BMS Advanced Course

Stochastic Processes III/ Wahrscheinlichkeitstheorie IV

Michael Scheutzow

Lecture Notes

Technische Universität Berlin

Wintersemester 2018/19

PRELIMINARY VERSION
APRIL 2ND 2019

Contents

Chapter 1

Girsanov's Theorem

1.1 Preparation for Girsanov's Theorem

Before stating the important Girsanov theorem we recall Lévy's characterization of Brownian motion and provide a proof of the statement (which we skipped in WTIII). We start with a lemma (see Lemma 6.2.13 in [KS91]). Here, i is the complex number $\sqrt{-1}$.

Lemma 1.1. *Let X, Y be $\mathbb{R}^d$-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. Assume that for each $u \in \mathbb{R}^d$ we have*

$$\mathbb{E}(\exp\{i\langle u, Y\rangle\}|\mathcal{G}) = \mathbb{E}\exp\{i\langle u, X\rangle\} \text{ a.s.}$$

Then the laws of X and Y coincide and Y and $\mathcal{G}$ are independent.

Proof. The first statement is clear since the characteristic function determines a probability measure uniquely. To see the second statement, we represent the conditional expectation via a regular conditional probability (see the chapter "Bedingte Erwartungen und Wahrscheinlichkeiten" in WT1):

$$\mathbb{E}(\exp\{i\langle u, Y\rangle\}|\mathcal{G})(\omega) = \int_{\mathbb{R}^d} \exp\{i\langle u, y\rangle\}\, Q(\omega, \mathrm{d}y), \text{ a.s.}$$

Since a characteristic function is continuous, we can assume that the exceptional set does not depend on $u \in \mathbb{R}^d$. The assumption in the lemma now shows that for almost all $\omega \in \Omega$, the characteristic functions of $Q(\omega, .)$ and the law of X coincide and therefore, the conditional law of Y given $\mathcal{G}$ and the law of X coincide for almost all $\omega \in \Omega$, so for each Borel set $B \subset \mathbb{R}^d$, we have $\mathbb{P}(Y \in B|\mathcal{G}) = \mathbb{P}(X \in B)$ which immediately implies that Y and $\mathcal{G}$ are independent. $\square$

Now we are ready to state Lévy's characterization of Brownian motion.

Theorem 1.2. *Let $M_t = (M_t^1, ..., M_t^d)$ be a continuous $\mathbb{R}^d$-valued local martingale starting at 0 on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. If*

$$\langle M^j, M^k\rangle_t = \delta_{jk} t;\ j, k \in \{1, ..., d\},$$

then M is a d-dimensional $\mathbb{F}$-Brownian motion.

Proof. We basically follow [KS91], Theorem 3.3.16. We need to show that for $0 \le s < t$, the random vector $M_t - M_s$ is independent of $\mathcal{F}_s$ and has distribution $\mathcal{N}(0,(t-s)I)$. Thanks to the previous lemma it is enough to prove that for any $u = (u_1, ..., u_d) \in \mathbb{R}^d$ and $0 \le s \le t$, we have

$$\mathbb{E}\big(\exp\{i\langle u, M_t - M_s\rangle\}|\mathcal{F}_s\big) = \exp\big\{-\frac{1}{2}|u|^2(t-s)\big\}.$$

Fix $u \in \mathbb{R}^d$ and $s \ge 0$ and let $f(x) := \mathrm{e}^{i\langle u,x\rangle}$, $x \in \mathbb{R}^d$. Applying Itô's formula to (the real and imaginary part of) f, we get

$$\mathrm{e}^{i\langle u,M_t\rangle} = \mathrm{e}^{i\langle u,M_s\rangle} + i\sum_{j=1}^d u_j \int_s^t \mathrm{e}^{i\langle u,M_v\rangle}\,\mathrm{d}M_v^j - \frac{1}{2}\sum_{j=1}^d u_j^2 \int_s^t \mathrm{e}^{i\langle u,M_v\rangle}\,\mathrm{d}v. \tag{1.1.1}$$

By Burkholder's inequality (Theorem 2.19 in WTIII) with $p = 2$, we have, using $\langle M^j\rangle_t = t$,

$$\mathbb{E}\sup_{0\le s\le t}|M_s^j| \le \mathbb{E}\sup_{0\le s\le t}(M_s^j)^2 + 1 \le C_2 t + 1 < \infty$$

showing that M^j is a martingale (Chapter 1 in WTIII) which is even square integrable.

Further, the real and imaginary parts of $\int_0^t \mathrm{e}^{i\langle u,M_v\rangle}\,\mathrm{d}M_v^j$ are (square integrable) martingales (since the integrand is bounded) and therefore

$$\mathbb{E}\Big(\int_s^t \mathrm{e}^{i\langle u,X_v\rangle}\,\mathrm{d}M_v^j\big|\mathcal{F}_s\Big) = 0 \text{ a.s.}$$

For any $A \in \mathcal{F}_s$, we get from (1.1.1)

$$\mathbb{E}\big(\mathrm{e}^{i\langle u,M_t-M_s\rangle}1_A\big) = \mathbb{P}(A) - \frac{1}{2}|u|^2\int_s^t \mathbb{E}\big(\mathrm{e}^{i\langle u,M_v-M_s\rangle}1_A\big)\,\mathrm{d}v.$$

This integral equation for the function $t \mapsto \mathbb{E}\big(\mathrm{e}^{i\langle u,M_t-M_s\rangle}1_A\big)$ (with s fixed and $t \ge s$) has a unique solution, namely

$$\mathbb{E}\big(\mathrm{e}^{i\langle u,M_t-M_s\rangle}1_A\big) = P(A)\exp\{-\frac{1}{2}|u|^2(t-s)\},$$

so the claim follows. □

Remark 1.3. In WTIII we defined the stochastic integral with respect to a continuous local martingale M for integrands f for which there exists a progressive process g which is square integrable with respect to the Doleans measure μ_M such that $\|f-g\|_M = 0$, i.e. f and g are in the same equivalence class in that space. We mention without proof (see e.g. [KS91]) the fact that if the process $t \mapsto \langle M\rangle_t$ is almost surely absolutely continuous, then this holds for every adapted process f in $L^2(\mu_M)$. Note that absolute continuity of $t \mapsto \langle M\rangle_t$ holds in particular if M is Brownian motion.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions and let $W = \{W_t = (W_t^{(1)}, ..., W_t^{(d)}),\, t \ge 0\}$ be a d-dimensional $\mathbb{F}$-Brownian motion on that space. Let $X = \{X_t = (X_t^{(1)}, ..., X_t^{(d)}),\, t \ge 0\}$ be progressive (or just jointly measurable and adapted) and satisfy

$$\mathbb{P}\Big(\int_0^T (X_t^{(i)})^2\,\mathrm{d}t < \infty\Big) = 1,\ 1 \le i \le d,\ 0 \le T < \infty.$$

Then, the process

$$Z_t := \exp\Big\{ \sum_{i=1}^d \int_0^t X_s^{(i)}\,\mathrm{d}W_s^{(i)} - \frac{1}{2}\int_0^t |X_s|^2\,\mathrm{d}s \Big\} \tag{1.1.2}$$

is well-defined (here, $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^d$). Writing $N_t := \sum_{i=1}^d \int_0^t X_s^{(i)}\,\mathrm{d}W_s^{(i)}$, the formula becomes

$$Z_t = \exp\Big\{ N_t - \frac{1}{2}\langle N\rangle_t \Big\} \tag{1.1.3}$$

and Z is called *stochastic exponential* of N (even in cases where N is an arbitrary continuous local martingale starting at 0), in short: $Z = \mathcal{E}(N)$. Further,

$$Z_t = 1 + \sum_{i=1}^d \int_0^t Z_s X_s^{(i)}\,\mathrm{d}W_s^{(i)} \Big(= 1 + \int_0^t Z_s\,\mathrm{d}N_s \Big). \tag{1.1.4}$$

This follows from Itô's formula applied to (1.1.3). On the other hand, there is only one solution Z of equation (1.1.4) (we will show this in class; hint: let Z and $\tilde{Z}$ be two solutions, define $Y_t := Z_{t\wedge\tau} - \tilde{Z}_{t\wedge\tau}$ for a suitable stopping time τ and apply Burkholder-Davis-Gundy's inequality with $p = 2$).

Note that (1.1.4) shows that the process Z is a continuous local martingale with $Z_0 = 1$. We will see later that it is of great interest to find conditions under which Z is even a martingale (in which case $\mathbb{E}Z_t = 1$ for all $t \geq 0$). It is not hard to see that a (rather strong) sufficient condition for Z to be a martingale is that there exists a deterministic function $L(t)$, $t \geq 0$ such that $\mathbb{P}(\sup_{s\leq t} |X_s| \leq L(t)) = 1$ for every $t \geq 0$. A weaker condition is *Novikov's condition* which we will provide in Theorem 1.7. We will see in class that for Z as defined in (1.1.3), the condition $\mathbb{E}Z_t = 1$ for all $t \geq 0$ is not only necessary but also sufficient for Z to be a martingale (hint: a nonnegative local martingale is a supermartingale [use Fatou's lemma for conditional expectations] and a supermartingale is a martingale iff its expected value is constant). If Z is a martingale, then we define, for each $T \in [0, \infty)$, a probability measure $\tilde{\mathbb{P}}_T$ on $(\Omega, \mathcal{F}_T)$ by

$$\tilde{\mathbb{P}}_T(A) := \mathbb{E}\big(1_A Z_T\big),\ A \in \mathcal{F}_T.$$

Note that the probability measures $\tilde{\mathbb{P}}_T$ and $\mathbb{P}|_{\mathcal{F}_T}$ are mutually absolutely continuous with density $\mathrm{d}\tilde{\mathbb{P}}_T/\mathrm{d}\mathbb{P}|_{\mathcal{F}_T} = Z_T$.

In the following *Girsanov* theorem, we use the symbol $\mathbb{F}_T$ to denote the restricted filtration $\mathcal{F}_t$, $0 \leq t \leq T$. It should be clear what we mean by an $\mathbb{F}_T$-Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{F}_T, \tilde{\mathbb{P}}_T)$, so we refrain from providing a precise definition.

Theorem 1.4 (Girsanov (1960))**.** *Assume that Z defined by* (1.1.2) *is a martingale. Define*

$$\tilde{W}_t^{(i)} := W_t^{(i)} - \int_0^t X_s^{(i)}\,\mathrm{d}s,\ i \in \{1, ..., d\},\ t \geq 0.$$

Then, for each fixed $T \geq 0$, the process $\{\tilde{W}_t\}, t \in [0, T]$ is an $\mathbb{F}_T$-Brownian motion on $(\Omega, \mathcal{F}_T, \mathbb{F}_T, \tilde{\mathbb{P}}_T)$.

For the proof we require two lemmas. We denote by $\tilde{\mathbb{E}}_T$ the expectation with respect to $\tilde{\mathbb{P}}_T$.

Lemma 1.5. *Let $T > 0$ and assume that Z is a martingale. If $0 \leq s \leq t \leq T$ and Y is real-valued $\mathcal{F}_t$-measurable satisfying $\tilde{\mathbb{E}}_T|Y| < \infty$, then*

$$\tilde{\mathbb{E}}_T\big(Y|\mathcal{F}_s\big) = \frac{1}{Z_s}\mathbb{E}\big(YZ_t|\mathcal{F}_s\big),\ a.s.\ w.r.t.\ \mathbb{P}\ and\ \tilde{\mathbb{P}}_T.$$

Proof. We have, for $A \in \mathcal{F}_s$,

$$\tilde{\mathbb{E}}_T\big(1_A Y\big) = \mathbb{E}\big(1_A Y Z_T\big) = \mathbb{E}\big(1_A Y \mathbb{E}(Z_T|\mathcal{F}_t)\big) = \mathbb{E}\big(1_A Y Z_t\big) = \mathbb{E}\big(1_A \mathbb{E}\big(Y Z_t|\mathcal{F}_s\big)\big)$$

and

$$\tilde{\mathbb{E}}_T\Big(1_A \frac{1}{Z_s}\mathbb{E}\big(Y Z_t|\mathcal{F}_s\big)\Big) = \mathbb{E}\Big(1_A \frac{1}{Z_s}\mathbb{E}\big(Y Z_t|\mathcal{F}_s\big) Z_T\Big) = \mathbb{E}\Big(1_A \frac{1}{Z_s}\mathbb{E}\big(Y Z_t|\mathcal{F}_s\big) Z_s\Big) = \mathbb{E}\big(1_A \mathbb{E}\big(Y Z_t|\mathcal{F}_s\big)\big).$$

□

In the following proposition, we denote by $\mathcal{M}_{loc,T}$ the set of continuous local $\mathbb{F}$-martingales on $[0,T]$ with initial condition $M_0 = 0$ with respect to the original measure $\mathbb{P}$ and by $\tilde{\mathcal{M}}_{loc,T}$ the corresponding set with $\mathbb{P}$ replaced by $\tilde{\mathbb{P}}_T$.

Proposition 1.6. *Fix $T > 0$ and assume that Z defined as in* (1.1.2) *is a martingale and $M \in \mathcal{M}_{loc,T}$. Then*

$$\tilde{M}_t := M_t - \sum_{i=1}^d \int_0^t X_s^{(i)} \mathrm{d}\langle M, W^{(i)}\rangle_s;\ 0 \le t \le T$$

is in $\tilde{\mathcal{M}}_{loc,T}$. If $N \in \mathcal{M}_{loc,T}$ and

$$\tilde{N}_t := N_t - \sum_{i=1}^d \int_0^t X_s^{(i)} \mathrm{d}\langle N, W^{(i)}\rangle_s;\ 0 \le t \le T,$$

then

$$\langle \tilde{M}, \tilde{N}\rangle_t = \langle M, N\rangle_t;\ 0 \le t \le T,\ \text{a.s.}$$

Proof. We first assume that M and N are bounded martingales with bounded quadratic variation and that Z and $\int_0^t \big(X_s^{(i)}\big)^2 \mathrm{d}s$ are bounded in t and ω (the general case can then be shown by stopping). Kunita-Watanabe's inequality (WTIII, Proposition 2.8 a)) shows that

$$\Big|\int_0^t X_s^{(i)}\, \mathrm{d}\langle M, W^{(i)}\rangle_s\Big|^2 \le \langle M\rangle_t \cdot \int_0^t (X_s^{(i)})^2 \mathrm{d}s,$$

so $\tilde{M}$ is also bounded. The integration-by-parts formula (WTIII, Proposition 2.12) implies

$$Z_t \tilde{M}_t = \int_0^t Z_s\, \mathrm{d}M_s + \sum_{i=1}^d \int_0^t \tilde{M}_s X_s^{(i)} Z_s\, \mathrm{d}W_s^{(i)},$$

which is a martingale under $\mathbb{P}$. Now Lemma 1.5 immediately implies that $\tilde{M} \in \tilde{\mathcal{M}}_{loc,T}$.

The last claim in the statement of the proposition follows since M and $\tilde{M}$ (resp. N and $\tilde{N}$) differ only by a process of locally finite variation. □

Proof of Theorem 1.4. We show that the process $\tilde{W}$ in the statement of the theorem satisfies the assumptions of Theorem 1.2. Setting $M = W^{(j)}$ in the previous proposition, we see that $\tilde{M} = \tilde{W}^{(j)}$ is in $\tilde{\mathcal{M}}_{loc,T}$. Setting $N = W^{(k)}$, we see that

$$\langle \tilde{W}^{(j)}, \tilde{W}^{(k)}\rangle = \langle W^{(j)}, W^{(k)}\rangle = \delta_{j,k} t;\ 0 \le t \le T,\ \text{a.s.},$$

so the claim follows. □

1.2 Novikov's Condition

Theorem 1.7. *If N is a continuous local martingale with $N_0 = 0$ such that*

$$\mathbb{E} \exp\Big\{\frac{1}{2}\langle N\rangle_T\Big\} < \infty,\ 0 \le T < \infty,$$

then the process Z defined by (1.1.3) *is a martingale. In particular, if X, W and Z are defined as in* (1.1.2) *and*

$$\mathbb{E}\Big(\exp\Big\{\frac{1}{2}\int_0^T |X_s|^2\,\mathrm{d}s\Big\}\Big) < \infty,\ 0 \le T < \infty,$$

then Z is a martingale.

Proof. Following [Eb16] we prove the result only under the slightly stronger condition

$$\mathbb{E}\Big(\exp\Big\{\frac{p}{2}\langle N\rangle_T\Big\}\Big) < \infty,\ 0 \le T < \infty,$$

for some $p > 1$. This simplifies the proof considerably. For the general case see [KS91] or [Bo18].

Since Z is a continuous local martingale, there exists a localizing sequence of stopping times T_n, $n \in \mathbb{N}$ such that $T_n \to \infty$ almost surely and $t \mapsto Z_{t\wedge T_n}$ is a (continuous) martingale for each $n \in \mathbb{N}$. If, for each $t \ge 0$, the sequence $Z_{t\wedge T_n}$ is uniformly integrable, then (by Satz 1.39 in WT2) $\lim_{n\to\infty} Z_{t\wedge T_n} = Z_t$ in L^1 and therefore the martingale property of $Z_{t\wedge T_n}$ is inherited by Z letting $n \to \infty$.

