On the Support of the Laws of Diffusion Processes with Irregular Drift

by

Matthew Dunlop

Thesis

Submitted for the degree of

Master of Science

Mathematics Institute

The University of Warwick

September 2013

Contents

Acknowledgments

I wish to express my thanks to my supervisors Dr. Xue-Mei Li and Professor Martin Hairer for their guidance and support throughout the writing of this thesis. I would like to thank Professor David Elworthy for his guidance also.

Additionally I wish to thank my fellow members in the MASDOC cohort whose help has been invaluable in getting me through this introduction to graduate level research.

Furthermore I would like to acknowledge the Engineering and Physical Science Research Council for providing the funds for the Mathematics and Statistics Doctoral Training Centre at the University of Warwick.

Finally a special thanks goes to my parents for their continued support allowing me to follow my chosen path.

This thesis was typeset with LaTeX 2_ε[1] by the author.

[1] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The style package *warwickthesis* was used.

Declarations

The work described in this thesis is the author's own, except where results are explicitly cited.

This thesis is submitted to the University of Warwick in partial fulfilment of the requirements for the degree of Master of Science in Mathematics and Statistics. It has not been submitted for another degree at the University of Warwick or any other university.

Abstract

This thesis discusses the support of the laws of stochastic processes that arise as solutions to stochastic differential equations. In the case when the drift coefficient is Lipschitz continuous it is known that the support may be characterised as the closure of the space of solutions to some approximating ODEs. We explore whether this result holds when there is less regularity on the drift. In these cases we need to generalise what we mean by a solution to the approximating ODEs.

We consider three cases: when the drift coefficient is bounded measurable, when it has linear spatial growth, and when it lies in $L_p^q(T)$ with $d/p + 2/q < 1$.

Some numerical experiments are performed to gain insight into the behaviour of some processes who support is more difficult to determine analytically.

Introduction

Given a Brownian motion W on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a solution x to the SDE

$$\mathrm{d}x_t = \sigma(t, x_t)\,\mathrm{d}W_t + b(t, x_t)\,\mathrm{d}t, \quad x_0 \in \mathbb{R}, \quad t \in [0, T]$$

induces a measure $\mu := x_*\mathbb{P}$, the law of x, on the space of sample paths. With sufficient regularity on the coefficients σ and b the sample paths will almost surely lie in $C_{x_0}([0, T]; \mathbb{R}^d)$, the space of continuous paths started from x_0 valued in $\mathbb{R}^d$. The process may not explore the whole of this space, however. The subset of paths upon which the measure really 'lives' is called the support of the measure. It is rigourously defined as the set of paths for which all neighbourhoods have positive measure, or equivalently the smallest closed set with full measure.

In 1972, Stroock and Varadhan provided a characterisation of the support of the law of a diffusion process. Suppose that $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded continuous with bounded continous derivatives, and that $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous. Then given any piecewise linear function $u : [0, T] \to \mathbb{R}^d$ started from 0, the approximating ordinary differential equation

$$\mathrm{d}x_t^u = \sigma(t, x_t^u)\,\mathrm{d}u_t + b(t, x_t^u)\,\mathrm{d}t$$

has a unique solution in light of the famous Picard-Lindelöf theorem. It turns out that the support of the law of the process x is given by the closure of the set of all such solutions x^u with respect to the topology of uniform convergence. Since this result was proved, analogous results have been shown in different topologies, and for jump processes, SPDEs, and Hilbert space valued processes.

When the diffusion is uniformly elliptic, i.e. when the matrix $\sigma\sigma^T$ is uniformly positive definite, the support is often the whole space $C_0([0, T]; \mathbb{R}^d)$ even when the drift isn't Lipschitz. This is the case when for example, b is such that either

(i) there exists $m \in L^2([0,T])$ with $|b(t,x)| \leq m(t)(1+|x|)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$; or

(ii) $b \in L^q_p(T) := L^q_p([0,T]; L^p(\mathbb{R}^d;\mathbb{R}^d))$ with $d/p + 2/q < 1$ and $p \geq 2$.

If the diffusion has some degeneracy then the support may be smaller. We can still try to obtain a Stroock-Varadhan type result in the above cases, assuming solutions exist, though we run into some trouble with the approximating ODEs. If we retain spatial continuity of the drift then we may still retain existence in the Carathéodory sense, however we lose uniqueness. Decreasing the regularity any further means that we don't even have existence, and so we need to extend what we mean by solutions to the ODEs.

The first approach we take is *Filippov regularisation*. For this we map the right hand side of the ODE to a set-valued function, resulting in a *differential inclusion*. These are a strict generalisation of differential equations where instead of equality we only insist that the derivative belongs to a set. In the case (i) above we then obtain existence of solutions to the approximating equations in this new sense. For case (ii) this approach is of no use since it requires local boundedness of the right hand side, whilst $b \in L^q_p(T)$ allows for singular drift.

The other approach we take is *regularisation with noise*. If we add a small amount of noise εB_t to the right hand side then in several cases we obtain existence and uniqueness of solutions in a stochastic sense. In particular this holds in the case that the right hand side lies in $L^q_p(T)$ with $d/p+2/q < 1$. Since the laws of the perturbed equations will have full support for all $\varepsilon > 0$, we ask what happens when we send ε to zero. In case (i) above the limiting law turns out to be supported on the set of Filippov solutions to the equation. Less is known in case (ii), however.

Numerical simulations can help us gain insight into the behaviour of diffusions which are difficult to approach analytically. We perform some such simulations in Chapter 5, exploring cases such as when the equation is elliptic with singular drift, and when the equation is degenerate with bounded drift.

Chapter 1

Support Theorems

1.1 The support of a measure

We first define what is meant by the (topological) support of a measure.

Definition 1.1 (Topological support)**.** *Given a measure μ on a topological space E, the* topological support $\text{supp}(\mu)$ *of μ is the closed subset of E defined as*

$$\text{supp}(\mu) = \{y \in E \mid \mu(U) > 0 \text{ for all open neighbourhoods } U \text{ of } y\}$$

The support of μ can informally be thought of as the set upon which the measure really 'lives'. If μ is a finite measure with total mass m, then an equivalent formulation of the support is

$$\text{supp}(\mu) = \bigcap_{\substack{K \text{closed} \\ \mu(K)=m}} K$$

i.e. it is the smallest closed set of full measure.

Remark. *We lose nothing by only considering open neighbourhoods of the form $B_\delta(y)$, where this denotes the open ball of radius δ centred at y. Clearly if $\mu(U) > 0$ for all open sets U containing a point, then it holds in particular for all open balls containing that point. Conversely, if we only assume $\mu(B_\delta(y)) > 0$ for all $\delta > 0$, then since any open set containing y contains some ball around y, the measure of any open set containing y is positive.*

1.1.1 Examples

1. Consider Euclidean space $\mathbb{R}^d$ with the Lebesgue measure. Then since all balls have positive measure, the support of the Lebesgue measure is the whole of $\mathbb{R}^d$.

2. Let δ_x denote the Dirac measure at x on some topological space, i.e. $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise. Then $\mathrm{supp}(\delta_x) = \{x\}$.

3. Let W be a Brownian motion on $\mathbb{R}$ defined on Wiener space ($C_0([0,T];\mathbb{R})$ equipped with the uniform topology) with law $\mathbb{P}$ (the Wiener measure). We know that Brownian paths are almost-surely continuous and start at 0, so $\mathbb{P}(C_0([0,T];\mathbb{R})) = 1$. In fact the support is the whole of this space, which follows from the strict positivity of the Wiener measure.

   Note that even though it is known that Brownian paths are almost-surely α-Hölder continuous of order $\alpha < 1/2$, the support of $\mathbb{P}$ turned out to be the whole of $C_0([0,T];\mathbb{R})$. This is a consequence of the topology we put on the space. It can be shown that if we consider the same space with the α-Hölder topology instead of the uniform topology, the support is $C_0^{0,\alpha}([0,T];\mathbb{R})$, $\alpha < 1/2$. See [AGL94] for a more general result.

4. Following the same setup as the previous example, let $x_t := |W_t|$ and denote by μ its law. We claim that the support of μ is the space C_0^+ given by

$$C_0^+ := \{\sigma \in C_0([0,T];\mathbb{R}) \,|\, \sigma(t) \geq 0 \text{ for all } t \in [0,T]\}$$

   First note that $X_t \in C_0([0,T];\mathbb{R})$ and $X_t \geq 0$ everywhere, so $\mu(C_0^+) = 1$. Since C_0^+ is closed, the support of μ can be no bigger than C_0^+.

   We now show that given any $\sigma \in C_0^+$, $\mu(B_\delta(\sigma)) > 0$ for all $\delta > 0$. First note that for any real numbers $a < b$, we have the relation

$$\{x \,|\, |x| \in (a,b)\} = \begin{cases} (a,b) \cup (-b,-a) & a,b > 0 \\ (-b,b) & a \leq 0, b > 0 \\ \emptyset & a,b \leq 0 \end{cases}$$

Therefore given $\sigma \in C_0^+$ and $\delta > 0$ we have

$$\begin{aligned}
\mu(B_\delta(\sigma)) &= \mathbb{P}(\{\omega \,|\, |\omega| \in B_\delta(\sigma)\}) \\
&= \mathbb{P}(\{\omega \,|\, \omega(t) \in B_\delta(\sigma(t)) \cup B_\delta(-\sigma(t)) \text{ for all } t \text{ s.t. } \sigma(t) > \delta, \\
&\qquad\qquad \omega(t) \in (-\sigma(t) - \delta, \sigma(t) + \delta) \text{ for all } t \text{ s.t. } \sigma(t) \leq \delta\}) \\
&\geq \mathbb{P}(\{\omega \,|\, \omega(t) \in B_\delta(\sigma(t)) \cup B_\delta(-\sigma(t)) \text{ for all } t \text{ s.t. } \sigma(t) > \delta, \\
&\qquad\qquad \omega(t) \in B_\delta(\sigma(t)) \cup B_\delta(-\sigma(t)) \text{ for all } t \text{ s.t. } \sigma(t) \leq \delta\}) \\
&= \mathbb{P}(B_\delta(\sigma) \cup B_\delta(-\sigma)) \\
&> 0
\end{aligned}$$

by the strict positivity of the Wiener measure.

The following result is clear and will turn out to be useful.

Proposition 1.2. *Let μ and ν be two equivalent measures. Then* $\operatorname{supp}(\mu) = \operatorname{supp}(\nu)$.

The converse does not hold: if two measures have the same support then they need not be equivalent. Consider for example two discrete measures μ and ν on $\mathbb{R}$ such that μ assigns positive mass only to all elements of $\mathbb{Q}$, and ν assigns positive mass only to all elements of $\sqrt{2} + \mathbb{Q}$. Then the support of both of these measures is all of $\mathbb{R}$, whilst they are clearly singular. Many more examples exist in infinite dimensions, where singularity of different measures is almost expected.

1.2 Diffusion measures

The measures we are interested in are those that arise as the laws of diffusion processes. Given maps $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^m \otimes \mathbb{R}^d$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and an m-dimensional Brownian motion W defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a diffusion x is defined by the SDE[1]

$$dx_t = \sigma(t, x_t)\, \mathrm{d}W_t + b(t, x_t)\, \mathrm{d}t, \quad x_0 \in \mathbb{R}^d$$

Under suitable regularity conditions on σ and b, there exist solutions to this SDE which almost surely take values in the space $C_{x_0}([0,T]; \mathbb{R}^d)$, see for example [IW81].

[1] For simplicity, from now on we will assume that $m = d$, since we can always zero out the appropriate rows/columns in a square matrix to effectively obtain a rectangular matrix.

A solution x therefore induces a measure $\mu := x_*\mathbb{P}$ on this space. Our question is what precisely the support of μ is.

In the case that σ is the identity matrix and the drift b is zero, the solution to the SDE is a Brownian motion, whose law we know has full support on $C_0([0,T];\mathbb{R}^d)$ from the previous section. In that section we also considered the absolute value of one-dimensional Brownian motion. Tanaka's formula tells us that this is formally defined by the SDE

$$dx_t = \operatorname{sgn}(x_t)\,\mathrm{d}W_t + \delta_0(x_t)\,\mathrm{d}t, \quad x_0 = 0$$

where δ_0 is the Dirac delta function at 0. In this case the support only turns out to be the space C_0^+. The smaller support is related to the reduced regularity on σ and b. Note in particular that the SDE above is not elliptic since $\operatorname{sgn}(0) = 0$; 0 is a 'barrier' to the diffusion.

In the uniformly elliptic case we often have that the support of the law is all of $C_{x_0}([0,T];\mathbb{R}^d)$.

Proposition 1.3. *Suppose that $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded, continuous and $\sigma\sigma^T$ is uniformly positive definite, so that there exist constants $\lambda, \Lambda > 0$ such that for all $\xi \in \mathbb{R}^d$ and all $(t,x) \in [0,T] \times \mathbb{R}^d$,*

$$\lambda|\xi|^2 \le \langle \xi, \sigma\sigma^T(t,x)\xi\rangle_{\mathbb{R}^d} \le \Lambda|\xi|^2$$

Let $b : [0,T] \times \mathbb{R}^d$ be such that either

- *(i) there exists $m \in L^2([0,T])$ with $|b(t,x)| \le m(t)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$;*
- *(ii) there exists $m \in L^2([0,T])$ with $|b(t,x)| \le m(t)(1+|x|)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$; or*
- *(iii) $b \in L_p^q(T) := L^q([0,T];L^p(\mathbb{R}^d;\mathbb{R}^d))$, where $d/p + 2/q < 1$ and $p \ge 2$.*

In case (iii), assume also that $\sigma\sigma^T$ is differentiable in its spatial component and its derivatives are bounded. Then the support of the law μ of the solution to the SDE

$$\mathrm{d}x_t = \sigma(t,x_t)\,\mathrm{d}W_t + b(t,x_t)\,\mathrm{d}t$$

started from $x_0 \in \mathbb{R}^d$ is the whole space $C_{x_0}([0,T];\mathbb{R}^d)$.

Proof. We will assume without loss of generality that $x_0 = 0$, and use Girsanov's theorem to transform away the drift term. Let y be the process defined by the SDE

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

and let ν denote its law. Define the density

$$\begin{aligned} Z_t &= \mathcal{E}\left(\int_0^{\cdot} \sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s)\,\mathrm{d}s\right)_t \\ &= \exp\left\{\int_0^t \langle \sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s),\,\mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t |\sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s)|^2\,\mathrm{d}s\right\} \end{aligned}$$

We first look at case (i). From the uniform positive-definiteness of $\sigma\sigma^T$ we know that its (matrix) inverse is uniformly bounded. Using the boundedness of σ also, we see that

$$|\sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s)| \le C|b(s, y_s)| \le Cm(s)$$

for all s. It follows that

$$\mathbb{E}\left(\exp\left\{\frac{1}{2}\int_0^T |\sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s)|^2\,\mathrm{d}s\right\}\right) < \infty \qquad (1.1)$$

This bound is Novikov's condition, so Z is a martingale with respect to ν and we have $\mathrm{d}\nu = Z\,\mathrm{d}\mu$. The measures μ and ν are hence equivalent and we have that $\mathrm{supp}(\mu) = \mathrm{supp}(\nu)$ by Proposition 1.2.

