Topics on regularization by noise

Franco Flandoli
University of Pisa

July 16, 2013

Abstract

In the first part we discuss existence and uniqueness of SDEs with non-degenerate additive noise and singular drift. In the second part we introduce the problem of the zero-noise limit for such singular SDEs.

1 Lecture 1: SDEs with rough drift

1.1 Introduction

Consider the SDE in $\mathbb{R}^d$

$$dX_t = b(t, X_t)\, dt + dW_t, \qquad X_0 = x$$

with $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ at least measurable and $W = (W_t)_{t \geq 0}$ a BM in $\mathbb{R}^d$.

If

$$\begin{aligned} |b(t, x) - b(t, y)| &\leq L|x - y| \\ |b(t, x)| &\leq C(1 + |x|) \end{aligned}$$

then there exists a pathwise unique strong solution (classical Cauchy-Lipschitz theorem; see definitions below). When b is also $C^{1,\alpha}$, there exists a stochastic flow of diffeomorphisms [23].

Let us recall that most results come from the following simple computation. Let $X_t^{(i)}$, $i = 1, 2$, be two solutions, with initial conditions $x^{(i)}$ and let $Y_t = X_t^{(1)} - X_t^{(2)}$. Then

$$Y_t = \int_0^t \left(b\left(X_s^{(1)}\right) - b\left(X_s^{(2)}\right)\right) ds$$

whence

$$|Y_t| \leq L \int_0^t |Y_s|\, ds$$

and now, by Gronwall lemma,

$$|Y_t| \leq e^{Lt} \left| x^{(1)} - x^{(2)} \right|.$$

This gives us uniqueness and, by Kolmogorov regularity theorem, a C^α flow, for any $\alpha \in (0, 1)$.

Remark 1 *These results remain valid under multiplicative noise, with similar assumptions on the diffusion coefficients.*

Remark 2 *When the flow $\varphi_t(\omega,\cdot)$ is established, it is also Lipschitz continuous, because*

$$|\varphi_t(\omega,x)-\varphi_t(\omega,y)|\leq e^{Lt}|x-y|.$$

However, This latter conclusion would fail for multiplicative noise. Moreover, full differentiability of the flow requires $b\in C^{1,\alpha}$, not just C^1, since one has to apply Kolmogorov test to the variational equation.

Plan of the talk: discuss the case when b is much weaker than Lipschitz continuous. Precisely:

1. recall loss of well-posedness without noise
2. the case $d=1$
3. any $d\in\mathbb{N}$, Hölder b
4. other results (L^p drift, infinite dimensional case, etc.).

More details can be found in [12].

Remark 3 *Let us mention in advance that the theory we are going to describes extends to some multiplicative noise but of regular and non-degenerate type, at the price of complicate regularity theory of PDEs. In this direction, see [29], [30], [14].*

Remark 4 *On the contrary, we do not know how to extend it to degenerate-noise cases and, very important, we do not know how to extend it to random drifts ($b=b(t,x,\omega)$, with poor regularity in x).*

Remark 5 *Moreover, we do not know similar tricks for rough diffusion coefficients.*

1.2 Definitions

Denote by $\{F_t^W\}_{t\geq0}$ the filtration of the Brownian motion (F_t^W is the σ-algebra generated by the family of r.v. $\{W_s, s\in[0,t]\}$).

Definition 6 *Given (Ω,F,P) and a d-dimensional Brownian motion W, a strong solution of the SDE is a stochastic process $X_t(\omega)$, $t\geq0$, $\omega\in\Omega$ such that*
i) the trajectories $t\mapsto X_t(\omega)$ are continuous, for P-a.e. $\omega\in\Omega$;
ii) X_t is adapted to F_t^W
iii) the identity

$$X_t(\omega)=x_0+\int_0^t b(s,X_s(\omega))\,ds+W_t(\omega),\quad t\geq0$$

holds for P-a.e. $\omega\in\Omega$.

Definition 7 *We say that there is pathwise uniqueness when, given (Ω,F,P), a filtration $\{F_t\}_{t\geq0}$, a Brownian motion W and two continuous adapted solutions $X_t^{(i)}$, $i=1,2$, then*

$$P\left(X_t^{(1)}=X_t^{(2)},\quad t\geq0\right)=1.$$

The full definition of stochastic flow is not relevant here; we use a reduces definition. We say that the SDE has a continuous stochastic flow if there is a measurable family of continuous maps

$$\varphi_t(\omega, \cdot) : \mathbb{R}^d \to \mathbb{R}^d$$

such that $\varphi_t(\cdot, x)$ is a solution of the SDE (in other words, we ask that the random field X_t^x has a continuous version in x). The definition of Hölder flow etc. is similar.

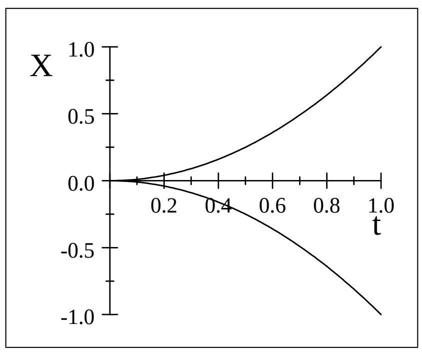
1.3 Loss of well-posedness without noise

When the drift is essentially weaker than Lipschitz, uniqueness may be lost for the deterministic equation (also existence in extreme cases, but we shall not address this problem). Let us first mention two elementary and well known examples.