It remains to show that the sequence $Z_{t\wedge T_n}$ is uniformly integrable for each fixed $t > 0$. Let $c > 0$ and define $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then, using Hölder's inequality, we get

$$\begin{aligned}
\mathbb{E}\big(Z_{t\wedge T_n} 1_{\{Z_{t\wedge T_n}\ge c\}}\big) &= \mathbb{E}\Big(\exp\big\{N_{t\wedge T_n} - \frac{p}{2}\langle N\rangle_{t\wedge T_n}\big\}\exp\big\{\frac{p-1}{2}\langle N\rangle_{t\wedge T_n}\big\}1_{\{Z_{t\wedge T_n}\ge c\}}\Big)\\
&\le \Big(\mathbb{E}\exp\big\{pN_{t\wedge T_n} - \frac{p^2}{2}\langle N\rangle_{t\wedge T_n}\big\}\Big)^{1/p}\cdot\Big(\mathbb{E}\big(\exp\big\{q\frac{p-1}{2}\langle N\rangle_{t\wedge T_n}\big\}1_{\{Z_{t\wedge T_n}\ge c\}}\big)\Big)^{1/q}\\
&\le \Big(\mathbb{E}\big(\exp\big\{\frac{p}{2}\langle N\rangle_t\big\}1_{\{Z_{t\wedge T_n}\ge c\}}\big)\Big)^{1/q}.
\end{aligned}$$

To show uniform integrability we have to show that if we first take the supremum over all $n \in \mathbb{N}$ and then the limit as $c \to \infty$, then this limit is 0. Lemma 1.38 in WT2 shows that it is actually sufficient to take the $\limsup_{n\to\infty}$ instead of $\sup_{n\in\mathbb{N}}$ (since $\mathbb{N}$ is countable). Observe that

$$\limsup_{n\to\infty} \mathbb{E}\big(Z_{t\wedge T_n}1_{\{Z_{t\wedge T_n}\ge c\}}\big) \le \Big(\mathbb{E}\big(\exp\big\{\frac{p}{2}\langle N\rangle_t\big\}1_{\{Z_t\ge c\}}\big)\Big)^{1/q}$$

by dominated convergence since $\exp\big\{\frac{p}{2}\langle N\rangle_t\big\}$ is integrable and hence the limit as $c \to \infty$ is 0, so the proof is complete. □

Remark 1.8. The previous theorem is known to be wrong if $1/2$ is replaced by any number $p < 1/2$ (see Remark 5.17 in [KS91]), even in case $d = 1$.

1.3 Application: Brownian Motion with Drift

We show how Girsanov's theorem can be employed to compute the density of the first passage times for one-dimensional Brownian motion with drift.

For a one-dimensional Brownian motion W and $b > 0$ let $\tau_b := \inf\{t \geq 0 : W_t = b\}$ denote the first passage time to level b. Note that $\mathbb{P}(\tau_b \leq t) = \mathbb{P}\big(\sup_{0 \leq s \leq t} W_s \geq b\big)$. Note that we computed the latter quantity in WT2 by applying Donsker's invariance principle. It is therefore easy to show that τ_b has the density

$$f_b(t) = \frac{b}{\sqrt{2\pi t^3}} \exp\Big\{-\frac{b^2}{2t}\Big\};\ t > 0.$$

Can we also compute the density $f_b^{(\mu)}$ of τ_b for Brownian motion with drift μ? Indeed we can and Girsanov's theorem helps us doing this. Let $\mu \neq 0$. Let $d = 1$ and define $\tilde{W}_t := W_t - t\mu$, $t \geq 0$. This corresponds to the choice $X_s = \mu$ in Girsanov's theorem. Clearly, Novikov's condition is satisfied and hence $Z_t := \exp\big\{\mu W_t - \frac{1}{2}\mu^2 t\big\}$, $t \geq 0$ is a martingale. Fix $T > 0$. By Girsanov's theorem $\tilde{W}$ is a Brownian motion on $[0, T]$ under $\tilde{\mathbb{P}}_T$ and hence $W_t = \tilde{W}_t + \mu t$, $0 \leq t \leq T$ is a Brownian motion with drift μ under $\tilde{\mathbb{P}}_T$. Denoting by τ_b the first passage time of W to b, we get for $t \in [0, T]$

$$\begin{aligned}
\tilde{\mathbb{P}}_T(\tau_b \leq t) &= \int 1_{\tau_b \leq t} \mathrm{d}\, \tilde{\mathbb{P}}_T \\
&= \int 1_{\tau_b \leq t} Z_T \mathrm{d}\, \mathbb{P} = \int 1_{\tau_b \leq t} Z_{\tau_b} \mathrm{d}\, \mathbb{P} \\
&= \mathbb{E}\big(1_{\tau_b \leq t} \exp\{\mu b - \frac{1}{2}\mu^2 \tau_b\}\big) \\
&= \int_0^t \exp\{\mu b - \frac{1}{2}\mu^2 s\} f_b(s)\, \mathrm{d}s,
\end{aligned}$$

where we applied the optional sampling theorem for the continuous martingale Z. It follows that the first passage time τ_b for Brownian motion with drift has a density given by

$$f_b^{(\mu)}(s) = \exp\{\mu b - \frac{1}{2}\mu^2 s\} f_b(s) = \frac{b}{\sqrt{2\pi s^3}} \exp\Big\{-\frac{(b - \mu s)^2}{2s}\Big\},\ s > 0.$$

Note that $f_b(\mu)$ is a true density (i.e. has integral 1) iff $\mu > 0$. If $\mu < 0$, then the integral is strictly smaller than 1 which is due to the fact that Brownian motion with drift $\mu < 0$ has a positive probability of never reaching level $b > 0$.

Chapter 2

Local Time

In this section we introduce the concept of local time of a real-valued continuous semimartingale X. Is $|X|$ a semimartingale? Itô's formula can certainly not be applied, but we will see that the answer is nevertheless *yes*.

To show this we define $f(x) = |x|$, $x \in \mathbb{R}$, approximate f by smoother functions f_n, apply Itô's fomula to $f_n(X)$ and then take the limit $n \to \infty$. Specifically, we choose $f_n \in C^2(\mathbb{R}, \mathbb{R})$ as follows:

(i) $f_n(x) = |x|$ for $|x| \geq 1/n$,

(ii) $f_n(x) = f_n(-x), \ x \in \mathbb{R}$,

(iii) $f_n''(x) \geq 0, \quad x \in \mathbb{R}$.

Hence, f_n converges to f uniformly and $f_n'(x)$ converges to $\operatorname{sgn}(x)$ pointwise (we define $\operatorname{sgn}(x) = 1$ if $x > 0$, $\operatorname{sgn}(x) = -1$ if $x < 0$ and $\operatorname{sgn}(0) = 0$). In addition $\int_{-1}^{1} f_n''(x) \, \mathrm{d}x = 2$ for all n. Itô's formula applied to f_n yields

$$f_n(X_t) - f_n(X_0) = \int_0^t f_n'(X_s) \, \mathrm{d}X_s + \frac{1}{2} \int_0^t f_n''(X_s) \, \mathrm{d}\langle X \rangle_s. \tag{2.0.1}$$

In order to see that the stochastic integral converges as $n \to \infty$ we need the following lemma (note that Lemma 2.16 in WT3 about ucp-convergence does not apply here). We formulate the lemma in a way which meets our demands (but not more).

Lemma 2.1. *Let H^n, $n \in \mathbb{N}$ be a sequence of uniformly bounded progressive processes, i.e. there exists a number $C \in \mathbb{R}$ such that $\left|H_t^n(\omega)\right| \leq C$ for all $t \geq 0$, $n \in \mathbb{N}$ and $\omega \in \Omega$ and assume that $H_t(\omega) := \lim_{n\to\infty} H_t^n(\omega)$ exists for all $t \geq 0$ and $\omega \in \Omega$. Let Y be a continuous semimartingale such that $Y_0 = 0$. Then*

$$\int_0^{\cdot} H_s^n \, \mathrm{d}Y_s \to \int_0^{\cdot} H_s \, \mathrm{d}Y_s \ \textit{ucp}.$$

Proof. We follow Section 8.3 of [WW90]. Let $Y = M + A$ be the unique decomposition of Y such that $M \in \mathcal{M}_{loc}^0$ and $A \in \mathcal{A}^0$. Clearly we have (by dominated convergence)

$$\int_0^{\cdot} H_s^n \, \mathrm{d}A_s \to \int_0^{\cdot} H_s \, \mathrm{d}A_s \text{ ucp},$$

so it remains to show the claim in case $Y = M$. Let $T_k := \inf\{t \geq 0 : |M_t| + \langle M \rangle_t \geq k\} \wedge k$, $k \in \mathbb{N}$. The BDG inequality implies

$$\mathbb{E} \sup_{t \leq k} \Big(\Big| \int_0^t 1_{[0,T_k]}(s)(H_s^n - H_s) \, \mathrm{d}M_s \Big|^2 \Big) \leq C_2 \mathbb{E} \Big(\int_0^k 1_{[0,T_k]}(s)(H_s^n - H_s)^2 \, \mathrm{d}\langle M \rangle_s \Big) \to 0, \text{ as } n \to \infty$$

by dominated convergence. In particular we have, for fixed $k \in \mathbb{N}$,

$$\int_0^\cdot 1_{[0,T_k]} H_s^n \, \mathrm{d}M_s \to \int_0^\cdot 1_{[0,T_k]} H_s \, \mathrm{d}M_s \text{ ucp.}$$

Letting $k \to \infty$ the claim follows since $T_k \to \infty$ almost surely (check this!). □

Now we are able to pass to the limit in (2.0.1). The previous lemma shows that

$$\int_0^\cdot f_n'(X_s) \, \mathrm{d}X_s \to \int_0^\cdot \operatorname{sgn}(X_s) \, \mathrm{d}X_s \text{ ucp}$$

(taking $C = 1$). Therefore, the second integral in (2.0.1) converges ucp to a process

$$L_t := \lim_{n \to \infty} \frac{1}{2} \int_0^t f_n''(X_s) \, \mathrm{d}\langle X \rangle_s = |X_t| - |X_0| - \int_0^t \operatorname{sgn}(X_s) \, \mathrm{d}X_s. \tag{2.0.2}$$

The first equality shows that $L \in \mathcal{A}^+$. The second representation shows that the process L does not depend on the particular choice of f_n. L is called *local time at 0* of the semimartingale X (the reason for this will become clear in the next theorem). Note that we have shown the following formula for $f(x) = |x|$:

$$f(X_t) = f(X_0) + \int_0^t \operatorname{sgn}(X_s) \, \mathrm{d}X_s + L_t,$$

showing that $f(X_t)$ is indeed a continuous semimartingale.

Theorem 2.2. *Let X be a continuous semimartingale with local time L given by* (2.0.2). *Then*

$$L_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1_{|X_s| \leq \varepsilon} \, \mathrm{d}\langle X \rangle_s$$

in probability. In particular, for an $\mathbb{F}$-Brownian motion $X = W$,

$$L_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \lambda\{s \leq t : |W_s| \leq \varepsilon\},$$

where λ denotes Lebesgue measure on $[0, \infty)$.

Proof. The second claim clearly follows from the first since $\langle W \rangle_s = s$. To show the first claim we choose f_n in a particular way namely such that (in addition to the previous assumptions)

$$n 1_{\left[-\frac{1}{n+1}, \frac{1}{n+1}\right]} \leq f_n'' \leq (n+1) 1_{\left[-\frac{1}{n}, \frac{1}{n}\right]}.$$

Then

$$\frac{n}{2} \int_0^t 1_{X_s \in \left[-\frac{1}{n+1}, \frac{1}{n+1}\right]} \mathrm{d}\langle X \rangle_s \leq \frac{1}{2} \int_0^t f_n''(X_s) \, \mathrm{d}\langle X \rangle_s \leq \frac{n+1}{2} \int_0^t 1_{X_s \in \left[-\frac{1}{n}, \frac{1}{n}\right]} \mathrm{d}\langle X \rangle_s$$

which implies the first claim (check this! Note that $L_t = \lim_{n \to \infty} \ldots$ implies $L_t = \lim_{\varepsilon \to 0} \ldots$ by monotonicity). □

Chapter 3

Weak Solutions and Martingale Problems

3.1 Weak solutions of stochastic differential equations

We will basically follow the approach presented in [KS91]. Consider the following *stochastic differential equation*

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sum_{k=1}^{m} \sigma_k(t, X_t)\,\mathrm{d}W_t^k, \tag{3.1.1}$$

where $W^1, ..., W^m$ are independent Brownian motions and b and $\sigma_1, ..., \sigma_m$ are measurable functions mapping $[0, \infty) \times \mathbb{R}^d$ to $\mathbb{R}^d$.

We start with the definition of a *weak solution* of (3.1.1).

Definition 3.1. A *weak solution* of (3.1.1) is a tuple (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where

(i) $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a FPS which satisfies the usual conditions,

(ii) X is a continuous, adapted $\mathbb{R}^d$-valued process and W is an m-dimensional $\mathbb{F}$-Brownian motion,

(iii) $\int_0^t |b(s, X_s)| + \sum_{k=1}^m |\sigma_k(s, X_s)|^2 \,\mathrm{d}s < \infty$ almost surely for every $t \geq 0$,

(iv)
$$X_t = X_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + \sum_{k=1}^{m} \int_0^t \sigma_k(s, X_s)\,\mathrm{d}W_s^k;\ 0 \leq t < \infty$$
holds almost surely.

Remark 3.2. Occasionally, we fix $T > 0$ and consider weak solutions on the interval $[0, T]$ instead of $[0, \infty)$. It should be clear what we mean by this, so we do not provide a formal definition.

Note that it might well happen that a weak solution (X, W) exists but that X is not adapted to the augmented filtration generated by W (here "augmented" means the smallest filtration which contains the filtration generated by W which satisfies the usual conditions). We will see an example of this kind soon. Note that in WT3 we established a stronger form of solution to

(3.1.1): we showed in Theorem 2.27 under appropriate assumptions (called (H)) that the sde has a (unique) solution with initial condition $x \in \mathbb{R}^d$ on *any* FPS carrying an $\mathbb{F}$-Brownian motion W. In particular, we can choose the filtration $\mathbb{F}$ to be the augmented filtration generated by W. Such a solution is called a *strong solution*. We will often allow the initial condition $X_0 = \xi$ to be random. In this case ξ is $\mathcal{F}_0$-measurable and therefore independent of the filtration generated by W. Then a process X on a given FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a given $\mathbb{F}$-Brownian motion W and a given $\mathcal{F}_0$-measurable ξ is called a strong solution, if (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a weak solution which is adapted to the augmented filtration generated by W and ξ and which satisfies the initial condition $X_0 = \xi$ almost surely. By definition, every strong solution is a weak solution. We will see that the converse is not true.

One may ask which of the two concepts of a solution is the more natural one. There is no clear answer to this question. It depends on which kind of phenomenon one would like to model by an sde. If we think of W as an input to a system and X as the output, then one can argue that the output should be a function of the input and in this case strong solutions seem more natural. On the other hand one can take a pragmatic viewpoint by saying that one just wants to describe some random phenomenon via an sde and one does not care whether the solution is a function of the input. In this case weak solutions are more appropriate.

After having defined the concept of a weak solution, we define what we mean by uniqueness of solutions.

Definition 3.3. We say that *pathwise uniqueness* holds for (3.1.1) if, whenever (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $(\tilde{X}, W)$, $(\Omega, \mathcal{F}, \tilde{\mathbb{F}}, \mathbb{P})$ are two weak solutions on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ (with possibly different filtrations) such that $\mathbb{P}(X_0 = \tilde{X}_0) = 1$, then the processes X and $\tilde{X}$ are indistinguishable (i.e. $\mathbb{P}(X_t = \tilde{X}_t;\ 0 \le t < \infty) = 1$).

Remark 3.4. Some authors define pathwise uniqueness slightly differently by assuming that $\mathbb{F}$ and $\tilde{\mathbb{F}}$ coincide. Formally, our definition is stronger (i.e. more restrictive). In fact, the two definitions can be shown to be equivalent (see [IW89], Remark IV.1.3.).

Remark 3.5. The proof of Theorem 2.27 in WT3 shows that pathwise uniqueness holds under the assumptions (H) stated there.

Definition 3.6. We say that *uniqueness in law* or *weak uniqueness* holds for (3.1.1) if, for any two weak solutions (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $(\tilde{X}, \tilde{W})$, $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ with the same initial distribution (i.e. $\mathcal{L}(X_0) = \mathcal{L}(\tilde{X}_0)$) the processes X and $\tilde{X}$ have the same law.

The following classical example due to H. Tanaka shows that a weak solution may not be a strong solution and that uniqueness in law does not imply pathwise uniqueness.

Example 3.7. Consider the one-dimensional sde

$$\mathrm{d}X_t = \mathrm{sign}(X_t)\,\mathrm{d}W_t;\ 0 \le t < \infty, \tag{3.1.2}$$

where

$$\mathrm{sign}(x) = \begin{cases} 1; & x > 0, \\ -1; & x \le 0. \end{cases}$$

If (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a weak solution, then

$$B_t := \int_0^t \mathrm{sign}(X_s)\,\mathrm{d}W_s$$

is a continuous local $\mathbb{F}$-martingale starting at 0 with quadratic variation $\langle B\rangle_t = \int_0^t \operatorname{sign}^2(X_s)\,\mathrm{d}s = t$. Therefore, B is an $\mathbb{F}$-Brownian motion by Lévy's theorem. Note that $X_t = X_0 + B_t$, $t \geq 0$ and that the process B is independent of $\mathcal{F}_0$. Since X_0 is $\mathcal{F}_0$-measurable, B and X_0 are independent, so the law of X is uniquely determined by that of X_0 and so weak uniqueness holds for (3.1.2).

Let us now show existence of a weak solution. Let X be an $\mathbb{F}$-Brownian motion on some FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathcal{F}_0$-measurable initial condition X_0 and define

$$W_t := \int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s.$$

Then W is an $\mathbb{F}$-Brownian motion by Lévy's theorem and

$$\int_0^t \operatorname{sign}(X_s)\,\mathrm{d}W_s = \int_0^t \operatorname{sign}^2(X_s)\,\mathrm{d}X_s = X_t - X_0,$$

so (X, W) is a weak solution!

Next we show that there is no pathwise uniqueness in case the initial condition is $X_0 = 0$. Let (X, W), $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be the weak solution which we just constructed (with $X_0 = 0$). Then $(-X, W)$, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is also a weak solution (in spite of the slight asymmetry in the definition of the function sign)! Therefore, pathwise uniqueness does not hold for (3.1.2).

If (X, W) is *any* weak solution of (3.1.2) with initial condition 0, then

$$\int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s = \int_0^t \operatorname{sign}^2(X_s)\,\mathrm{d}W_s = W_t,$$

so W is measurable with respect to the filtration generated by X. We will see in a few seconds that X is *not* measurable with respect to the filtration generated by W, so X is *not* a strong solution. Since any strong solution is a weak solution this shows that no strong solution of (3.1.2) with initial condition 0 exists.

Since the process X is an $\mathbb{F}$-Brownian motion (for some filtration $\mathbb{F}$) it can be represented as

$$|X_t| = \int_0^t \operatorname{sgn}(X_s)\,\mathrm{d}X_s + L_t = \int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s + L_t = W_t + L_t,$$

by the results of the last chapter, where L is the local time of the semimartingale X at 0. The second part of Theorem 2.2 shows that L is not only $\mathbb{F}$-adapted but even adapted to the filtration generated by $|X_t|$, $t \geq 0$. Therefore the same holds true for W. Obviously, the filtration $\mathcal{G}_t$, $t \geq 0$ generated by $|X|$ is strictly smaller than that generated by X (the event $\{X_1 > 0\}$ is for example contained $\sigma(X_s, s \leq 1)$ but not in $\mathcal{G}_1$).

