T] \times \mathbb{R}^d \to S_d$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be such that

$$b, \nabla \cdot b \in L^1_{\mathrm{loc}}\left([0,T] \times \mathbb{R}^d\right) \quad \text{and} \quad a, \nabla \cdot a, \nabla^2 : a \in L^1_{\mathrm{loc}}\left([0,T] \times \mathbb{R}^d\right).$$

Let $u \in L^\infty_{\mathrm{loc}}\left([0,T] \times \mathbb{R}^d\right)$ be a solution of FPE(L) then u is a *renormalized solution* if for any convex $\beta \in C^2(\mathbb{R};\mathbb{R})$ and any non-negative $f \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ holds

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f_t \beta(u_t(x))\,\mathrm{d}x \leq \int \left(\partial_t f_t + L_t f_t\right) \beta(u_t(x))\,\mathrm{d}x - \int f_t \left(\beta'(u_t)u_t - \beta(u_t)\right)(\operatorname{div} L_t)\,\mathrm{d}x, \tag{4.26}$$

Lemma 4.22 (Comparison principle). *Assume, that for some $p, q \in [1,\infty]$ holds*

$$\frac{a}{1+|x|^2} \in L^1_t L^p_x, \qquad \frac{b}{1+|x|} \in L^1_t L^q_x$$
$$(\operatorname{div} L)_- \in L^1_{\mathrm{loc}}\left([0,T] \times \mathbb{R}^d\right).$$

Assume, that any solution u of FPE(L) *is renormalized, then the comparison principle holds: If u, v are solutions and $u_0 \leq v_0$ then $u_t \leq v_t$.*

In particular, solutions, if they exist, are unique.

Proof. Approximate $(a)_+$ by $\beta_\varepsilon(s) = \frac{1}{2}\left(\sqrt{s^2+\varepsilon^2} + s - \varepsilon\right)$ and use a spatial cutoff χ_R for f in (4.26) pass to the limit by first letting $\varepsilon \to 0$ and then $R \to \infty$. See [9, Lemma 4.10]. □

The above Lemma shows that renormalized solutions lead to a solution concept, which allows us to use the chain rule and deduce a priori estimates by suitable Gronwall estimates. For brevity, we restrict directly to the validity of (4.26) for the choices $x \mapsto |x|^r$, which suffices to proof uniqueness for FPE(L).

4.7.2. *Sobolev spaces and heat semigroup.* For $p, q \in [1,\infty]$, the space $W^{1,p}_t(L^q_x) \subset L^p_t(L^q_x)$ is defined as the space of all $u \in L^p_t(L^q_x)$, which have distributional derivative $\partial_t u$, which is represented as the unique function $g \in L^p_t(L^q_x)$ such that

$$\int_0^T \int_{\mathbb{R}^d} \partial_t f_t \; u_t \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int f_t g_t \,\mathrm{d}x\,\mathrm{d}t, \qquad \text{for all } f \in C_c^{1,2}\left([0,T] \times \mathbb{R}^d\right).$$

The space becomes a Banach space with the norm $\|u\|_{L^p_t L^q_x} + \|\partial_t u\|_{L^p_t L^q_x}$.

Lemma 4.23 (Density and chain-rule). *The functional space $C_b^{1,2}\left([0,T] \times \mathbb{R}^d\right)$ is dense in $W^{1,p}_t(L^q_x)$ for all $p, q < \infty$. Moreover, for any $\beta \in C^1_b(\mathbb{R})$, $f \in C_b^{1,2}\left([0,T] \times \mathbb{R}^d\right)$ and $u \in W^{1,p}_t(L^q_x)$ holds*

$$\partial_t \left(f\beta(u)\right) = \partial_t f \beta(u) + f\beta'(u)\partial_t u.$$

Proof. Let $\rho \in C_c^\infty(\mathbb{R})$, non-negative, support in $[-1,1]$ and $\int \rho = 1$. For $\varepsilon > 0$, let $\rho^\varepsilon = \varepsilon^{-1}\rho(x\varepsilon^{-1})$. Therewith, let $u^\varepsilon = \rho^\varepsilon * u$ with u extended constantly to the right and left. We will first show with

$$\partial_t u^\varepsilon = \rho^\varepsilon * \partial_t u,$$

where u is the weak derivative.

$$\begin{aligned}\partial_t u^\varepsilon(t) &= \int \partial_t \rho^\varepsilon(t-s)u(s)\,\mathrm{d}s = -\int \left(\partial_s \rho^\varepsilon(t-s)\right)u(s)\,\mathrm{d}s \\ &= \int \rho^\varepsilon(t-s)\partial_s u(s)\,\mathrm{d}s = \left(\rho^\varepsilon * \partial_t u\right)(t).\end{aligned}$$

Now, we use the standard result on convolution, that for any $f \in L^p$ with $p \in [1,\infty)$ it holds $\rho^\varepsilon * f \to f$ in L^p_{loc}.