The most well-known example is perhaps:

$$d = 1, \quad b(x) = 2sign(x)\sqrt{|x|}, \quad x_0 = 0$$

which has infinitely many solutions: $X(t) = 0$, $X(t) = t^2$, $X(t) = -t^2$, and others. The function b of this example is Hölder continuous.

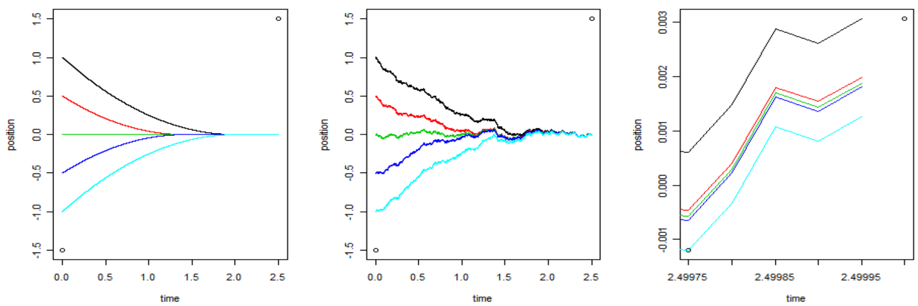


Extremal solutions of the equation $X' = sign(X)\sqrt{|X|}$.

In the case:

$$d = 1, \quad b(x) = -2sign(x)\sqrt{|x|}, \quad x_0 = 0$$

there is uniqueness (it can be proved) but infinitely many solutions coalesce in finite time with $X(t) = 0$. We mention this example mainly for graphical reasons. Indeed, it is difficult to see graphically that noise restores uniqueness; on the contrary, one can see the effect of noise on coalescence. The next three pictures show the deterministic case, the stochastic one (apparently coalescing) and a blow-up of the last few time steps of the stochastic simulation where the different trajectories are still visible.



In $d=1$, deterministic case, the only way, at least for a continuous drift, to produce singular behavior is to be equal to zero at some point. For instance we have:

Proposition 8 *If b is continuous in a neighbor U of x_0 and $b(x_0) \neq 0$, then there is at most one solution, locally in time, of the Cauchy problem $X_t' = b(X_t)$, $X_0 = x_0$.*

Proof. If X_t is a local solution, being continuous, we have $X_t \in U$ and thus $b(X_t) \neq 0$ on some interval $[0,\tau]$ with $\tau > 0$. Then

$$\frac{X_t'}{b(X_t)} = 1$$

on $[0,\tau]$ and therefore, integrating on $[0,\tau]$ and changing variable, we get

$$\int_{x_0}^{X_t} \frac{dx}{b(x)} = t \qquad \text{for all } t \in [0,\tau].$$

The integration is performed on an interval contained in U, where $b \neq 0$. If $H(x)$ denotes the function $H(x) = \int_{x_0}^{x} \frac{dx'}{b(x')}$, then we have $H(X_t) = t$. Since b does not change sign in U, H is strictly monotone hence invertible and thus $X_t = H^{-1}(t)$. Thus X_t is uniquely determined by b and x_0, on a small time interval. ■

In the same way one can prove local existence. The result can be extended also to more rough b with proper non-zero and non-explosion assumptions.

In dimension $d > 1$ one can produce counterexamples to uniqueness of very wild form, see [1].

In spite of the great restriction in $d=1$, let us mention that one can accumulate a lot of singular points. An extreme case is $b(x) =$ Dirichlet function:

$$b(x) = \begin{cases} 1 & \text{if} \quad x \notin \mathbb{Q} \\ 0 & \text{if} \quad x \in \mathbb{Q} \end{cases}.$$

For every initial condition x_0 and any time interval $[0,\tau]$ with $\tau > 0$, there are infinitely many solutions in $[0,\tau]$. By solution, we mean any continuous function $x(t)$ such that

$$x(t) = x_0 + \int_0^t b(x(s))\, ds.$$

If $x_0 \in \mathbb{Q}$, then $x_-(t) \equiv x_0$ is a solution. But also $x_+(t) = x_0 + t$: indeed $x_+(s) \notin \mathbb{Q}$ for a.e. $s \geq 0$ and thus $b(x_+(s)) = 1$ for a.e. $s \geq 0$; it follows that $\int_0^t b(x_+(s))\, ds = t$, hence x_+ is a solution.

If $x_0 \notin \mathbb{Q}$, $x_+(t) = x_0 + t$ is still a solution, as above (the constant function x_0 is no more a solution). But given any $\tau > 0$, there is $t_0 \in (0, \tau)$ such that $x_+(t_0) \in \mathbb{Q}$. Then we may consider the new function

$$x_+^{t_0}(t) = \begin{cases} x_+(t) & \text{if} \quad t \in [0, t_0] \\ x_+(t_0) & \text{if} \quad t > t_0 \end{cases}.$$

This is a new solution. Since there are infinitely many such points $t_0 \in (0, \tau)$, there are infinitely many such solutions. Of course, from each one of them we may bifurcate again along parallel to the diagonal and again in horizontal way, many times. The set of all solutions, from any x_0, is wild and it is such as close as we like to $t = 0$.

We shall come back to some of these examples in Lecture 2 on the zero-noise limit.