3.2 Weak solutions via Girsanov's theorem

The following proposition shows how Girsanov's theorem can be used to show existence of a weak solution to a certain class of stochastic differential equations.

Proposition 3.8. *Fix $T > 0$ and consider the stochastic differential equation*

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \mathrm{d}W_t;\ 0 \leq t \leq T, \tag{3.2.1}$$

where W is d-dimensional Brownian motion and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is jointly measurable and bounded. Then, for any $x \in \mathbb{R}^d$, (3.2.1) has a weak solution with initial condition $X_0 = x$ almost surely.

Proof. Fix $x \in \mathbb{R}^d$ and let $(X_t)_{t\geq 0}$ be a d-dimensional $\mathbb{F}$-Brownian motion starting at x on some FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. By Theorem 1.7, the process

$$Z_t := \exp\Big\{\sum_{j=1}^d \int_0^t b_j(s, X_s)\,\mathrm{d}X_s^{(j)} - \frac{1}{2}\int_0^t |b(s,X_s)|^2\,\mathrm{d}s\Big\},\ t \geq 0$$

is a continuous martingale, so Girsanov's theorem implies that under Q given by $\mathrm{d}Q/\mathrm{d}\mathbb{P}|_{\mathcal{F}_T} = Z_T$, the process

$$W_t := X_t - x - \int_0^t b(s, X_s)\,\mathrm{d}s;\ 0 \leq t \leq T$$

is an $\mathbb{F}_T$-Brownian motion (starting at 0). Therefore,

$$X_t = x + \int_0^t b(s, X_s)\,\mathrm{d}s + W_t;\ 0 \leq t \leq T,$$

and hence $(X, W), (\Omega, \mathcal{F}_T, \mathbb{F}_T, Q)$ is a weak solution of (3.2.1) with initial condition $Q(X_0 = x) = 1$. □

Remark 3.9. The assumption that b is bounded is stronger than necessary but simplifies the argument. For a more general result (which also allows the initial condition to be random), see [KS91], Proposition 5.3.6 and the remark following that proposition. Uniqueness in law for (3.2.1) is discussed in Proposition 5.3.10 in [KS91].

3.3 Martingale Problems

We continue to consider stochastic differential equations of the form

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sum_{k=1}^m \sigma_k(t, X_t)\,\mathrm{d}W_t^k, \tag{3.3.1}$$

with measurable coefficients as before. The $d \times d$ (symmetric and nonnegative definite) matrix

$$a_{ij}(t,x) := \sum_{k=1}^m (\sigma_k(t,x))_i(\sigma_k(t,x))_j;\ i,j \in \{1,...,d\}$$

is called the *diffusion matrix* associated to (3.3.1). We denote by $C_c^k(\mathbb{R}^d)$ the space of real-valued k-times continuously differentiable functions on $\mathbb{R}^d$ with compact support, and $C_c^\infty(\mathbb{R}^d) := \bigcap_{k\in\mathbb{N}} C_c^k(\mathbb{R}^d)$. Then the following holds.

Proposition 3.10. *Let $(X, W), (\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a weak solution of (3.3.1). Then for all $f \in C^2(\mathbb{R}^d)$, the process*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \Big(\sum_{i=1}^d b_i(s, X_s)\,\partial_i f(X_s) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(s, X_s)\partial_{ij}^2 f(X_s)\Big)\,\mathrm{d}s$$

is a continuous local $\mathbb{F}$-martingale with $M_0^f = 0$. If b and $\sigma_1, ..., \sigma_k$ are locally bounded (in both variables) and $f \in C_c^2(\mathbb{R}^d)$, then M^f is a continuous L^2-martingale.

Proof. This is a straightforward application of Itô's formula. □

The idea of Stroock and Varadhan ([SV79]) was to *characterize* a weak solution as the solution of a (local) martingale problem. Let

$$\mathcal{S}_d := \{A \in \mathbb{R}^{d\times d} : A \text{ is symmetric and nonnegative definite}\}$$

and let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $a : [0,\infty) \times \mathbb{R}^d \to \mathcal{S}_d$ be measurable. For $t \geq 0$ we define the second order differential operator $\mathcal{A}_t$ by

$$\begin{aligned}\mathcal{A}_t f(x) &:= \sum_{i=1}^d b_i(t,x)\,\partial_i f(x) + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(t,x)\partial_{ij}^2 f(x)\\ &= \langle b(t,x), \nabla f(x)\rangle + \frac{1}{2}\mathrm{trace}\big(a(t,x)\mathrm{Hess}_f(x)\big), \qquad f \in C^2(\mathbb{R}^d).\end{aligned}$$

Definition 3.11. Let b and a be as above. A continuous $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t\geq 0}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a solution to the *local martingale problem* for (a, b) or for the associated family of operators $(\mathcal{A}_t)_{t\geq 0}$ if for each $f \in C_c^\infty(\mathbb{R}^d)$, the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{A}_s f(X_s)\,\mathrm{d}s \tag{3.3.2}$$

is a continuous local martingale with respect to the filtration generated by X. In short, we say that X solves $\mathrm{LMP}(a, b)$. If X has initial distribution $\mu = \mathbb{P}X_0^{-1}$, then we say that X solves $\mathrm{LMP}(a, b, \mu)$. If (3.3.2) is a martingale for each $f \in C_c^\infty(\mathbb{R}^d)$, then we say that X solves the associated *martingale problem* and write $\mathrm{MP}(a, b)$ and $\mathrm{MP}(a, b, \mu)$.

Finally, we say that the solution to $\mathrm{LMP}(a, b, \mu)$ or $\mathrm{MP}(a, b, \mu)$ is *unique* if for any two solutions X and $\tilde{X}$ (on possibly different spaces) the laws of X and $\tilde{X}$ coincide.

Remark 3.12. If X solves $\mathrm{LMP}(a, b, \mu)$, then we can transfer X to the following *canonical* set-up: consider the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}) := (C([0,\infty), \mathbb{R}^d), \mathcal{B}(C([0,\infty), \mathbb{R}^d)), \mathcal{L}(X))$, where $\mathcal{L}(X)$ denotes the law of X and $C([0,\infty), \mathbb{R}^d)$ is equipped with the usual topology (one can easily find a complete metric which makes the space Polish). Define the canonical process $\pi_t(\tilde{\omega}) := \tilde{\omega}_t$, $t \geq 0$. Then π has the same law as X and therefore π also solves $\mathrm{LMP}(a, b, \mu)$. In particular, we can identify a solution X to $\mathrm{LMP}(a, b, \mu)$ with the probability measure $\mathcal{L}(X)$ on the measurable space $(C([0,\infty), \mathbb{R}^d), \mathcal{B}(C([0,\infty), \mathbb{R}^d)))$. Analogous statements hold for $\mathrm{LMP}(a, b)$, $\mathrm{MP}(a, b)$, and $\mathrm{MP}(a, b, \mu)$. Therefore, we will often use formulations like "Let P be a solution of $\mathrm{LMP}(a, b, \mu)$" if P is the law of a solution of $\mathrm{LMP}(a, b, \mu)$.

We know already from the previous proposition that weak solutions solve the corresponding local martingale problem. Theorem 3.15 shows that the converse is also true. Before we formulate that result, we provide a representation theorem due to Doob of an $\mathbb{R}^d$-valued local martingale as a stochastic integral with respect to a Brownian motion on a possibly enlarged probability space. We will need that result in the proof of Theorem 3.15 but it is also of interest otherwise.

Definition 3.13. If $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a FPS and $m \in \mathbb{N}$ is fixed, then we define an *m-extension* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ as follows. Let W be an m-dimensional $\hat{\mathbb{F}}$-Brownian motion on a FPS $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$ and define $\tilde{\Omega} := \Omega \times \hat{\Omega}$, $\check{\mathcal{F}} := \mathcal{F} \otimes \hat{\mathcal{F}}$, $\check{\mathcal{F}}_t := \mathcal{F}_t \otimes \hat{\mathcal{F}}_t$, $\check{\mathbb{P}} := \mathbb{P} \otimes \hat{\mathbb{P}}$. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ be the smallest FPS which contains (or extends) $(\tilde{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{\mathbb{P}})$ and which satisfies the usual conditions. Define $\tilde{W}_t(\tilde{\omega}) = \tilde{W}_t(\omega, \hat{\omega}) := W_t(\hat{\omega})$, $t \geq 0$. Note that $\tilde{W}$ is a $\tilde{\mathbb{F}}$-Brownian motion. For any adapted process A on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ we define $\tilde{A}_t(\tilde{\omega}) = \tilde{A}_t(\omega, \hat{\omega}) := A_t(\omega)$, $t \geq 0$. Note that $\tilde{A}$ is $\tilde{\mathbb{F}}$-adapted and $\tilde{A}$ is $\tilde{\mathbb{F}}$-progressive if A is $\mathbb{F}$-progressive. We will drop the tildes whenever there is no danger of confusion.

Theorem 3.14. *Let M be a continuous $\mathbb{R}^d$-valued local $\mathbb{F}$-martingale with $M_0 = 0$ on some FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ such that*

$$\langle M^i, M^j \rangle_t = \sum_{k=1}^{m} \int_0^t V_s^{ik} V_s^{jk} \, \mathrm{d}s, \quad t \geq 0 \tag{3.3.3}$$

for progressive processes $(V_t^{ik})_{t \geq 0}$, $1 \leq i \leq d$, $1 \leq k \leq m$. Then on any m-extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ there exists an $\mathbb{R}^m$-valued $\tilde{\mathbb{F}}$-Brownian motion B such that

$$M_t^i = \sum_{k=1}^{m} \int_0^t V_s^{ik} \mathrm{d}B_s^k.$$

Proof. For each $t \geq 0$ we regard $V_t := \left(V_t^{ik}\right)_{i=1,\ldots,d;k=1,\ldots,m}$ as a linear map from $\mathbb{R}^m$ to $\mathbb{R}^d$ (or a $d \times m$-matrix). Let $N_t \subseteq \mathbb{R}^m$ be its null space and $R_t \subseteq \mathbb{R}^d$ its range. Let $V_t^{-1} : R_t \to N_t^{\perp}$ be the (bijective) inverse of the restriction of V_t to $N_t^{\perp}$ and denote the orthogonal projection from a Euclidean space to a subspace U by π^U. Let W be an m-dimensional Brownian motion as in the previous definition (in particular W is independent of M) and define

$$B_t := \int_0^t V_s^{-1} \pi^{R_s} \, \mathrm{d}M_s + \int_0^t \pi^{N_s} \, \mathrm{d}W_s, \quad t \geq 0,$$

where we interpret $V_s^{-1} \pi^{R_s}$ as an $m \times d$-matrix (or linear map from $\mathbb{R}^d$ to $\mathbb{R}^m$). Note that the integrands are progressive and the stochastic integrals are well-defined! Then the covariation (matrix) of the continuous local martingale B is

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle B^k, B^l \rangle = \sum_{i,j=1}^{d} \left(V_t^{-1} \pi^{R_t}\right)_{ki} \left(V_t^{-1} \pi^{R_t}\right)_{kj} \frac{\mathrm{d}}{\mathrm{d}t} \langle M^i, M^j \rangle + \sum_{\nu=1}^{m} \left(\pi^{N_t}\right)_{k\nu} \left(\pi^{N_t}\right)_{l\nu}.$$

Inserting (3.3.3) we obtain (in matrix notation with $.^T$ denoting the transpose)

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle B \rangle_t = V_t^{-1} \pi^{R_t} V_t V_t^T \left(V_t^{-1} \pi^{R_t}\right)^T + \pi^{N_t} \left(\pi^{N_t}\right)^T = \pi^{N_t^{\perp}} \left(\pi^{N_t^{\perp}}\right)^T + \pi^{N_t} \left(\pi^{N_t}\right)^T = \pi^{N_t^{\perp}} + \pi^{N_t} = I_m.$$

Therefore, Lévy's theorem shows that B is an m-dimensional Brownian motion. Further, using the fact that $V_s \pi^{N_s} \equiv 0$ and for $N \in \mathcal{M}_{loc}^0$, $\langle N \rangle \equiv 0$ implies $N \equiv 0$,

$$\begin{aligned}
\int_0^t V_s \, \mathrm{d}B_s &= \int_0^t V_s V_s^{-1} \pi^{R_s} \, \mathrm{d}M_s + \int_0^t V_s \pi^{N_s} \, \mathrm{d}W_s \\
&= \int_0^t \pi^{R_s} \, \mathrm{d}M_s + 0 \\
&= \int_0^t \pi^{R_s} \, \mathrm{d}M_s + \int_0^t \pi^{R_s^{\perp}} \, \mathrm{d}M_s = M_t,
\end{aligned}$$

so the proof is complete. $\square$

Theorem 3.15. *Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $a : [0,\infty) \times \mathbb{R}^d \to \mathcal{S}_d$ be measurable. Suppose that X is a solution to LMP(a, b) on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $m \in \mathbb{N}$ and let $\sigma : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be a measurable function such that $a = \sigma\sigma^T$ (note that such a function always exist when $m \geq d$). Let $\mathbb{F}$ be the complete filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ generated by X. Then, on any m-extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ of the FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ there exists an $\mathbb{R}^m$-valued $\tilde{\mathbb{F}}$-Brownian motion B such that $(X, B), (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ is a weak solution of* (3.3.1).

Proof. Fix $i \in \{1, ..., d\}$. For each $n \in \mathbb{N}$, choose $f_n^{(i)} \in C_c^\infty(\mathbb{R}^d)$ with $f_n^{(i)}(x) = x_i$ for all $|x| \leq n$. For each $n \in \mathbb{N}$, choose a localizing sequence $(T_m^{(n)})_{m\in\mathbb{N}}$ of stopping times for the continuous local martingale $\big(M_t^{f_n^{(i)}}\big)_{t\geq 0}$. We can and will assume that $T_1^{(n+1)} \geq T_n^{(n)} \vee T_{n+1}^{(1)}$ for all $n \in \mathbb{N}$. Then $\big(M_{t\wedge T_m^{(n)}}^{f_n^{(i)}}\big)_{t\geq 0}$ is a martingale for every $m, n \in \mathbb{N}$. Define the stopping time $\tau_n := \inf\{t \geq 0 : |X_t| \geq n\}$. Then

$$M_t^{f_n^{(i)}} = X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(s, X_s)\,\mathrm{d}s, \qquad t \leq \tau_n.$$

Observe that $S_n := \tau_n \wedge T_n^{(n)}$ is an increasing sequence of stopping times with $S_n \uparrow \infty$ almost surely and that

$$M_{t\wedge S_n}^{f_n^{(i)}} = X_{t\wedge S_n}^{(i)} - X_0^{(i)} - \int_0^{t\wedge S_n} b_i(s, X_s)\,\mathrm{d}s$$

is a martingale for each n. Consequently,

$$M_t^{(i)} := X_t^{(i)} - X_0^{(i)} - \int_0^t b_i(s, X_s)\,\mathrm{d}s \tag{3.3.4}$$

is a continuous local martingale. This formula also shows that X is a continuous semimartingale.

Next, fix $i, j \in \{1, ..., d\}$ and choose $f_n^{(i,j)} \in C_c^\infty(\mathbb{R}^d)$ such that $f_n^{(i,j)}(x) = x_i x_j$ whenever $|x| \leq n$. As above, one can show that

$$M_t^{(i,j)} := X_t^{(i)} X_t^{(j)} - X_0^{(i)} X_0^{(j)} - \int_0^t \big(X_s^{(i)} b_j(s, X_s) + X_s^{(j)} b_i(s, X_s) + a_{ij}(s, X_s)\big)\,\mathrm{d}s$$

is a continuous local martingale. Integrating by parts and using (3.3.4), we obtain

$$\begin{aligned} M_t^{(i,j)} &= \int_0^t X_s^{(i)}\,\mathrm{d}X_s^{(j)} + \int_0^t X_s^{(j)}\,\mathrm{d}X_s^{(i)} + \langle X^{(i)}, X^{(j)}\rangle_t \\ &\quad - \int_0^t \big(X_s^{(i)} b_j(s, X_s) + X_s^{(j)} b_i(s, X_s) + a_{ij}(s, X_s)\big)\,\mathrm{d}s \\ &= \int_0^t X_s^{(i)}\,\mathrm{d}M_s^{(j)} + \int_0^t X_s^{(j)}\,\mathrm{d}M_s^{(i)} + \langle M^{(i)}, M^{(j)}\rangle_t - \int_0^t a_{ij}(s, X_s)\,\mathrm{d}s. \end{aligned}$$

Therefore, the process $\langle M^{(i)}, M^{(j)}\rangle_t - \int_0^t a_{ij}(s, X_s)\,\mathrm{d}s$ is a continuous local martingale starting at 0 which is of locally finite variation, so Theorem 1.57 in WT3 implies that the process is 0 almost surely.

We have shown that the processes $M^{(i)}$ are continuous local martingales starting at 0 such that $\langle M^{(i)}, M^{(j)}\rangle_t = \int_0^t a_{ij}(s, X_s)\,\mathrm{d}s$. Therefore, the claim follows from Theorem 3.14. □

Remark 3.16. Note that the previous theorem establishes a one-to-one correspondence between weak solutions and solutions of martingale problems (up to the non-unique choice of the matrix σ when a is given). In particular: the sde (3.3.1) with coefficients b and σ has a weak solution if and only if $\mathrm{LMP}(\sigma\sigma^T, b)$ has a solution and we have weak uniqueness if and only $\mathrm{LMP}(\sigma\sigma^T, b, \mu)$ has at most one solution $P \in \mathcal{M}_1(C([0,\infty), \mathbb{R}^d))$ for every $\mu \in \mathcal{M}(\mathbb{R}^d)$. Note that the law of a weak solution depends on σ only via $\sigma\sigma^T$. Note also that the local martingale formulation does not say much about pathwise uniqueness. Indeed, for equation (3.1.2), we have $\sigma(x) = \mathrm{sign}(x)$ and therefore $a(x) = 1$. We showed that weak uniqueness holds and that a weak solution exists for any initial law μ. Therefore the associated $\mathrm{LMP}(1, 0, \mu)$ has a unique solution (namely standard Brownian motion with initial distribution μ). Note that for the particular square root $\sigma(x) = \mathrm{sign}(x)$ of $a = 1$ pathwise uniqueness does not hold and we do not have a strong solution (at least not if the initial condition is 0), while for the root $\tilde{\sigma}(x) = 1$ pathwise uniqueness holds and we have a strong solution.