The chain rule follows by similar arguments. So first, we have

$$|\beta(u) - \beta(u^\varepsilon)| \leq \|\beta\|_{C^1} |u - u^\varepsilon|$$

Similarly, since β' is bounded, we get $\beta'(u^\varepsilon)\partial_t u^\varepsilon \to \beta'(u)\partial_t u$ by dominated convergence. □

Any $u \in W_t^{1,p})(L_x^q)$ has an absolutely continuous representative and in particular we have traces for all $t \in [0,T]$, i.e. the mapping $t \mapsto u_t$ is linear and continuous from $W_t^{1,p}(L_x^q)$.

Definition 4.24 (Sobolev spaces of diffusion operators)**.** For $p \in [1,\infty]$ and $a, b \in L_t^1 L_x^p$ define the space $D^p(L)$

$$D^p(L) := \left\{ f \in L_t^1(L_x^p) : \exists f_n \in C_b^{1,2}([0,T]\times\mathbb{R}^d) \ \text{ s.t. } \ \|f_n - f\|_{D^p(L)} \to 0 \text{ as } n \to \infty \right\},$$

with

$$\|f\|_{D^p(L)} := \|f\|_{L_t^1 L_x^p} + \|Lf\|_{L_t^1 L_x^p}\,.$$

Likewise, define $D^p(L, \langle\nabla, a\nabla\rangle)$ by the completion of $C_b^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ wrt. to the norm

$$\|f\|_{D^p(L,\langle\nabla,a\nabla\rangle)} := \| |f| + |Lf| + \langle \nabla f, a\nabla f\rangle \|_{L_t^1(L_x^p)}\,.$$

Lemma 4.25 (Approximation in $D^p(L)$ and chain rule)**.** *Let $f \in W_t^{1,1}(L_x^p) \cap D^p(L)$, then there exists $f_n \in C_b^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ such that*

$$f_n \to f \quad \text{in } W_t^{1,1}(L_x^p) \qquad \text{and} \qquad f_n \to f \quad \text{in } D^p(L). \tag{4.27}$$

Moreover, for any $\gamma \in C^2(\mathbb{R})$ with γ', γ'' uniformly bounded and any $u \in D^p(L, \langle\nabla, a\nabla\rangle)$ holds

$$L\left(\gamma(u)\right) = \gamma'(u)L(u) + \gamma''(u)\left\langle \nabla u, a\nabla u\right\rangle.$$

Proof. Let $g_n \in C_b^{1,2}\left([0,T]\times\mathbb{R}^d\right)$ such that $g_n \to f \in D^p(L)$. Then consider $g_{n,m} = \rho_m * g_n$ with g_n suitable extended outside of $[0,T]$ and $\rho_m := \rho^{\varepsilon_m}$ as before with $\varepsilon_m \to 0$ as $m \to \infty$. The convolution acts as an contraction on $D^p(L)$, i.e. for any $h \in D^p(L)$, it holds $\|\rho_m * h\|_{D^p(L)} \leq \|h\|$. Hence, also $g_{n,m} \to f$ in $D^p(L)$ whenever $n, m \to \infty$. Now, for any fixed m and any $h \in L_t^1(L_x^p)$ holds $\|\rho_m * h\|_{W_t^{1,1}(L_x^p)} \leq \|\rho_m\|_\infty \|h\|_{L_t^1(L_x^p)}$ in analog to the previous proof. Now, from a diagonal argument, we can extract a sequence f_n such that (4.27) holds.

The chain rules follows by the same procedure as in the proof of Lemma 4.23. □

The above results allows to reformulate the weak solution concept for FPE(L). Assume $u \in L_t^\infty(L_x^r)$ with $r > 1$ is a weak solution to FPE(L), then the weak formulation (4.3) extends with r' s.t. $\frac{1}{r} + \frac{1}{r'} = 1$ to

$$\forall f \in W_t^{1,1}(L_x^{r'}) \cap D^{r'}(L) : \int_0^T \int [(\partial_t + L_t) f] \, u_t \, \mathrm{d}x \, \mathrm{d}t = \int f_T u_T \, \mathrm{d}x - \int f_0 u_0 \, \mathrm{d}x. \quad (4.28)$$

Definition 4.26 (Distributional divergence)**.** The operator $\operatorname{div} L$ is defined in the sense of distributions as linear operator $C_c^{1,2}\left([0,T] \times \mathbb{R}^d\right) \mapsto \int_0^T \int L_t f \, \mathrm{d}x \, \mathrm{d}t$. Moreover, $\operatorname{div} L \in L_t^1(L_x^p)$ if there exists $g \in L_t^1(L_x^p)$ such that

$$\forall f \in C_c^{1,2}\left([0,T] \times \mathbb{R}^d\right) : \quad \int_{(0,T)\times\mathbb{R}^d} Lf \, \mathrm{d}x \, \mathrm{d}t = -\int_{(0,T)\times\mathbb{R}^d} fg \, \mathrm{d}x \, \mathrm{d}t.$$

Likewise, $(\operatorname{div} L)_- \in L_t^1(L_x^p)$ if the above identity holds with $\leq$ and $f \geq 0$.