1.4 The case $d = 1$

Assume $b = b(x)$ for simplicity. Consider the variational equation

$$\frac{d}{dt} V_t = b'(X_t) V_t, \qquad V_0 = 1.$$

When b is regular, this is the equation for $\frac{\partial}{\partial x} X_t^x$ and its solution is

$$V_t = \exp\left(\int_0^t b'(X_s)\, ds \right).$$

When b is not regular, a priori this is not well defined. However, again under some regularity, if $B(x)$ is a function such that $B' = b$, then

$$\begin{aligned} dB(X_t) &= b(X_t)\, dX_t + \frac{1}{2} b'(X_t)\, dt \\ &= b^2(X_t)\, dt + b(X_t)\, dW_t + \frac{1}{2} b'(X_t)\, dt \end{aligned}$$

whence

$$\int_0^t b'(X_s)\, ds = 2B(X_t) - 2B(x) - 2\int_0^t b^2(X_s)\, ds - 2\int_0^t b(X_s)\, dW_s.$$

The idea is then:

- this formula allows us to define $\int_0^t b'(X_s)\, ds$ even when b is not differentiable; it is sufficient that b is L^∞ (even less is sufficient)
- the estimate we get for the variational equation will imply uniqueness and existence of a stochastic flow.

The following result is taken from [15] and it can be extended to more rough b's.

Theorem 9 *If $b \in C_b$ then we have pathwise uniqueness and existence of a β-Hölder stochastic flow, for any $\beta \in (0, 1)$. If $b \in C_b^\theta$ then the flow is differentiable.*

Proof. Let $X_t^{(i)}$, $i = 1, 2$, be two solutions, with initial conditions $x^{(i)}$ and let $Y_t = X_t^{(1)} - X_t^{(2)}$. Then

$$\frac{d}{dt} Y_t = b\left(X_t^{(1)}\right) - b\left(X_t^{(2)}\right).$$

Let b_ε be usual smooth approximations of b. Then

$$\frac{d}{dt} Y_t = \lim_{\varepsilon \to 0} \left(b_\varepsilon\left(X_t^{(1)}\right) - b_\varepsilon\left(X_t^{(2)}\right)\right) = \lim_{\varepsilon \to 0} \left(\int_0^1 b'_\varepsilon \left(\alpha X_t^{(1)} + (1-\alpha) X_t^{(2)} \right) d\alpha Y_t \right).$$

Set $X_t^\alpha = \alpha X_t^{(1)} + (1-\alpha) X_t^{(2)}$. If we prove that $\lim_{\varepsilon \to 0} \int_0^1 b'_\varepsilon (X_t^\alpha)\, d\alpha$ exists a.s., for any given $t \geq 0$, and is continuous in t, then

$$\frac{d}{dt} Y_t = Y_t \left(\lim_{\varepsilon \to 0} \int_0^1 b'_\varepsilon (X_t^\alpha)\, d\alpha \right)$$

whence

$$X_t^{(1)} - X_t^{(2)} = \left(x^{(1)} - x^{(2)} \right) \exp \int_0^t \left(\lim_{\varepsilon \to 0} \int_0^1 b'_\varepsilon (X_s^\alpha)\, d\alpha \right) ds.$$

This gives us immediately pathwise uniqueness (when $x^{(1)} = x^{(2)}$) and the flow properties with some additional work.

Notice that

$$dX_t^\alpha = \left(\alpha b\left(X_t^{(1)}\right) + (1-\alpha)\, b\left(X_t^{(2)}\right) \right) dt + dW_t.$$

If B_ε is the analogous approximation of b_ε, we have

$$\begin{aligned} dB_\varepsilon (X_t^\alpha) &= b_\varepsilon (X_t^\alpha)\, dX_t^\alpha + \frac{1}{2} b'_\varepsilon (X_t^\alpha)\, dt \\ &= b_\varepsilon (X_t^\alpha) \left(\alpha b\left(X_t^{(1)}\right) + (1-\alpha)\, b\left(X_t^{(2)}\right) \right) dt + b_\varepsilon (X_t^\alpha)\, dW_t + \frac{1}{2} b'_\varepsilon (X_t^\alpha)\, dt \end{aligned}$$

and thus

$$\int_0^t b'_\varepsilon (X_s^\alpha)\, ds = 2B_\varepsilon (X_t^\alpha) - 2B_\varepsilon (x) - 2 \int_0^t b_\varepsilon (X_s^\alpha)\, b_s^{(\alpha)} ds - 2 \int_0^t b_\varepsilon (X_s^\alpha)\, dW_s.$$

where $b_s^{(\alpha)} := \alpha b\left(X_s^{(1)}\right) + (1-\alpha)\, b\left(X_s^{(2)}\right)$ Then

$$\lim_{\varepsilon \to 0} \int_0^t b'_\varepsilon (X_s^\alpha)\, ds = 2B (X_t^\alpha) - 2B (x) - 2 \int_0^t b (X_s^\alpha)\, b_s^{(\alpha)} ds - 2 \int_0^t b (X_s^\alpha)\, dW_s$$

which is continuous in t.