We show that (1.1) holds in the more general case (ii) also. By Jensen's inequality and Fubini's theorem we have

$$\mathbb{E}\left(\exp\left\{\frac{1}{2}\int_0^T |\sigma^T(\sigma\sigma^T)^{-1}(s, y_s)b(s, y_s)|^2\,\mathrm{d}s\right\}\right) \le C\exp\left\{C\int_0^T |m(t)|^2\mathbb{E}|y_s|^2\,\mathrm{d}s\right\}$$

so it suffices to show that $\mathbb{E}|y_s|^2$ is bounded. This follows via the Itô isometry:

$$\mathbb{E}|y_t|^2 = \mathbb{E}\left|\int_0^t \sigma(s, y_s)\,\mathrm{d}W_s\right|^2 \le C\int_0^T \mathbb{E}|\sigma(s, y_s)|^2\,\mathrm{d}s < \infty$$

For case (iii), in the case that $\sigma \equiv I_d$, we use a result from the proof of Lemma 3.2 in [KR05] which tells us that for any $\kappa > 0$,

$$\mathbb{E}\left(\exp\left\{\kappa\int_0^T |b(s, x_s)|^2\,\mathrm{d}s\right\}\right) < \infty \qquad (1.2)$$

and so in particular it holds for $\kappa = K^2/2$ where $K = \sup_{t,x}|\sigma^T(\sigma\sigma^T)^{-1}(t, x)|$.

Therefore (1.1) holds in this case also.

We wish to extend the bound (1.2) to the general uniformly elliptic case. We show that the proof in [KR05] still holds with the Brownian motion replaced by the process y as defined above.

First note that Khas'minskii's lemma still holds with y in place of a Brownian motion since y still satisfies the Markov property. Next note that the density ρ_t of y_t satisfies the Fokker-Planck equation

$$\frac{\partial \rho_t}{\partial t} = \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j} \left((\sigma\sigma^T)_{ij} \rho_t \right), \quad \rho_0 = \delta_{x_0}$$

From the uniform ellipticity of $\sigma\sigma^T$, and the boundedness of $\sigma\sigma^T$ and its derivatives, we can use the main result of [Aro67] to deduce that ρ_t satisfies the Gaussian upper bound

$$\rho_t(x) \leq K t^{-d/2} \exp\left(-\frac{1}{2t} \delta |x|^2 \right)$$

for some constants $K, \delta > 0$. The right hand side is a rescaled heat kernel, and so we see that the arguments in the proof of Lemma 3.2 in [KR05] still hold[2]. In particular (1.1) holds and we have $\text{supp}(\mu) = \text{supp}(\nu)$ again.

We therefore only need to look at the support of ν. We use the following lemma from [SV72].

Lemma 1.4. *Let $\varphi : [0,T] \to \mathbb{R}^d$ be once continuously differentiable such that $\varphi_0 = 0$. Then for all $\varepsilon > 0$, $\nu(\{\omega \mid \|\omega - \varphi\|_\infty < \varepsilon\}) > 0$.*

Since the functions φ considered are dense in $C_0([0,T];\mathbb{R}^d)$, the result follows. □

The result of (*iii*) is perhaps surprising since it allows for singular drift. These singularities could believably be strong enough such that they act as barriers for the process, with the process almost-surely being forced away from them by the high intensity drift.

Consider the one-dimensional case in which the drift is time-independent. Then sending $q \to \infty$ in the condition $d/p + 2/q < 1$, we see that we have free choice of $p > 1$. It is standard that the function $b(x) = |x|^{-\alpha} \mathbb{1}_{(0,1)}(x)$ lies in $L^p(\mathbb{R})$ for $\alpha < 1/p < 1$. Fix $\alpha = 1/2$, say. Then the proposition tells us that the process x

[2]See the appendix in [FF10] for a more detailed proof of the lemma in [KR05].

given by

$$dx_t = dW_t + \frac{1}{\sqrt{|x_t|}} \mathbb{1}_{(0,1)}(x_t)\,\mathrm{d}t, \quad x_0 > 0$$

can reach the whole of $\mathbb{R}$. This could be due to the requirement $\alpha < 1$, which means that the drift can't grow *too* fast near zero. It isn't unreasonable to think that for larger α the drift is strong enough to prevent the process from passing the origin, and that perhaps the result of (iii) is somewhat sharp. In Chapter 5 we will simulate these processes numerically to gain an insight into their behaviour.

In light of Proposition 1.3, the case when the equation isn't elliptic is arguably much more interesting since the support of the law can potentially be more exotic. For example, consider *Brownian motion on* $[0,1]$ *with sticky boundary*[3], given by the SDE

$$\mathrm{d}x_t = \mathbb{1}_{(0,1)}(x_t)\,\mathrm{d}W_t + \theta\left(\mathbb{1}_{\{0\}}(x_t) - \mathbb{1}_{\{1\}}(x_t)\right)\,\mathrm{d}t, \quad x_0 \in [0,1]$$

where $\theta \in (0,\infty)$ is a positive constant. This doesn't fall into the class of SDEs considered in Proposition 1.3 due to the discontinuous and non-elliptic diffusion coefficient, though the drift is bounded and hence okay. In this case the support of the law of x is given by

$$\{\sigma \in C_{x_0}([0,T];\mathbb{R}) \,|\, 0 \le \sigma(t) \le 1 \text{ for all } t \in [0,T]\}.$$

Note that in all the examples of supports of measures on $C_{x_0}([0,T];\mathbb{R}^d)$ we've looked at, the restrictions come in which sets the sample paths can reach rather than the regularity of the paths. This is a consequence of the uniform topology on the space and the closedness of the support: higher regularity cannot be expected to be preserved under limits in the uniform topology.

For cases when the equation isn't uniformly elliptic, the support theorems may be able to give a description of the support.

1.3 Existing theorems

In certain cases, the support of the law of a diffusion process can be characterised by the closure of the set of solutions of a family of approximating ODEs. Such a characterisation can be used to prove strong maximum principles for the PDEs associated with the generators of the diffusions.

[3]More information about Brownian motion with sticky boundary can be found in [DT94]

The classical Stroock-Varadhan support theorem as originally given in [SV72] is as follows:

Theorem 1.5 (Stroock-Varadhan support theorem)**.** *Let $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be bounded measurable functions. Assume also that $\sigma_{ij} \in C_b^{1,2}([0,T] \times \mathbb{R}^d)$ and b is uniformly Lipschitz continuous in x. Define the differential operator*

$$\begin{aligned} L_t &= \frac{1}{2}\sigma^T \nabla_x \cdot \sigma^T \nabla_x + b \cdot \nabla_x \\ &= \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^T)_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} \left(b + \frac{1}{2}\sigma'\sigma \right)_i \frac{\partial}{\partial x_i} \end{aligned}$$

Define also the space $\mathcal{S}_{\sigma,b}(t_0, x_0)$ as the class of $\varphi \in C([t_0,T]; \mathbb{R}^d)$ for which there exists a piecewise linear $u : [t_0, T] \to \mathbb{R}^d$ such that

$$\varphi_t = x_0 + \int_{t_0}^{t} \sigma(r, \varphi_r)\dot{u}_r \,\mathrm{d}r + \int_{t_0}^{t} b(r, \varphi_r)\,\mathrm{d}r \tag{1.3}$$

The the support of the unique solution P_{t_0,x_0} to the martingale problem for L_t is given by

$$\operatorname{supp}(P_{t_0,x_0}) = \overline{\mathcal{S}_{\sigma,b}(t_0, x_0)}$$

where both the support and closure are taken in the uniform topology.

Remarks.

1. *This theorem remains true when instead we require that the control functions u belong to $L_0^{2,1}([0,T]; \mathbb{R}^d)$, the Cameron-Martin space for the Wiener measure. Indeed this is the space that is used in most modern adaptations of the theorem, and the space that we will be using later. It may be interesting to study which other spaces we may substitute in place, however.*

2. *P_{t_0,x_0} is the law of the solution to the Stratonovich SDE*

$$x_t = x_0 + \int_{t_0}^{t} \sigma(s, x_s) \circ \mathrm{d}W_s + \int_{t_0}^{t} b(s, x_s)\,\mathrm{d}s \tag{SDE}$$

   with the $\sigma'\sigma$ term appearing in L_t being the Itô-Stratonovich correction. It is this SDE that we shall be interested in from now on (with $t_0 = 0$).

3. *The spatial continuity and temporal measurability of the integrands on the right hand side of (1.3) ensure that these ODEs have solutions.*

The theorem as stated above is fairly old, and there have been developments since it was published back in 1972. It turns out that it is also true in the finer α-Hölder topologies for $\alpha \in [0, 1/2)$, see [AGL94] [MSS94]. The modern tools of rough path theory have not only been used to give a succinct proof of the theorem in the uniform topology, but also provide similar theorems for certain classes of SPDEs [FV10] [LQZ02]. Other extensions include to Hilbert space valued SDEs and to jump processes, see for example [Nak04] [Sim00].

1.3.1 Sketch proof in the uniform topology

We sketch the proof only in the degenerate case, since the non-degenerate case is more straightforward. We work in the space $C_0([0,T];\mathbb{R}^d)$ equipped with the uniform norm. This is the proof given in [SV72], corresponding to Theorem 1.5. The proofs in the Hölder topology follow the same general structure, though there are more technical details involved.

Without loss of generality we assume that the process is started from 0 at time 0, and write $\mu = P_{0,0}$, $\mathcal{S}_{\sigma,b} = \mathcal{S}_{\sigma,b}(0,0)$. There are two inclusions that we must show. The first, $\mathrm{supp}(\mu) \subseteq \overline{\mathcal{S}_{\sigma,b}}$, is shown using a Wong-Zakai approximation to the SDE. The other inclusion is proved by showing that, conditional on the control function being close to the Brownian motion, the solution to the SDE is close to that of the control problem with high probability.

$\mathrm{supp}(\mu) \subseteq \overline{\mathcal{S}_{\sigma,b}}$

This inclusion is typically regarded as the easy one. Take an approximation $W^{(n)}$ of the Brownian motion, piecewise linear on the dyadic partition of $[0,T]$. The approximating process then given by

$$x_t^{(n)} = \int_0^t \sigma(r, x_r^{(n)}) \dot{W}_r^{(n)} \,\mathrm{d}r + \int_0^t b(r, x_r^{(n)}) \,\mathrm{d}r$$

Wong-Zakai's theorem tells us that the law of $x^{(n)}$ converges weakly to μ, the law of x. We have that $x_t^{(n)} \in \overline{\mathcal{S}_{\sigma,b}}$ for all sample paths and all n. This is preserved under the weak limit thanks to the following lemma, and so we are done.

Lemma 1.6. *Let $(\nu_n)_{n\geq 1}$ be a family of probability measures on a space X such that $\nu_n \to \nu$ weakly. If $A \subseteq X$ is a closed set such that* $\mathrm{supp}(\nu_n) \subseteq A$ *for all n, then* $\mathrm{supp}(\nu) \subseteq A$.

Proof. Suppose the result is false, and there exists $x \in \operatorname{supp}(\mu)$ such that $x \notin A$. Then since A is closed, $d(x, A) = \delta > 0$. Choose $O = B_{\delta/2}(x)$. Then $\mu_n(O) = 0$ for all n since $O \cap \operatorname{supp}(\mu_n) = \emptyset$ for all n, but $\mu(O) > 0$ since $x \in \operatorname{supp}(\mu)$. Therefore by the lower semi-continuity of weak convergence on open sets we have

$$0 = \liminf_{n\to\infty} \mu_n(O) \geq \mu(O) > 0$$

which is a contradiction. □

$\overline{\mathcal{S}_{\sigma,b}} \subseteq \operatorname{supp}(\mu)$

It suffices to show that, for all φ in a dense subset of $\overline{\mathcal{S}_{\sigma,b}}$ and all $\varepsilon > 0$,

$$\mu(\{\omega \mid \|\omega - \varphi\|_\infty < \varepsilon\}) > 0 \tag{1.4}$$

The dense set considered will be $\varphi = \varphi^u$ of the form

$$\varphi^u_t = \int_0^t \sigma(r, \varphi^u_r)\dot{u}_r \,\mathrm{d}r + \int_0^t b(r, \varphi^u_r)\,\mathrm{d}r$$

where $u \in C^2([0,T] \times \mathbb{R}^d)$ with $u_0 = 0$. (1.4) is then proved by showing that

$$\lim_{\delta\downarrow 0} \mathbb{P}(\|x - \varphi^u\|_\infty < \varepsilon \mid \|W - u\|_\infty < \delta) = 1 \tag{1.5}$$

This is done by first considering the case $u = 0$. An application of Itô's formula shows that it is sufficient to show that

$$\lim_{\delta\downarrow 0} \mathbb{P}(\|\Delta\|_\infty \mid \|W\|_\infty < \delta) = 1$$

where Δ_t is a sum of stochastic integrals of derivatives of σ. This follows from a lengthy calculation involving several applications of Itô's formula and Gaussian estimates. An application of Girsanov's theorem then yields (1.5).

1.4 Extension to irregular drift

In all of the above theorems the drift coefficient b is assumed to be Lipschitz continuous. We ask what happens when we have less regularity on b. More specifically we ask whether a support theorem holds when $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ satisfies any of:

(i) b is bounded measurable;

(ii) there exists $m \in L^2([0,T])$ with $|b(t,x)| \le m(t)(1+|x|)$ for all $t \in [0,T]$ and $x \in \mathbb{R}^d$; or

(iii) b lies in the space $L^q_p(T) := L^q([0,T]; L^p(\mathbb{R}^d;\mathbb{R}^d))$, where $d/p + 2/q < 1$.

As in the above theorem we will treat the elliptic and degenerate cases separately. The general degenerate case is particularly difficult to approach because we aren't even guaranteed the existence of solutions to the SDE!

Another issue is that the control ODEs do not necessarily have solutions in the classical sense. We recall Carathéodory's well-known existence theorem for solutions to first order ODEs:

Theorem 1.7 (Carathéodory)**.** *Consider the differential equation*

$$y_t = y_0 + \int_0^t f(s, y_s)\,\mathrm{d}s$$

with f defined on the rectangular domain $R = \{(t,y) \,|\, |t-t_0| \le a, |y-y_0| \le b\}$. If the function f satisfies the following three conditions:

(i) $f(t,y)$ is continuous in y for each fixed t

(ii) $f(t,y)$ is measurable in t for each fixed y

(iii) there is a Lebesgue-integrable function $m(t)$, $|t-t_0| \le a$, such that $|f(t,y)| \le m(t)$ for all $(t,y) \in R$

then the differential equation has a solution in a neighbourhood of the initial condition.

Our control ODEs will be of the form

$$x^u_t = x_0 + \int_0^t \left(\sigma(s, x^u_s)\dot{u}_s + b(s, x^u_s)\right)\,\mathrm{d}s \tag{1.6}$$

and so because of the lack of spatial continuity of b, condition (i) in the theorem fails. In the $L^q_p(T)$ case, condition (iii) also fails due to the lack of boundedness.

We remedy some of these issues by extending what we mean for a function to be a solution to (1.6).

Chapter 2

Regularisation of ODEs

Our main problem so far is that solutions to the control problems don't necessarily exist in the classical sense. We address this in two different ways. The first approach is to instead look for a generalised type of solutions, known as *Filippov solutions*. These are solutions to *differential inclusions*, a generalisation of differential equations to set-valued right hand sides. The second approach is to perturb the ODEs with noise, yielding existence and uniqueness in a stochastic sense. We then consider the limit as the intensity of the noise goes to zero.

In the case of bounded measurable drift, both of these approaches turn out to be strongly related.