Lemma 4.27 (A priori estimate)**.** *For $u \in D^p(L)$ and assume $(\operatorname{div} L)_- \in L_t^1(L_x^p)$, then for any $\beta \in C^2(\mathbb{R})$ convex, $\beta(0) = 0$ and $\beta^{(i)}(u), \beta'(u)u - \beta(u) \in L_t^\infty(L_x^{p'})$ for $i = 0, 1, 2$ holds*

$$\int_{(0,T)\times\mathbb{R}^d} L\left(\beta'(u)\right) u \, \mathrm{d}x \, \mathrm{d}t \leq \int_{(0,T)\times\mathbb{R}^d} \left(\beta'(u)u - \beta(u)\right) (\operatorname{div} L)_- \, \mathrm{d}x \, \mathrm{d}t. \quad (4.29)$$

Proof. Let ρ be a smooth convolution kernel on $\mathbb{R}^d$ and $\rho^m(x) = m^d \rho(mx)$. Let L^m be the operator with convoluted coefficients $a^m = \rho^m * a$ and $b^m = \rho^m * b$. Then for $u \in C_b^{1,2}\left([0,T] \times \mathbb{R}^d\right)$ holds by Lemma 4.20 the estimate (4.25) and since $\partial_t\left(\beta(u_z)\right) = \beta'(u_t) L_t^* u_t$, we get (4.29) after integration by parts. Now, by approximation, we can let $m \to \infty$ and choose $u^n \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ with $u^n \to u$ in $D^p(L)$. □

The last estimates involve the usual Sobolev spaces. For $p \in [1,\infty]$, define

$$W_x^{1,p} := \{f \in L_x^p : \nabla f \in L_x^p\} \qquad \text{and} \qquad W_x^{2,p} := \left\{f \in L_x^p : \nabla^2 f \in L_x^p\right\}.$$

Lemma 4.28 (Smoothing properties of the heat semigroup)**.** *For $f \in L_x^1$ define the heat semigroup $\mathsf{P}^\alpha : L_x^1 \to L_x^1$ by*

$$\mathsf{P}^\alpha f(x) = \int g_d(\alpha, x-y) f(y) \, \mathrm{d}y \qquad \text{with} \quad g_d(t,z) := g_d^t(z) := \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{|z|^2}{2t}}. \quad (4.30)$$

Then, it holds for any $\alpha > 0$, any $p \in [1,\infty]$ and any $k = 0, 1, 2, \ldots$:

(i) $\mathsf{P}^\alpha : W_x^{k,p} \to W_x^{k,p}$ *is a contraction.*

(ii) $\alpha^{\frac{k}{2}} \left\| \nabla^{(k)} \mathsf{P}^\alpha f \right\|_{L^p_x} \leq C(k,p,d) \left\| f \right\|_{L^p_x}$.

Proof. The property (i) is a general convolution property, since $g_d(\alpha,\cdot) \in C^\infty(\mathbb{R}^d) \cap \mathcal{P}(\mathbb{R}^d)$ for any $\alpha > 0$. Now, we take $k = 1$ and assume $p \in (1,\infty)$, then we get

$$\left| \nabla \mathsf{P}^\alpha f(x) \right| \leq \frac{1}{\alpha} \int |x-y|\, g_d(\alpha, x-y) f(y) \,\mathrm{d}y \leq \frac{1}{\alpha} \left(\int |z|^q \, g_d(\alpha, \mathrm{d}z) \right)^{\frac{1}{q}} \left| g_d^\alpha * f^p \right|^{\frac{1}{p}} .$$

Since $f^p \in L^1_x$, we get by (i) that $\| g_d^\alpha * f^p \|_{L^1_x} \leq \| f^p \|_{L^1_x}$. Then, taking the p-th power, integrating and taking the p-th square root, leads to

$$\| \nabla \mathsf{P}^\alpha f \|_{L^p_x} \leq \frac{\left(\int |z|^q \, g_d(\alpha, \mathrm{d}z) \right)}{\alpha} \| f \|_{L^p_x}$$

So, we have to estimate:

$$\begin{aligned}
\int |z|^q \, g_d(\alpha, \mathrm{d}z) &= \frac{\left| S^{d-1} \right|}{(2\pi\alpha)^{\frac{d}{2}}} \int_0^\infty r^{q+d-1} \exp\left(-\frac{r^2}{2\alpha} \right) \mathrm{d}r \\
&= \frac{\left| S^{d-1} \right|}{(2\pi\alpha)^{\frac{d}{2}}} \int_0^\infty (2\alpha s)^{\frac{q+d-1}{2}} \exp\left(-s \right) \frac{\sqrt{\alpha}}{\sqrt{2s}} \,\mathrm{d}s \\
&= \alpha^{\frac{q}{2}} \frac{\left| S^{d-1} \right|}{(2\pi)^{\frac{d}{2}}} \int_0^\infty (2s)^{\frac{q+d-2}{2}} \exp\left(-s \right) \mathrm{d}s \\
&=: \alpha^{\frac{q}{2}} \, C(1,p,d) \leq \sqrt{\alpha} \, C(1,p,d),
\end{aligned}$$

which concludes the claim in this case. For the case $p = \infty$, we write

$$\begin{aligned}
\operatorname*{ess\,sup}_{x \in \mathbb{R}^d} \left| \nabla \mathsf{P}^\alpha f \right|(x) &= \operatorname*{ess\,sup}_{x \in \mathbb{R}^d} \left| \int \frac{x-y}{\alpha} g_d(\alpha, x-y) f(y) \,\mathrm{d}y \right| \\
&\leq \| f \|_{L^\infty_x} \frac{1}{\alpha} \int |x-y| \, g_d(\alpha, x-y) \,\mathrm{d}y.
\end{aligned}$$