To obtain the flow properties one has to prove a few exponential estimates, that we omit. ■

Remark 10 *The Lipschitz or C^1 regularity of the flow under weaker assumptions than $b \in C_b^\theta$ is open, except for results on particular drifts, see [2], [25].*

1.5 Any dimension

There are generalizations of these results to any dimensions, some of them older than [15] and other new. Let us state the classical result of Veretennikov [28]:

Theorem 11 *If $b \in L^{\infty}$, then there is strong existence and pathwise uniqueness.*

Concerning flows, one of the most recent results is [24]:

Theorem 12 *If $b \in L^{\infty}$, then there exists a β-Hölder stochastic flow, for any $\beta \in (0,1)$, which is also of class $W^{1,p}_{loc}$ for every $p \geq 1$.*

Two other results on the subject, by [13] and [11], are:

Theorem 13 *If $b \in L^{\infty}\left([0,T];C^{\alpha}_{b}\right)$ then there exists a $C^{1,\alpha'}$ flow, for any $\alpha' \in (0,\alpha)$.*

Theorem 14 *If $b \in L^{q}\left(0,T;L^{p}\right)$ with $p,q > 2$,*

$$\frac{d}{p} + \frac{1}{q} < 1$$

then there exists a a β-Hölder stochastic flow, for any $\beta \in (0,1)$, which is also of class $W^{1,p}_{loc}$ for every $p \geq 1$.

Let us discuss the proof of Theorem 13 only. How to repeat the argument of the 1D case? The variational equation has no explicit solution that could be extended to non-smooth drift. We follow a different idea, from [13] (in fact we realized later that it was used before in stochastic homogenization). Instead of using Itô formula to "regularize" $\int_0^t b'\left(X_s\right)ds$, we use it to "regularize" $\int_0^t b\left(X_s\right)ds$, the term which appears in the formulation itself of the SDE. But we cannot simply ask $B' = b$ here.

Proof. Let us prove Theorem 13. Let us consider the backward parabolic vector-valued PDE

$$\frac{\partial U}{\partial t} + \frac{1}{2}\Delta U + b \cdot \nabla U = -b + \lambda U, \qquad U\left(T\right) = 0.$$

Assume $b \in C\left([0,T];C^{\alpha}_{b}\right)$ (the case $b \in L^{\infty}\left([0,T];C^{\alpha}_{b}\right)$ requires more details). Then there exists a unique solution with the regularity

$$\frac{\partial U}{\partial t}, \quad D^{2}U \in C\left([0,T];C^{\alpha}_{b}\right)$$

and such that

$$\left\|DU\right\|_{\infty} \leq C\left(\lambda\right)$$

with $\lim_{\lambda\to\infty} C\left(\lambda\right) = 0$. This is classical regularity theory of PDEs, (see [20]). Set $\mathcal{L}U = \frac{\partial U}{\partial t} + b \cdot \nabla U + \frac{1}{2}\Delta U$ and recall that $\mathcal{L}U = -b + \lambda U$. By Itô formula

$$\begin{aligned} dU\left(t,X_t\right) &= \mathcal{L}U\left(t,X_t\right)dt + DU\left(t,X_t\right)dW_t \\ &= \left(-b+\lambda U\right)\left(t,X_t\right)dt + DU\left(t,X_t\right)dW_t \end{aligned}$$

whence

$$\int_0^t b(s, X_s)\,ds = U(0,x) - U(t, X_t) + \lambda \int_0^t U(s, X_s)\,ds + \int_0^t DU(s, X_s)\,dW_s$$

and thus, from

$$X_t = x + \int_0^t b(s, X_s)\,ds + W_t$$

we get

$$X_t = x + U(0,x) - U(t, X_t) + \lambda \int_0^t U(s, X_s)\,ds + \int_0^t (1 + DU(s, X_s))\,dW_s.$$

The fundamental advantage of this new formulation, although apparently longer, is that U is twice more regular than b and ∇U is once more regular.

Let us prove pathwise uniqueness. If $X^{(i)}(t)$, $i = 1, 2$, are two solutions, then

$$\begin{aligned} X_t^{(1)} - X_t^{(2)} &= U\left(t, X_t^{(2)}\right) - U\left(t, X_t^{(1)}\right) \\ &+ \lambda \int_0^t \left(U\left(s, X_s^{(1)}\right) - U\left(s, X_s^{(2)}\right)\right) ds \\ &+ \int_0^t \left(\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right) dW_s. \end{aligned}$$

Since

$$\left|U\left(t, X_t^{(2)}\right) - U\left(t, X_t^{(1)}\right)\right| \leq \frac{1}{2}\left|X_t^{(1)} - X_t^{(2)}\right|$$

for $\lambda \geq 0$ large enough, we get

$$\begin{aligned} \frac{1}{2}\left|X_t^{(1)} - X_t^{(2)}\right| &\leq \lambda \int_0^t \left|U\left(s, X_s^{(1)}\right) - U\left(s, X_s^{(2)}\right)\right| ds \\ &+ \left|\int_0^t \left(\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right) dW_s\right|. \end{aligned}$$

Let us go to the squares:

$$\begin{aligned} \frac{1}{4}\left|X_t^{(1)} - X_t^{(2)}\right|^2 &\leq \frac{\lambda\sqrt{t}}{2} \int_0^t \left|X_s^{(1)} - X_s^{(2)}\right|^2 ds \\ &+ 2\left|\int_0^t \left(\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right) dW(s)\right|^2 \end{aligned}$$

where we have used again the inequality above. By isometry formula we get

$$\begin{aligned} \frac{1}{4}E\left[\left|X_t^{(1)} - X_t^{(2)}\right|^2\right] &\leq \frac{\lambda\sqrt{t}}{2} \int_0^t E\left[\left|X_s^{(1)} - X_s^{(2)}\right|^2\right] ds \\ &+ 2E \int_0^t \left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds \end{aligned}$$

and finally, using the boundedness of the second derivatives of U,

$$\frac{1}{4}E\left[\left|X_t^{(1)}-X_t^{(2)}\right|^2\right] \leq \frac{\lambda\sqrt{t}}{2}\int_0^t E\left[\left|X_s^{(1)}-X_s^{(2)}\right|^2\right]ds$$
$$+2\left\|D^2U\right\|_\infty E\int_0^t E\left[\left|X_s^{(1)}-X_s^{(2)}\right|^2\right]ds.$$