2.1 Differential inclusions and Filippov solutions

Definition 2.1 (Differential inclusion). *Let $F : [0,T] \times \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \emptyset$ be a set-valued map. We say that a function $x : [0,T] \to \mathbb{R}^d$ is a solution to the differential inclusion*

$$\dot{x}_t \in F(t, x_t), \quad x_0 \in \mathbb{R}^d$$

if x is absolutely continuous and the above inclusion holds for almost all $t \in [0,T]$.

It can be seen that if F is single-valued then this reduces to the definition of an ordinary differential equation (in the extended sense).

As a simple example of a differential inclusion, consider the case when F is given by the set-valued sign function:

$$\dot{x}_t \in \begin{cases} 1 & x_t > 0 \\ [-1,1] & x_t = 0 \\ -1 & x_t < 0 \end{cases}$$

If we start solutions from $x_0 \neq 0$, then they will just move linearly in the direction away from 0. If we start the solutions from $x_0 = 0$, they may move linearly in either direction immediately, or they may remain at 0 for an arbitrary period of time before leaving in the same manner.

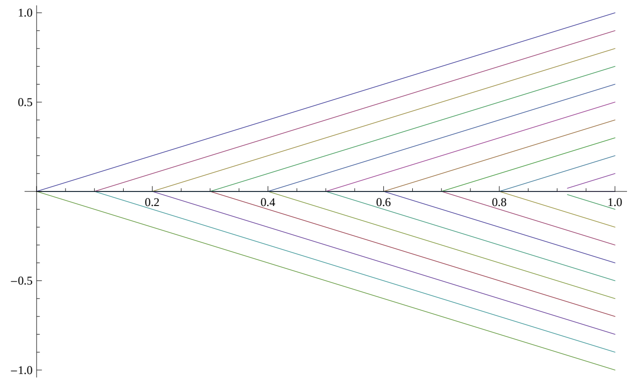

Figure 2.1: Plots of 20 different solutions to the above differential inclusion.

If we flip the sign of $\dot{x}$ in the above inclusion, then we achieve uniqueness: all solutions will immediately converge linearly to the origin and then stay there.

A more exotic example would be given by $\dot{x}_t \in \{-1, 1\}$. In this case the space of solutions contains, for example, the space of linearly interpolated sample paths of a simple random walk on $\mathbb{Z}$.

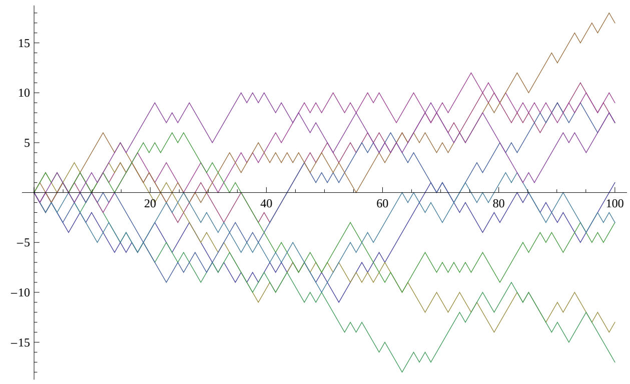

Figure 2.2: Examples of some solutions to the inclusion $\dot{x}_t \in \{-1, 1\}$.

As expected, some assumptions on the map F are required to ensure existence of solutions. There are quite a few different existence theorems with competing degrees of regularity, but the one we are mainly interested in is Theorem 4.7 in [Smi00]. Before we state this we will need a definition, again taken from [Smi00]:

Definition 2.2 (Upper semi-continuity)**.** *Let F be a set-valued map defined on a topological space X. F is said to be* upper semi-continuous at $x_0 \in X$ *if for any open set M containing $F(x_0)$ there exists a neighbourhood Ω of x_0 such that $F(\Omega) \subseteq M$. F is said to be* upper semi-continuous *if it is so at every point $x_0 \in X$.*

As an example, consider the set-valued map on $\mathbb{R}$ given by $F(x) = [-|x|, |x|]$. Without loss of generality we can assume that any open set containing $F(x_0)$ is of the form $M = (-|x_0| - \varepsilon, |x_0| + \varepsilon)$ for some $\varepsilon > 0$. If we then choose $\Omega = (x_0 - \varepsilon/2, x_0 + \varepsilon/2)$, we see that $F(\Omega) \subseteq M$ and so F is upper semi-continuous.

For another example, again on $\mathbb{R}$, consider the map given by

$$F(x) = \begin{cases} \{0\} & x \in \mathbb{Q} \\ \mathbb{R} & x \notin \mathbb{Q} \end{cases}$$

Then F is not upper semi-continuous on the rationals: if $q \in \mathbb{Q}$ then $F(q) = \{0\}$, but $F(\Omega) = \mathbb{R}$ for any neighbourhood Ω of q. Upper semi-continuity on the irrationals is clear.

Theorem 2.3 (Existence)**.** *Let $F : [0, T] \times \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \emptyset$ be a set-valued map with closed convex values. Assume that*

(i) the set-valued map $x \mapsto F(t, x)$ is upper semi-continuous for almost all $t \in [0, T]$;

(ii) for any $x \in \mathbb{R}^d$ there exists a measurable function $t \mapsto f(t, x)$ satisfying $f(t, x) \in F(t, x)$; and

(iii) there exists a function $m \in L^1([0, T])$ such that $|f(t, x)| \leq m(t)$, $t \in [0, T]$.

Then for any $x_0 \in \mathbb{R}^d$ there exists a solution x to the differential inclusion

$$\dot{x}_t \in F(t, x_t), \quad t \in [0, T]$$

started from x_0.

Compare this with Caratheodory's existence theorem for ODEs given in the previous chapter: if $F(t,x) = \{f(t,x)\}$ is single-valued and f is continuous in the spatial component we have the same result.

Remark. *This theorem does require that the map $x \mapsto f(t,x)$ is globally bounded, though this can be relaxed to local boundedness as will be discussed later.*

We will denote by $\mathcal{S}(F, x_0)$ the set of solutions to the inclusion $\dot{x}_t \in F(t, x_0)$ started from $x_0 \in \mathbb{R}^d$. In the situations we will be interested in this set will be compact: we prove this in Proposition 2.10.

In some cases it will be more convenient to work with the integral form of differential inclusions. We must first however define what me mean by the integral of a set-valued function.

Definition 2.4. *Let $F : [0,T] \times \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \emptyset$ be upper semi-continuous in the second component, and let $x : [0,T] \to \mathbb{R}^d$ be absolutely continuous. We define the set-valued function $(s,t) \mapsto \int_s^t F(u, x_u)\,\mathrm{d}u$ by*

$$\int_s^t F(u, x_u)\,\mathrm{d}u = \left\{ \int_s^t f(u)\,\mathrm{d}u \,\middle|\, f(u) \in F(u, x_u) \text{ for all } u \in [s,t] \text{ and } f \text{ integrable} \right\}$$

Remark. *Such functions f in this definition are referred to as* selections *from $F(\cdot, x_\cdot)$.*

We can now quote Lemma 2.1.1 from [AC84].

Lemma 2.5. *Let F be an upper semi-continuous map from $[0,T] \times \mathbb{R}^d$ into the compact*[1] *convex subsets of $\mathbb{R}^d$. Then a continuous function x is a solution on $[0,T]$ to the inclusion*

$$\dot{x}_t \in F(t, x_t)$$

if and only if for every pair (s,t),

$$x_t \in x_s + \int_s^t F(u, x_u)\,\mathrm{d}u$$

In the case that F is single valued, this lemma just tells us the equivalence between differential and integral equations.

[1]The proof of this lemma does not appear to use the compactness assumption directly - it is just used to guarantee existence of solutions to the inclusion.

2.1.1 Filippov regularisation

In the problems we are interested in we don't have a set-valued right hand side, we just have an $\mathbb{R}^d$-valued function. We therefore need a natural way to obtain a set-valued map corresponding to our function. A suitable choice is given by the so-called Filippov regularisation [BOQ09]:

Definition 2.6 (Filippov regularisation). *Given a function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, define the set-valued map $F_f : [0,T] \times \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \emptyset$ by*

$$F_f(t,x) := \bigcap_{|N|=0} \bigcap_{\delta>0} \overline{\text{co}} \{ f(t, B_\delta(x) \setminus N) \}$$

F_f is called the Filippov regularisation *of f. Here $B_\delta(x)$ denotes the d-dimensional ball of radius δ centred at x, $\overline{\text{co}}\{A\}$ denotes the closed convex hull of a set A, and the first intersection is taken over all sets of Lebesgue measure zero.*

This definition is such that solutions to the differential inclusion $\dot{x}_t \in F_f(t,x_t)$ can be thought of as limits of solutions to $\dot{x}^\delta_t = f_\delta(t, x^\delta_t)$, where $f_\delta(t,x)$ is obtained by averaging $f(t,y)$ over $y \in B_\delta(x)$. Indeed suppose that f is locally integrable, so that all of its points are Lebesgue points. Define the average f_δ of f over B_δ via

$$f_\delta(t,x) := \frac{1}{|B_\delta|} \int_{B_\delta(x)} f(t,y) \, \mathrm{d}y$$

Then Lebesgue's differentiation theorem tells us that $f_\delta(t,x) \to f(t,x)$ as $\delta \downarrow 0$. Another application of the theorem tells us that for all t, the map $x \mapsto f_\delta(t,x)$ is continuous, and so by Carathéodory's existence theorem we have existence of solutions to the ODE $\dot{x}^\delta_t = f_\delta(t, x^\delta_t)$ for all t and each $\delta > 0$. It follows that any solution x^δ satisfies

$$\dot{x}^\delta_t \in \overline{\text{co}} \left\{ f(t, B_\delta(x^\delta_t)) \right\}$$

for all t and each $\delta > 0$. Taking the intersection over all $\delta > 0$, we can deduce that any limit z of solutions (if it exists) will satisfy

$$\dot{z}_t \in \bigcap_{\delta>0} \overline{\text{co}} \{ f(t, B_\delta(z_t)) \} .$$

Since we have $F_f(t,x) \subseteq \bigcap_{\delta>0} \overline{\text{co}} \{ f(t, B_\delta(x)) \}$, it follows that any Filippov solutions will also solve the same inclusion as the limits of the approximations x^δ.

We ideally want F_f to satisfy the assumptions of Theorem 2.3 so that we can infer existence of solutions to the corresponding differential inclusion. This is the main reason why the closed convex hull appears in the definition. There are some useful properties of the regularisation which are collected in the following proposition. This a minor modification of Proposition 2 in [BOQ09], where instead we allow for explicit time dependence.

Proposition 2.7 (Properties I). *Let $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a measurable function such that $x \mapsto f(t,x)$ is locally bounded for almost all $t \in [0,T]$. Then*

(i) There exists a set $N_f(t)$ negligible under the Lebesgue measure such that for any $(t,x) \in \mathbb{R} \times \mathbb{R}^d$,

$$F_f(t,x) = \bigcap_{\delta>0} \overline{\text{co}} \left\{ f(t, B_\delta(x) \setminus N_f(t)) \right\}.$$

(ii) For almost all $(t,x) \in [0,T] \times \mathbb{R}^d$, we have $f(t,x) \in F_f(t,x)$.

(iii) For each t, the set-valued map $F_f(t,\cdot)$ is the smallest upper semi continuous set-valued map F^t with closed convex values such that $f(t,x) \in F^t(x)$ for almost all $x \in \mathbb{R}^d$.

(iv) The map $(t,x) \mapsto F_f(t,x)$ is single-valued if and only if there exists a continuous function g which coincides almost everywhere with f. In this case we have $F_f(t,x) = \{g(t,x)\}$ for almost all $(t,x) \in [0,T] \times \mathbb{R}^d$.

(v) If a function $\tilde{f}$ coincides almost everywhere with f, then $F_f(t,x) = F_{\tilde{f}}(t,x)$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$.

(vi) There exists a function $\bar{f}$ which is equal almost everywhere to f and such that

$$F_f(t,x) = \bigcap_{\delta>0} \overline{\text{co}} \left\{ \bar{f}(t, B_\delta(x)) \right\}$$

(vii) We have

$$F_f(t,x) = \bigcap_{\tilde{f}=f \text{ a.e.}} \bigcap_{\delta>0} \overline{\text{co}} \left\{ \tilde{f}(t, B_\delta(x)) \right\}$$

where the first intersection is taken over all functions $\tilde{f}$ being equal to f almost everywhere.

Proof. The proof is identical to that in [BOQ09]: we just treat t as a parameter. □

Corollary 2.8. *Let $f : [0,T] \times \mathbb{R}^d$ be such that for all $x \in \mathbb{R}^d$, $t \mapsto f(t,x)$ is measurable and there exists $m \in L^1([0,T])$ such that $|f(t,x)| \leq m(t)$ for $t \in [0,T]$. Then there exists a solution to the differential inclusion*

$$\dot{x}_t \in F_f(t, x_t), \quad x(0) = x_0$$

Remark. *We say that x is a solution to the differential equation $\dot{x}_t = f(t, x_t)$ in the Filippov sense, or a Filippov solution, if it is a solution to the corresponding differential inclusion $\dot{x}_t \in F_f(t, x_t)$.*

If F, G are set valued functions satisfying the assumptions of the existence theorem, and $F(t,x) \subseteq G(t,x)$ for all $t \in [0,T]$, $x \in \mathbb{R}^d$, then it's clear that $\mathcal{S}(F, x_0) \subseteq \mathcal{S}(G, x_0)$. It may therefore be tempting to think that if f is continuous almost everywhere and a classical solution exists to the problem

$$\dot{x}(t) = f(t, x(t)), \quad x_0 \in \mathbb{R}^d$$

then it is a Filippov solution to the corresponding differential inclusion. This is incorrect, since although the regularisation F_f will be single valued, it may differ from f on a null set. Consider for example the ODE

$$\dot{x}_t = \mathbb{1}_{\mathbb{R}\setminus\{0\}}(x_t), \quad x_0 = 0$$

Then $x_t = 0$ is a classical solution, but it is not a Filippov solution since we have $F_f(t,x) = \{1\}$ everywhere.

We will also need some additional properties related to the algebra of the Filippov regularisations so that we can more easily calculate specific regularisations.

Proposition 2.9 (Properties II). *(i) Let $f, g : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$. Then*

$F_{f+g}(t,x) = F_f(t,x) + F_g(t,x),$
$F_{f \cdot g}(t,x) = F_f(t,x) \cdot F_g(t,x)$

(ii) Let $g : \mathbb{R}^d \to \mathbb{R}^d$ be C^1 with $\operatorname{rank}(Dg(x)) = d$ and let $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be locally bounded. Set $h(t,x) = f(t, g(x))$. Then $F_h(t,x) = F_f(t, g(x))$

(iii) If $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies the bound $|f(t,x)| \leq h(t,x)$ for some $h : [0,T] \times \mathbb{R}^d \to \mathbb{R}^+$ non-negative and continuous, then for all elements $\alpha(t,x) \in F_f(t,x)$, $|\alpha(t,x)| \leq h(t,x)$.

Proof. (i) This follows from a more general result for families of sets. Let (A_α)

and (B_α) be families of sets parameterised by α. Then

(a) $\bigcap_\alpha A_\alpha + \bigcap_\alpha B_\alpha = \bigcap_\alpha (A_\alpha + B_\alpha)$

(b) $\bigcap_\alpha A_\alpha \cdot \bigcap_\alpha B_\alpha = \bigcap_\alpha (A_\alpha \cdot B_\alpha)$

To see this in the summation case, note that

$$\begin{aligned}\bigcap_\alpha A_\alpha + \bigcap_\alpha B_\alpha &= \{x \mid x \in A_\alpha \text{ for all } \alpha\} + \{y \mid y \in B_\alpha \text{ for all } \alpha\}\\ &= \{x + y \mid x \in A_\alpha, y \in B_\alpha \text{ for all } \alpha\}\\ &= \bigcap_\alpha (A_\alpha + B_\alpha)\end{aligned}$$

The product case follows similarly.