The remaining integral can be estimated by as above with $q = 1$. The case $p = 1$ becomes after taking Fubini the same estimate

$$\int \left| \nabla \mathsf{P}^\alpha f \right|(x) \,\mathrm{d}x \leq \int |f(y)| \int \frac{|x-y|}{\alpha} g_d(\alpha, x-y) \,\mathrm{d}x \,\mathrm{d}y.$$

The estimate for $k \geq 2$ follows along the same lines, however producing k-th moments of $g_d(\alpha,\cdot)$, which result in the scaling $\alpha^{\frac{k}{2}}$. □

02.01.16

4.8. **Existence in degenerate case.**

Theorem 4.29 (Existence of FPE with degenerate Sobolev coefficients)**.** *Let $p \in (1,\infty]$, $r \geq 2p/(p-1)$ and a, b satisfy*

$$a \in L^1_t(W^{2,p}_x), \qquad b \in L^1_t(W^{1,p}_x), \qquad \textit{and} \qquad (\operatorname{div} L)_- \in L^1_t L^\infty_x .$$

Then for every probability density $\bar{u} \in L^r_x$, there exists a uniquely weakly continuous solution $(u_t)_{t\in[0,T]}$ of* FPE(L) *with $u_0 = \bar{u}$ and $u \in L^\infty_t(L^r_x)$.*

The existence statement is obtained by an approximation procedure as already performed several times in the course. So we concentrate on the uniqueness, which will follow from the following Gronwall estimate of smoothed approximation in combination with a commutator Lemma afterwards.

Lemma 4.30 (Energy estimate)**.** *In the setting of Theorem 4.29 and with $u_0 \equiv 0$, then it holds for all $\alpha > 0$ the a priori estimate*

$$\|u^\alpha\|_{L^\infty_t L^2_x} \leq \exp\left(\left\|(\operatorname{div} L_t)_-\right\|_{L^1_t L^\infty_x}\right)\left(\int_0^T \left|\int u\,[\mathsf{P}^\alpha, \partial_t + L_t]\,u^\alpha\,\mathrm{d}x\right|\mathrm{d}t\right)^{\frac{1}{2}}. \tag{4.31}$$

Proof. By Lemma 4.28, a smoothing in t and the density of $C^{1,2}_b$ in $W^{1,1}_t\left(L^{r'}_x\right) \cap D^{r'}(L)$ from Lemma 4.25 follows that P^α maps $W^{1,1}_t\left(L^{r'}_x\right) \cap D^{r'}(L)$ into itself. Therefore, we can use $\mathsf{P}^\alpha f$ for any $f \in W^{1,1}_t\left(L^{r'}_x\right) \cap D^{r'}(L)$ as a test function in the weak formulation of FPE(L) (see (4.28))

$$\int_0^T \int \left[(\partial_t + L_t)\,f\right] u^\alpha_t\,\mathrm{d}x\,\mathrm{d}t = \int f^\alpha_T u_T\,\mathrm{d}x - \int f^\alpha_0 u_0\,\mathrm{d}x + \int_0^T \int u\,[\mathsf{P}^\alpha, \partial_t + L_t]\,f\,\mathrm{d}x\,\mathrm{d}t, \tag{4.32}$$

where we used that $\int f^\alpha u\,\mathrm{d}x = \int f u^\alpha\,\mathrm{d}x$. Similarly, we can conclude from Lemma 4.28, that $u^\alpha \in D^{r'}(L)$, since

$$\|Lu^\alpha\|_{L^1_t L^{r'}_x} \leq \||a| + |b|\|_{L^1_t L^p_x}\,\|u^\alpha\|_{L^\infty_t W^{2,\theta}}$$

with

$$\theta = r'\left(\frac{p}{r'}\right)' = r'\left(1 - \frac{r}{p(r-1)}\right)^{-1} = \frac{r}{r-1}\frac{p(r-1)}{p(r-1)-r} = \frac{rp}{r(p-1)-p}.$$

From the condition $\theta \leq r$, we deduce the assumption $r \geq 2p/(p-1)$. If, we take $f \in C^{1,2}_c$ as test function in (4.32), it directly follows since P^α commutes with ∂_t and L on smooth functions, that $\partial_t u^\alpha = L^* u^\alpha$ in the sense of distribution. In addtion, the latter is explicitly given in terms of

$$L^* u^\alpha = \frac{1}{2}\nabla^2 : (a u^\alpha) + \nabla\cdot(b u^\alpha) = -\left(\operatorname{div} L\right) u^\alpha + \frac{1}{2}\left(\nabla\cdot a\right)\cdot\nabla u^\alpha \in L^1_t L^{r'}_x,$$

by a similar Hölder estimate as before. Hence $\partial_t u^\alpha \in L^1_t L^{r'}_x$ and it has an absolutely continuous representative, which then has to be the same as the one obtained by smoothing u_t for all $t \in (0,T)$. We are now in the position to use u^α in the weak formulation (4.32). By choosing the test-function $f(t)u^\alpha$ with $f \in C^1_c$, we obtain

$$\int_0^T \int \left((\partial_t f)u^\alpha_t + f\left(\partial_t u^\alpha_t\right) + L_t u^\alpha_t\right) u^\alpha_t\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int f u\,[\mathsf{P}^\alpha, \partial_t + L_t]\,u^\alpha_t\,\mathrm{d}x\,\mathrm{d}t.$$