3.4 Martingale Problems: Existence of solutions

We already stated an existence result for the solution of a local martingale problem in case $a \equiv I_d$ using Girsanov's theorem. Our aim in this section is to prove a similar statement for more general functions a. Before we formulate and prove an existence result for a solution of $\mathrm{LMP}(a, b)$ we present a useful result about weak convergence (see [K02], Theorem 4.27) and then a tightness criterion for probability measures on the space $C([0,\infty), \mathbb{R}^d)$.

Proposition 3.17. *Let E and $\tilde{E}$ be metric spaces, let $\mu, \mu_1, ... \in \mathcal{M}_1(E)$ such that $\mu_n \Rightarrow \mu$ and let $f, f_1, f_2, ... : E \to \tilde{E}$ be Borel-measurable mappings. Assume that there exists a set $C \in \mathcal{B}(E)$ such that $\mu(C) = 1$ and $f_n(x_n) \to f(x)$ whenever $x_n \to x \in C$. Then $\mu_n f_n^{-1} \Rightarrow \mu f^{-1}$.*

Remark 3.18. In WT2 we treated the special case in which all functions f_n equal f and f is continuous at μ-almost every point in E.

Proof of Proposition 3.17. Fix an open set $G \subset \tilde{E}$ and let $x \in f^{-1}(G) \cap C$ in case the set in non-empty. Then there exists an open neighborhood N of x and some $m \in \mathbb{N}$ such that $f_k(x') \in G$ for all $k \geq m$ and all $x' \in N$. Thus, $N \subset \bigcap_{k \geq m} f_k^{-1}(G)$, and so

$$f^{-1}(G) \cap C \subset \bigcup_m \Big(\bigcap_{k \geq m} f_k^{-1}(G) \Big)^\circ,$$

where A° denotes the interior of the set A. Using the Portmanteau theorem from WT2, we get

$$\begin{aligned}
\mu\big(f^{-1}(G)\big) &\leq \mu\Big(\bigcup_m \Big(\bigcap_{k \geq m} f_k^{-1}(G) \Big)^\circ \Big) = \sup_m \mu\Big(\Big(\bigcap_{k \geq m} f_k^{-1}(G) \Big)^\circ \Big) \\
&\leq \sup_m \liminf_{n \to \infty} \mu_n \Big(\bigcap_{k \geq m} f_k^{-1}(G) \Big) \leq \liminf_{n \to \infty} \mu_n \Big(f_n^{-1}(G) \Big).
\end{aligned}$$

Again using the Portmanteau theorem the claim follows. □

Consider the space $\mathcal{C}_d := C([0,\infty), \mathbb{R}^d)$ equipped with the metric

$$\rho(f, g) := \sum_{k=1}^{\infty} 2^{-k} \Big(\max_{x \in [0,k]} |f(x) - g(x)| \wedge 1 \Big).$$

It is easy to see that $(\mathcal{C}_d,\rho)$ is complete and separable (and hence Polish) and that a sequence of elements of $\mathcal{C}_d$ converges with respect to ρ iff the sequence converges uniformly on every compact subset of $[0,\infty)$. For $t \geq 0$, define the projection maps $\pi_t : \mathcal{C}_d \to C([0,t],\mathbb{R}^d)$ by $(\pi_t(f))(s) := f_s$, $s \in [0,t]$ and $\tilde{\pi}_t : \mathcal{C}_d \to \mathbb{R}^d$ by $\tilde{\pi}_t(f) := f_t$. Clearly, π_t and $\tilde{\pi}_t$ are continuous if $C([0,t],\mathbb{R}^d)$ is equipped with the supremum norm. The Borel-σ-algebra $\mathcal{B}(\mathcal{C}_d)$ on $\mathcal{C}_d$ coincides with the σ-algebra $\mathcal{B}$ respectively $\tilde{\mathcal{B}}$ generated by the maps π_t, $t \geq 0$ respectively $\tilde{\pi}_t$, $t \geq 0$: the fact that continuous maps are measurable shows that $\mathcal{B} \subset \mathcal{B}(\mathcal{C}_d)$ and $\tilde{\mathcal{B}} \subset \mathcal{B}(\mathcal{C}_d)$. It is easy to see that $\tilde{\mathcal{B}} \subset \mathcal{B}$. To see the inclusion $\mathcal{B}(\mathcal{C}_d) \subset \tilde{\mathcal{B}}$ consider all sets of the form $\{g \in \mathcal{C}_d : \sup_{x\in[0,N]} |g(x) - f(x)| < \varepsilon\}$, where $\varepsilon > 0$, $N \in \mathbb{N}$ and $f \in \mathcal{C}$. These sets are open and are in $\tilde{\mathcal{B}}$ (cf. the second part of the proof of Lemma 4.28 in WT2) and every open set in $\mathcal{C}_d$ is a countable union of countable intersections of such sets and thus in $\tilde{\mathcal{B}}$. We define the *modulus of continuity* of $f \in \mathcal{C}_d$ on $[0,t]$ by $w_t^f(h) := \sup\{|f_r - f_s|;\, r,s \in [0,t], |r-s| \leq h\}$, $h > 0$.

Proposition 3.19. *The set $\Gamma \subset \mathcal{M}_1(\mathcal{C}_d)$ is relatively compact (and hence tight) iff the following two conditions hold:*

(i) For each $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \mathbb{R}^d$ such that $\mu(\{f \in \mathcal{C}_d : f_0 \in K_\varepsilon\}) \geq 1-\varepsilon$ for every $\mu \in \Gamma$.

(ii) For each $T > 0$, $\varepsilon > 0$ and $\delta > 0$ there exists some $h > 0$ such that

$$\mu\big(\{f \in \mathcal{C}_d : w_T^f(h) \geq \delta\}\big) \leq \varepsilon \text{ for all } \mu \in \Gamma.$$

Proof. This follows easily from the Arzelà-Ascoli theorem. We will show the details in class. □

We will use the following sufficient tightness criterion on $\mathcal{C}_d$.

Proposition 3.20. *Let Y^i, $i \in I$ be a family of $\mathcal{C}_d$-valued random variables on possibly different spaces $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$. A sufficient condition for tightness of $\Gamma := \{\mathcal{L}(Y^i),\, i \in I\}$ is that the following two conditions hold:*

(i) For each $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \mathbb{R}^d$ such that $\mathbb{P}_i(\{Y_0^i \in K_\varepsilon\}) \geq 1-\varepsilon$ for every $i \in I$.

(ii) For each $T > 0$ there exist $a_T, b_T, c_T > 0$ such that

$$\mathbb{E}_i\big(|Y_t^i - Y_s^i|^{a_T}\big) \leq c_T|t-s|^{1+b_T};\ s,t \in [0,T],\, i \in I.$$

Proof. Clearly (i) of Proposition 3.19 is the same as (i) of Proposition 3.20. To see that (ii) of Proposition 3.19 holds, fix $T > 0$ and define $Z_t^i := Y_{tT}^i$, $t \geq 0$. Then, for $s,t \in [0,1]$,

$$\mathbb{E}_i\big(|Z_t^i - Z_s^i|^{a_T}\big) = \mathbb{E}_i\big(|Y_{tT}^i - Y_{sT}^i|^{a_T}\big) \leq c_T T^{1+b_T}|t-s|^{1+b_T}.$$

Choose $\kappa_T \in (0, b_T/a_T)$. Then Kolmogorov's continuity theorem as e.g. in Theorem 5.2 in the WT3 Lecture notes implies

$$\mathbb{P}_i\big(w_T^{Y^i}(h) \geq \delta\big) = \mathbb{P}_i\big(w_1^{Z^i}(h/T) \geq \delta\big) \leq \frac{1}{\delta^{a_T}}\mathbb{E}_i\Big(\big(w_1^{Z^i}(h/T)\big)^{a_T}\Big) \leq \frac{1}{\delta^{a_T}} h^{\kappa_T a_T}\xi_T,$$

for some $\xi_T > 0$ (which does not depend on $i \in I$ nor on δ and h), so the claim follows from the previous proposition. □

Theorem 3.21. *Let $a : \mathbb{R}^d \to \mathcal{S}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ be bounded and continuous and let $x \in \mathbb{R}^d$. Then LMP(a, b, δ_x) has a solution.*

Proof. The idea of the proof is to approximate the functions a and b by smoother functions a_n and b_n such that we know that $\mathrm{LMP}(a_n, b_n, \delta_x)$ has a solution $P_n \in \mathcal{M}_1(\mathcal{C}_d)$ and to show that the sequence P_n, $n \in \mathbb{N}$ is tight. Since $\mathcal{C}_d$ is a Polish space, we know from Prohorov's theorem (WT2) that there exists a subsequence of P_n, $n \in \mathbb{N}$ which converges weakly to some $P \in \mathcal{M}_1(\mathcal{C}_d)$. Then it remains to show that P solves $\mathrm{LMP}(a, b, \delta_x)$.

We choose $b_n : \mathbb{R}^d \to \mathbb{R}^d$ such that $b_n \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and $b_n \to b$ uniformly on $\mathbb{R}^d$ (it is clear that such a sequence exists). To define a_n, we first choose a continuous map $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ such that $a(x) = \sigma(x)\sigma^T(x)$ (note that σ is automatically bounded). Choose $\sigma_n \in C^\infty(\mathbb{R}^d, \mathbb{R}^{d\times d})$ such that $\sigma_n \to \sigma$ uniformly in $\mathbb{R}^d$ and define $a_n := \sigma_n \sigma_n^T$. Without loss of generality we can assume that $\sup_{x\in\mathbb{R}^d} \big(|b_n(x)| + \|a_n(x)\|\big) \le B < \infty$ for some $B \in \mathbb{R}$. Denote the k-th column of σ_n by σ_{nk}. We saw in WT3 that a stochastic differential equation with bounded (in fact linearly bounded is enough) and locally Lipschitz continuous coefficients b_n, $\sigma_{n1}, ..., \sigma_{nd}$ and deterministic initial condition $x \in \mathbb{R}^d$ has a (unique) strong solution X^n defined on any FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ which carries a d-dimensional Brownian motion W. Note that we can choose the same space and the same Brownian motion for every $n \in \mathbb{N}$. In particular, the equation has a weak solution for each $n \in \mathbb{N}$ and thus $\mathrm{LMP}(a_n, b_n, \delta_x)$ has a solution $P_n \in \mathcal{M}_1(\mathcal{C}_d)$, namely the law of X^n. We show that $\Gamma := \{P_n\}_{n\in\mathbb{N}}$ is tight by verifying the assumptions of Proposition 3.20. Condition (i) holds since $P_n(f \in \mathcal{C}_d : f(0) = x) = 1$. To show (ii) in Proposition 3.20, fix $T > 0$. Then, for $0 \le s < t \le T$ and $p \ge 2$ and using Burkholder's inequality, we have

$$\begin{aligned}
\mathbb{E}\big(|X_t^n - X_s^n|^p\big) &= \mathbb{E}\Big(\Big|\int_s^t b_n(X_u^n)\,\mathrm{d}u + \sum_{k=1}^d \int_s^t \sigma_{nk}(X_u^n)\,\mathrm{d}W_u^k\Big|^p\Big) \\
&\le (d+1)^{p-1}\Big(\mathbb{E}\Big(\Big|\int_s^t b_n(X_u^n)\,\mathrm{d}u\Big|^p\Big) + \sum_{k=1}^d \mathbb{E}\Big(\Big|\int_s^t \sigma_{nk}(X_u^n)\,\mathrm{d}W_u^k\Big|^p\Big)\Big) \\
&\le (d+1)^{p-1}\Big(B^p(t-s)^p + C_p d^{1+\frac{p}{2}} B^{p/2}(t-s)^{p/2}\Big),
\end{aligned}$$

so choosing $p > 2$ Proposition 3.20 shows that the set Γ is tight. Prohorov's theorem (WT2) implies that there is some $P \in \mathcal{M}_1(\mathcal{C}_d)$ and a subsequence of (P_n) which converges to P weakly. Without loss of generality we assume that the sequences b_n and a_n are such that $P_n \Rightarrow P$. Let X be the canonical process on $(\mathcal{C}_d, \mathcal{B}(\mathcal{C}_d), P)$. It remains to verify that X solves $\mathrm{LMP}(a, b, \delta_x)$. Since a and b are bounded this is the same as showing that X solves $\mathrm{MP}(a, b, \delta_x)$ which is equivalent to showing that for every $0 \le s < t < \infty$, any $f \in C_c^\infty(\mathbb{R}^d)$ and any bounded and continuous function $g : C([0, s], \mathbb{R}^d) \to \mathbb{R}$

$$\mathbb{E}\Big(\Big(f(X_t) - f(X_s) - \int_s^t \mathcal{A}_r f(X_r)\,\mathrm{d}r\Big) g(X_u, u \in [0, s])\Big) = 0 \tag{3.4.1}$$

(by an application of the monotone class theorem), where E denotes the expectation with respect to P. We fix s, t and f and write equation (3.4.1) as $E\varphi(\pi_t \circ X) = 0$ where $\varphi : C([0, t], \mathbb{R}^d) \to \mathbb{R}$ is bounded and continuous. We know that the corresponding equation $\mathbb{E}\varphi_n(\pi_t \circ X^n) = \int_{\mathcal{C}_d} \varphi_n(\pi_t(h))\,\mathrm{d}P_n(h) = 0$ holds where φ_n is defined like φ but with $\mathcal{A}_r$ replaced by $\mathcal{A}_r^n$. Note that $\varphi_n \to \varphi$ uniformly, so Proposition 3.17 implies $E\varphi(X) = 0$, so P (or X) solve $\mathrm{(L)MP}(a, b, \delta_x)$. $\square$

Remark 3.22. The previous theorem can be generalized considerably (see for example [K02]): the functions a and b can be allowed to be time-dependent, random, and to depend also on the past values of the solution process. Further, boundedness can be weakened to a linear growth condition and δ_x can be replaced by an arbitrary $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

3.5 The Yamada-Watanabe results

In this section we will not always supply complete proofs. Before we state and prove the Yamada-Watanabe theorem let us show that two weak solutions $(X^j, W^j), (\Omega^j, \mathcal{F}^j, \mathbb{F}^j, \mathbb{P}^j)$ of equation (3.3.1) with the same given initial law $\mathcal{L}(X_0^j) = \mu$, $j = 1, 2$ can be defined on a common probability space. For $j = 1, 2$ we define $Y_t^j := X_t^j - X_0^j$, $t \geq 0$. Let $\Theta := \mathbb{R}^d \times \mathcal{C}_m \times \mathcal{C}_d$ be equipped with its Borel σ-algebra $\mathcal{B}(\Theta)$ and let $P_j \in \mathcal{M}_1(\Theta)$ be the law of the triple (X_0^j, W^j, Y^j), $j = 1, 2$. It may well be that $P_1 \neq P_2$ (even if we have weak uniqueness (which we didn't assume)!), but we know that the projection of P_j on the first two coordinates is $\mu \otimes \mathcal{W}$ for both $j = 1, 2$ where $\mathcal{W}$ is the law of an m-dimensional Brownian motion. Since we only have weak solutions, the third coordinate of the canonical process on $(\Theta, \mathcal{B}(\Theta), P_j)$ will generally not be a function of the first two, so the conditional law of the third coordinate given the first two will be non-trivial (i.e. not a Dirac measure) in general. Since the space Θ is Polish the important theorem on the existence of regular conditional distributions (e.g. Satz und Definition 8.19 in the Chapter *Bedingte Erwartungen und Wahrscheinlichkeiten* of WT1 or [IW89], p.12-16) tells us that for the σ-algebra $\mathcal{G}$ generated by the first two coordinates of the canonical process on $(\Theta, \mathcal{B}(\Theta), P_j)$, there exists a Markov kernel Q_j from $(\Theta, \mathcal{G})$ to $(\Theta, \mathcal{B}(\Theta))$ such that $Q_j(\theta, A) = P_j(A|\mathcal{G})(\theta)$, P_j-almost surely for every $A \in \mathcal{B}(\Theta)$. Since Q_j depends on the first two coordinates of θ only, we write $Q_j((x, w), A)$ instead of $Q_j((x, w, y), A)$ for $\theta = (x, w, y) \in \Theta$.

The impatient reader may ask if this means that we have found a weak solution of our sde on the probability space $(\Theta, \mathcal{B}(\Theta), P_j)$. Strictly speaking it doesn't because we have not defined a filtration on that space. Before taking care of these things we look for a single space carrying both weak solutions. We choose the space $\Omega := \Theta \times \mathcal{C}_d$ equipped with its Borel σ-algebra. We define

$$\mathbb{P}(\mathrm{d}\omega) := Q_1(x, w, \mathrm{d}y_1)\, Q_2(x, w, \mathrm{d}y_2)\, \mathrm{d}\mu(x)\, \mathrm{d}\mathcal{W}(w),$$

where $\omega = (x, w, y_1, y_2) \in \Omega$. Let $\mathcal{F}$ be the $\mathbb{P}$-completion of $\mathcal{B}(\Omega)$ and let $\mathbb{F}$ be the augmented filtration generated by the coordinates. By construction, the first three coordinates have law P_1 and the first two together with the last one have law P_2. It is not hard to show that these triples correspond to the weak solutions of the original sde (but we will skip the proof; for example one has to check that the second coordinate is really an $\mathbb{F}$-Brownian motion).

Here is the first part of the Yamada-Watanabe theorem.

Theorem 3.23. *Pathwise uniqueness implies uniqueness in law.*

Proof. Consider two weak solutions $(X^j, W^j), (\Omega^j, \mathcal{F}^j, \mathbb{F}^j, \mathbb{P}^j)$ with given initial law $\mathcal{L}(X_0^j) = \mu$, $j = 1, 2$ of equation (3.3.1) as above and construct the FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ as above. Pathwise uniqueness implies that the sum of the first and third and the sum of the first and fourth coordinate are indistinguishable, so $P_1 = P_2$ which implies (even more than) weak uniqueness. □

Here is the second part of the Yamada-Watanabe theorem.