(ii) See Theorem 1 in [PS87] and treat t as a parameter.

(iii) Given any $y \in B_\delta(x)$, we have that $|f(t,y)| \leq h(t,y)$ for all t. Thus there exists $y_\delta \in \overline{B_\delta(x)}$ such that $|f(t,y)| \leq h(t,y_\delta)$ for all $y \in B_\delta(x)$. Taking the intersection over all $\delta > 0$ and $\tilde{f} = f$ a.e., we see from the continuity of h that any element $\beta(t,x)$ of $F_{|f|}(t,x)$ satisfies $\beta(t,x) \leq h(t,x)$.

Now given any set A, suppose that $z \in |\overline{\text{co}}\{A\}| := \{|v| \mid v \in \overline{\text{co}}\{A\}\}$. Then there exist $p, q \in A$ and $\lambda \in [0,1]$ such that

$$x = |\lambda p + (1-\lambda)q| \leq \lambda|p| + (1-\lambda)|q| \in \overline{\text{co}}\{|A|\}$$

Therefore we have that $|F_f(t,x)| \subseteq F_{|f|}(t,x)$, and so given any $\alpha(t,x) \in F_f(t,x)$, $|\alpha(t,x)| \in F_{|f|}(t,x)$. Taking $\beta(t,x) = |\alpha(t,x)|$ we see that the result is proved.

□

We now prove that the set $\mathcal{S}(F_f, x_0)$ of Filippov solutions is compact under certain assumptions on the function f. This result will be useful later on.

Proposition 2.10 (Compactness of the solution set). *Suppose that $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is measurable and satisfies, for almost all $t \in [0,T]$ and all $x \in \mathbb{R}^d$,*

$$|f(t,x)| \leq m(t)(1 + |x|)$$

where $m \in L^2([0,T])$. Let $(x^k)_{k\geq 1} \subseteq \mathcal{S}(F_f, x_0)$ be a sequence of Filippov solutions.

Then there exists a subsequence $(x^{k_p})_{p\geq 1}$ which converges uniformly to a solution $x \in \mathcal{S}(F_f, x_0)$.

Proof. Assume we have existence of solutions, otherwise the result is clear. We first show that the sequence $(x^k)_{k\geq 1}$ is uniformly bounded. From Proposition 2.9(iii) we know that any selection $\alpha(t,x)$ from $F_f(t,x)$ satisfies the bound $|\alpha(t,x)| \leq m(t)(1+|x|)$, and so we have that

$$\begin{aligned}
|x_t^k| &\leq \int_0^t |\dot{x}_s^k| \,\mathrm{d}s \\
&\leq \int_0^t m(s)(1+|x_s^k|) \,\mathrm{d}s \\
&\leq C + \int_0^t m(s)|x_s^k| \,\mathrm{d}s
\end{aligned}$$

since $m \in L^2([0,T]) \subseteq L^1([0,T])$. Therefore by Gronwall's lemma we obtain the uniform bound

$$|x_t^k| \leq C \exp\left(\int_0^T m(s)\,\mathrm{d}s\right) < \infty$$

We now show that the $(x^k)_{k\geq 1}$ is uniformly equicontinuous. We have that

$$\begin{aligned}
|x_t^k - x_s^k| &= \left|\int_0^t \dot{x}_u^k \,\mathrm{d}u - \int_0^s \dot{x}_u^k \,\mathrm{d}u\right| \\
&\leq \int_s^t |\dot{x}_u^k| \,\mathrm{d}u \\
&\leq \int_s^t m(u)(1+|x_u^k|) \,\mathrm{d}u \\
&\leq (t-s)\int_s^t m(u)^2(1+|x_u^k|)^2 \,\mathrm{d}u \\
&\leq C(t-s)
\end{aligned}$$

using Cauchy-Schwarz, the uniform boundedness of $|x^k|$ and the fact $m \in L^2([0,T])$. The result now follows from the Arzela-Ascoli theorem. □

Remark. *The result is trivially true if $\mathcal{S}(F_f, x_0)$ is empty. As we will see later however, this is never the case when the assumptions of the proposition are satisfied.*

2.1.2 Examples

1. Let $A \subseteq [-1,1]$ be a dense set such that $0 \in A$ and $|A \cap [a,b]| \in (0, b-a)$ for all $-1 \leq a < b \leq 1$. Define $f : [-1,1] \to \mathbb{R}$

$$f(x) = \begin{cases} 1 & x \in A \\ -1 & x \notin A \end{cases}$$

Then the differential equation

$$\dot{x}(t) = f(x(t)), \quad x(0) = 0$$

can be seen to have no classical solution. Indeed suppose that it did: immediately the solution would move away from zero, and then immediately after this it would move back to zero. This would repeat, suggesting that the zero function is the only possible candidate for a solution. This does not solve the equation, however.

The fact that $A \cap (a,b)$ is dense in (a,b) without full measure ensures that we have, for each x,

$$\bigcap_{\delta > 0} \overline{\mathrm{co}} \{ f(B_\delta(x) \} = [-1,1]$$

The condition on the measure of intersections of A with intervals means that the above is also true for any $\tilde{f}$ such that $\tilde{f} = f$ almost everywhere, and therefore $F_f(x) = [-1,1]$. The corresponding differential inclusion is hence given by

$$\dot{x}(t) \in [-1,1]$$

which has a very large number of solutions: given any measurable $y : [-1,1] \to \mathbb{R}$ such that $|y| \leq 1$,

$$x^y(t) := \int_0^t y(s) \, \mathrm{d}s$$

solves the inclusion.

2. Now that we have a suitable existence result we return to the control problem for (SDE), given by

$$x_t^u = x_0 + \int_0^t \sigma(s, x_s^u) \dot{u}_s \, \mathrm{d}s + \int_0^t b(s, x_s^u) \, \mathrm{d}s$$

where $\dot{u} \in L^2([0,T]; \mathbb{R}^d)$, $\sigma \in BC^1([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$ and $b \in L^\infty([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$.

Defining $f(t,x) = \sigma(t,x)\dot{u}(t) + b(t,x)$, it is clear that $f(t,\cdot)$ is bounded for all t. We also have that for all $x \in \mathbb{R}^d$,

$$\begin{aligned}|f(t,x)| &\leq \|\sigma(\cdot,\cdot)\|_{L^\infty([0,T]\times\mathbb{R}^d)}|\dot{u}_t| + \|b(\cdot,\cdot)\|_{L^{[0,T]\times\infty}(\mathbb{R}^d)} \\ &= C|\dot{u}_t| + D \\ &=: m(t)\end{aligned}$$

Since we have $|\dot{u}| \in L^2([0,T]) \subseteq L^1([0,T])$, $m \in L^1([0,T])$. The assumptions of Corollary 2.8 are therefore satisfied and a solution to the control problem exists in the Filippov sense.

Using Proposition 2.9 and the continuity of $\sigma(t,\cdot)\dot{u}_t$ we see that the regularisation $F_f(t,x)$ is given by

$$F_f(t,x) = \sigma(t,x)\dot{u}_t + F_b(t,x)$$

and F_b is compact and convex valued.

2.2 Regularisation by noise

We are not always guaranteed solutions to differential inclusions, with the main obstacle in our case being the local boundedness requirement on the right hand side. One approach to remedy this situation is to perturb the ODEs by some non-degenerate noise such as Brownian motion. In certain cases the resulting SDEs are known to have both unique strong solutions, as well as unique 'path-by-path' solutions.

To start with, we consider the simple example given by the ODE

$$x_t = 2\int_0^t \operatorname{sgn}(x_s)\,\mathrm{d}s \tag{2.1}$$

Then three solutions to this equation are given by $x_t = 0$ and $x_t = \pm t^2$. Moreover, the non-zero solutions can be connected to the zero solution via

$$x_t = \begin{cases} 0 & t \leq C \\ \pm(t-C)^2 & t > C \end{cases}$$

resulting in an uncountable family of solutions.

Now consider the stochastic perturbation of (2.1) given by

$$x_t^\varepsilon = 2\int_0^t \operatorname{sgn}(x_s^\varepsilon)\,\mathrm{d}s + \varepsilon W_t \tag{2.2}$$

for some $\varepsilon > 0$, where W is a Brownian motion. Since $b(x) := 2\operatorname{sgn}(x)$ is bounded and the SDE is elliptic, existence of a unique strong solution follows from a theorem of Veretennikov [Ver80]. The question is how this strong solution compares to the solutions of the deterministic equation (2.1). It makes sense to look for some kind of convergence of x_t^ε as $\varepsilon \downarrow 0$.

This limit has been studied previously, see for example [Tre13]. If μ^ε denotes the law of the solution to (2.2), then it is known that the weak limit of μ^ε as $\varepsilon \downarrow 0$ is given by

$$\mu := \frac{1}{2}\delta_{t^2} + \frac{1}{2}\delta_{-t^2}$$

That is, the limiting process chooses one of the extremal paths $\pm t^2$ with equal probability. In particular, the limiting measure assigns full mass to the space of solutions to (2.1).

We wish to apply this technique to our control ODEs, which unlike (2.1) may not have solutions in a classical sense. In this case we would aim to describe the support of the law of (SDE) in terms of the limits of the laws of the stochastically perturbed ODEs rather than the solutions to the ODEs themselves. The controlled problems we are interested in are of the form

$$x_t^u = x_0 + \int_0^t \sigma(s, x_s^u)\dot{u}_s\,\mathrm{d}s + \int_0^t b(s, x_s^u)\,\mathrm{d}s \tag{2.3}$$

where $u \in H := L_0^{2,1}([0,T];\mathbb{R}^d)$, the Cameron-Martin space of the Wiener measure given by

$$L_0^{2,1}([0,T];\mathbb{R}^d) := \left\{ u \in C_0([0,T];\mathbb{R}^d) \,\middle|\, u_t = \int_0^t \varphi_s\,\mathrm{d}s \text{ for some } \varphi \in L^2([0,T];\mathbb{R}^d) \right\}$$

In the case that b is not locally bounded, we do not have an existence theorem even for Filippov solutions. We do however have existence theorems for the stochastic perturbation of (2.3) in certain cases.

As before, let W be a Brownian motion and let $\varepsilon > 0$. Consider the SDE

$$x_t^{u,\varepsilon} = x_0 + \int_0^t \sigma(s, x_s^{u,\varepsilon})\dot{u}_s\,\mathrm{d}s + \int_0^t b(s, x_s^{u,\varepsilon})\,\mathrm{d}s + \varepsilon W_t \tag{2.4}$$

Existence and uniqueness of solutions to this equation in different senses have been shown, depending upon which space $f := \sigma\dot{u} + b$ lies in:

(i) (Veretennikov, [Ver80]) f bounded measurable implies existence and uniqueness of a strong solution.

(ii) (Davie, [Dav07] [Dav10] [Fla11]) f bounded measurable almost-surely implies existence and uniqueness of path-by-path solutions. That is, if a sample path of W is fixed, then almost surely the resulting (random) ODE has a unique solution.

(iii) (Krylov-Röckner, [KR05] [FF10]) $f \in L^q_p(T) := L^q([0,T]; L^p(\mathbb{R}^d; \mathbb{R}^d))$ with $d/p + 2/q < 1$ implies existence and uniqueness of a strong solution.

2.2.1 The zero-noise limit

Although we have existence of solutions to (2.4) in the cases described above, much less is known about the limit of these solutions as the intensity of the noise ε tends to zero. This problem was originally studied by Bafico and Baldi in their 1982 paper [BB82]. Their result concerned the one-dimensional case when f is continuous. The problem has since been studied in the multi-dimensional case with fewer restrictions on f, for example Buckdahn et al. studied the case when f has at most linear growth. The following theorem is taken from [BOQ09]:

Theorem 2.11. *Suppose that $f : \mathbb{R}^d \to \mathbb{R}^d$ is Lebesgue measurable and satisfies*

$$|f(x)| \leq M(1 + |x|), \quad \forall x \in \mathbb{R}^d$$

For any $\varepsilon > 0$, let x^ε be the solution to

$$\mathrm{d}x^\varepsilon_t = f(x^\varepsilon_t)\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t, \quad x_0 = x \tag{2.5}$$

Then there exists $\varepsilon_n \to 0$ such that x^ε converges in law, as $\varepsilon_n \to 0$, to some x which belongs almost surely to the set of Filippov solutions[2] *to*

$$\dot{x}_t = f(x_t), \quad x_0 = x$$

Furthermore, any cluster point of x^ε is also almost surely in the set of Filippov solutions.

[2] Theorem 10.1 in [Dei92] tells us that this set is non-empty.

Let μ^ε denote the law of the solution to (3.2). Then the theorem tells us that, given any sequence $\varepsilon_n \to 0$ such that $(\mu^{\varepsilon_n})_{n\geq 1}$ converges weakly, the limiting measure μ will satisfy $\mu(\mathcal{S}(F_f, x)) = 1$. Unfortunately it is possible that different subsequences will result in different limiting measures, even though all such measures will assign full mass to the set of Filippov solutions.

A more specific result is proved in [Zha12]. In this paper, under additional assumptions on the function f, the author shows that the limiting process in fact belongs almost surely to the subset of Filippov solutions that leave the starting point immediately. This was the case in the example discussed above, where only the extremal solutions $\pm t^2$ were reachable by the limiting process.

Chapter 3

Elliptic case

We first concentrate on the case when the diffusion matrix is uniformly elliptic. In this case the support of the law of the diffusion is given by the whole space $C_{x_0}([0,T];\mathbb{R}^d)$, as shown in Proposition 1.3. We therefore only need to show that the space of solutions to the control problems is dense in $C_{x_0}([0,T];\mathbb{R}^d)$.

3.1 Bounded measurable drift

Assume without loss of generality that $x_0 = 0$ and let $H = L_0^{2,1}([0,T];\mathbb{R}^d)$. Choose $u \in H$ and consider the ODE

$$x_t^u = \int_0^t \sigma(s, x_s^u)\dot{u}_t \,\mathrm{d}u + \int_0^t b(s, x_s^u)\,\mathrm{d}s$$

where $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is BC^1 and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded measurable. We interpret this in the Filippov sense, so that

$$x_t^u \in \int_0^t \sigma(s, x_s^u)\dot{u}_t \,\mathrm{d}u + \int_0^t F_b(s, x_s^u)\,\mathrm{d}s \tag{3.1}$$

where as usual F_b denotes the Filippov regularisation of b. In this case the boundedness of b means that F_b is compact valued.

Denote by $\mathcal{S}(u)$ the set of solutions to (3.1). We aim to show that

$$\overline{\bigcup_{u \in H} \mathcal{S}(u)} = C_0([0,T];\mathbb{R}^d)$$

where the closure is taken in the uniform topology.

We first note that each solution is of the form

$$x_t^u = u_t + \int_0^t y_s^u \, \mathrm{d}s$$

where y^u is a selection from $F_b(\cdot, x_\cdot^u)$, and so in particular is bounded measurable. Now given any family of uniformly bounded measurable functions $Z = (z^u)_{u \in H}$, we define the set

$$A_Z := \left\{ u_t + \int_0^t z_s^u \, \mathrm{d}s \,\middle|\, u \in H \right\}$$

Since z^u is bounded, it is clear that $A_Z \subseteq H$ for all Z, and so $\overline{A_Z} \subseteq C_0([0,T];\mathbb{R}^d)$. As a consequence, we have

$$\overline{\bigcup_{u \in H} \mathcal{S}(u)} \subseteq C_0([0,T];\mathbb{R}^d)$$

by choosing the appropriate Z.