By Gronwall lemma we deduce $E\left[\left|X_t^{(1)}-X_t^{(2)}\right|^2\right]=0$ for every $t\geq 0$ and thus, by continuity of trajectories, $P\left(X_t^{(1)}=X_t^{(2)},\quad t\geq 0\right)=1$. The proof of uniqueness is complete. To prove the flow property notice that by similar estimates as above one gets

$$E\left[\left|X_t^{(1)}-X_t^{(2)}\right|^p\right]\leq C_p\left|x^{(1)}-x^{(2)}\right|^p$$

and thus Kolmogorov criterion can be applied. For the differentiability more work is needed, that we omit in detail, see [13]. However, the idea is that we can differentiate in the initial condition and get (denote $\partial_h X_t^x=\lim_{\epsilon\to 0}\epsilon^{-1}\left(X_t^{x+\epsilon h}-X_t^x\right)$)

$$\partial_h X_t^x = h+\partial_h U(0,x) - DU(t,X_t^x)\,\partial_h X_t^x + \lambda\int_0^t DU(s,X_s^x)\,\partial_h X_s^x ds + \int_0^t D^2U(s,X_s^x)\,\partial_h X_s^x dW_s$$

where all terms are well defined and still allow to apply Kolmogorov test to prove that $x\mapsto D_h X_t^x$ has an Hölder continuous modification. ■

1.6 Miscellanea

We mention here a few other directions.

One is the interpretation, the intuitive reason behind these uniqueness results in spite of the singularities of the drift. In our opinion, the intuitive reason is the *regularity of the occupation measure*. The measure distributed by single trajectories of diffusions in $\mathbb{R}^d$ (the push-forward of Lebesgue measure on bounded time intervals) is singular with respect to Lebesgue measure (except for $d=1$, when it has a density, the *local time*), but still very regular and diffused with respect to the occupation measure of solutions to deterministic ODEs. This regularity smooths out the singularities of the drift, opposite to the deterministic case in which the solution may persist on the singularities. See [12] for more details.

Another one is the generalization to infinite dimensions. Doing carefully a probabilistic proof of the PDE theorem behind Theorem 13, one can show that the main constants are independent of n under proper assumptions on the drift. This allows one to extend the result to suitable equations in Hilbert spaces. See again [12] and the more recent extension to L^∞ drift in [8].

Using the reformulation of the SDE made above by means of the PDE:

$$X_t = x+U(0,x)-U(t,X_t)+\lambda\int_0^t U(s,X_s)\,ds+\int_0^t(1+DU(s,X_s))\,dW_s$$

one can differentiate X_t with respect to parameters and other variables. Let us mention the possibility to do Malliavin calculus (see [16]): if D_u denotes Malliavin derivative, $u\geq 0$, one can prove that

$$\begin{aligned} D_u X_t^x &= -DU\left(t, X_t^x\right) D_u X_t^x + \lambda \int_0^t DU\left(s, X_s^x\right) D_u X_s^x ds \\ &\quad + \int_0^t D^2 U\left(s, X_s^x\right) D_u X_s^x dW_s + 1_{u \le t} DU\left(u, X_u\right). \end{aligned}$$

Finally, in a separate subsection, let us discuss the proof of Theorem 14.

1.6.1 Integrable drift

Let us briefly discuss Theorem 14 since it contains a few interesting technical ideas.

see point **??**. Assume $b \in L^q\left(0, T; L^p\right)$ with $p, q > 2$,

$$\frac{d}{p} + \frac{1}{q} < 1.$$

The following theorem about the PDE is known [21]: there exists a unique solution with the regularity

$$\frac{\partial U}{\partial t}, \quad D^2 U \in L^q\left(0, T; L^p\right)$$

and such that

$$\left\| DU \right\|_\infty \le C\left(\lambda\right)$$

with $\lim_{\lambda \to \infty} C\left(\lambda\right) = 0$. Notice that similar statements are true under weaker integrability conditions except that DU is no more bounded, in such cases.

Go back to the proof of pathwise uniqueness, where we proved:

$$E\left[\left|X_t^{(1)} - X_t^{(2)}\right|^2\right] \le C \int_0^t E\left[\left|X_s^{(1)} - X_s^{(2)}\right|^2\right] ds + 2E \int_0^t \left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds.$$

Now we can bound

$$\begin{aligned} & E \int_0^t \left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds \\ \le \quad & E \int_0^t \left|\int_0^1 D^2 U\left(s, \alpha X_s^{(1)} + \left(1 - \alpha\right) X_s^{(2)}\right) d\alpha\right|^2 \left|X_s^{(1)} - X_s^{(2)}\right|^2 ds \end{aligned}$$

but then, how to decuple the expectation, to apply Gronwall lemma? Moreover, $D^2 U$ is not bounded, just integrable. One can solve these two difficulties with two tricks.