Since,

$$\int_0^T \int \left((\partial_t f)\, u_t^\alpha + f\,(\partial_t u_t^\alpha) \right) u_t^\alpha \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int \left(-f \partial_t \,(u^\alpha)_t^2 + \tfrac{1}{2} f \partial_t \,(u^\alpha)_t^2 \right) \mathrm{d}x\,\mathrm{d}t$$
$$= -\frac{1}{2} \int_0^T \int f \partial_t (u^\alpha)_t^2 \,\mathrm{d}x\,\mathrm{d}t$$

we obtain the identity

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} \int (u^\alpha)_t^2 \,\mathrm{d}x = \int L\,(u^\alpha)_t^{r-1}\, u_t^\alpha \,\mathrm{d}x - \int_0^T \int f u \,[\mathsf{P}^\alpha, \partial_t + L_t]\, u_t^\alpha \,\mathrm{d}x\,\mathrm{d}t$$
$$\leq \left\| |\operatorname{div} L|_- \right\|_{L_x^\infty} \int (u^\alpha)_t^2 \,\mathrm{d}x + \left| \int u \,[\mathsf{P}^\alpha, \partial_t + L_t]\, u_t^\alpha \,\mathrm{d}x \right|,$$

where we applied Lemma 4.27 for $p = r'$ and $\beta(z) = z^2$. The conclusion follows from a Gronwall estimate. □

The conclusion follows, if we can show that the commutator on the rhs. of the estimate (4.31) converges to 0 as $\alpha \to 0$, which is the content of the next both Lemmata. The first one is the commutator with the vectorfield and the second handles the commutator with the diffusion matrix.

Lemma 4.31 (Commutator estimate for vectorfield)**.** *For $q \in (1,\infty]$, $r, s \in (1,\infty)$ such that $q^{-1} + r^{-1} + s^{-1} = 1$. Let $b \in L_t^1(W_x^{1,q}), u \in L_t^\infty(L_x^r)$ and $f \in L_r^\infty(W_x^{1,s})$, then it holds for $\alpha \in (0,1)$*

$$\int_0^T \left| \int u_t \,[\mathsf{P}^\alpha, b_t \cdot \nabla]\, f_t \,\mathrm{d}x \right| \mathrm{d}t \leq c \,\|\nabla b\|_{L_t^1 L_x^q} \,\|u\|_{L_t^\infty L_x^r} \,\|f\|_{L_t^\infty L_x^s},$$

where $c > 0$ only depends on d

Lemma 4.32 (Commutator estimate for diffusion matrix)**.** *For $q \in (1,\infty]$, $r, s \in (1,\infty)$ such that $q^{-1} + r^{-1} + s^{-1} = 1$. Let $a \in L_t^1(W_x^{2,q}), u \in L_t^\infty(L_x^r)$ and $f \in L_t^\infty(W_x^{2,s})$. For $\alpha \in (0,1)$, it holds*

$$\int_0^T \left| \int u_t \left[\mathsf{P}^\alpha, a_t : \nabla^2\right] f_t \,\mathrm{d}x - \alpha \int u_t \left[\Delta, a_t : \nabla^2\right] \mathsf{P}^\alpha f_t \,\mathrm{d}x \right| \mathrm{d}t \tag{4.33}$$
$$\leq \left\|\nabla^2 a\right\|_{L_t^\infty L_x^q} \|u\|_{L_t^\infty L_x^r} \|f\|_{L_x^\infty L_t^s}.$$

for some c only depending on d. Moreover, for $u \in L_t^\infty(L_x^r \cap L_x^s)$, it holds

$$\left| \int u \left[\mathsf{P}^\alpha, a : \nabla^2\right] (\mathsf{P}^\alpha u) \,\mathrm{d}x \right| \to 0, \qquad \text{in } L^1(0,T), \text{ as } \alpha \to 0. \tag{4.34}$$

Corollary 4.33. *In the setting of Theorem 4.29: If $u \in L_t^\infty(L_x^r)$ is a solution to* FPE(L) *and let $u_t^\alpha := \mathsf{P}^\alpha u_t$ with P^α the heat semigroup as defined in (4.30). Then, it holds*

$$\int_0^T \left| \int u \,[\mathsf{P}^\alpha, \partial_t + L_t]\, u^\alpha \,\mathrm{d}x \right| \mathrm{d}t \to 0 \qquad \text{as } \alpha \to 0.$$