We now need to show that A_Z is dense. It suffices to show that zero is a limit point of A_Z. Moreover, we need only to consider the 'worst case' for Z, specifically if b is uniformly bounded by K, a choice such that $|u_t - z_t^u| \geq K$ for all $u \in L^2([0,T];\mathbb{R}^d)$ and all $t \in [0,T]$. In one dimension, this may be given by

$$z_t^u = \begin{cases} -K & \dot{u}_t \leq 0 \\ +K & \dot{u}_t > 0 \end{cases}$$

In higher dimensions, z_t^u can take the value $(\pm K, \ldots, \pm K)/\sqrt{d}$ in the same orthant as $\dot{u}_t$. We have that, for example,

$$\begin{aligned} \left| \int_0^t \dot{u}_s \, \mathrm{d}s + \int_0^t z_s^u \, \mathrm{d}s \right| &= \left| \int_{\{\dot{u}_s \leq 0\} \cap [0,t]} (\dot{u}_s - K) \, \mathrm{d}s + \int_{\{\dot{u}_s > 0\} \cap [0,t]} (\dot{u}_s + K) \, \mathrm{d}s \right| \\ &= \left| \int_0^t \dot{u}_s \, \mathrm{d}s - K |\{\dot{u}_s \leq 0\} \cap [0,t]| + K |\{\dot{u}_s > 0\} \cap [0,t]| \right| \end{aligned}$$

which can be made arbitrarily small in the uniform norm by the appropriate choice of $\dot{u}$.

Alternatively, we could follow the same approach as described in the sketch proof earlier, using a Wong-Zakai approximation. The approximation still holds with our regularity on the coefficients (at least in the continuous case), but we defer the proof until the next chapter, see Corollary 4.3.

It follows that

$$C_0([0,T];\mathbb{R}^d) \subseteq \overline{\bigcup_{u\in H} \mathcal{S}(u)}$$

and so we have the desired equality.

3.1.1 Regularisation with noise

We could instead perturb the control ODEs by noise, yielding

$$\mathrm{d}x_t^{u,\varepsilon} = \sigma(t, x_t^{u,\varepsilon})\dot{u}_t \,\mathrm{d}t + b(t, x_t^{u,\varepsilon})\,\mathrm{d}t + \varepsilon \,\mathrm{d}B_t$$

As discussed in Chapter 2 there exists a unique strong solution to this SDE. Furthermore, letting $\mu^{u,\varepsilon}$ denote its law and given any $\varepsilon_n \downarrow 0$ such that μ^{u,ε_n} converges weakly, the limiting measure μ^u satisfies $\mu^u(\mathcal{S}(u)) = 1$. Note that this limiting measure isn't necessarily unique across all subsequences.

We now consider the set

$$\mathcal{T}(u) := \left\{ y \in C_0([0,T];\mathbb{R}^d) \,\middle|\, \exists \varepsilon_n \downarrow 0 \text{ s.t. } \liminf_{\varepsilon_n \downarrow 0} \mu^{u,\varepsilon_n}(B_\delta(y)) > 0 \text{ for all } \delta > 0 \right\}$$

We show that this contains at least one element of $\mathcal{S}(u)$ for each u. Given any $y \in \mathcal{S}(u)$ and any $\delta > 0$, we have by the lower semi-continuity of weak convergence of measures on open sets that

$$\liminf_{\varepsilon_n \downarrow 0} \mu^{u,\varepsilon_n}(B_\delta(y)) \geq \mu^u(B_\delta(y))$$

First suppose that $|\mathcal{S}(u)| < \infty$. Then since $\mu^u(\mathcal{S}(u)) = 1$ it follows that there exists $y \in \mathcal{S}(u)$ such that $\mu^u(\{y\}) > 0$, and so the same holds for all balls containing that y. It follows that $y \in \mathcal{T}(u)$.

If instead $|\mathcal{S}(u)| = \infty$, we use the compactness of $\mathcal{S}(u)$ and cover it with $(B_\delta(y))_{y\in\mathcal{S}(u)}$. We take a finite subcover $(B_\delta(y_n))_{n=1}^k$ and note that

$$1 = \mu^u(\mathcal{S}(u)) \leq \sum_{n=1}^{k} \mu^u(B_\delta(y_n))$$

At least one of the terms in the sum must be positive, and so we have $y_n \in \mathcal{T}(u)$ for some n.

Now using the same argument as in the previous section, we can conclude that

$$\overline{\bigcup_{u \in H} \mathcal{T}(u)} = C_0([0,T];\mathbb{R}^d) = \mathrm{supp}(\mu)$$

This is the result we aim to obtain for the case of $L_p^q(T)$ drift, where we don't have existence of Filippov solutions.

We can summarise our results for bounded measurable drift in the following proposition.

Proposition 3.1. *Suppose that $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded measurable, and let $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be continuous, bounded and such that $\sigma\sigma^T$ is uniformly positive definite. Let μ denote the law of a solution to the SDE*

$$\mathrm{d}x_t = \sigma(t, x_t)\,\mathrm{d}W_t + b(t, x_t)\,\mathrm{d}t$$

Then retaining the same notation as before, we have

$$\overline{\bigcup_{u \in H} \mathcal{S}(u)} = \overline{\bigcup_{u \in H} \mathcal{T}(u)} = C_0([0,T];\mathbb{R}^d) = \mathrm{supp}(\mu)$$

3.2 Drift with linear growth

We can in fact generalise this result slightly to the case where the drift has at most linear growth in space and an L^1 dependence on time, thanks to an existence result for inclusions by Tolstonogov [Tol88]:

Theorem 3.2. *Let the mapping $F : [0,T] \times \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \emptyset$ be measurable and satisfy*

(i) for almost all $t \in [0,T]$, for each $x \in \mathbb{R}^d$ the mapping $F(t, \cdot)$ has a closed graph at the point x and the set $F(t,x)$ is convex; and

(ii) there exists a non-negative $m \in L^1([0,T])$ such that

$$F(t,x) \cap \{y \in \mathbb{R}^d \mid |y| \leq m(t)(1 + |x|)\} \neq \emptyset$$

almost everywhere on $[0,T]$ for any x.

Then there exists a solution on $[0,T]$ to the differential inclusion $\dot{x}_t \in F(t, x_t)$ started from any $x_0 \in \mathbb{R}^d$.

Any upper semi-continuous map with closed values has a closed graph, see for example Proposition 2.1 in [Smi00]. It is therefore not hard to see that the conditions on F required in the theorem are satisfied by the Filippov regularisation F_f of a measurable function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ which satisfies the bound $|f(t,x)| \leq m(t)(1+|x|)$ for some $m \in L^2([0,T]) \subseteq L^1([0,T])$. This condition is similar to that in Theorem 2.11, except now there is an explicit time-dependence. Nonetheless, the conclusion of that theorem remains true in this case:

Theorem 2.11′. *Suppose that $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is measurable and satisfies, for almost all $t \in [0,T]$ and all $x \in \mathbb{R}^d$,*

$$|f(t,x)| \leq m(t)(1+|x|)$$

where $m \in L^2([0,T])$. For any $\varepsilon > 0$, let x^ε be the solution to

$$\mathrm{d}x_t^\varepsilon = f(t, x_t^\varepsilon)\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t, \quad x_0 = x \tag{3.2}$$

Then there exists $\varepsilon_n \to 0$ such that x^ε converges in law, as $\varepsilon_n \to 0$, to some x which belongs almost surely to the set of Filippov solutions to

$$\dot{x}_t = f(t, x_t), \quad x_0 = x$$

Furthermore, any cluster point of x^ε is also almost surely in the set of Filippov solutions.

Proof. The previous theorem gives us the existence of the Filippov solutions. Proposition 2 in [BOQ09] remains true when we add time-dependence: it is Proposition 2.7 here. Most other parts of the proof of Theorem 4 in [BOQ09] remain virtually identical in light of this. The only small change required is in showing that the processes $\widetilde{Y}'_{\varepsilon_n}$ as defined in their paper remain uniformly bounded in $L^2([0,T] \times \Omega; \mathbb{R}^d)$. The processes satisfy the bound

$$|\widetilde{Y}'_{\varepsilon_n}(t)| \leq m(t)(1 + |\widetilde{X}_{\varepsilon_n}(t)|)$$

where $\widetilde{X}_{\varepsilon_n}(t)$ is a weak solution to the SDE

$$\widetilde{X}_{\varepsilon_n}(t) = x + \int_0^t f(s, \widetilde{X}_{\varepsilon_n}(s))\,\mathrm{d}s + \varepsilon_n \widetilde{W}_{\varepsilon_n}(t)$$

for some Brownian motion $\widetilde{W}_{\varepsilon_n}$. Therefore, using the bound on f and the fact that

$(a+b)^2 \leq 2(a^2+b^2)$,

$$|\widetilde{X}_{\varepsilon_n}(t)|^2 \leq C\left(|x|^2 + \varepsilon_n^2|\widetilde{W}_{\varepsilon_n}(t)|^2 + \int_0^t m(s)^2(1+|\widetilde{X}_{\varepsilon_n}(s)|^2)\,\mathrm{d}s\right)$$

and so

$$\mathbb{E}\left(\sup_{s\leq t}|\widetilde{X}_{\varepsilon_n}(s)|^2\right) \leq C\bigg(|x|^2 + \varepsilon_n^2 \sup_{s\leq T}|\widetilde{W}_{\varepsilon_n}(s)|^2 + \int_0^T m(s)^2\,\mathrm{d}s \\ + \int_0^t m(s)^2\mathbb{E}\left(\sup_{r\leq s}|\widetilde{X}_{\varepsilon_n}(r)|^2\right)\mathrm{d}s\bigg)$$

since m^2 is non-negative. The first three terms on the right hand side can be bounded independently of n, and so using Gronwall's lemma we deduce that

$$\mathbb{E}\left(\sup_{s\leq T}|\widetilde{X}_{\varepsilon_n}(s)|^2\right) \leq C\exp\left(\int_0^T m(s)^2\,\mathrm{d}s\right) < \infty$$

independently of n. It follows that

$$\begin{aligned}\mathbb{E}\left(\int_0^T |\widetilde{Y}'_{\varepsilon_n}(t)|^2\,\mathrm{d}t\right) &\leq C\int_0^T m(t)^2\left(1+\mathbb{E}\left(\sup_{s\leq T}|\widetilde{X}_{\varepsilon_n}(s)|^2\right)\right)\mathrm{d}s \\ &\leq C\int_0^T m(t)^2\,\mathrm{d}s \\ &< \infty\end{aligned}$$

Thus the processes $\widetilde{Y}'_{\varepsilon_n}(t)$ are uniformly bounded in $L^2([0,T]\times\Omega;\mathbb{R}^d)$. We are therefore still able to deduce the weak $H^1([0,T]\times\Omega;\mathbb{R}^d)$ convergence of the $\widetilde{Y}_{\varepsilon_n}$, and the rest of the proof proceeds in the same way. □

Using the above theorem we are able to characterise the supports of the laws of some more diffusion processes:

Proposition 3.3. *Suppose that $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ is measurable and satisfies, for almost all $t\in[0,T]$ and all $x\in\mathbb{R}^d$,*

$$|b(t,x)| \leq m(t)(1+|x|)$$

where $m\in L^2([0,T])$. Let $\sigma:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ be continuous, bounded and such that $\sigma\sigma^T$ is uniformly positive definite. Let μ denote the law of a solution to *the SDE*

$$\mathrm{d}x_t = \sigma(t, x_t)\,\mathrm{d}W_t + b(t, x_t)\,\mathrm{d}t$$

Then retaining the same notation as before, we have

$$\overline{\bigcup_{u\in H} \mathcal{S}(u)} = \overline{\bigcup_{u\in H} \mathcal{T}(u)} = C_0([0,T];\mathbb{R}^d) = \mathrm{supp}(\mu)$$

Proof. The fact that the support of μ is all of $C_0([0,T];\mathbb{R}^d)$ follows from Proposition 1.3.

Let $u \in H$ and define $f(t,x) = \sigma(t,x)\dot{u}_t + b(t,x)$. Then it's easy to see that there exists $m \in L^1([0,T])$ such that $|f(t,x)| \le m(t)(1+|x|)$ for almost all $t \in [0,T]$ and all $x \in \mathbb{R}^d$. Therefore by Theorem 3.2 there exists a solution x^u to the inclusion $\dot{x}^u_t \in F_f(t, x_t)$, and from the algebra of the Filippov regularisations it can be seen to satisfy

$$x^u_t = x_0 + \int_0^t \sigma(s, x^u_s)\dot{u}_s\,\mathrm{d}s + \int_0^t y^u_s\,\mathrm{d}s$$

for a selection y^u from $F_b(\cdot, x^u_\cdot)$. Since we have $|b(t,x)| \le m(t)(1+|x|)$, it follows from Proposition 2.9(iii) that y^u satisfies

$$|y^u_t| \le m(t)(1 + |x^u_t|)$$

Since x^h_t is continuous and we're working on a compact interval, it is bounded. From this we can see that $y^h_t \in L^2([0,T])$. Then similarly to the bounded measurable case we see that

$$\overline{\bigcup_{u\in H} \mathcal{S}(u)} = \overline{\bigcup_{u\in H} \mathcal{T}(u)} = C_0([0,T];\mathbb{R}^d)$$

∎

3.3 $L^q_p(T)$ drift

The case when we have $L^q_p(T)$ drift is more complicated since we don't have any existence theorems for the control ODEs or inclusions: we are forced to regularise by noise. However as remarked earlier, little is known about the zero noise limit in this case.

In [Fla08], the author explains that there is some progress in this direction. Assume that $f : \mathbb{R}^d \to \mathbb{R}^d$ has no time dependence and lies in $L^p(\mathbb{R}^d;\mathbb{R}^d)$. Assume further that f is continuous and satisfies a growth condition such as $\langle f(x), x\rangle \le C(|x|^2+1)$ for all $x \in \mathbb{R}^d$. Then the family of laws $(\mu^\varepsilon)_{\varepsilon>0}$ of solutions to the SDEs

$$x_t^\varepsilon = f(x_t^\varepsilon)\,\mathrm{d}t + \varepsilon\,\mathrm{d}W_t, \quad x_0 \in \mathbb{R}$$

is tight and each limit point assigns full mass to the space of solutions to the corresponding ODE.

The conditions on f are particularly restrictive, ruling out the possibility of singular drift. However, the note [Fla08] does not make use of the fact that there exists a unique strong solution in the more general non-autonomous case $f \in L_p^q(T)$ [KR05] [FF10], a result which may be of use.