Preliminary, we rewrite

$$X_t = x + U\left(0, x\right) - U\left(t, X_t\right) + \lambda \int_0^t U\left(s, X_s\right) ds + \int_0^t \left(1 + DU\left(s, X_s\right)\right) dW_s$$

as

$$\begin{aligned} Y_t &= Y_0 + \lambda \int_0^t U\left(s, \Phi_s^{-1}\left(Y_s\right)\right) ds + \int_0^t \left(1 + DU\left(s, \Phi_s^{-1}\left(Y_s\right)\right)\right) dW_s \\ dY_t &= U\left(t, \Phi_t^{-1}\left(Y_t\right)\right) dt + \left(1 + DU\left(t, \Phi_t^{-1}\left(Y_t\right)\right)\right) dW_t \end{aligned}$$

for the new process

$$Y_t = \Phi_t(X_t)$$

where

$$\Phi_t(x) = x + U(t,x).$$

For $\lambda \geq 0$ so large that $\|DU\|_\infty \leq 1/2$ the transformation $x \longmapsto \Phi_t(x)$ is smoothly invertible. Let $X_t^{(i)}$, $i = 1,2$, be two solutions and let $Y_t^{(i)} = \Phi_t\left(X_t^{(i)}\right)$. This transformation could be applied from the very beginning and reduces the initial SDE to one with smooth coefficients (for instance, one can apply immediately the theory of stochastic flows of [23]). This transformation however does not solve the two problems above, since we get (M_t is martingale)

$$\left|Y_t^{(1)} - Y_t^{(2)}\right|^2 \leq C\int_0^t \left|Y_s^{(1)} - Y_s^{(2)}\right|^2 ds + M_t + \int_0^t \left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds$$

and the same difficulties are met.

The first trick is probably due to [28]. Introduce

$$A_t = \int_0^t \frac{\left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2}{\left|X_s^{(1)} - X_s^{(2)}\right|^2} 1_{X_s^{(1)} \neq X_s^{(2)}} ds.$$

We may prove

$$E[A_t] < \infty.$$

We omit the interesting proof (see [22], [11]) conceptually based on the so called Krylov estimates.

The second trick consists in applying Itô formula to $e^{-A_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^2$, not only to $\left|Y_t^{(1)} - Y_t^{(2)}\right|^2$ as done above. We get

$$\begin{aligned} e^{-A_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^2 \quad &\leq \quad -\int_0^t e^{-A_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^2 dA_s + C\int_0^t e^{-A_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^2 ds + M_t \\ &\quad + \int_0^t e^{-A_s}\left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds \end{aligned}$$

and now,

$$\begin{aligned} &\int_0^t e^{-A_s}\left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds \\ = &\int_0^t e^{-A_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^2 dA_s. \end{aligned}$$

Thus the negative term cancels the bad term $\int_0^t e^{-A_s}\left|\nabla U\left(s, X_s^{(1)}\right) - \nabla U\left(s, X_s^{(2)}\right)\right|^2 ds$. We deduce

$$E\left[e^{-A_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^2\right] \leq C\int_0^t E\left[e^{-A_s}\left|Y_s^{(1)} - Y_s^{(2)}\right|^2\right] ds$$

which implies

$$E\left[e^{-A_t}\left|Y_t^{(1)} - Y_t^{(2)}\right|^2\right] \leq e^{CT}\left|x^{(1)} - x^{(2)}\right|^2$$

by Gronwall. This allows to prove pathwise uniqueness.

Since one can even prove

$$E\left[e^{\lambda A_t}\right]<\infty$$

for all $\lambda>0$, by Hölder inequality one can also deduce

$$E\left[\left|Y_t^{(1)}-Y_t^{(2)}\right|^2\right]\leq C\left|x^{(1)}-x^{(2)}\right|^2$$

and similarly

$$E\left[\left|Y_t^{(1)}-Y_t^{(2)}\right|^p\right]\leq C_p\left|x^{(1)}-x^{(2)}\right|^p$$

with some more work, so that Kolmogorov test applies and one gets an Hölder flow.

2 Lecture 2: zero noise limit for SDEs with rough drift

2.1 Introduction

Consider, as in the first lecture, the SDE in $\mathbb{R}^d$

$$dX_t^\varepsilon=b\left(X_t^\varepsilon\right)dt+\varepsilon dW_t,\qquad X_0^\varepsilon=x$$

with $b:\mathbb{R}^d\to\mathbb{R}^d$ at least measurable, $W=(W_t)_{t\geq 0}$ a BM in $\mathbb{R}^d$ on a probability space $(\Omega,\mathcal{F},P)$ and $\varepsilon\in[0,1]$. We have seen that under very general conditions on b (much less than Lipschitz continuous) there exists a unique solution when $\varepsilon>0$. Sometimes we shall emphasize the dependence on the initial condition and write $X_t^{x,\varepsilon}$ for the solution.

The deterministic equation ($\varepsilon=0$) however may have non-uniqueness, as we have seen in simple examples. Thus it is interesting to investigate the limit of X_t^ε as $\varepsilon\to 0$.

When b is Lipschitz continuous and thus there is a unique solution X_t^0 also for $\varepsilon=0$, one can show that

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,T]}\left|X_t^\varepsilon-X_t^0\right|=0$$

a.s. and the relevant theory becomes LDP theory, where roughly speaking it is proved that

$$\begin{aligned}P\left(X_{[0,T]}^\varepsilon\in A\right)&\sim\exp\left(-\varepsilon^2\inf_A I_{[0,T]}\right)\\ I_{[0,T]}(g)&=\int_0^T\left|g'(t)-b(g(t))\right|^2dt\end{aligned}$$

and other relevant results on exit times from domains and so on, see [17].