Proof. From Lemma 4.31, the commutator operator $L_t^\infty(W_x^{1,s}) \ni f \mapsto [\mathsf{P}^\alpha, b\cdot\nabla] f \in L_t^1 L_x^{r'}$ extends to a linear continuous operator on $L_t^\infty L_x^s$. Then, by the uniform boundedness principle follows for any $f \in L_t^\infty L_x^s$

$$[\mathsf{P}^\alpha, b\cdot\nabla] f \to 0 \qquad \text{strongly in } L_t^1(L_x^{r'}) \text{ as } \alpha \to 0. \tag{4.35}$$

Note, that if we choose in Lemmate 4.31 and 4.32: $q = p$, then we get $s \le (2p)/(p+1) = (2p/(p-1))' \le 2 \le r$ and so we are allowed to choose $f = u^\alpha$. So the conclusion follows from (4.35) and (4.34). □

The proof of uniqueness in Theorem 4.29 is now an immediate consequence of a combination of Lemma 4.30 and Corollary 4.33.

Proof of Lemma 4.31. For proofing this kind of inequalities, it is always enough to first proof them for functions smooth enough and then argue by approximation with the help of convolutions. Hence, lets assume $f, b^i \in C_b^{1,2}$ and $u \in C_c^{1,2}$. Moreover, we fix $t \in (0,T)$ and obtain the final inequality after integration.

Now, we introduce $F(\beta) := \int u^\beta b\cdot\nabla f^{\alpha-\beta}\,\mathrm{d}x$, which is $C_b^1((0,\alpha))$. Therewith the commutator $\int [u_t\mathsf{P}^\alpha, b_t\cdot\nabla] f_t\,\mathrm{d}x$ can be written and estimated by $|F(\alpha) - F(0)| \le \int_0^\alpha \left|\frac{\mathrm{d}}{\mathrm{d}\beta}F(\beta)\right|\mathrm{d}\beta$. Hence, we have to estimate the derivative of $F(\beta)$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\beta}F(\beta) &= \int \left(\Delta u^\beta\right) b\cdot\nabla f^{\alpha-\beta} - u^\beta b\cdot\nabla\left(\Delta f^{\alpha-\beta}\right)\mathrm{d}x\\ &= \int -\nabla u^\beta\cdot\nabla b\nabla f^{\alpha-\beta} - \nabla u^\beta\cdot\nabla^2 f^{\alpha-\beta} b + \left(\nabla u^\beta\cdot b + u^\beta\nabla\cdot b\right)\Delta f^{\alpha-\beta}\,\mathrm{d}x.\end{aligned}$$

Now, we also have

$$\int \Delta f^{\alpha-\beta}\nabla u^2\cdot b = \int -\nabla u^2\cdot(\nabla b)^T\nabla f^{\alpha-\beta} - b\cdot\nabla^2 u^2\nabla f^{\alpha-\beta}$$

as well as

$$-\int b\cdot\nabla^2 u^\beta\nabla f^{\alpha-\beta} = \int (\nabla\cdot b)\nabla u^\beta\cdot\nabla f^{\alpha-\beta} + \nabla u^2\cdot\nabla^2 f^{\alpha-\beta} b.$$

Hence, in total

$$\frac{\mathrm{d}}{\mathrm{d}\beta}F(\beta) = \int -\left\langle\nabla u^\beta, \left(\nabla b + (\nabla b)^T\right)\nabla f^{\alpha-\beta}\right\rangle + (\nabla\cdot b)\left(\nabla u^\beta\cdot\nabla f^{\alpha-\beta} + u\Delta f^{\alpha-\beta}\right)\mathrm{d}x.$$

Let $D^{\mathrm{sym}}b := \frac{1}{2}\left(\nabla b + (\nabla b)^T\right)$, then the first term can be estimated after integration in β as follows

$$\begin{aligned}\int_0^\alpha\int \left|\left\langle\nabla u^\beta, \left(\nabla b + (\nabla b)^T\right)\nabla f^{\alpha-\beta}\right\rangle\right|\mathrm{d}x\,\mathrm{d}\beta &\le 2\left\|D^{\mathrm{sym}}b\right\|_{L_x^q}\int_0^\alpha \left\|\nabla u^\beta\right\|_{L_x^r}\left\|\nabla f^{\alpha-\beta}\right\|_{L_x^s}\mathrm{d}\beta\\ &\le 2c\left\|D^{\mathrm{sym}}b\right\|_{L_x^q}\|u\|_{L_x^r}\|f\|_{L_x^s}\int_0^\alpha \frac{\mathrm{d}\beta}{\sqrt{\beta(\alpha-\beta)}}.\end{aligned}$$

The term involving $(\nabla \cdot b) \left(\nabla u^\beta \cdot \nabla f^{\alpha-\beta}\right)$ follows along the same lines. We rewrite the last term

$$\int_0^\alpha (\nabla \cdot b)\ u^s \Delta f^{\alpha-\beta} \,\mathrm{d}\beta = -(\nabla \cdot b) \int_0^\alpha u^\beta \frac{\mathrm{d}}{\mathrm{d}\beta} f^{\alpha-\beta} \,\mathrm{d}\beta.$$