Theorem 3.24. *Weak existence and pathwise uniqueness imply strong existence.*

Proof. The idea of the proof is simple: we take some weak solution $(X^1, W^1), (\Omega^1, \mathcal{F}^1, \mathbb{F}^1, \mathbb{P}^1)$ and take another copy of it which we equip with indices 2 everywhere. Then perform the same construction as above and denote $Q := Q_1 = Q_2$. Let $\Delta := \{(f, f) : f \in \mathcal{C}_d\}$ be the diagonal in $\mathcal{C}_d \times \mathcal{C}_d$. Since $\mathcal{C}_d$ is a Polish space Δ is a measurable subset of $\mathcal{C}_d \times \mathcal{C}_d$. By pathwise uniqueness, we have $Q((x, w), .) \otimes Q((x, w), .)(\Delta) = 1$ for almost all (x, w) which means that $Q((x, w), .)$ is a Dirac measure for almost all (x, w), but this means that the third coordininate y is a (measurable) function F of (x, w) and so the process $x_t := x + y_t$, $t \geq 0$ is a strong solution with respect to the Brownian motion w. It is not hard to see that this implies that the original process X^1 can be written as the same function F applied to the pair (X_0^1, W^1) showing that X^1 is also a strong solution. □

3.6 Martingale Problems: Uniqueness of solutions and the strong Markov property

This section is rather sketchy. We will state results without proofs (they can be found in [K02], [KS91] or [SV79]).

Definition 3.25. A martingale problem (or an sde) is said to be *well posed* if for every initial condition $x \in \mathbb{R}^d$ it admits a unique (weak) solution.

Until further notice we will now assume that the coefficients b and σ (or a) are time-homogeneous. One remarkable result (see [K02], Theorem 21.11 or Theorem 4.20 in [KS91] [under the additional assumption that a and b are locally bounded] for a precise formulation and proof) is that well posedness of LMP(a, b) implies the strong Markov property of the induced Markov family P_x, $x \in \mathbb{R}^d$ in $\mathcal{M}_1(\mathcal{C}_d)$.

The proof of the following proposition can be found in [KS91], Proposition 4.27.

Proposition 3.26. *Suppose that for every $x \in \mathbb{R}^d$ any two solutions P and $\tilde{P}$ of LMP(a, b, δ_x) have the same marginal distribution, i.e. $\tilde{P}\tilde{\pi}_t^{-1} = P\tilde{\pi}_t^{-1}$ for all $t \geq 0$. Then $P = \tilde{P}$.*

This proposition can be used to show uniqueness using PDE methods (see Section 5.4.E in [KS91]).

A sufficient condition for uniqueness of LMP(a, b) is the following result of Stroock and Varadhan.

Theorem 3.27. *Let a be continuous and b be measurable and locally bounded. Assume that a is uniformly elliptic in the following sense: there exists $\lambda > 0$ such that*

$$\langle a(x)\xi, \xi\rangle = \sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2,\ x, \xi \in \mathbb{R}^d.$$

Then uniqueness of (L)MP(a, b) holds.

Remark 3.28. Note that the assumptions in the previous theorem do not guarantee existence of a solution of LMP(a, b) (since solutions can *blow up* in finite time).

Chapter 4

Kunita-type Stochastic Differential Equations and Stochastic Flows

4.1 Kunita-type stochastic differential equations

When we investigated stochastic differential equations so far, we fixed an initial condition $x \in \mathbb{R}^d$ and studied solutions with this fixed initial condition. We did not study the *joint* distribution of solutions starting at different initial conditions (except in a short section in WT3 where we studied continuity of the solution with respect to the initial condition). If we think of many light (non-interacting) particles on the surface of a fluid, then one can try to model their joint motion by a single stochastic differential equation. We can, for example, consider *all* initial conditions simultaneously. For such applications it seems unnatural to drive the equation just by finitely many Brownian motions. It is more natural to assume that particles which are far apart are driven by (almost) independent Brownian motions and this forces us to consider stochastic differential equations driven by an infinite number of Brownian motions. This is not a severe problem; one just has to be careful with infinite sums. As an alternative (suggested and developed by Kunita in [Ku90]) one can introduce semimartingale fields $F(t,x)$, $t \geq 0$, $x \in \mathbb{R}^d$ depending on the spatial parameter x which can be thought of as describing the local behaviour of a solution if a particle happens to be at location x. Formally, an associated Kunita-type sde is written as

$$\mathrm{d}X_t = F(\mathrm{d}t, X_t). \tag{4.1.1}$$

We will soon impose assumptions on the field F which guarantee that (4.1.1) does not only have a unique strong solution for every fixed initial condition but that it even generates a stochastic flow of homeomorphisms (which we will define shortly).

In the following we assume that all stochastic processes are defined on a common FPS $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ which satisfies the usual conditions. The following will be our minimal assumptions on the local martingale part M of F in what follows:

Assumption 4.1.

- $M : [0,\infty) \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is (jointly) measurable,
- $t \mapsto M(t,x,\omega)$ is continuous for all $x \in \mathbb{R}^d$ and $\omega \in \Omega$,
- $M(.,x) \in \mathcal{M}^0_{loc}$ for each fixed $x \in \mathbb{R}^d$.

We will sketch the definition of a stochastic integral with respect to M which generalizes the approach presented in WT3 but is largely analogous. For details see [Ku90], p.80ff. Let

$f : [0, \infty) \times \Omega \to \mathbb{R}^d$ be a *simple process*, i.e. there exist $\Delta = \{0 = t_0 < t_1 < ...\}$ such that $t_i \to \infty$ and $\mathcal{F}_{t_i}$-measurable $\mathbb{R}^d$-valued ξ_i, $i = 0, 1, ...$ such that

$$f_t(\omega) = \xi_0(\omega) 1_{\{0\}}(t) + \sum_{i=0}^{\infty} \xi_i(\omega) 1_{(t_i, t_{i+1}]}(t);\ t \geq 0,\ \omega \in \Omega. \tag{4.1.2}$$

For M as above we define the stochastic integral of a simple process f as follows:

$$M_t(f) := \sum_{i=0}^{\infty} \big(M(t \wedge t_{i+1}, \xi_i) - M(t \wedge t_i, \xi_i)\big). \tag{4.1.3}$$

Note that for each t at most one of the summands in the infinite sums in (4.1.2) and (4.1.3) is non-zero. Note further that the process $M_t(f)$ is adapted, has continuous paths and satisfies $M_0(f) = 0$. It is not hard to show that $M_t(f)$ is a local martingale. We will now impose an additional condition on the joint quadratic variation of the local martingales $M(., x)$, $x \in \mathbb{R}^d$ (which is a bit stronger than the condition in [Ku90]).

Assumption 4.2. There exists a continuous function $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\langle M(., x), M(., y)\rangle_t = t\, a(x, y),\ t \geq 0,\ x, y \in \mathbb{R}^d,$$

and that there exists some $\delta > 0$ such that for every compact set K in $\mathbb{R}^d$

$$\sup_{x,y,x',y' \in K; x \neq x', y \neq y'} \frac{|a(x, y) - a(x', y) - a(x, y') + a(x', y')|}{|x - x'|^\delta |y - y'|^\delta} < \infty.$$

Before we proceed with the definition of $M_t(f)$ for more general processes f, we provide an example which shows how this new stochastic integral is related to the one defined in WT3.

Example 4.3. Let $W_t^1, ..., W_t^m$ be independent standard $\mathbb{F}$-Brownian motions and let $\sigma_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, ..., m$ be measurable. Define

$$M(t, x) := \sum_{k=1}^{m} \sigma_k(x) W_t^k.$$

Clearly, M satisfies Assumption 4.1. Further, we have

$$\langle M(., x), M(., y)\rangle_t = t \sum_{k=1}^{m} \sigma_k(x) \sigma_k(y),$$

so the first part of Assumption 4.2 holds with $a(x, y) = \sum_{k=1}^{m} \sigma_k(x)\sigma_k(y)$. The second part of Assumption 4.2 holds in case all functions σ_k are locally δ-Hölder continuous for some $\delta > 0$. If f is a simple process as in (4.1.2), then

$$M_t(f) = \sum_{k=1}^{m} \int_0^t \sigma_k(f_s)\, \mathrm{d}W_s^k,$$

where the right hand side denotes the usual Itô integral (introduced in WT3).

We provide the following theorem (which is a special case of results in [Ku90]) without proof.

Theorem 4.4. *Let Assumption 4.2 hold and let $f : [0,\infty)\times\Omega \to \mathbb{R}^d$ be adapted with continuous paths. For a (deterministic) partition $\Delta := \{0 = t_0 < t_1 < ...\}$ such that $t_i \to \infty$ define*

$$f_t^{\Delta}(\omega) := \sum_{i=0}^{\infty} f_{t_i}(\omega) 1_{(t_i,t_{i+1}]}(t);\ t \geq 0,\ \omega \in \Omega.$$

Then,

a) $M_t(f) := \int_0^t M(\mathrm{d}s, f_s) := \lim_{|\Delta|\to 0} M_t(f^{\Delta})$ exists, where the limit is understood in the sense of uniform convergence on compact subsets of $[0,\infty)$ in probability (ucp),

b) $M_.(f) \in \mathcal{M}^0_{loc}$,

c) if g is another adapted process with continuous paths, then

$$\langle M_.(f), M_.(g)\rangle_t = \int_0^t a(f_s, g_s)\,\mathrm{d}s, \quad a.s.$$

We are now in a position to formulate a theorem about existence and uniqueness of solutions of Kunita-type stochastic differential equations of the form

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + M(\mathrm{d}t, X_t),\ X_0 = x \in \mathbb{R}^d. \tag{4.1.4}$$

We will impose the following assumptions on b and $M = (M^1, ..., M^d)$.

Assumption 4.5. a) $b : \mathbb{R}^d \to \mathbb{R}^d$ is continuous.

b) M^i satisfies Assumption 4.1, $i = 1, ..., d$.

c) $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuous and $\langle M^i(.,x), M^j(.,y)\rangle_t = t\,a_{ij}(x,y)$, $i,j \in \{1,...,d\}$.

d) Define

$$\mathcal{A}(x,y) := a(x,x) - a(x,y) - a(y,x) + a(y,y),\ x,y \in \mathbb{R}^d,$$

then, for each $R > 0$ there exists some $K_R > 0$ such that

$$2\langle b(x) - b(y), x - y\rangle + \mathrm{tr}\big(\mathcal{A}(x,y)\big) \leq K_R |x-y|^2$$

whenever $|x|, |y| \leq R$.

e) There exists $\tilde{K}$ such that $2\langle b(x), x\rangle + \mathrm{tr}\big(a(x,x)\big) \leq \tilde{K}\big(|x|^2 + 1\big)$ for all $x \in \mathbb{R}^d$.

Remark 4.6. Note that Assumption 4.5 d) implies that Assumption 4.2 holds for each M^i and a_{ii} with $\delta = 1/2$ (I will explain this in class; if you want to try by yourself, here is a hint: use the Kunita-Watanabe inequality). This guarantees that all stochastic integrals in (4.1.4) are defined.

Remark 4.7. Note that

$$\langle M^i(.,x) - M^i(.,y), M^j(.,x) - M^j(.,y)\rangle_t = t\mathcal{A}_{ij}(x,y),$$

so $\mathcal{A}(x,y)$ is symmetric. In fact $\mathcal{A}(x,y)$ is even non-negative definite since for $z \in \mathbb{R}^d$

$$0 \leq \langle \sum_i z_i\big(M^i(.,x) - M^i(.,y)\big)\rangle_t = \sum_{i,j} z_i z_j t\mathcal{A}_{ij}(x,y).$$

Theorem 4.8. *Under Assumption 4.5, the sde* (4.1.4) *has a unique (strong) solution for every initial condition* $x \in \mathbb{R}^d$.

Proof. The proof (via an Euler approximation) is (almost) exactly like that of Theorem 2.27 in WT3. Therefore we omit it. $\square$

Remark 4.9. The reader interested in seeing a proof of the previous theorem under even slightly weaker assumptions (some guaranteeing only local existence and uniqueness) is referred to [SS17].

The following proposition is the same as Proposition 2.35 in WT3 (and hence the proof – based on Kolmogorov's continuity theorem – is identical).

Proposition 4.10. *Assume that, in addition to the assumptions in Theorem 4.8, there exist* $p > d$ *such that* $p \geq 2$ *and* $K \geq 0$ *such that*

$$2\langle b(x) - b(y), x - y\rangle + \operatorname{tr}\big(\mathcal{A}(x,y)\big) + (p-2)\|\mathcal{A}(x,y)\| \leq K|x-y|^2$$

for all $x, y \in \mathbb{R}^d$, *where* $\|.\|$ *denotes the matrix norm induced by the Euclidean norm on* $\mathbb{R}^d$, *then there exists a modification* φ *of the solution map* $\phi : \mathbb{R}^d \times [0,\infty) \times \Omega \to \mathbb{R}^d$ *of* (4.1.4) *which is jointly continuous in* (t,x).

Remark 4.11. Recall that in the proof of the previous proposition we used Lemma 2.34 in WT3 which established the bound

$$\mathbb{E} \sup_{0 \leq t \leq T} |\phi_t(x) - \phi_t(y)|^q \leq \big(c_{q/p} + 1\big)|x-y|^q \exp\{KqT/2\}, \tag{4.1.5}$$

for any $T > 0$, $x, y \in \mathbb{R}^d$ and $q \in (0,p)$, where c_r is a constant whose numerical value was given in the WT3 notes. We then applied Kolmogorov's continuity theorem (choosing $q \in (d,p)$) to show the existence of a continuous modification φ of ϕ. In fact Kolmogorov's continuity theorem as in the appendix of WT3 can not only be used to prove the existence of a continuous modification φ but also to provide an upper bound on the growth of the diameter of $\varphi_T(A)$ for a bounded set $A \subset \mathbb{R}^d$. If, for example, we choose $A = [0,1]^d$, then we obtain for each $T > 0$, $\kappa \in (0, \frac{q-d}{q})$, and $u > 0$,

$$\mathbb{P}\Big(\sup_{0 \leq t \leq T} \sup_{x,y \in [0,1]^d} |\varphi_t(x) - \varphi_t(y)| \geq u\Big) \leq C(\kappa, q, d) \exp\{KqT/2\} u^{-q},$$

for some function $C(\kappa, q, d)$. Inserting $u = \exp\{\gamma T\}$, we see that

$$\limsup_{T \to \infty} \frac{1}{T} \log \mathbb{P}\Big(\sup_{0 \leq t \leq T} \sup_{x,y \in [0,1]^d} |\varphi_t(x) - \varphi_t(y)| \geq \exp\{\gamma T\}\Big) \leq \frac{Kq}{2} - \gamma q,$$

which is negative when $\gamma > \frac{K}{2}$. Under additional assumptions much sharper estimates can be obtained (see [SS17]). We mention that the claim of Proposition 4.10 is generally untrue even for an sde of the form

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}W_t,$$

when σ is bounded and C^∞ (but not globally Lipschitz). An example with $d = 2$ can be found in [LS11].

4.2 Stochastic flows generated by stochastic differential equations

We start by defining the concept of a stochastic (semi-)flow. We restrict our study to $\mathbb{R}^d$-valued continuous-time flows but of course this concept can be generalized to other state spaces and more general index sets.

Definition 4.12. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathbb{T} \in \{[0,\infty), \mathbb{R}\}$. A measurable map $\varphi : \mathbb{T}^2 \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is called a *stochastic flow* (of homeomorphisms) if there exists a set $N \in \mathcal{F}$ such that $\mathbb{P}(N) = 0$ and for all $\omega \notin N$ we have

i) $\varphi_{s,u}(.,\omega) = \varphi_{t,u} \circ \varphi_{s,t}(.,\omega),\ s,t,u \in \mathbb{T}$

ii) $\varphi_{s,s}(.,\omega) = \mathrm{id}_{\mathbb{R}^d},\ s \in \mathbb{T}$,

iii) $(s,t,x) \mapsto \varphi_{s,t}(x,\omega)$ is continuous.

If $\Delta := \{(s,t) \in \mathbb{T}^2 : s \le t\}$, then a measurable map $\varphi : \Delta \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is called a *(stochastic) semi-flow* if ii) holds and i) and iii) hold for all $(s,t),(t,u) \in \Delta$.

Remark 4.13. Note that if φ is a stochastic flow, then $\varphi_{s,t}(.,\omega)$ is a homeomorphism of $\mathbb{R}^d$ for all $s,t \in \mathbb{T}$ and all $\omega \notin N$. Further, the formula $\varphi_{t,s}(x,\omega) = \varphi_{s,t}^{-1}(x,\omega)$ holds for all $s,t \in \mathbb{T}$ and all $\omega \notin N$.

It is natural to conjecture that under the conditions provided in the previous section, the solution of an sde generates a stochastic semi-flow, i.e. one can find a modification φ of the solution map ϕ which is a stochastic semi-flow (see [SS17] for precise statements and proofs). Note that we cannot expect that an sde as above generates a stochastic *flow* since even in the deterministic case the solution map may not be one-to-one (we will provide an example in class). Rather than showing the semi-flow property for sdes as above, we impose slighly stronger conditions and show that under these conditions an sde even generates a stochastic flow.

Theorem 4.14. *Assume that $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\mathcal{A} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are Lipschitz in the following sense: there exist (deterministic) constants L_b and κ such that, for all $x,y \in \mathbb{R}^d$,*

$$|b(x) - b(y)| \le L_b|x-y|, \qquad \|\mathcal{A}(x,y)\| \le \kappa|x-y|^2.$$

Let $\mathbb{T} := [0,\infty)$ and let $\psi : \Delta \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ be the solution map associated to the sde

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + M(\mathrm{d}t, X_t),$$

i.e. $\psi_{s,t}(x)$, $(s,t) \in \Delta$, $x \in \mathbb{R}^d$ is a solution of

$$\psi_{s,t}(x) = x + \int_s^t b(\psi_{s,u}(x))\,\mathrm{d}u + \int_s^t M(\mathrm{d}u, \psi_{s,u}(x)).$$

Then there exists a stochastic flow φ such that for every $s \ge 0$ and $x \in \mathbb{R}^d$ there exists a set $N_{s,x}$ of measure 0 such that

$$\varphi_{s,t}(x,\omega) = \psi_{s,t}(x,\omega),\ t \ge s,\ x \in \mathbb{R}^d,\ \omega \notin N_{s,x}.$$

The proof is a bit long and will be split up into the following 5 steps. We basically follow the proof of Theorem 4.5.1 in [Ku90].