When non-uniqueness for $\varepsilon=0$ may happen, a priori we may expect that the family $\left(P_\varepsilon^x\right)_{\varepsilon\in(0,1)}$ of laws of $X_{[0,T]}^\varepsilon$ on $C\left([0,T];\mathbb{R}^d\right)$ has more than one limit point. We would like to have criteria such that there is a unique limit point P^x and characterize the support of P^x, which would represent a *criterion of selection for the deterministic equation.*

The topics treated in the next subsections are taken only from [5], [9], [27]; however, see also [4], [6], [7], [10], [18], [19], [24], [31] for other results.

2.2 Examples

In $d=1$, we have seen that non-uniqueness for $\varepsilon=0$ may happen only at points x where $b(x)=0$. We shall consider then the simplest situation where b has only one such point, $x=0$ say, where it is Hölder continuous but not Lipschitz continuous; and we assume it is smooth for $x\neq 0$. Cases considerably different from this one has not been investigated (except little extensions like $b(x)=sign(x)$). The paradigm is the function

$$b(x)=\begin{cases} A^{+}|x|^{\alpha} & \text{for} \quad x\geq 0 \\ -A^{-}|x|^{\alpha} & \text{for} \quad x<0 \end{cases}$$

for $\alpha\in(0,1)$ and $A^{+},A^{-}>0$. We recall that it has the two extremal solutions

$$x_t^{+}=C_{\alpha,A^{+}}t^{\frac{1}{1-\alpha}}, \quad x_t^{-}=-C_{\alpha,A^{-}}t^{\frac{1}{1-\alpha}},$$

with $C_{\alpha,A^{\pm}}=(A^{\pm}(1-\alpha))^{1/(1-\alpha)}$ and infinitely many other solutions, given by $x=0$ for all times or over some $[0,t_0]$ after which it holds $\pm C(t-t_0)^{\frac{1}{1-\alpha}}$ with C as above.

In $d>1$ one of the problems is to have a precise problem in mind. It is not so clear which b should we investigate, which are the paradigmatic examples. For this reason, I shall restrict myself to a particular example that I met in a specific application, and mention another example that I cannot treat yet. The specific example is

$$\begin{aligned} dX_t^{\varepsilon} &= V_t^{\varepsilon}dt \\ dV_t^{\varepsilon} &= sign(X_t^{\varepsilon})dt+\varepsilon dW_t \end{aligned}$$

with

$$X_0^{\varepsilon}=0, \qquad V_0^{\varepsilon}=0.$$

It comes from two particles interacting with Vlasov-Poisson kernel in 1D, with independent BM on each one. The "extremal" solutions for $\varepsilon=0$ are $\pm\left(\frac{t^2}{2},t\right)$ and the others are equal to zero for ever or for a while, before diverging as these two ones.

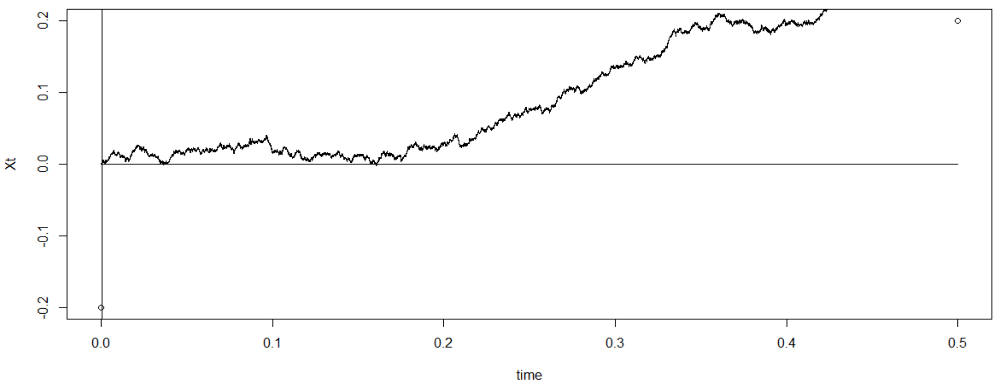
Another example in $d>1$, that however we still do not know how to treat, is

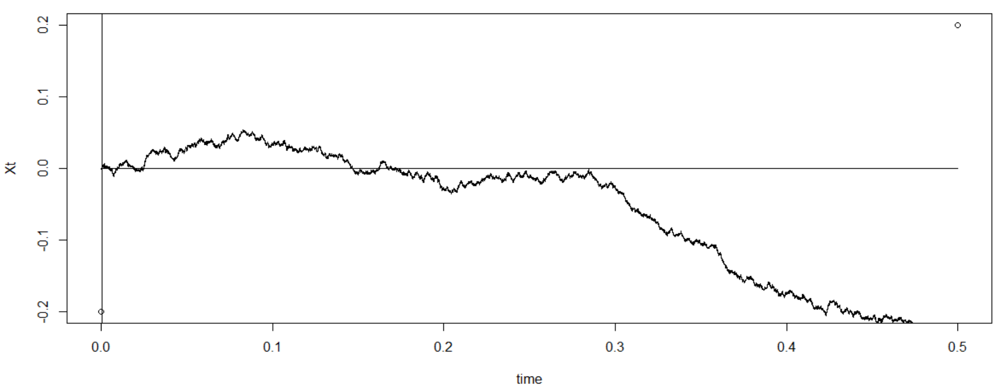
$$b(x)=|x|^{\alpha(x)-1}x$$

where $\alpha(x)$ depends for instance on $\frac{x}{|x|}\in S^{d-1}$. If α is constant, we may reduce the problem to the one-dimensional case, but for non-constant α, for instance in $d=2$ equal to a value in the first quadrant and to a larger value in the others, the answer is not completely clear even from an intuitive or numeric viewpoint.