Now, since $f^0 - f^\alpha - \int_0^\alpha \frac{\mathrm{d}}{\mathrm{d}\beta} f^{\alpha-\beta} \,\mathrm{d}\beta = 0$ follows

$$\left| \int_0^\alpha u^s \frac{\mathrm{d}}{\mathrm{d}s} f^{\alpha-s} \,\mathrm{d}s \right| \leq \left| u^\alpha \left(f^0 - f^\alpha\right) \right| + \int_0^\alpha \left| \left(u^\beta - u^\alpha\right) \Delta f^{\alpha-\beta} \right| \mathrm{d}\beta$$

and the conclusion follows after integration by parts from Hölder's inequality. □

Proof of Lemma 4.32. Again, we assume that all functions involved are smooth enough and for transparency, we calculate in coordinates for fixed $i, j \in \{1, \ldots, d\}$. Let us introduce the function

$$[0, \alpha] \ni \beta \mapsto F(\beta) := \int u^\beta a^{i,j} \partial_{i,j}^2 f^{\alpha-\beta} \,\mathrm{d}\beta.$$

with derivative

$$F'(\beta) = \int u^s \left[\Delta, a^{i,j} \partial_{i,j}^2\right] f^{\alpha-\beta} \,\mathrm{d}x = \int u^\beta \left[\Delta, a^{i,j}\right] \partial_{i,j}^2 f^{\alpha-\beta} \,\mathrm{d}\beta,$$

where we used that derivatives commute. We introduce $h^{\alpha-\beta} := \partial_{i,j}^2 f^{\alpha-\beta} = \left(\partial_{i,j}^2 f\right)^{\alpha-\beta}$ and define $b := \nabla a^{i,j}$. Then, since $\Delta(ab) = b\Delta a + 2\nabla a \cdot \nabla b + a\Delta b$ follows

$$F'(\beta) = 2 \int u^\beta b \cdot \nabla h^{\alpha-\beta} \,\mathrm{d}x + \int u^\beta \left(\Delta a^{i,j}\right) h^{\alpha-\beta} \,\mathrm{d}x. \tag{4.36}$$

Hence, we differentiate once more and obtain

$$F''(\beta) = 2 \int u^\beta \left[\Delta, b \cdot \nabla\right] h^{\alpha-\beta} \,\mathrm{d}x + \int u^\beta \left[\Delta, \left(\Delta a^{i,j}\right)\right] h^{\alpha-\beta} \,\mathrm{d}x.$$

Now, we obtain a bound by doing a second order Taylor expansion

$$\left|F(\alpha) - F(0) - \alpha F'(\alpha)\right| \leq \int_0^\alpha \left|F''(\beta)\right| (\alpha - \beta) \,\mathrm{d}\beta.$$

Then, upto an integration on $(0, T)$ the left hand side of (4.33) is given by the left hand side above.

Let us first conclude the case $\nabla \cdot b = \Delta a^{i,j} = 0$, in which we have the identity

$$F''(\beta) = 2 \int u^\beta \left[\Delta, b \cdot \nabla\right] h^{\alpha-\beta} \,\mathrm{d}x = -4 \int \left\langle \nabla u^\beta, \left(\nabla^2 a^{i,j}\right) \nabla h^{\alpha-\beta} \right\rangle \mathrm{d}x,$$

which allows to estimate

$$\left|F''(\beta)\right| \leq 4 \left\|\nabla^2 a^{i,j}\right\|_{L_x^q} \left\|\nabla u^s\right\|_{L_x^r} \left\|\nabla h^{\alpha-\beta}\right\|_{L_x^s} \leq \frac{c}{\sqrt{\beta(\alpha-\beta)^3}} \left\|\nabla^2 a^{i,j}\right\|_{L_x^q} \|u\|_{L_x^r} \|f\|_{L_x^s}$$

The right hand side is integrable in β on $(0, \alpha)$ with the factor $(\alpha - \beta)$.

Let us skip the general case and refer to [15, Lemma 3.5].

Now, we turn to (4.34), which follows if we show that $\alpha \int u\,[\Delta, a : \nabla^2]\, u^{2\alpha}\,\mathrm{d}x$ vanishes as $\alpha \to 0$. The first argument is a series of integration by parts independently of the particular choice $f = u^\alpha$ and starts from the identity (4.36)

$$\begin{aligned}\int u\left[\Delta, a^{i,j}\partial^2_{i,j}\right] f^\alpha\,\mathrm{d}x &= 2\int ub\cdot\nabla\partial^2_{i,j}f^\alpha\,\mathrm{d}x + \int u\left(\Delta a^{i,j}\right)\partial^2_{i,j}f^\alpha\,\mathrm{d}x\\ &= -2\int (b\cdot\nabla u)\partial^2_{i,j}f^\alpha\,\mathrm{d}x - \int u\left(\Delta a^{i,j}\right)\partial^2_{i,j}f^\alpha\,\mathrm{d}x\\ &= -2\int\left(\partial^2_{i,j}f\right)\mathsf{P}^\alpha\,(b\cdot\nabla u)\,\mathrm{d}x - \int u\left(\Delta a^{i,j}\right)\partial^2_{i,j}f^\alpha\,\mathrm{d}x\\ &= -2\int\left(\partial^2_{i,j}f\right)\left([\mathsf{P}^\alpha, b\cdot\nabla]\,u + (b\cdot u^\alpha)\right)\mathrm{d}x - \int u\left(\Delta a^{i,j}\right)\partial^2_{i,j}f^\alpha\,\mathrm{d}x.\end{aligned}$$