Step 1: There exists a continuous modification $\tilde{\psi} : [0,\infty) \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ of $\psi_{0,.}$.

Step 2: $\tilde{\psi}$ can be chosen in such a way that $\tilde{\psi}_t(.,\omega) : \mathbb{R}^d \to \mathbb{R}^d$ is one-to-one for all $t \geq 0$, $\omega \in \Omega$.

Step 3: $\tilde{\psi}$ can be chosen in such a way that $\tilde{\psi}_t(.,\omega) : \mathbb{R}^d \to \mathbb{R}^d$ is - in addition - onto for all $t \geq 0$, $\omega \in \Omega$.

Step 4: For $\tilde{\psi}$ as in Step 3 the map $(s,x) \mapsto \tilde{\psi}_s^{-1}(x,\omega)$ is continuous for every $\omega \in \Omega$.

Step 5: Define $\varphi_{s,t}(x,\omega) := \tilde{\psi}_t(\tilde{\psi}_s^{-1}(x,\omega),\omega)$, $s,t \geq 0$ and show that φ satisfies the claim in Theorem 4.14.

Note that we proved Step 1 already: it is a special case of Proposition 4.10. Nevertheless, a new proof will be given here.

Lemma 4.15. *There exists some $c \geq 0$ such that for all $x, y \in \mathbb{R}^d$ there exists a standard Brownian motion W (possibly on an enlarged probability space) such that for every $t \geq 0$, we have*

$$\sup_{0\leq s\leq t} |\psi_{0,s}(x) - \psi_{0,s}(y)| \leq |x-y| \exp\{ct + \sqrt{\kappa}\, W_t^*\}, \ \textit{a.s. and}$$
$$\inf_{0\leq s\leq t} |\psi_{0,s}(x) - \psi_{0,s}(y)| \geq |x-y| \exp\{-ct + \sqrt{\kappa}\, {}^*W_t\}, \ \textit{a.s.},$$

where $W_t^* := \sup_{0\leq s\leq t} W_s$ and ${}^*W_t := \inf_{0\leq s\leq t} W_s$.

Proof. Fix $x \neq y$ and $\varepsilon > 0$. Define

$$D_t := \psi_{0,t}(x) - \psi_{0,t}(y),\ Z_t^{(\varepsilon)} := \frac{1}{2}\log(|D_t|^2 + \varepsilon),$$

so

$$Z_t^{(\varepsilon)} = f_\varepsilon(D_t) \text{ for } f_\varepsilon(z) := \frac{1}{2}\log(|z|^2 + \varepsilon).$$

Using Itô's formula, we get

$$\begin{aligned} \mathrm{d}Z_t^{(\varepsilon)} =& \frac{D_t \cdot (M(\mathrm{d}t, \psi_{0,t}(x)) - M(\mathrm{d}t, \psi_{0,t}(y)))}{|D_t|^2 + \varepsilon} + \frac{D_t^T (b(\psi_{0,t}(x)) - b(\psi_{0,t}(y)))}{|D_t|^2 + \varepsilon}\mathrm{d}t \\ &+ \frac{1}{2}\frac{1}{|D_t|^2+\varepsilon}\mathrm{tr}\big(\mathcal{A}(\psi_{0,t}(x), \psi_{0,t}(y))\big)\,\mathrm{d}t - \sum_{i,j} \frac{D_t^i D_t^j}{(|D_t|^2+\varepsilon)^2}\mathcal{A}_{i,j}(\psi_{0,t}(x), \psi_{0,t}(y))\,\mathrm{d}t. \end{aligned}$$

Note that

$$N_t^{(\varepsilon)} := \int_0^t \frac{D_s}{|D_s|^2+\varepsilon} \cdot \big(M(\mathrm{d}s, \psi_{0,s}(x)) - M(\mathrm{d}s, \psi_{0,s}(y))\big)$$

is a continuous local martingale and we get

$$Z_t^{(\varepsilon)} = Z_0^{(\varepsilon)} + N_t^{(\varepsilon)} + \int_0^t \alpha_\varepsilon(s,\omega)\,\mathrm{d}s,$$

where

$$c := \sup_{x,y} \sup_{s} \sup_{\varepsilon>0} \operatorname{esssup}_{\omega} |\alpha_\varepsilon(s,\omega)| < \infty$$

and

$$\mathrm{d}\langle N^{(\varepsilon)}\rangle_t = \sum_{i,j} \frac{D_t^i D_t^j}{(|D_t|^2+\varepsilon)^2} \mathcal{A}_{i,j}(\psi_{0,t}(x), \psi_{0,t}(y))\,\mathrm{d}t \le \kappa \mathrm{d}t.$$

By the Dambis-Dubins-Schwarz theorem (Theorem 2.23 and Remark 2.24 in WT3), we can find a Wiener process $W^{(\varepsilon)}$ (possibly on an enlarged probability space) such that $N_t^{(\varepsilon)} = \sqrt{\kappa} W_{\tau(t)}^{(\varepsilon)}$ such that $\tau(t) \le t$ for all $t \ge 0$ almost surely. Therefore,

$$\begin{aligned} Z_t^{(\varepsilon)} &\ge Z_0^{(\varepsilon)} + \sqrt{\kappa} W_{\tau(t)}^{(\varepsilon)} - ct \\ &\ge \log(|x-y|) + \sqrt{\kappa} \inf_{0\le s\le t} W_s^{(\varepsilon)} - ct. \end{aligned}$$

The law of the right hand side does not depend on ε. Defining

$$Z_t := \log |D_t| = \lim_{\varepsilon\to 0} Z_t^{(\varepsilon)},$$

we therefore get

$$\mathbb{P}\Big(\inf_{0\le s\le t} Z_s = -\infty\Big) = \mathbb{P}\Big(\inf_{0\le s\le t} |D_s| = 0\Big) = 0 \text{ for all } t \ge 0.$$

Therefore, all previous calculations extend to the case $\varepsilon = 0$. Letting $W_t := W_t^{(0)}$, we get

$$Z_t \le \log|x-y| + \sqrt{\kappa}\, W_t^* + ct$$

and

$$Z_t \ge \log|x-y| + \sqrt{\kappa}\, {}^*W_t - ct$$

and the claim in the lemma follows after applying exp to both sides of the equations. □

Remark 4.16. The constant c in the previous proof satisfies

$$c \le L_b + \frac{1}{2} d\kappa.$$

To see this, fix $x \ne y$. Since the trace of $\mathcal{A}(x,y)$ is the sum of its eigenvalues and $\|\mathcal{A}(x,y)\|$ is the largest eigenvalue, we get $\operatorname{tr}\big(\mathcal{A}(x,y)\big) \le d\|\mathcal{A}(x,y)\| \le d\kappa|x-y|^2$. We mention without proof that c in the statement of Lemma 4.15 even holds for $c = L_b + \frac{1}{2}(d-1)\kappa$.

Lemma 4.17. *For every $p \in \mathbb{R}$ and $T \ge 0$ there exists some $\tilde{c} = \tilde{c}(p,T)$, such that for every $x \ne y$ we have*

$$\mathbb{E} \sup_{0\le t\le T} |\psi_{0,t}(x) - \psi_{0,t}(y)|^p \le \tilde{c}|x-y|^p.$$

Proof. For $p \ge 0$ the first conclusion in the previous lemma implies

$$\mathbb{E} \sup_{0\le t\le T} |\psi_{0,t}(x) - \psi_{0,t}(y)|^p \le |x-y|^p \mathbb{E} \exp\{pcT + p\sqrt{\kappa}\, W_T^*\}$$

and for $p \le 0$, the second conclusion implies

$$\mathbb{E} \sup_{0\le t\le T} |\psi_{0,t}(x) - \psi_{0,t}(y)|^p \le |x-y|^p \mathbb{E} \exp\{-pcT + p\sqrt{\kappa}\, {}^*W_T\}.$$

Since the expected values on the right hand side are finite the assertion follows. □

Note that Lemma 4.17 allows us to apply Kolmogorov's continuity theorem, thus completing Step 1. We denote the continuous modification of $\psi_{0,\cdot}$ by $\tilde{\psi}$ (as in the statement of Step 1).

In order to finish the proof of Step 2, we need the following lemma

Lemma 4.18. *Define $\eta_t(x,y) := |\psi_{0,t}(x) - \psi_{0,t}(y)|^{-1}$, $t \geq 0$, $x \neq y$. For every $p \geq 0$ and $T \geq 0$ there exists some $\check{c} = \check{c}(p,T)$ such that*

$$\mathbb{E}\Big(\sup_{0\leq u\leq T} |\eta_u(x,y) - \eta_u(x',y')|^p\Big) \leq \check{c}|x-y|^{-p}|x'-y'|^{-p}|(x,y)-(x',y')|^p$$

for all $x, y, x', y' \in \mathbb{R}^d$ such that $x \neq y$, $x' \neq y'$.

Proof. Note that for every $x \neq y$ we have $\eta_t(x,y) \neq 0$ for all $t \geq 0$ almost surely by the second conclusion in Lemma 4.15. We have

$$\begin{aligned}
|\eta_u(x,y) - \eta_u(x',y')| &= \Big|\frac{1}{|\psi_{0,u}(x) - \psi_{0,u}(y)|} - \frac{1}{|\psi_{0,u}(x') - \psi_{0,u}(y')|}\Big| \\
&= \eta_u(x,y)\eta_u(x',y')\Big||\psi_{0,u}(x') - \psi_{0,u}(y')| - |\psi_{0,u}(x) - \psi_{0,u}(y)|\Big| \\
&\leq \eta_u(x,y)\eta_u(x',y')\big|\psi_{0,u}(x') - \psi_{0,u}(y') - \psi_{0,u}(x) + \psi_{0,u}(y)\big| \\
&\leq \eta_u(x,y)\eta_u(x',y')\big(\big|\psi_{0,u}(x) - \psi_{0,u}(x')\big| + \big|\psi_{0,u}(y) - \psi_{0,u}(y')\big|\big).
\end{aligned}$$

Applying Hölder's inequality and Lemma 4.17, we get

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq u\leq T} |\eta_u(x,y) - \eta_u(x',y')|^p\Big) \\
&\qquad \leq \Big(\mathbb{E}\sup_{0\leq u\leq T} |\eta_u(x,y)|^{3p}\Big)^{1/3}\Big(\mathbb{E}\sup_{0\leq u\leq T} |\eta_u(x',y')|^{3p}\Big)^{1/3} \\
&\qquad\qquad \Big(\mathbb{E}\sup_{0\leq u\leq T} (|\psi_{0,u}(x) - \psi_{0,u}(x')| + |\psi_{0,u}(y) - \psi_{0,u}(y')|)^{3p}\Big)^{1/3} \\
&\qquad \leq \tilde{c}(-3p,T)^{2/3}|x-y|^{-p}|x'-y'|^{-p}2^p\tilde{c}(3p,T)^{1/3}\big(|x-x'|^p + |y-y'|^p\big).
\end{aligned}$$

(note that the last inequality holds for every $p > 0$). The proof of the lemma is complete. □

Now we can complete Step 2 (the one-to-one property). Define $I := \{(x,y) \in \mathbb{R}^{2d} : x \neq y\}$ (which is an open subset of $\mathbb{R}^{2d}$). Note that Kolmogorov's continuity theorem implies that there exists a process $\tilde{\eta} : I \to C([0,\infty),\mathbb{R})$ which is continuous for all $\omega \in \Omega$ and which is a modification of η in the sense that $\mathbb{P}\big(\tilde{\eta}(z) \equiv \eta(z)\big) = 1$ for each $z \in I$. In particular, $\tilde{\eta}$ does not attain the value ∞ on I. Unfortunately this does not yet show the one-to-one property since we do not yet know if the modification $\tilde{\eta}$ corresponds to a modification of $\psi_{0,\cdot}$. To close this gap we argue as follows: To define η, we use the continuous modification $\tilde{\psi}$ instead of ψ. Then η is a continuous process which takes values in $[0,\infty]$. Then we define $\tilde{\eta}$ as above. Since both η and $\tilde{\eta}$ are continuous on I (η taking values in $[0,\infty]$ and $\tilde{\eta}$ taking values in $[0,\infty)$) they actually agree identically almost surely. This implies that there exists a set N of measure 0 such that $\tilde{\psi}_t(\omega,x) \neq \tilde{\psi}_t(\omega,y)$ for all $t \geq 0$, all $x \neq y$, and all $\omega \notin N$. Redefining $\tilde{\psi}$ as the identity map on N, we ensure that $\tilde{\psi}$ satisfies the properties stated in Step 2.

To complete the proof of Step 3 we need two more lemmas.

Lemma 4.19. *For each $T \geq 0$ and $p \in \mathbb{R}$ there exists some $\hat{c} = \hat{c}(p,T)$ such that*

$$\mathbb{E} \sup_{0 \leq u \leq T} (1 + |\psi_{0,u}(x)|)^p \leq \hat{c}(1 + |x|)^p$$

for every $x \in \mathbb{R}^d$.

Proof. The proof is very similar to that of Lemmas 4.15 and 4.17. Fix $x \in \mathbb{R}^d$ and define

$$D_t := \psi_{0,t}(x), \; Z_t := \frac{1}{2} \log(|D_t|^2 + 1).$$

As in the proof of Lemma 4.15 Itô's formula implies

$$\begin{aligned} \mathrm{d}Z_t = & \frac{D_t \cdot (M(\mathrm{d}t, \psi_{0,t}(x))}{|D_t|^2 + 1} + \frac{D_t^T b(\psi_{0,t}(x))}{|D_t|^2 + 1} \, \mathrm{d}t + \frac{1}{2} \frac{1}{|D_t|^2 + 1} \mathrm{tr}\big(a(\psi_{0,t}(x), \psi_{0,t}(x))\big) \, \mathrm{d}t \\ & - \sum_{i,j} \frac{D_t^i D_t^j}{\big(|D_t|^2 + 1\big)^2} a_{ij}(\psi_{0,t}(x), \psi_{0,t}(x)) \, \mathrm{d}t. \end{aligned}$$

Defining the continuous local martingale

$$N_t := \int_0^t \frac{D_s}{|D_s|^2 + 1} \cdot M(\mathrm{d}s, \psi_{0,s}(x)),$$

we get

$$Z_t = Z_0 + N_t + \int_0^t \alpha(s, \omega) \, \mathrm{d}s,$$

where α is bounded by a deterministic constant c_1 (uniformly in x) as in the proof of Lemma 4.15. The reader is encouraged to show this and to show that the derivative of the quadratic variation of N is uniformly bounded by a deterministic constant c_2. Hint: show (using the Kunita-Watanabe inequality and the definition and properties of $\mathcal{A}$) that there exists a constant c_3 such that $|a_{ii}(y,y)| \leq c_3(1 + |y|^2)$ for every $i = 1, ..., d$ and $y \in \mathbb{R}^d$. As in the proof of Lemma 4.15 we see that there exists a standard Brownian motion (on a possibly enlarged probability space) and constants $c_1, c_2 \geq 0$ such that, for every $t \geq 0$,

$$\begin{aligned} \sup_{0 \leq u \leq t} \big(|\psi_{0,u}(x)|^2 + 1\big)^{1/2} &\leq \big(|x|^2 + 1\big)^{1/2} \exp\{c_1 t + c_2 \, W_t^*\}, \\ \inf_{0 \leq u \leq t} \big(|\psi_{0,u}(x)|^2 + 1\big)^{1/2} &\geq \big(|x|^2 + 1\big)^{1/2} \exp\{-c_1 t + c_2 \, {}^*W_t\}. \end{aligned}$$

Now the claim follows as in the proof of Lemma 4.17. □

Lemma 4.20. *For $t \geq 0$ and $x \in \mathbb{R}^d$ define*

$$\zeta_t(x) := \begin{cases} \Big(1 + \big|\psi_{0,t}\big(\frac{x}{|x|^2}\big)\big|\Big)^{-1}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

Then, for every $T > 0$ and $p > 0$ there exists some $\mathring{c} = \mathring{c}(p,T)$ such that

$$\mathbb{E}\Big(\sup_{0 \leq u \leq T} |\zeta_u(x) - \zeta_u(y)|^p\Big) \leq \mathring{c}|x - y|^p$$

for all $x, y \in \mathbb{R}^d$.

Proof. Let us first assume that $x, y \neq 0$. Then

$$|\zeta_u(x) - \zeta_u(y)| \le \zeta_u(x)\zeta_u(y)\Big|\psi_{0,u}\Big(\frac{x}{|x|^2}\Big) - \psi_{0,u}\Big(\frac{y}{|y|^2}\Big)\Big|.$$

Applying Hölder's inequality, we get - as in the proof of Lemma 4.18 and using Lemmas 4.19 and 4.17 -

$$\mathbb{E}\Big(\sup_{0\le u\le T}\big|\zeta_u(x) - \zeta_u(y)\big|^p\Big) \le \hat{c}(-3p,T)^{2/3}|x|^p|y|^p\tilde{c}(3p,T)^{1/3}\Big|\frac{x}{|x|^2} - \frac{y}{|y|^2}\Big|^p$$

Assuming $|y| \ge |x|$ (without loss of generality), we obtain

$$\begin{aligned}\Big|\frac{x}{|x|^2} - \frac{y}{|y|^2}\Big| &= \Big|\frac{x|y|^2 - x|x|^2 + x|x|^2 - |x|^2 y}{|x|^2|y|^2}\Big| \\ &\le \frac{|x-y|}{|y|^2} + |x|\frac{\big||y|^2 - |x|^2\big|}{|x|^2|y|^2} \\ &\le \frac{|x-y|}{|x||y|} + |x|\frac{2|y||x-y|}{|x|^2|y|^2} = 3\frac{|x-y|}{|x||y|}\end{aligned}$$

and therefore

$$\mathbb{E}\Big(\sup_{0\le u\le T}\big|\zeta_u(x) - \zeta_u(y)\big|^p\Big) \le 3^p\hat{c}(-3p,T)^{2/3}\tilde{c}(3p,T)^{1/3}|x-y|^p$$

as claimed. The case in which x or y are 0 is similar (in fact easier). □

The previous lemma together with Kolmogorov's continuity theorem imply that the process ζ has a continuous modification $\tilde{\zeta}$, i.e. $x \mapsto \tilde{\zeta}_.(x)$ is continuous as a map from $\mathbb{R}^d$ to $C([0,\infty),\mathbb{R}^d)$. In particular, the map is continuous at $x = 0$ and therefore, for each $T > 0$,

$$\lim_{|x|\to\infty}\inf_{0\le u\le T}\big|\tilde{\psi}_u(x,\omega)\big| = \infty, \text{ a.s.}$$

Redefine $\tilde{\psi}$ as the identity on the set of measure zero on which the previous property does not hold for every $T > 0$. Note that $\tilde{\psi}_0(x) = x$ almost surely and the exceptional set of measure zero can be chosen the same for all $x \in \mathbb{R}^d$ by continuity of $\tilde{\psi}_0$. Again, by the same kind of redefinition, we can and will assume that $\tilde{\psi}_0(x) = x$ for *every* $x \in \mathbb{R}^d$ and $\omega \in \Omega$. Let $\overline{\mathbb{R}^d}(= \mathbb{R}^d \cup \{\infty\})$ be the one-point compactification of $\mathbb{R}^d$ and extend (temporarily) the domain of definition of $\tilde{\psi}$ to $[0,\infty) \times \overline{\mathbb{R}^d} \times \Omega$ by

$$\tilde{\psi}_t(\infty,\omega) := \infty.$$

Then we have, for every $\omega \in \Omega$,

a) $(t,x) \mapsto \tilde{\psi}_t(x)$ is continuous,

b) $\tilde{\psi}_t(.,\omega) : \overline{\mathbb{R}^d} \to \overline{\mathbb{R}^d}$ is one-to-one for each $t \ge 0$,

c) $\tilde{\psi}_t(\infty,\omega) = \infty$ for each $t \ge 0$,

d) $\tilde{\psi}_0(x,\omega) = x$ for all $x \in \overline{\mathbb{R}^d}$.