Suppose that $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ lies in $L_p^q(T)$ with $d/p + 2/q < 1$. We will need some decay assumption on $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d\otimes\mathbb{R}^d$: we assume it is bounded and has compact support K (or at least strong decay at infinity). Assume also that $\dot{u} \in L^2([0,T];\mathbb{R}^d)$. Then we can use Hölder's inequality to deduce that $f := \sigma\dot{u} + b$ lies in $L_p^q(T)$ also:

$$\begin{aligned}
\|f\|_{L_p^q(T)}^q &\le \int_0^T \left(\int_K |\sigma(t,x)\dot{u}(t)|^p\,\mathrm{d}x\right)^{q/p} \mathrm{d}t + \|b\|_{L_p^q(T)} \\
&\le C\int_0^T |\dot{u}(t)|^q\,\mathrm{d}t + \|b\|_{L_p^q(T)} \\
&\le C\|1\|_{L^{q'}(T)}\|\dot{u}\|_{L^2(T)} + \|b\|_{L_p^q(T)}
\end{aligned}$$

where $q' = q/(q-2)$. This is well-defined since by assumption $q > 2$. Therefore under these assumptions we have existence of a unique strong solution to the perturbed control ODE

$$x_t^{u,\varepsilon} = x_0 + \int_0^t \sigma(s, x_s^{u,\varepsilon})\dot{u}_s\,\mathrm{d}s + \int_0^t b(s, x_s^{u,\varepsilon})\,\mathrm{d}s + \varepsilon W_t$$

We aim to investigate the set $\mathcal{T}(u)$ corresponding to this SDE, defined earlier. However due to the lack of Filippov solutions, we must investigate this set directly.

Chapter 4

Degenerate case

4.1 Bounded measurable drift

We assume throughout this section that $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ lies in BC^1 and that $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded measurable.

4.1.1 A Wong-Zakai type result

We wish to prove that if we replace the Brownian motion in (SDE) by an approximation, we get convergence of the approximate solutions to the true solution in some sense. Under higher regularity on the drift, the result is known for weak convergence [SV72], L^2 convergence [IW81] and convergence in probability [GP90].

Let W be a Brownian motion on $\mathbb{R}$ and let $W^{(n)}$ be a polygonal approximation to it such that

$$\dot{W}^{(n)}(t) = 2^n(W(t_n^+) - W(t_n))$$

where

$$t_n := \frac{\lfloor 2^n t \rfloor}{2^n}, \quad t_n^+ := \frac{\lfloor 2^n t \rfloor + 1}{2^n}$$

Then $\dot{W}^{(n)} \in L^2([0,T])$ for all $T > 0$. Set $f_n(t,x) := \sigma(t,x)\dot{W}^{(n)}(t) + b(t,x)$, so that we have existence of solutions to the differential inclusion

$$x_t^n \in x_0 + \int_0^t F_{f_n}(s, x_s^n)\,\mathrm{d}s$$

We are interested in how these solutions relate to those of the Stratonovich SDE

$$x_t = x_0 + \int_0^t \sigma(s, x_s) \circ \mathrm{d}W_s + \int_0^t b(s, x_s)\,\mathrm{d}s \tag{SDE}$$

Since we do not have an existence theorem for this SDE in the general degenerate case, we simply assume existence. We first need to see what $F_{f_n}(t,x)$ looks like. From the continuity of σ in space and the algebra of Filippov regularisations, we have

$$F_{f_n}(t,x) = \sigma(t,x)\dot{W}^{(n)}(t) + F_b(t,x)$$

The only term of concern is hence the drift b. We first concentrate on the one-dimensional case because it is more transparent what is going on. Define

$$b^-(t,x) := \lim_{\delta \downarrow 0} \inf_{y \in B_\delta(x)} b(t,y), \quad b^+(t,x) := \lim_{\delta \downarrow 0} \sup_{y \in B_\delta(x)} b(t,y)$$

Then it's clear that almost everywhere we have the equality

$$F_b(t,x) = \bigcap_{\tilde{b}=b \text{ a.e.}} \bigcap_{\delta>0} \overline{\mathrm{co}} \left\{ \tilde{b}(B_\delta(t,x) \right\} = [b^-(t,x), b^+(t,x)] \tag{4.1}$$

and this interval is finite due to the boundedness of b. Therefore we have that

$$F_{f_n}(t,x) = \left[\sigma(t,x)\dot{W}^{(n)}(t) + b^-(t,x), \sigma(t,x)\dot{W}^{(n)}(t) + b^+(t,x) \right]$$

and so

$$x_t^n \in x_0 + \int_0^t \left[\sigma(s,x_s^n)\dot{W}^{(n)}(s) + b^-(s,x_s^n), \sigma(s,x_s^n)\dot{W}^{(n)}(s) + b^+(s,x_s^n) \right] \mathrm{d}s$$

Equivalently, by the absolute continuity of x^n, we have that for all s,

$$\sigma(s,x_s^n)\dot{W}^{(n)}(s) + b^-(s,x_s^n) \le \dot{x}_s^n \le \sigma(s,x_s^n)\dot{W}^{(n)}(s) + b^+(s,x_s^n)$$

In higher dimensions we don't have a nice analogue of (4.1), but we do at least know that it is a compact convex set due to the boundedness of b. We can therefore write

$$\dot{x}_s^n \in \sigma(s,x_s^n)\dot{W}^{(n)}(s) + F_b(s,x_s^n)$$

where where addition is done pointwise, and F_b is known to be a compact and convex set valued function.

We wish to show that each solution x^n converges weakly to a solution x. The lack of uniqueness both for the approximations and the original SDE is an issue however: we need to ensure that we pick the correct sequence of solutions from the sets of solutions to each approximation, otherwise we may have no hope for convergence.

We first show that the family of laws of x^n is relatively weakly compact. For convenience we will use the following notation:

$$\begin{aligned}
\alpha^{(n)}(t) &= \sigma(t, x^n(t_n))\dot{W}^{(n)}(t) \\
(\sigma'\sigma)_i^{j,l}(t,x,s,y) &= \frac{\partial \sigma^{ij}}{\partial x_k}(t,x)\sigma^{kl}(s,y) \\
(\sigma' b)_i^{j}(t,x,s,y) &= \frac{\partial \sigma^{ij}}{\partial x_k}(t,x)b_k(s,y)
\end{aligned}$$

Lemma 4.1. *Let $(x^n)_{n\geq 1}$, $x^n : \mathbb{R}\times\Omega \to \mathbb{R}^d$, be a sequence of processes such that $x^n(\cdot,\omega) \in \mathcal{S}(W^{(n)}(\omega))$ for each $n \geq 1$ and $\omega \in \Omega$. Then the family of laws $(\mu_n)_{n\geq 1}$ of $(x^n)_{n\geq 1}$ is relatively weakly compact.*

Proof. We follow the idea of the proof in [SV72]. It suffices to prove that

$$\sup_n \mathbb{E}^{\mu_n}\left(|\omega(t)-\omega(s)|^4\right) \leq C_T|t-s|^2, \quad 0 \leq s \leq t \leq T, \quad T > 0$$

where $t \mapsto \omega(t)$ is the evaluation map on $C_0([0,T];\mathbb{R}^d)$. First note that[1]

$$\begin{aligned}
\mathbb{E}^{\mu_n}|\omega(t)-\omega(s)|^4 &= \mathbb{E}^W\left|\int_s^t \dot{x}_t^n \,\mathrm{d}r\right|^4 \\
&\leq C\left(\mathbb{E}^W\left|\int_s^t \sigma(r,x_r^n)\dot{W}^{(n)}(r)\,\mathrm{d}r\right|^4 + \mathbb{E}^W\left|\int_s^t |F_b(r,x_r^n)|\,\mathrm{d}r\right|^4\right) \\
&\leq C\mathbb{E}^W\left|\int_s^t \sigma(r,x_r^n)\dot{W}^{(n)}(r)\,\mathrm{d}r\right|^4 + C|t-s|^4
\end{aligned}$$

so we just need to deal with the first term. By the Fundamental Theorem of Calculus

[1]For a set $A \subseteq \mathbb{R}^d$, we define $|A| := \sup_{x\in A}|x|$

we have

$$
\begin{aligned}
\sigma^{ij}(t,x_t^n) &= \sigma^{ij}(t,x_{t_n}^n) + \int_{t_n}^t \left\langle \nabla_x \sigma^{ij}(t,x_w^n), \dot{x}_w^n \right\rangle \mathrm{d}w \\
&\in \sigma^{ij}(t,x_{t_n}^n) + \int_{t_n}^t \left\langle \nabla_x \sigma^{ij}(t,x_w^n), \sigma(w,x_w^n)\dot{W}_w^{(n)} \right\rangle \mathrm{d}w \\
&\quad + \int_{t_n}^t \left\langle \nabla_x \sigma^{ij}(t,x_w^n), F_b(w,x_w^n) \right\rangle \mathrm{d}w
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\left[\int_s^t \sigma(r,x_r^n)\dot{W}^{(n)}(r)\,\mathrm{d}r\right]_i \in \int_s^t \alpha_i^n(u)\,\mathrm{d}u
&+ \int_s^t \int_{u_n}^u \frac{\partial \sigma^{ij}}{\partial x_k}(u,x_w^n)\sigma^{kl}(w,x_w^n)\dot{W}_l^{(n)}(w)\dot{W}_j^{(n)}(u)\,\mathrm{d}w\,\mathrm{d}u \\
&+ \int_s^t \int_{u_n}^u \frac{\partial \sigma^{ij}}{\partial x_k}(u,x_w^n)(F_b)_k(w,x_w^n)\dot{W}_u^{(n)}\,\mathrm{d}w\,\mathrm{d}u
\end{aligned}
$$

and so

$$
\begin{aligned}
\int_s^t \sigma(r,x_r^n)\dot{W}^{(n)}(r)\,\mathrm{d}r \in \int_s^t \alpha^n(u)\,\mathrm{d}u
&+ \int_s^t \int_{u_n}^u (\sigma'\sigma)^{j,l}(u,x_w^n,w,x_w^n)\dot{W}_l^{(n)}(w)\dot{W}_j^{(n)}(u)\,\mathrm{d}w\,\mathrm{d}u \\
&+ \int_s^t \int_{u_n}^u (\sigma' K^b)^j(u,x_w^n,w,x_w^n)\dot{W}_k^{(n)}(u)\,\mathrm{d}w\,\mathrm{d}u \\
&=: I_1 + I_2 + I_3
\end{aligned}
$$

Now

$$
\begin{aligned}
\mathbb{E}^W|I_1|^4 &= \mathbb{E}^W \left| \int_s^t \sigma(u,x_{u_n}^n)\dot{W}^{(n)}(u)\,\mathrm{d}u \right|^4 \\
&= \mathbb{E}^W \left| \int_{s_n}^{t_n} \sigma^n(u)\,\mathrm{d}W(u) \right|^4 \\
&\le C(t-s)^2
\end{aligned}
$$

where

$$
\sigma^{(n)}(u) = 2^n \int_{u_n \vee s}^{u_n^+ \wedge t} \sigma(v,x_{u_n}^n)\,\mathrm{d}v
$$

For the second term we use Jensen's inequality and the fact that $\sigma \in BC^1$:

$$\mathbb{E}^W|I_2|^4 \le (t-s)^3\mathbb{E}^W\left[\int_s^t \left|\int_{u_n}^u (\sigma'\sigma)^{j,l}(u,x_w^n,w,x_w^n)\dot{W}_l^{(n)}(w)\dot{W}_j^{(n)}(u)\,\mathrm{d}w\right|^4 \mathrm{d}u\right]$$
$$\le C(t-s)^3\int_s^t (u-u_n)^3\int_{u_n}^u \mathbb{E}^W\left|\sum_{l,l'=1}^d \dot{W}_l^{(n)}(w)\dot{W}_j^{(n)}(u)\right|^4 \mathrm{d}w\,\mathrm{d}u$$

Since $w \in [u_n, u]$ it follows that $\dot{W}^{(n)}(w) = \dot{W}^{(n)}(u)$, and so by the distributional properties of Brownian motion,

$$\mathbb{E}^W|I_2|^4 \le C(t-s)^3 \cdot 2^{8n}\int_s^t (u-u_n)^4\mathbb{E}^W|W(u_n^+)-W(u_n)|^8\,\mathrm{d}u$$
$$= C(t-s)^3 \cdot \sum_{k=\lfloor 2^n s\rfloor}^{\lfloor 2^n t\rfloor} 2^{4n}\int_{k/2^n}^{(k+1)/2^n}\left(u-\frac{k}{2^n}\right)^4\mathrm{d}u$$
$$\le C(t-s)^3\cdot 2^{4n}\cdot 2^n \cdot 2^{-5n} = C(t-s)^3$$

The third term is dealt with similarly to the second:

$$\mathbb{E}^W|I_3|^4 \le C(t-s)^3 \cdot 2^{4n}\int_s^t (u-u_n)^4\mathbb{E}^W|W(u_n^+)-W(u_n)|^4\,\mathrm{d}u$$
$$= C(t-s)^3 \cdot \sum_{k=\lfloor 2^n s\rfloor}^{\lfloor 2^n t\rfloor} 2^{2n}\int_{k/2^n}^{(k+1)/2^n}\left(u-\frac{k}{2^n}\right)^4\mathrm{d}u$$
$$= C(t-s)^3\cdot 2^{-2n} \le C(t-s)^3$$

Combining everything, we see that we have the uniform bound

$$\mathbb{E}^{\mu_n}|x(t)-x(s)|^4 \le C_T|t-s|^2$$

and so the result is proved. □

We now need to check that the weak limit of (a subsequence of) $(\mu_n)_{n\ge1}$ converges to what we want, namely, the law μ of a solution to the original SDE.

Proposition 4.2. *Let x be a solution to (SDE) with law μ. Then there exists a sequence of processes $(y^n)_{n\ge1}$, $y^n : \mathbb{R}\times\Omega \to \mathbb{R}^d$, with $y^n(\cdot,\omega) \in \mathcal{S}(W^{(n)}(\omega))$ for all $n \ge 1$ and $\omega \in \Omega$, such that their laws μ_n converge weakly to μ.*

Proof. From the previous lemma we know that the family of laws of any processes satisfying the above is relatively weakly compact, and so has a weakly convergent subsequence. There are two issues we must deal with: firstly we must be able to choose the 'correct' solution to each control ODE, and secondly we must be able to show that the limit of this subsequence is indeed what we want it to be. For the latter we will use a martingale argument, similar to [SV72].

First of all, note that for each path ω and each $n \geq 1$, there exists a solution to the differential inclusion

$$\dot{y}_t^{(n)}(\omega) \in \sigma(t, y_t^{(n)}(\omega))\dot{W}_t^{(n)}(\omega) + F_b(t, y_t^{(n)}(\omega))$$

and each solution satisfies

$$\dot{y}_t^{(n)}(\omega) = \sigma(t, y_t^{(n)}(\omega))\dot{W}_t^{(n)}(\omega) + b_t^{(n)}(\omega)$$

for some $b^{(n)}(\omega) : [0, T] \to \mathbb{R}^d$ with $b_t^{(n)}(\omega) \in F_b(t, y_t^{(n)}(\omega))$ for all t.

In order to proceed it seems necessary to assert that b is continuous almost everywhere in the spatial component, though it may be possible to relax this at a later date. In this case we have that $F_b(t, x) = \{b(t, x)\}$ almost everywhere, and so $b_\cdot^{(n)}(\omega) = b(\cdot, y_\cdot^n(\omega))$ almost surely.