This identity holds by approximation for $u \in L^\infty_t(L^r_x)$ and $f \in L^\infty_t(W^{2,s}_x)$. Now, we insert the special choice $f = u^\alpha$. The first term can be estimated by Lemma 4.31, since $\alpha\partial^2_{i,j}u^\alpha \in L^\infty_1(L^r_x)$ by Lemma 4.28 and we have

$$\alpha\left|\int\left(\partial^2_{i,j}u^\alpha_t\right)\left([\mathsf{P}^\alpha, b\cdot\nabla]\,u_t\right)\right| \to 0 \qquad \text{in } L^1(0,T) \text{ as } \alpha\to 0.$$

Likewise, the third term can be estimated by Hölder as

$$\alpha\left|\int u(\Delta a^{i,j})\Delta f^\alpha\,\mathrm{d}x\right| \le \left\|\Delta a^{i,j}\right\|_{L^q_x}\|u\|_{L^r_x}\left\|\alpha\partial^2_{i,j}u^{2\alpha}\right\|_{L^s_x} \to 0.$$

Now, recalling that $b = \nabla a^{i,j}$, we obtain

$$\int(\partial^2_{i,j}u^\alpha)b\cdot\nabla u^\alpha\,\mathrm{d}x = -\sum_{k=1}^d\int\left(\partial_i u^\alpha\left(\partial^2_{j,k}a^{i,j}\right)\partial_k u^\alpha + (\partial_i u^\alpha)\left(\partial_k a^{i,j}\right)\partial^2_{k,j}u^\alpha\right)\mathrm{d}x.$$

The first term multiplied by α is again bounded by a similar Hölder estimate as before. The second term can be symmetrized in i, j and rewritten as

$$\begin{aligned}&\alpha\int\left((\partial_i u^\alpha)\left(\partial_k a^{i,j}\right)\partial^2_{k,j}u^\alpha + (\partial_j u^\alpha)\left(\partial_k a^{i,j}\right)\partial^2_{k,i}u^\alpha\right)\mathrm{d}x\\ &= \frac{\alpha}{2}\int\left(\partial_k a^{i,j}\right)\partial_k\left((\partial_i u^\alpha + \partial_j u^\alpha)^2 - (\partial_i u^\alpha)^2 - (\partial_j u^\alpha)^2\right)\mathrm{d}x.\end{aligned}$$

Again a Hölder estimate allows to conclude

$$\alpha\left\|\nabla^2 a\right\|_{L^q_x}\|\nabla u^\alpha\|_{L^r_x}\|\nabla u^\alpha\|_{L^2_x} \to 0 \qquad \text{in } L^1((0,T)) \text{ as } \alpha\to 0.$$

□

09.02.16

References

[1] Fabrice Baudoin. Stochastic calculus lectures. https://fabricebaudoin.wordpress.com/category/stochastic-calculus-lectures/.

[2] Patrick Billingsley. *Convergence of Probability Measures*. Wiley, wiley seri edition, 1968. doi:10.1137/1011078.

[3] Anton Bovier. Lecture: Markov Processes WS09/10. http://wt.iam.uni-bonn.de/bovier/teaching/past-teaching-pages/markov-processes-ws0910/.

[4] José A. Cañizo. The lemma of de la Vallée-Poussin. http://canizo.org/tex/vallee-poussin.pdf.

[5] R. J. DiPerna and P. L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones Mathematicae*, 98(3):511–547, oct 1989. URL: http://www.springerlink.com/index/10.1007/BF01393835.

[6] J.L. Doob. *Stochastic Processes*. John Wiley & Sons, 1953.

[7] Andreas Eberle. Lecture: Stochastic Analysis WS11/12. http://wt.iam.uni-bonn.de/fileadmin/WT/Inhalt/people/Andreas_Eberle/StoAn1112/StochasticAnalysisFinal.pdf.

[8] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes - Characterization and Convergence*. 2005.

[9] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *Journal of Functional Analysis*, 254(1):109–153, 2008. http://cvgmt.sns.it/paper/1292/.

[10] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1964. URL: https://books.google.de/books?id=XN0pAQAAMAAJ.

[11] G. Kersting. Lecture: Schwache Konvergenz und Martingalprobleme. http://www.math.uni-frankfurt.de/~ismi/kersting/lecturenotes/SchwacheKonvergenz.pdf, 2013.

[12] Bernt Øksendal. *Stochastic Differential Equations*. Universitext. Springer-Verlag, Berlin, 6 edition, 2003.

[13] L.C.G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. URL: http://ebooks.cambridge.org/ref/id/CBO9780511805141.

[14] D.W. Stroock and S.R.S. Varadhan. *Multidimensional Diffussion Processes*. Springer, 1979. URL: http://books.google.de/books?id=DuDsmoyqCy4C.

[15] Dario Trevisan. Well-posedness of Multidimensional Diffusion Processes with Weakly Differentiable Coefficients. 2015. arXiv:1507.01357.