Since $\overline{\mathbb{R}^d}$ is known to be homeomorphic to the unit sphere $\mathbb{S}^d$ in $\mathbb{R}^{d+1}$ and a well-known theorem in homotopy theory says that a continuous map $f : \mathbb{S}^d \to \mathbb{S}^d$ which is one-to-one is automatically onto, it follows from c) that the restriction of $\tilde{\psi}$ to $\mathbb{R}^d$ is also onto. Therefore we have finished Step 3.

Next we show Step 4, i.e. we show that the map $(s, x) \mapsto \tilde{\psi}_s^{-1}(x, \omega)$ is continuous on $[0, \infty) \times \overline{\mathbb{R}^d}$. Fix $\omega \in \Omega$, $(s, x) \in [0, \infty) \times \overline{\mathbb{R}^d}$ and a sequence (s_n, x_n) which converges to (s, x). Let $y_n := \tilde{\psi}_{s_n}^{-1}(x_n, \omega)$. By compactness of $\overline{\mathbb{R}^d}$ any subsequence of (s_n, x_n) has a further subsequence (s_{n_k}, x_{n_k}) such that y_{n_k} converges to some $y \in \overline{\mathbb{R}^d}$. Hence

$$x \leftarrow x_{n_k} = \tilde{\psi}_{s_{n_k}}(y_{n_k}, \omega) \to \tilde{\psi}_s(y, \omega),$$

so $\tilde{\psi}_s(y, \omega) = x$ and therefore $y = \tilde{\psi}_s^{-1}(x, \omega)$ showing that $y_n \to y$, so $(s, x) \mapsto \tilde{\psi}_s^{-1}(x, \omega)$ is continuous.

It remains to prove Step 5.

Define $\varphi_{s,t}(x, \omega) := \tilde{\psi}_t(\tilde{\psi}_s^{-1}(x, \omega), \omega)$, $s, t \geq 0$. We will show that φ satisfies the claim in Theorem 4.14 thus completing the proof of the theorem. The fact that φ is a stochastic flow is easy to see from the definition and the properties of $\tilde{\psi}$ above, so it remains to show that for each $s \geq 0$ and $x \in \mathbb{R}^d$, the process $\varphi_{s,t}(x)$, $t \geq s$ solves the sde

$$X_t = x + \int_s^t b(X_u)\,\mathrm{d}u + \int_s^t M(\mathrm{d}u, X_u),\ t \geq s.$$

This is clear for $s = 0$ since $\varphi_{0,t}(x) = \tilde{\psi}_t(x)$ and $\tilde{\psi}_t(x)$ agrees with the solution $\psi_{0,t}(x)$ up to a set of measure 0. In particular, we have, for any $0 \leq s \leq t$, and $z \in \mathbb{R}^d$,

$$\begin{aligned}
\varphi_{0,t}(z) &= z + \int_0^t b(\varphi_{0,u}(z))\,\mathrm{d}u + \int_0^t M(\mathrm{d}u, \varphi_{0,u}(z)),\\
\varphi_{0,s}(z) &= z + \int_0^s b(\varphi_{0,u}(z))\,\mathrm{d}u + \int_0^s M(\mathrm{d}u, \varphi_{0,u}(z)).
\end{aligned}$$

Subtracting these equations, we get

$$\varphi_{0,t}(z) = \varphi_{0,s}(z) + \int_s^t b(\varphi_{0,u}(z))\,\mathrm{d}u + \int_s^t M(\mathrm{d}u, \varphi_{0,u}(z)).$$

Now, we would like to insert $z := \varphi_{0,s}^{-1}(x, \omega)$. Then we get

$$\varphi_{s,t}(x) = x + \int_s^t b(\varphi_{s,u}(x))\,\mathrm{d}u + \int_s^t M(\mathrm{d}u, \varphi_{s,u}(x)),$$

provided the following formula holds almost surely:

$$\int_s^t M(\mathrm{d}u, \varphi_{0,u}(z))\Big|_{z=\varphi_{0,s}^{-1}(x)} = \int_s^t M(\mathrm{d}u, \varphi_{s,u}(x)).$$

This formula is not obvious but true and follows from Theorem 3.3.3 in [Ku90]. Note that the corresponding formula for the integral of b does not generate any problem. Therefore, the proof of Theorem 4.14 is complete.

4.3 Expansion of stochastic flows

In this section we will discuss how one can obtain an (asymptotic) upper bound for the diameter of the set $\varphi_{0,T}(S,\omega)$ under a stochastic flow φ for a bounded set $S \subset \mathbb{R}^d$. We will assume that φ is the stochastic flow generated by a stochastic differential equation as in the previous section. In addition, we assume that b and a are bounded. We will often write φ_T instead of $\varphi_{0,T}$. Here is a method how to obtain an upper bound for $\varphi_T(S,\omega)$: cover the set S with a minimal number $N_{T,\varepsilon}$ of cubes $C_1, ..., C_{N_{T,\varepsilon}}$ of side length $\varepsilon > 0$ with respective centers $x_1, ..., x_{N_{T,\varepsilon}}$. We choose $\varepsilon = \varepsilon(T) > 0$ so small that it is unlikely that the diameter of the image of any of the cubes under φ at time T is larger than 1. If none of the diameters of those cubes exceeds one, then the diameter of $\varphi_T(S,\omega)$ is at most $2(\max_i |\varphi_T(x_i)| + 1)$. The probability that this quantity is larger than u can then be estimated from above using a union bound and an estimate of $\mathbb{P}(|\varphi_T(x_i)| \geq u)$ for every i (which can be obtained easily using boundedness of a and b).

We first establish an asymptotic upper bound for the linear growth rate of the solution of a stochastic differential equation with bounded a and b. As before we assume that $\mathcal{A}$ and b are Lipschitz. In fact boundedness of b can be weakened a bit in the following proposition and in Theorem 4.24.

Proposition 4.21. *Assume that $\|a(x,x)\| \leq A^2$ for some $A > 0$ and all $x \in \mathbb{R}^d$ and that*

$$\limsup_{|x|\to\infty} \langle b(x), \frac{x}{|x|} \rangle \leq B$$

for some $B \geq 0$. Then, for each compact set $S \subset \mathbb{R}^d$ and $k > 0$, we have

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{x\in S} \mathbb{P}\Big(\sup_{0\leq t\leq T} |\varphi_t(x)| \geq kT \Big) \leq -\frac{1}{2A^2}(k-B)_+^2.$$

Proof. Let $S \subset \mathbb{R}^d$ be compact and $k > B$ (otherwise there is nothing to prove). Fix $\varepsilon \in (0, k-B)$ and let $r_0 > 1$ be so large that

$$\langle b(x), \frac{x}{|x|} \rangle + \frac{d-1}{2|x|} A^2 \leq B + \varepsilon \text{ for all } |x| \geq r_0$$

and that S is contained in a ball of radius r_0 around the origin. Let $h : \mathbb{R} \to \mathbb{R}$ be an even C^∞ function such that $h(y) = |y|$ for $|y| \geq 1$, $h'(y) = 0$ for $|y| \leq 1/2$, and $|h'(y)| \leq 1$ for all $y \in \mathbb{R}$. Applying Itô's formula to $\rho_t(x) := h(|\varphi_t(x)|)$, we get

$$\mathrm{d}\rho_t(x) = \mathrm{d}N_t + f(\varphi_t(x))\,\mathrm{d}t,$$

where (interpreting an integrand of the form 0/0 as 0)

$$N_t := \sum_{i=1}^d \int_0^t h'(\rho_s(x)) \frac{\varphi_s^i(x)}{\varphi_s(x)} M^i(\mathrm{d}s, \varphi_s(x)) \text{ and } f \in C(\mathbb{R}^d, \mathbb{R}) \text{ satisfies}$$

$$\begin{aligned} f(x) &= \langle b(x), \frac{x}{|x|} \rangle + \frac{1}{2|x|} \operatorname{tr} a(x,x) - \frac{1}{2|x|^3} x^T a(x,x) x \\ &\leq \langle b(x), \frac{x}{|x|} \rangle + \frac{d-1}{2|x|} A^2 \leq B + \varepsilon \text{ on the set} \{|x| \geq r_0\}. \end{aligned}$$

Then, for $t \geq s \geq 0$,

$$\langle N \rangle_t - \langle N \rangle_s \leq \int_s^t \frac{1}{|\varphi_u(x)|^2} \varphi_u^T(x) a(\varphi_u(x), \varphi_u(x)) \varphi_u(x)\, \mathrm{d}u \leq A^2(t-s).$$

Therefore, the continuous local martingale N can be represented (possibly on an enlarged space) in the form $N_t = AW_{\tau(t)}$ where W is a standard Brownian motion and $\tau(t)$ is a family of stopping times which satisfies $\tau(t) - \tau(s) \leq t - s$ whenever $0 \leq s \leq t$. For $|x| \leq r_0 < kT$ we get

$$\begin{aligned}
\mathbb{P}\Big(\sup_{0 \leq t \leq T} \big|\varphi_t(x)\big| \geq kT \Big) &\leq \mathbb{P}\big(\exists\, 0 \leq s \leq t \leq T :\ \rho_t(x) - \rho_s(x) \geq kT - r_0,\ \inf_{s \leq u \leq t} \rho_u(x) \geq r_0 \big) \\
&\leq \mathbb{P}\big(\exists\, 0 \leq s \leq t \leq T :\ A(W_{\tau(t)} - W_{\tau(s)}) + (B + \varepsilon)(t-s) \geq kT - r_0 \big) \\
&\leq \mathbb{P}\Big(\max_{0 \leq s \leq 1} W_s - \min_{0 \leq s \leq 1} W_s \geq \frac{k - B - \varepsilon}{A} \sqrt{T} - \frac{r_0}{A\sqrt{T}} \Big).
\end{aligned}$$

The density of the *range* $R := \max_{0 \leq s \leq 1} W_s - \min_{0 \leq s \leq 1} W_s$ is known (in terms of an infinite sum), see [F51] and from that formula one easily sees that $\mathbb{P}(R \geq u) \lesssim \exp\{-u^2/2\}$ (alternatively one can use the cruder bound $\mathbb{P}(R \geq u) \leq 2\mathbb{P}\big(\max_{0 \leq s \leq 1} W_s \geq u/2\big) \lesssim \exp\{-u^2/8\}$). Since $\varepsilon > 0$ is arbitrary this easily implies the statement of the proposition. □

Remark 4.22. Proposition 2.8 in [S09] is a generalization of the previous proposition in which B is allowed to be negative (in this case the formula has to be modified when $k < -B$).

Next, we establish an upper bound for the tails of the distribution of the diameter of a small cube under the flow generated by the stochastic differential equation. In the following proposition, we just assume the conditions of Theorem 4.14 and impose no additional boundedness conditions.

Proposition 4.23. *Let the assumptions of Theorem 4.14 be satisfied. Assume that $\kappa > 0$. For $\gamma > 0$ and $\Lambda := L_b + \frac{1}{2}(d-1)\kappa$ define*

$$I(\gamma) := \begin{cases} \frac{(\gamma - \Lambda)^2}{2\kappa} & \text{if } \gamma \geq \Lambda + \kappa d, \\ d(\gamma - \Lambda - \frac{1}{2}\kappa d) & \text{if } \Lambda + \frac{1}{2}\kappa d \leq \gamma \leq \Lambda + \kappa d, \\ 0 & \text{if } \gamma \leq \Lambda + \frac{1}{2}\kappa d. \end{cases}$$

Then, for each $u > 0$, we have

$$\limsup_{T \to \infty} \frac{1}{T} \sup_{\chi_T} \log \mathbb{P}\Big(\sup_{x, y \in \chi_T} \sup_{0 \leq t \leq T} |\varphi_t(x) - \varphi_t(y)| \geq u \Big) \leq -I(\gamma),$$

where $\sup_{\chi_T}$ *means that we take the supremum over all cubes χ_T in $\mathbb{R}^d$ with side length* $\exp\{-\gamma T\}$.

Proof. Fix $\gamma > 0$ and $T > 0$. Without loss of generality we assume that $\chi = \chi_T := [0, \mathrm{e}^{-\gamma T}]^d$. Define $Z_x(t) := \varphi_t(\mathrm{e}^{-\gamma T} x)$, $x \in \mathbb{R}^d$. Lemma 4.17 (or Lemma 4.15) and Remark 4.16 imply, for $q \geq 1$,

$$\Big(\mathbb{E}\big(\sup_{0 \leq t \leq T} |Z_x(t) - Z_y(t)|^q \big) \Big)^{1/q} \leq 2\mathrm{e}^{-\gamma T} |x - y| \exp\{(\Lambda + \frac{1}{2} q\kappa) T\},$$

i.e. the assumptions of Kolmogorov's continuity theorem are satisfied when $q > d$. Hence, for $\xi \in \big(0, \frac{q-d}{q}\big)$, we get

$$\mathbb{P}\Big(\sup_{x,y\in\chi}\sup_{0\le t\le T}|\varphi_t(x)-\varphi_t(y)|\ge u\Big)\le\Big(\frac{2d}{1-2^{-\xi}}\Big)^q\frac{2^q d 2^{a\xi-b}}{1-2^q d 2^{a\xi-b}}\exp\{(\Lambda-\gamma+\frac{1}{2}q\kappa)qT\}u^{-q}.$$

Taking logs, dividing by T, letting $T\to\infty$, and optimizing over $q>d$ yields the claim (I will provide more details in class; don't miss it!). □

Observe that the asymptotic bound in Proposition 4.21 depends on the bounds B and A only, while the bound in Proposition 4.23 only depends on the Lipschitz constants L_b and κ of b and $\mathcal{A}$. Now we combine both propositions to obtain a bound on the linear growth rate of a bounded set under a stochastic flow as above.

Theorem 4.24. *Under the conditions of Theorem 4.14 and Proposition 4.21 we have, for any bounded subset $S\subset\mathbb{R}^d$, $d\ge 1$,*

$$\limsup_{T\to\infty}\Big(\frac{1}{T}\sup_{x\in S}\sup_{0\le t\le T}|\varphi_t(x)|\Big)\le K,$$

where

$$K:=B+A\sqrt{2(d-1)\big(\Lambda+\kappa(d-1)+\sqrt{\kappa^2(d-1)^2+2(d-1)\Lambda\kappa}\big)}.$$

If the drift b is identically 0, then $L_b=B=0$ and we have

$$K=A(d-1)\sqrt{\kappa}\sqrt{3+2\sqrt{2}}.$$

Proof. Clearly, it suffices to prove the statement for sets of the form $S=[-M,M]^d$, $M\ge 1$. Let $N(M,r)$ be the minimal number of cubes of side length $r>0$ which cover ∂S. Clearly, $N(M,\mathrm{e}^{-\gamma T})\le 2d\big(3M\mathrm{e}^{\gamma T}\big)^{d-1}$ for $\gamma>0$ and for $T>0$ sufficiently large. Fix such γ and T and let $C_1,C_2,...,C_{N(M,\mathrm{e}^{-\gamma T})}$ be such a cover with centers in S. For $k>0$ we have (using the fact that φ_t is a homeomorphism for each $t\ge 0$)

$$\mathbb{P}\Big(\sup_{x\in S}\sup_{0\le t\le T}|\varphi_t(x)|\ge kT\Big)\le\Gamma_1+\Gamma_2,$$

where

$$\Gamma_1:=2d\big(3M\mathrm{e}^{\gamma T}\big)^{d-1}\sup_{x\in S}\mathbb{P}\Big(\sup_{0\le t\le T}|\varphi_t(x)|\ge kT-1\Big)$$

and

$$\Gamma_2:=2d\big(3M\mathrm{e}^{\gamma T}\big)^{d-1}\max_i\mathbb{P}\Big(\sup_{0\le t\le T}\operatorname{diam}\varphi_t(C_i)\ge 1\Big).$$

Therefore, using Propositions 4.21 and 4.23 and the formula $\log(\alpha+\beta)\le\log 2+\log(\alpha\vee\beta)$, $\alpha,\beta>0$, we get

$$\begin{aligned}\zeta(\gamma,k):=\limsup_{T\to\infty}\frac{1}{T}\log\mathbb{P}\Big(\sup_{x\in S}\sup_{0\le t\le T}|\varphi_t(x)|\ge kT\Big)&\le\gamma(d-1)+\Big[\Big(-\frac{1}{2A^2}(k-B)_+^2\Big)\vee\big(-I(\gamma)\big)\Big]\\&=\gamma(d-1)-\Big[\Big(\frac{1}{2A^2}(k-B)_+^2\Big)\wedge\big(I(\gamma)\big)\Big].\end{aligned}$$

If $d\ge 2$, then we let γ_0 be the unique strictly positive solution of $I(\gamma)=\gamma(d-1)$. Then

$$\gamma_0=\Lambda+\kappa(d-1)+\sqrt{2\Lambda\kappa(d-1)+\kappa^2(d-1)^2}.$$

If $d = 1$, then we define $\gamma_0 = \Lambda + \kappa$. In any case $\zeta(\gamma, k) < 0$ whenever $\gamma > \gamma_0$ and $k > k_0(\gamma) := B + A\sqrt{2\gamma(d-1)}$. Therefore, for every $\varepsilon > 0$, we get

$$\sum_{n=1}^{\infty} \mathbb{P}\Big(\frac{1}{n} \sup_{x \in S} \sup_{0 \le t \le n} |\varphi_t(x)| \ge B + A\sqrt{2\gamma_0(d-1)} + \varepsilon\Big) < \infty$$

and the first Borel-Cantelli lemma (together with an easy interpolation argument) implies the assertion of the theorem.