Now as in [SV72] we aim to show that

$$\mathbb{E}^\mu \left[F \cdot (f(\omega(t)) - f(\omega(s)))\right] = \mathbb{E}^\mu \left[F \cdot \int_s^t L_u f(\omega(u))\, \mathrm{d}u\right]$$

for all $f \in C_0^\infty(\mathbb{R}^d)$, $0 \leq s < t$ and bounded $\mathcal{F}_s$ measurable $F : \Omega \to \mathbb{R}$, where L_t is as defined in Theorem 1.5. It suffices to assume that F is continuous, and so by the Fundamental Theorem of Calculus we have

$$\begin{aligned}
\mathbb{E}^{\mu_n} \left[F \cdot (f(\omega(t)) - f(\omega(s)))\right] &= \mathbb{E}^{\mu_n} \left[F \cdot \int_s^t \langle \nabla_x f(\omega(u)), b(u, \omega(u))\rangle\, \mathrm{d}u\right] \\
&\quad + \mathbb{E}^{\mu_n} \left[F \cdot \int_s^t \langle \nabla_x f(\omega(u)), \sigma(u, \omega(u))\dot{W}^{(n)}(u)\rangle\, \mathrm{d}u\right] \\
&= I_1^{(n)} + I_2^{(n)}
\end{aligned}$$

From the boundedness of b, the convergence of $I_2^{(n)}$ is the same as in [SV72]. Similarly, from the fact the that integrand in $I_1^{(n)}$ is a bounded continuous function of

ω, the weak convergence of the μ_n implies that

$$\mathbb{E}^{\mu_n}\left[F\cdot\int_s^t\langle\nabla_x f(\omega(u)),b(u,\omega(u))\rangle\,\mathrm{d}u\right]\to\mathbb{E}^{\mu}\left[F\cdot\int_s^t\langle\nabla_x f(\omega(u)),b(u,\omega(u))\rangle\,\mathrm{d}u\right]$$

and so the result follows. □

We can now see that the first inclusion for the support theorem holds:

Corollary 4.3. *Suppose that* $\sigma:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ *is* BC^1 *and* $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ *is bounded continuous in space and* L^2 *in time. Let* $\mathcal{S}(W)$ *denote the space of solutions to (SDE), and let* μ_x *denote the law of* $x\in\mathcal{S}(W)$. *Then*

$$\bigcup_{x\in\mathcal{S}(W)}\operatorname{supp}(\mu_x)\subseteq\overline{\bigcup_{u\in H}\mathcal{S}(u)}$$

Chapter 5

Numerical Experiments

We can simulate SDE sample paths numerically to get a feel for how solutions should behave. This is particularly useful for gauging which sets appear to be reachable by sample paths, since we can simulate numerous paths and plot them together. Likewise, we can plot solutions to the control ODEs and see how they compare with either the simulated SDE sample paths, or the support if already known analytically.

5.1 Stochastic Runge-Kutta methods

We use general Runge-Kutta methods for the simulations. The implementation is essentially identical to the that for deterministic ODEs, just performed path-wise.

Suppose we wish to solutions to the SDE

$$\mathrm{d}x_t = \sigma(t, x_t)\,\mathrm{d}W_t + b(t, x_t)\,\mathrm{d}t, \quad x_0 = a \in \mathbb{R}^d$$

Fix a sequence of timesteps $(\tau_j)_{j=1}^{J}$, possibly all equal. Set $t_n = \sum_{j=1}^{n} \tau_j$. We estimate dW_{t_n} with $\omega_n := W_{t_n} - W_{t_{n-1}} \sim N(0, \tau_n I_d)$. Choose $s \in \mathbb{N}$ (number of stages in the method), along with a_{ij}, b_j and $c_i := \sum_{j=1}^{s} a_{ij} \in \mathbb{R}$ for each $i, j = 1, \ldots, s$.

The value of our approximation at time t_n will be denoted y^n, so that $y^n \approx x_{t_n}(\kappa)$ for some sample $\kappa \in \Omega$. Set $y^0 = a$. At each timestep we have three stages to perform:

1. Simulate $\omega_n \sim N(0, \tau_n I_d)$
2. Find a sequence of vectors $(p_i)_{i=1}^s$ such that

$$p_i = y^{n-1} + \sum_{j=1}^{s} a_{ij} \Big[b(t_{n-1} + c_j \tau_n, p_j)\tau_n + \sigma(t_{n-1} + c_j \tau_n, p_j)\omega_n \Big]$$

3. Set

$$y^n := y^{n-1} + \sum_{i=1}^{s} b_i \Big[b(t_{n-1} + c_i \tau_n, p_i)\tau_n + \sigma(t_{n-1} + c_i \tau_n, p_i)\omega_n \Big]$$

This is repeated until we reach the end of the time interval we are simulating over.

See [But87] for more detail on Runge-Kutta schemes for ODEs.

5.1.1 Examples of schemes

A Runge-Kutta scheme is determined by its Butcher tableau.

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{a} \\ \hline & \boldsymbol{b}^T \end{array} = \begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}$$

The method is explicit if $a_{ij} = 0$ for all $j \geq i$.

The most basic example of a scheme is the forward Euler method, a fully explicit one stage method. In this case results in the well-known Euler-Maruyama approximation for the SDE. Its Butcher tableau is given by

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

Similarly we have the backward Euler method, a fully implicit one stage method whose Butcher tableau is given by

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

This has the advantage of higher stability, allowing for greater timesteps.

For the non-explicit methods we need to invoke a non-linear solver to find the $(p_i)_{i=1}^s$, which requires knowledge of the spatial derivatives of the coefficients. Since we would rather not restrict ourselves to C^1 coefficients, we will stick to explicit schemes. The one we focus on is the 3-stage Heun method, a third order scheme given by the Butcher tableau

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 \\ 2/3 & 0 & 2/3 & 0 \\ \hline & 1/4 & 0 & 3/4 \end{array}$$

We implement the scheme in C++ and plot the solutions with gnuplot.

5.2 Simulations

We first consider a process that we considered earlier in Chapter 1, given by the SDE

$$dx_t = dW_t + \frac{1}{\sqrt{|x_t|}} \mathbb{1}_{(0,1)}(x_t)\,\mathrm{d}t, \quad x_0 = a > 0 \tag{5.1}$$

It was proved in Proposition 1.3 that the law of this process is supported on all of $C_0([0,T];\mathbb{R}^d)$. We provide numerical confirmation of this result.

Figure 5.1 shows an example of a sample path of the solution to (5.1), which can be seen to cross zero. Since the behaviour of a single path doesn't tell us much, in Figure 5.2 we plot the graphs of fifty sample paths. Whilst the vast majority of the paths remain positive, a significant number of them pass through the origin. Once a path passes the origin it acts as a Brownian motion since their is no drift there, from which it is more clear that it can reach any point on the negative real line.

We now look at the similar SDE,

$$dx_t = dW_t + \frac{1}{|x_t|^2} \mathbb{1}_{(0,1)}(x_t)\,\mathrm{d}t, \quad x_0 = a > 0 \tag{5.2}$$

The drift does not belong to L^p for any $p > 1$ in this case. We postulated that this drift could be strong enough such that the process cannot reach the negative real line.

Figure 5.3 shows an example of a single sample path, which remains positive. Figure 5.4 shows fifty sample paths, all of which also remain positive. This could suggest that the diffusion does in fact remain positive almost surely, despite the uniform ellipticity of the equation.

The next process we try is Brownian motion on the positive real line with sticky boundary. Let $\theta > 0$ be some constant, then the process is given by

$$\mathrm{d}x_t = \mathbb{1}_{(0,\infty)}(x_t)\,\mathrm{d}W_t + \theta \mathbb{1}_{\{0\}}(x_t)\,\mathrm{d}t, \quad x_0 = a \geq 0$$

This SDE is not elliptic since we turn off the noise when the process hits zero. The constant θ determines how quickly the process will leave the origin once it hits it. The case $\theta = 0$ gives absorbed Brownian motion, and the case $\theta = \infty$ can be understood as describing reflected Brownian motion, i.e. $x_t = |W_t|$. The SDE isn't elliptic since the diffusion coefficient vanishes outside of $(0, \infty)$.

We simulate the case $\theta = 1$. Looking at Figure 5.5, the sample path behaves as expected. From the plot of fifty sample paths together (Figure 5.6), it seems likely that the support of x is contained in the space $C^+_{x_0}$ defined in the first chapter.

Note that since the drift is bounded in the case $\theta < \infty$, replacing the diffusion coefficient by $\mathbb{1}_{[0,\infty)}$ would result in a process with full support.

We can also define Brownian motion on the interval $[0, 1]$ with sticky boundary. It is given by the SDE

$$\mathrm{d}x_t = \mathbb{1}_{(0,1)}(x_t)\,\mathrm{d}W_t + \theta\left(\mathbb{1}_{\{0\}}(x_t) - \mathbb{1}_{\{1\}}(x_t)\right)\mathrm{d}t, \quad x_0 = a \in [0, 1]$$

Plots of simulations of a single sample path and fifty sample paths are given in Figure 5.7 and Figure 5.8 respectively.

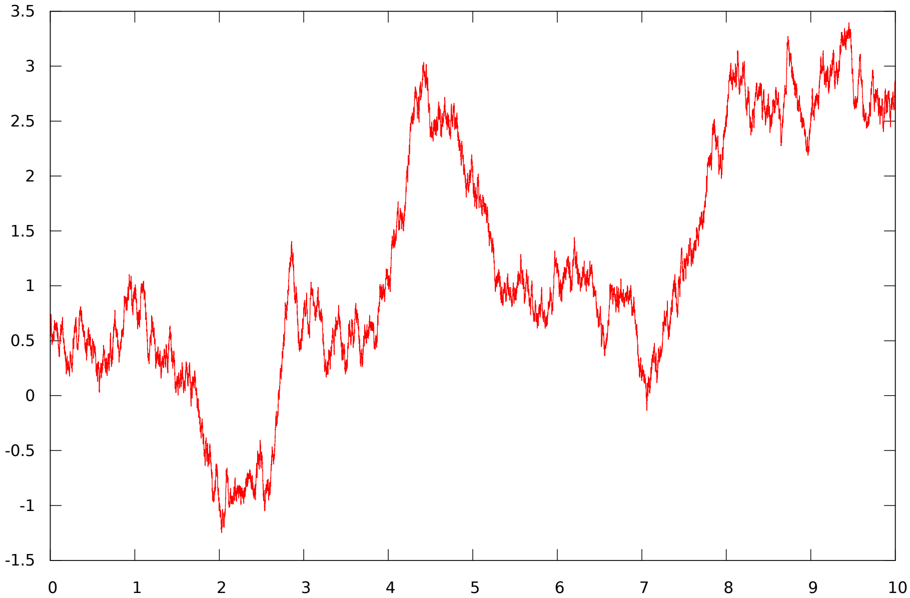

Figure 5.1: A sample path of the process given by (5.1).

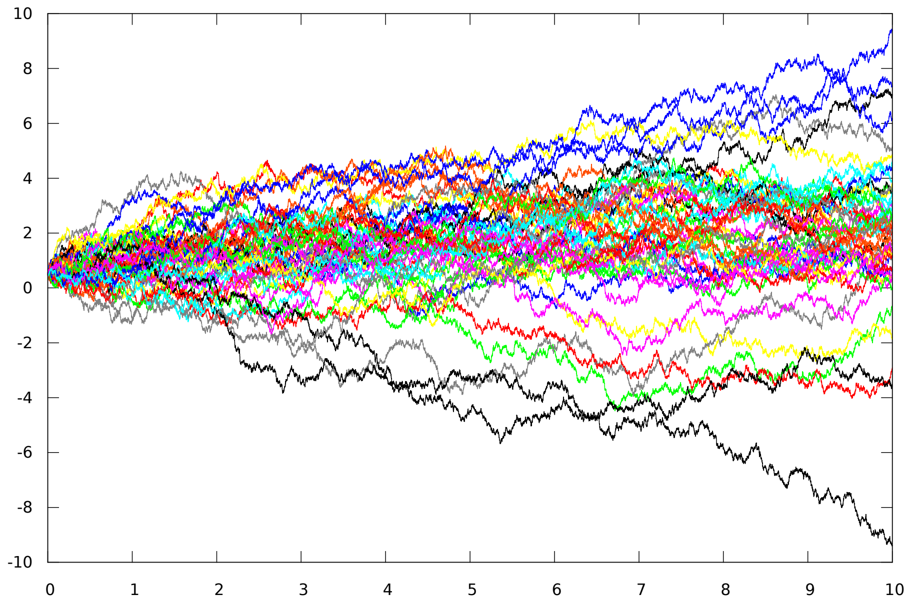

Figure 5.2: Fifty sample paths of the process given by (5.1).

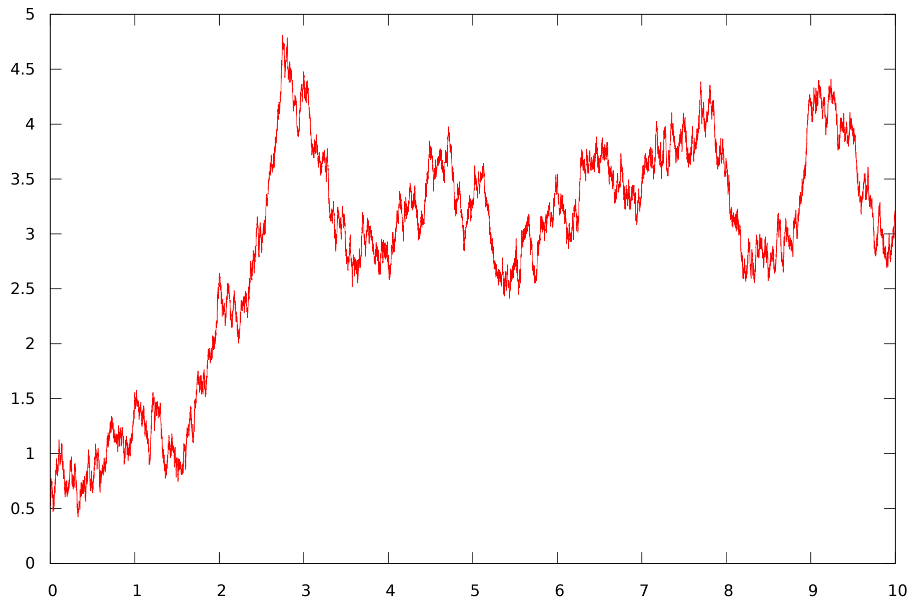

Figure 5.3: A sample path of the process given by (5.2).

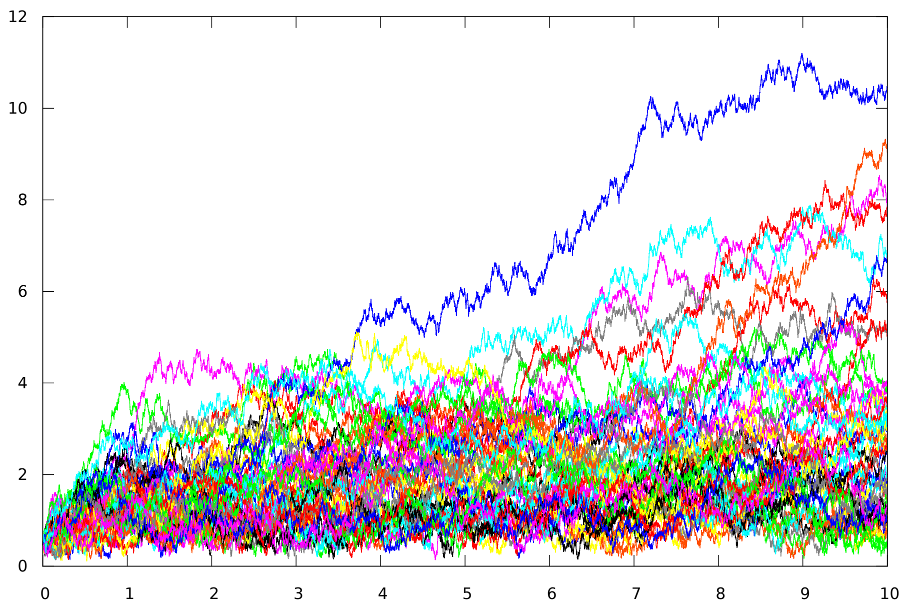

Figure 5.4: Fifty sample paths of the process given by (5.2).