The final statement in the theorem is clear. □

Remark 4.25. If the initial set S is of lower dimension than $d-1$ then the bound in the previous theorem can be sharpened (see Theorem 2.3. in [S09]).

Chapter 5

Random Dynamical Systems

5.1 Basic definitions

The basic reference for this chapter is the monograph by Ludwig Arnold [A98]. We start by defining the concept of a metric dynamical system (MDS) with respect to the group $(G,+)$ where $G \in \{\mathbb{Z}, \mathbb{R}\}$ and let $\mathcal{G}$ be the corresponding Borel σ-algebra on G. Further let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 5.1. Let $\theta : G \times \Omega \to \Omega$ satisfy

- θ is $(\mathcal{G} \otimes \mathcal{F}, \mathcal{F})$-measurable,
- $\theta_0(\omega) = \omega$ for all $\omega \in \Omega$,
- $\theta_{g+h}(\omega) = \theta_g(\theta_h(\omega))$ for all $g, h \in G$, $\omega \in \Omega$.
- $\mathbb{P}\theta_g^{-1} = \mathbb{P}$ for each $g \in G$.

Then, $(\Omega, \mathcal{F}, \mathbb{P}, \theta) := (\Omega, \mathcal{F}, \mathbb{P}, (\theta_g)_{g \in G})$ is called a *metric dynamical system* (MDS) with respect to G.

Next we define the concept ot a random dynamical system (RDS).

Definition 5.2. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_g)_{g \in G})$ be an MDS and $(E, \mathcal{E})$ a measurable (state) space. Let $\mathbb{T}$ be a subset of G which contains 0 and is closed with respect to addition. Let $\mathcal{T}$ be the trace σ-algebra of $\mathcal{G}$ on $\mathbb{T}$. Then $\varphi : \mathbb{T} \times E \times \Omega \to E$ is called a *cocycle* if

- φ is $(\mathcal{T} \otimes \mathcal{E} \otimes \mathcal{F}, \mathcal{E})$-measurable,
- $\varphi_0(., \omega) = \mathrm{id}_E$ for all $\omega \in \Omega$,
- $\varphi_{t+s}(., \omega) = \varphi_t(\varphi_s(., \omega), \theta_s(\omega))$ for all $s, t \in \mathbb{T}$, $\omega \in \Omega$.

In this case $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is called (E-valued) *random dynamical system*.

From now on, we will always deal with one of the following two cases:

1. Discrete time: $G = \mathbb{Z}$ and $\mathbb{T} = G$ or $\mathbb{T} = \mathbb{N}_0$.
2. Continuous time: $G = \mathbb{R}$ and $\mathbb{T} = G$ or $\mathbb{T} = [0, \infty)$.

Let us start with a simple class of examples.

Example 5.3. Let $X_0, X_1, \cdots$ be a stationary process defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ taking values in a measurable space $(K, \mathcal{K})$. Let $(E, \mathcal{E})$ be a measurable space and let $f : E \times K \to E$ be a measurable map. For each $x \in E$ we define recursively the sequence $Y_n^{(x)}(\tilde{\omega}) := f(Y_{n-1}^{(x)}(\tilde{\omega}), X_n(\tilde{\omega}))$ for $n \geq 1$ with initial condition $Y_0^{(x)}(\tilde{\omega}) = x \in E$. The sequences $Y_n^{(x)}$ are not yet a random dynamical system since there is no MDS so far. By changing the probability space we can however find an RDS which has the same joint distribution as $(Y_n^{(x)}, n \in \mathbb{N}_0, x \in E)$ as follows.

First we extend the sequence $X_0, X_1, \cdots$ to a two-sided stationary sequence: we consider a stationary sequence $\tilde{X}_i$, $i \in \mathbb{Z}$ such that the laws of $(X_0, X_1, \cdots)$ and $(\tilde{X}_0, \tilde{X}_1, \cdots)$ coincide. Note that such a sequence always exists in case the $(K, \mathcal{K})$ is a Polish space with its Borel-σ algebra (possibly on a different probability space) and that the law $\mathbb{P}$ of $\tilde{X}_i$, $i \in \mathbb{Z}$ is uniquely defined by the law of $(X_0, X_1, \cdots)$ (even if K is not Polish). It is not clear to me if there exists an example where $(K, \mathcal{K})$ is not Polish and such an extension does not exist. Note that $\mathbb{P}$ is a probability measure on the space $(\Omega, \mathcal{F})$, where $\Omega := K^{\mathbb{Z}}$ and $\mathcal{F} = \mathcal{K}^{\otimes \mathbb{Z}}$. Let $G := \mathbb{Z}$ and $(\theta_n(\omega))_k := \omega_{n+k}$, $n, k \in \mathbb{Z}$. Then $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ is an MDS. Next, we define $\hat{X}_n(\omega) := \omega_n$, $n \in \mathbb{Z}$ (or $n \in \mathbb{N}_0$) and, recursively, $\hat{Y}_n^{(x)}(\omega) := f(\hat{Y}_{n-1}^{(x)}(\omega), \hat{X}_n(\omega))$ for $n \geq 1$ with initial condition $\hat{Y}_0^{(x)}(\omega)) = x \in E$. Then we define $\varphi_n(x, \omega) := \hat{Y}_n^{(x)}(\omega)$, for $x \in E$ and $\omega \in \Omega$. It is straightforward to check that $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is a random dynamical system. Note that it is in general impossible to extend φ to an RDS with $\mathbb{T} = \mathbb{Z}$ (why?).

We will see later that a similar (but slightly more technical) procedure allows us to interpret solutions of a stochastic differential equation which is driven by a stationary increment process (like Brownian motion or a Lévy process) as a random dynamical system.

Definition 5.4. Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ be an E-valued RDS with respect to G and $\mathbb{T}$ as above. Define

$$\Theta_t(x, \omega) := \big(\varphi_t(x), \theta_t(\omega)\big),\ \omega \in \Omega,\ t \in \mathbb{T},\ x \in E.$$

Then, the family of maps $\Theta_t : \Omega \times E \to \Omega \times E$, $t \in \mathbb{T}$ is called the *skew-product flow* associated to the RDS.

Remark 5.5. It is easy to see that the skew-product flow satisfies the following properties

- Θ is (jointly) measurable,
- $\Theta_0(x, \omega) = x,\ \omega \in \Omega,\ x \in E$ and
- $\Theta_{t+s}(x, \omega) = \Theta_t(\Theta_s(x, \omega), \omega),\ s, t \in \mathbb{T},\ x \in E,\ \omega \in \Omega.$

Next, we investigate the relation between a stochastic semi-flow and an RDS. Define

$$\Delta := \{(s, t) \in G \times G :\ s \leq t\},$$

(this is similar but not identical to the notation in the previous chapter). Recall the definition of a stochastic semi-flow (here we do not insist on continuity properties and allow a more general state space and we do not allow exceptional sets of measure 0).

Definition 5.6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ a measurable space. Let G be as above. A map $\phi : \Delta \times E \times \Omega \to E$ is called *stochastic semi-flow* if the following properties hold:

- ϕ is jointly measurable,
- $\phi_{s,s}(.,\omega) = \mathrm{id}_E$, $s \in G$, $\omega \in \Omega$,
- $\phi_{s,u}(.,\omega) = \phi_{t,u}(.,\omega) \circ \phi_{s,t}(.,\omega)$, $s \le t \le u$, $s,t,u \in G$, $\omega \in \Omega$.

The following proposition shows that any RDS generates a stochastic semi-flow.

Proposition 5.7. *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ be an E-valued RDS with respect to G and $\mathbb{T} := G^+ := \{g \in G : g \ge 0\}$. Define $\phi : \Delta \times E \times \Omega \to E$ by*

$$\phi_{s,t}(x,\omega) := \varphi_{t-s}(x, \theta_s(\omega)),\ (s,t) \in \Delta,\ x \in E,\ \omega \in \Omega.$$

Then ϕ is a stochastic semi-flow which enjoys the additional property

$$\phi_{s+h,t+h}(x,\omega) = \phi_{s,t}(x, \theta_h(\omega)),\ x \in E,\ (s,t) \in \Delta,\ h \in G,\ \omega \in \Omega. \tag{5.1.1}$$

Proof. Straightforward. □

In general, a stochastic semi-flow ϕ does not correspond to an RDS. However the following holds.

Proposition 5.8. *Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ be an MDS with associated group G and assume that $\phi : \Delta \times E \times \Omega \to E$ is a stochastic semi-flow which enjoys property* (5.1.1). *Then $\varphi_t(x,\omega) := \phi_{0,t}(\omega)$, $t \ge 0$, $x \in E$, $\omega \in \Omega$ is a cocycle and hence $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is an RDS with $\mathbb{T} = G^+$.*

Proof. Straightforward. □

Remark 5.9. If ϕ is a semi-flow in the sense that $\phi_{s,t}$ is only defined for $s,t \ge 0$ (and all other properties in the definition hold), and if ϕ satisfies (5.1.1) for $h \ge 0$, then it is very easy to see that the claim of Proposition 5.8 still holds.

Definition 5.10. An RDS $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is called *continuous RDS* if the state space $(E, \mathcal{E})$ is Polish and the map $x \mapsto \varphi_t(x,\omega)$ is continuous for every $t \ge 0$, $\omega \in \Omega$.

5.2 SDEs and random dynamical systems

An important class of RDS with continuous time are those generated by stochastic differential equations. To simplify the presentation we do not consider Kunita-type equations here. We assume that the equation is driven by an m-dimensional Brownian motion. A natural (but not the only) way to set up a suitable MDS is the following.

Let Ω be the space of $\mathbb{R}^m$-valued continuous functions defined on $\mathbb{R}$ which are 0 at 0. We equip Ω with the corresponding Borel σ-field $\mathcal{F}$ (as in Chapter 3). Let $\mathbb{P}$ be two-sided m-dimensional Wiener measure on $(\Omega, \mathcal{F})$ (I will explain that in class) and define the shift θ by $(\theta_t(\omega))(s) := \omega_{t+s} - \omega_t$, $s,t \in \mathbb{R}$. Clearly, $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ is a MDS. Let ϕ be the solution flow of an SDE with Lipschitz coefficients (driven by m-dimensional Brownian motion) defined on the space $(\Omega, \mathcal{F}, \mathbb{P}, \theta)$ with Brownian motion $W_t(\omega) := \omega_t$, $t \ge 0$. If we knew that $\phi_{s+h,t+h}(x,\omega) = \phi_{s,t}(x, \theta_h(\omega))$ for all $h \ge 0$, $\omega \in \Omega$, and $0 \le s \le t$, then, by Proposition 5.8 and Remark 5.9, we get an associated RDS φ. Unfortunately, this equality does not necessarily hold. It is not hard to show that for every $s,h \ge 0$ the equality holds almost surely for all $t \ge s$ but the exceptional sets may depend on s,h, so the question arises if one can modify the semi-flow ϕ in such a way that the equality holds true for all $h \ge 0$, $\omega \in \Omega$, and $0 \le s \le t$. The answer is positive and the procedure to obtain such a modification is called *perfection technique* (see [A98] for details).

5.3 Invariant measures and random attractors

Definition 5.11. Let $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ be an E-valued RDS with respect to G and T as above. Let μ be a measure on $(E \times \Omega, \mathcal{E} \otimes \mathcal{F})$ with the property $\mu\pi_2^{-1} = \mathbb{P}$, i.e. the second marginal of μ equals the given measure $\mathbb{P}$ on Ω. Then μ is called an *invariant measure* of the RDS if μ is invariant under the associated skew-product Θ, i.e.

$$\mu\Theta_t^{-1} = \mu \text{ for all } t \geq 0.$$

From now on assume that the state space (E, d) is a Polish space and that $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ is a continuous RDS.

On the space $\mathcal{K}$ of non-empty and compact subsets of E we define the *Hausdorff metric* $d_{\mathcal{K}}$ as follows

$$d_{\mathcal{K}}(K_1, K_2) := \max\big\{ \sup_{x\in K_1} \inf_{y\in K_2} d(x,y), \sup_{y\in K_2} \inf_{x\in K_1} d(x,y)\big\}.$$

It is not hard to check that $(\mathcal{K}, d_{\mathcal{K}})$ is a metric space (this is even true if (E, d) is an arbitrary metric space). Further, $(\mathcal{K}, d_{\mathcal{K}})$ is complete and separable and hence Polish (this property uses the fact that (E, d) is Polish).

Definition 5.12. A random compact set $A(\omega)$ (i.e. A is a $\mathcal{K}$-valued random variable) is called a *pullback attractor* or just *attractor* or *random attractor* of the continuous RDS $(\Omega, \mathcal{F}, \mathbb{P}, \theta, \varphi)$ if

i) A is strictly invariant, i.e. $\varphi_t(A(\omega), \omega) = A(\theta_t(\omega))$ for all $t \geq 0$, $\omega \in \Omega$.

ii) A attracts all compact sets in the pullback sense, i.e. for each $K \in \mathcal{K}$ we have

$$\lim_{t\to\infty} \sup_{x\in K} \inf_{y\in A(\omega)} d(\varphi_t(x, \theta_{-t}(\omega)), y) = 0 \text{ almost surely.}$$

We call A a *weak attractor* if A is as above but with "almost surely" in ii) replaced by "in probability". We call A a *(pullback) point attractor* if A is as above but $\mathcal{K}$ is replaced by the family of all singletons in E. *Weak point attractors* are defined analogously.

Remark 5.13. One can also find slightly different definitions in the literature. Some authors (e.g. [CF94]) require that in (ii) "compact" is replaced by "bounded" (which is a stronger condition) or that certain classes of *random* sets are attracted by A.

Note that a pullback attractor is automatically a pullback point attractor (and the same for "weak" instead of "pullback"). There are examples for which the converse is not true. Not every continuous RDS admits a pullback attractor and it is of interest to find conditions which guarantee that a particular RDS admits a pullback or weak (point) attractor. If an RDS has an attractor (of whatever kind) then it may or may not be true that $A(\omega)$ is almost surely a singleton. This phenomenon is often called *synchronization*. A simple but interesting example is the SDE

$$\mathrm{d}X_t = \big(-X_t^3 + X_t\big)\,\mathrm{d}t + \sigma \mathrm{d}W_t,$$

where W is one-dimensional Brownian motion and $\sigma \geq 0$. The equation generates an $\mathbb{R}$-valued RDS (in spite of the fact that the drift is not Lipschitz) which admits a pullback attractor $A(\omega)$ for each value of $\sigma \geq 0$. In the deterministic case, we have $A(\omega) = [-1, 1]$ (no synchronization) while $A(\omega)$ is a (random) singleton in case $\sigma > 0$, i.e. we have *synchronization by noise*. Point attractors need not be unique but weak attractors (and hence pullback attractors) are (maybe I will show this in class).

Bibliography

[A98] L. Arnold, *Random Dynamical Systems*, Springer, Berlin, 1998.

[Bo18] A. Bovier, *Introduction to stochastic analysis*, Course Notes, Bonn, https://www.dropbox.com/s/1r0ys3xij2fb4ys/stochana.pdf?dl=0 2018.

[CF94] H. Crauel and F. Flandoli, Attractors for random dynamical systems, *Probab. Theory Relat. Fields*, **100**, 365–393, 1994.

[DZ96] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge Univ. Press, Cambridge, 1996.

[Eb16] A. Eberle, *Introduction to stochastic analysis*, Course Notes, Bonn, https://wt.iam.uni-bonn.de/eberle/skripten/ 2016

[EK86] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.

[F51] W. Feller, The asymptotic distribution of the range of sums of independent random variables, *Ann. Math. Stat.*, **22**, 427–432, 1951.

[IW89] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Ansterdam, 1989.

[K02] O. Kallenberg, *Foundations of Modern Probability*, second edition, Springer, New York, 2002.

[KS91] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, second edition, Springer, New York, 1991.

[Ku90] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, Cambridge, 1990.

[LS11] X. Li and M. Scheutzow, Lack of strong completeness for stochastic flows, *Ann. Probab.*, **39**, 1407–1421, 2011.

[Pr92] P. Protter, *Stochastic Integration and Stochastic Differential Equations*, second edition, Springer, New York, 1992.

[Re13] M. v.Renesse, *Stochastische Differentialgleichungen*, Course Notes, Leipzig, http://www.math.uni-leipzig.de/~renesse, 2013.

[RY94] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, second edition, Springer, 1994.

[RW194] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales*, Volume I: Foundations, second edition, Wiley, 1994.

[RW294] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales*, Volume 2: Itô Calculus, second edition, Wiley, 1994.

[S09] M. Scheutzow, *Chaining techniques and their application to stochastic flows*, in: Trends in Stochastic Analysis, eds: Blath, J., Mörters, P., Scheutzow, M., 35–63, Cambridge University Press, 2009.

[SS17] M. Scheutzow and S. Schulze, Strong completeness and semi-flows for stochastic differntial equations with monotone drift, *J. Math. Anal. Appl.*, **446**, 1555–1570, 2017.

[SV79] D. Stroock and S. Varadhan, *Multidimensional Diffusion Processes*, Springer, 1979.

[St13] T. Sturm, *Stochastische Analysis*, Course Notes, Bonn, http://www-wt.iam.uni-bonn.de/˜sturm/skript/stochana.pdf

[WW90] H. v.Weizsäcker and G. Winkler, *Stochastic Integrals*, Vieweg, 1990.