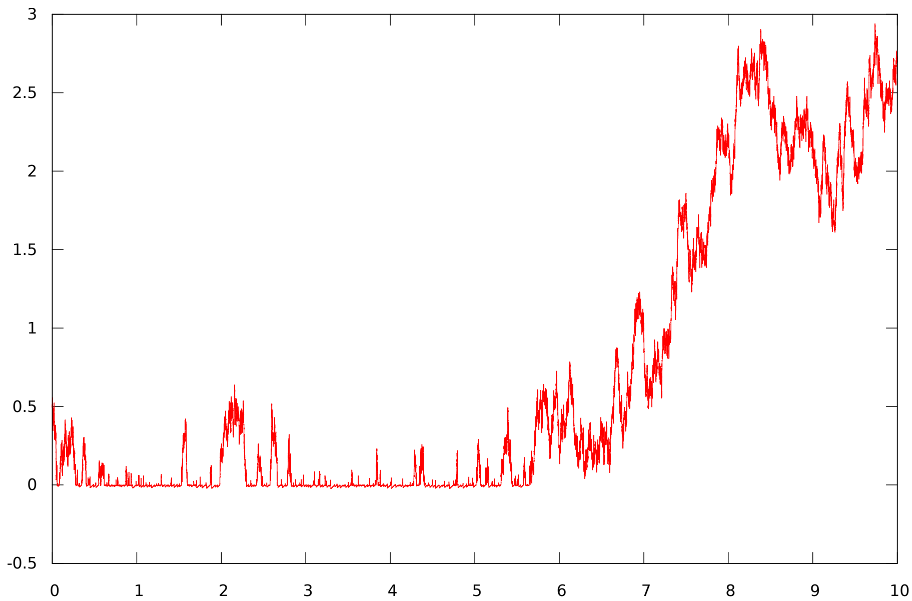

Figure 5.5: A sample path of Brownian motion on the positive real line with sticky boundary.

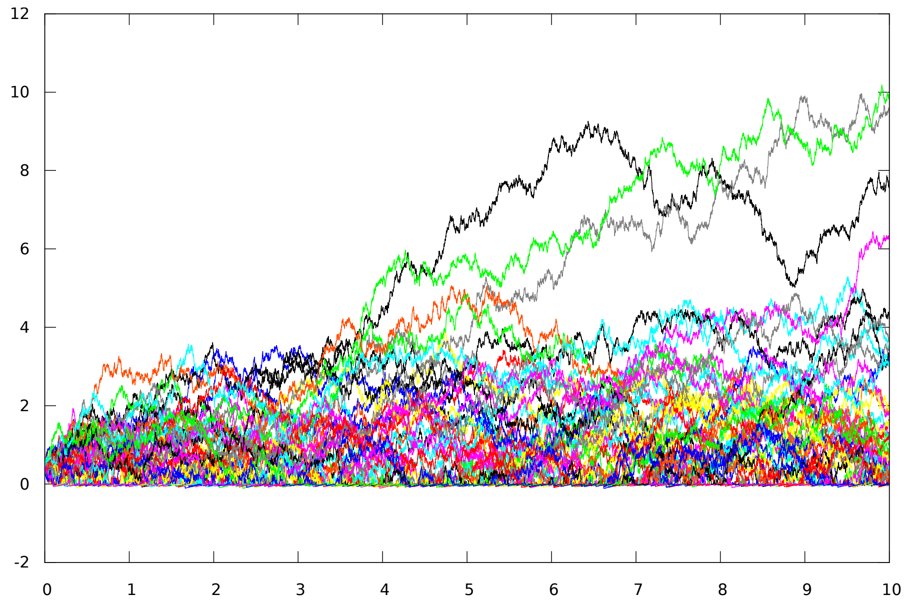

Figure 5.6: Fifty sample paths of Brownian motion on the positive real line with sticky boundary.

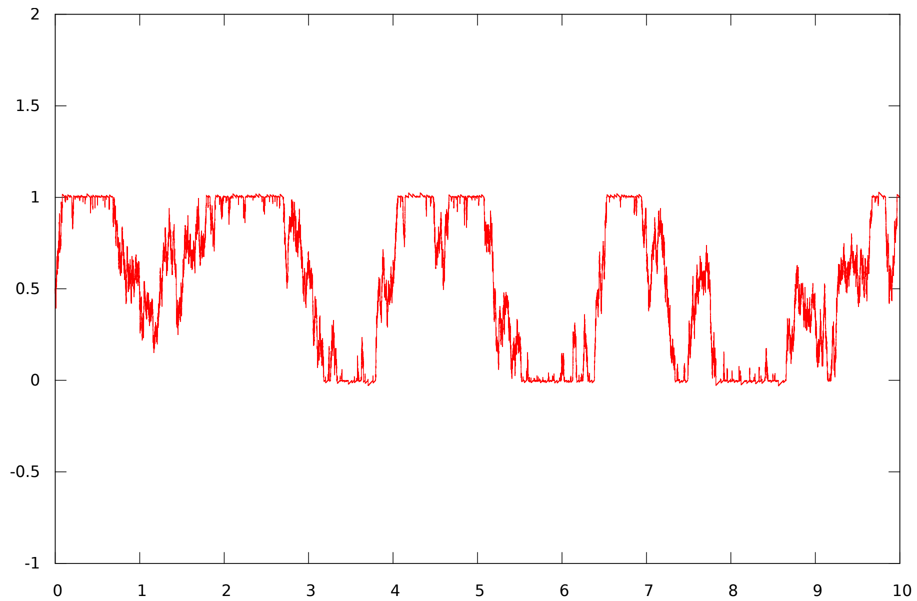

Figure 5.7: A sample path of Brownian motion on $[0, 1]$ with sticky boundary.

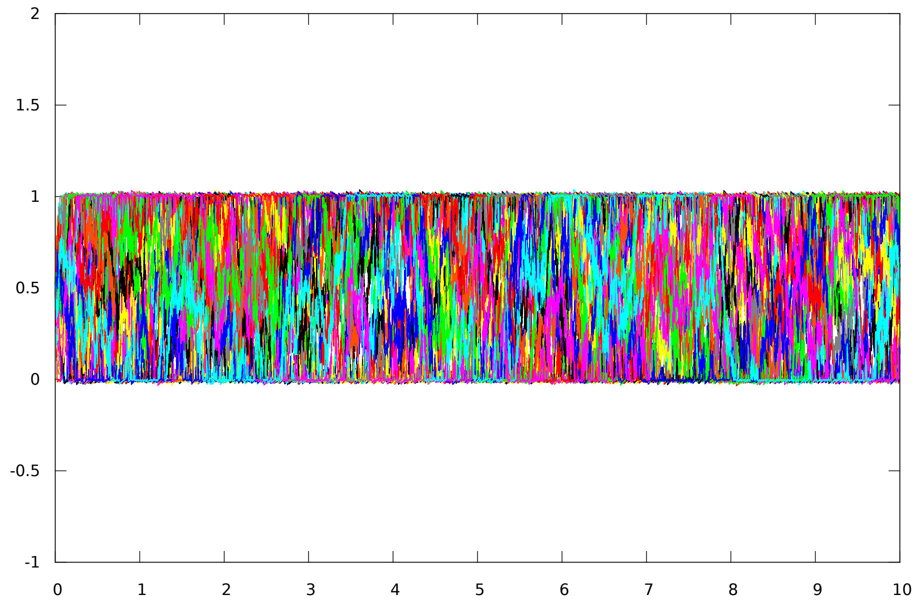

Figure 5.8: Fifty sample paths of Brownian motion on $[0, 1]$ with sticky boundary.

We now try an example with $d = 2$, since this leaves more freedom for degeneracy instead of just points where $\sigma = 0$. Let

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Consider the SDE

$$\mathrm{d}x_t = \left(\sigma_1 \mathbb{1}_{((0,\infty)\times\mathbb{R})\cup\{0,0\}}(x_t) - \sigma_2 \mathbb{1}_{\mathbb{R}\times[0,\infty)}(x_t)\right) \mathrm{d}W_t, \quad x_0 = a \in \mathbb{R}^2$$

so that the diffusion coefficient $\sigma : \mathbb{R}^2 \to \mathbb{R}^2 \otimes \mathbb{R}^2$ is given by

$$\sigma(x,y) = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} & x \leq 0, y \geq 0 \\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & x > 0, y \geq 0 \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & x > 0, y < 0 \text{ or } (x,y) = (0,0) \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & x \leq 0, y < 0 \end{cases}$$

Then clearly we have no degeneracy in one quadrant, some degeneracy in two and full degeneracy in the other. The effect of this can be seen in the simulation in Figure 5.9.

The support of this process appears to be sensitive to the starting point x_0 as expected. If started in the quadrant $x \leq 0, y < 0$, then the process cannot move so it's support is just the starting point. If it is started in the quadrant $x > 0, y < 0$, then the support appears to be paths restricted to the whole (closed) quadrant $x \geq 0, y \leq 0$.

The process becomes more interesting when started in the positive quadrant $x > 0, y \geq 0$. The process initially performs a one-dimensional Brownian motion parallel to the x-axis. Once it hits the y-axis, which occurs almost-surely in a finite time, it stays there and performs a one-dimensional Brownian motion along it. It will then almost surely hit the origin in a finite time. At this point there is a positive chance it will enter the quadrant $x > 0, y < 0$. Once it is there it is free to explore that whole quadrant. Therefore it is believable that the support in this case, when started from (x_0, y_0), is given by the set of paths restricted to take values in the set

$$([0,\infty) \times \{y_0\}) \cup (\{0\} \times [0,\infty)) \cup ([0,\infty) \times (-\infty, 0])$$

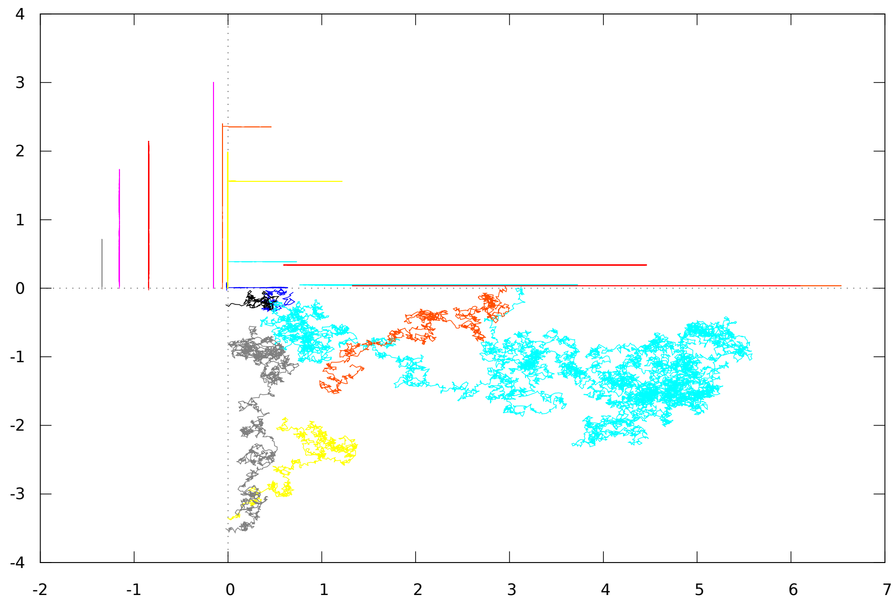

Figure 5.9: Plots of 20 sample paths, with $x_0 \sim N(0, I_2)$

Bibliography

[AC84] J P Aubin and A Cellina. Differential inclusions, 1984.

[AGL94] G Ben Arous, M Gradinaru, and M Ledoux. Hlder norms and the support theorem for diffusions. *Ann. Inst. H. Poincar*, (30):415–436, 1994.

[Aro67] D G Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc*, (73):890–896, 1967.

[BB82] R Bafico and P Baldi. Small random perturbations of peano phenomena. *Stochastics*, 6(3-4):279–292, 1982.

[BOQ09] R Buckdahn, Y Ouknine, and M Quincampoix. On limiting values of stochastic differential. *Bull. Sci. math*, (133):229–237, 2009.

[But87] J C Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods.* Wiley-Interscience, 1987.

[Dav07] A Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Notices*, (2007):124–126, 2007.

[Dav10] A Davie. Individual path uniqueness of solutions of stochastic differential equations. In D Crisan, editor, *Stochastic Analysis 2010*, pages 213–225, 2010.

[Dei92] K Deimling. Multivalued differential equations. walter de gruyter, 1992.

[DT94] M J Davies and A Truman. Brownian motion with a sticky boundary and point sources in quantum mechanics. *Mathl. Comput. Modelling*, 20(4/5):173–193, 1994.

[FF10] E Fedrizzi and F. Flandoli : Pathwise uniqueness and continuous dependence for sdes with nonregular drift, preprint arxiv, 2010.

[Fla08] F Flandoli. Remarks on uniqueness and strong solutions to deterministic and stochastic differential equations, to appear on. In *the proceedings of the 8 th GOCPS, Metrika, special*, 2008.

[Fla11] F Flandoli. Regularizing properties of brownian paths and a result of davie. *Stochastics and Dynamics*, 11(2/3):322–331, 2011.

[FV10] P Friz and N Victoir. Multidimensional stochastic processes as rough paths. theory and applications, 2010.

[GP90] I Gyongy and T Prohle. On the approximation of stochastic differential equations and on stroock-varadhan's support theorem. *Computers Math. Applic.*, 19(1):65–70, 1990.

[IW81] V Ikeda and S Watanabe. Stochastic differential equations and diffusion processes, north-holland, 1981.

[KR05] N V Krylov and M Rckner. Strong solutions of stochastic equations with singular time dependent drift. In *Probab. Theory Related Fields 131*, pages 154–196, 2005.

[LQZ02] M Ledoux, Z Qian, and T Zhang. Large deviations and support theorem for diffusion processes via rough paths. stochastic process. *Appl*, (102):265–283, 2002.

[MSS94] A Millet and M Sanz-Sol. A simple proof of the support theorem for diffusion processes. In *Sminaire de Probabilits XXVIII, Lecture Notes in Mathematics 1583*, pages 36–48, 1994.

[Nak04] T Nakayama. Support theorem for mild solutions of sde's in hilbert spaces. *Journal of Mathematical Sciences-University of Tokyo*, (11.3):245–312, 2004.

[PS87] B Paden and S S Sastry. A calculus for computing filippovs differential inclusion with application to the variable structure control of robot manipulators. *IEEE Transactions on Circuits and Systems*, (34):73–82, 1987.

[Sim00] T Simon. Support theorem for jump processes. *Stoch. Proc. and Appl*, (89):1–30, 2000.

[Smi00] G V Smirnov. Introduction to the theory of differential inclusions. *American Mathematical Society, Rhode Island*, 2000.

[SV72] D Stroock and S Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proc. 6th Berkeley Symp. Math. Stat. Prob*, pages 333–359, 1972.

[Tol88] A Tolstonogov. Solutions of a differential inclusion with unbounded right-hand side. *Siberian Math. J*, (29):857–868, 1988.

[Tre13] D Trevisan. Zero noise limits using local times. *Electronic Communications in Probability*, 18:1–7, 2013.

[Ver80] A Yu Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S)*, (111):480, 1980.

[Zha12] L Zhang. Random perturbation of some multi-dimensional non-lipschitz ordinary differential equations. *arXiv preprint arXiv:1202.4131*, 2012.