2.2.1 Simulations

The problem is not easy, as the following two picture show.





Both are typical samples from the equation

$$dX_t = 2sign\,(X_t)\,\sqrt{|X_t|}dt + \varepsilon dW_t, \qquad X_0 = 0$$

with

$$\varepsilon = 0.1$$

(for much smaller ε one can see the same kind of pictures just by rescaling time and space, as explained in Section 2.7 below). One should remember that Brownian motion intersects the line $x = 0$ infinitely many times on any interval $[0, \tau]$, $\tau > 0$. Moreover, for very small times, the "$\sqrt{t}$" contribution of W_t, even if multiplied by ε, is much stronger than the $o\,(t)$ contribution of $\int_0^t 2sign\,(X_s)\,\sqrt{|X_s|}ds$, in the decomposition

$$X_t = \int_0^t 2sign\,(X_s)\,\sqrt{|X_s|}ds + \varepsilon W_t.$$

Thus, close to $t = 0$, also X_t crosses the line $x = 0$ infinitely many times.

Thus, for a little while, X_t remains close to the line $x = 0$ and lives on both sides of it. When does it make the choice between the upper and the lower branch? Could it be that this choice is arbitrarily delayed?

2.3 General remarks

First, let us remark that it is not natural to look for results of convergence different from the weak one in law. Attanasio [3] has proved that convergence in probability does not hold, in the usual $|x|^\alpha$ example and more generally.

Given the initial condition x of the Cauchy problem and the time interval $[0,T]$, we shall denote by

$$\mathcal{S} \subset C\left([0,T];\mathbb{R}^d\right)$$

the set of all solutions of the deterministic problem. Under the next assumptions and others, it is non-empty (Peano Theorem) and closed, in $C\left([0,T];\mathbb{R}^d\right)$, hence it is a Borel set.

The purpose of the next pages is to treat the case when b is not Lipschitz continuous. For instance, let us consider the case

$$b \in C\left(\mathbb{R}^d, \mathbb{R}^d\right), \qquad |b(x)| \leq C(1+|x|). \tag{1}$$

Under this assumption, P^ε is unique for all $\varepsilon > 0$ (but not for $\varepsilon = 0$). The following results have a routine proof.

Lemma 15 *Under assumption (1), the family* $(P_\varepsilon^x)_{\varepsilon\in(0,1)}$ *is tight in* $\mathcal{S}$.

Lemma 16 *Under assumption (1), any weak limit point* P^x *of the family* $(P_\varepsilon^x)_{\varepsilon\in(0,1)}$ *in* $C\left([0,T];\mathbb{R}^d\right)$ *is concentrated on* $\mathcal{S}$ *:* $P^x(\mathcal{S}) = 1$.

The question addressed by the previous lemma is less obvious when b is not continuous (see for instance [7]). However, in the few discontinuous cases treated below, which essentially involve only the *sign* function, the proof of the same statement can be made by simple arguments.

Lemma 17 *Assume there exists a continuous map* $\mathcal{T} : C\left([0,T];\mathbb{R}^d\right) \to C\left([0,T];\mathbb{R}^d\right)$ *and a set* $S^* \subset \mathcal{S}$ *such that*

- $\mathcal{T}(P_\varepsilon^x) = P_\varepsilon^x$ *for all* $\varepsilon > 0$
- *any weak limit point* P^x *of* $(P_\varepsilon^x)_{\varepsilon\in(0,1)}$ *satisfies* $P^x(S^*) = 1$
- *there exists at most one* P^x *with the properties*

$$\begin{aligned} \mathcal{T}(P^x) &= P^x \\ P^x(S^*) &= 1. \end{aligned}$$

Then there is only one weak limit point P^x *of the family* $(P_\varepsilon^x)_{\varepsilon\in(0,1)}$ *and* $P_\varepsilon^x \rightharpoonup P^x$.

2.3.1 Separating classes

Let $S \subset C\left([0,\infty),\mathbb{R}^d\right)$ be the set of all solutions of the deterministic Cauchy problem from a given x_0.

We call (set-theretical) *selection* any (closed) subset $S^* \subset S$.

Definition 18 *We say that* $\{C(t)\}_{t\geq 0}$ *is a separating class for the selection* S^* *if*
i) $C(t)$ *are open sets of* $\mathbb{R}^d$
ii) $g \in (S^*)^c$ *if and only if there is* $t_0 > 0$ *such that* $g(t) \in C(t)$ *for all* $t \in (0,t_0)$.

Example 19 *1D (usual example):* $C(t) = t^{\frac{1}{1-\alpha}}(-\gamma,\gamma)$ *for* $\gamma > 0$ *small enough.*

Example 20 *Vlasov:* $C(t) = \left\{(x,v) : 2|x| + |v|^2 < t^2\right\}$.

Proposition 21 *Assume that, for some* $t_0 > 0$ *and every* $t \in (0,t_0)$

$$\lim_{\varepsilon\to 0} P_\varepsilon^x\left(g(t) \in C(t)\right) = 0. \tag{2}$$

Then any limit point P^x *is supported on* S^*.

Used in conjunction with Lemma 17 (when applicable!), it provides the simplest strategy to solve completely the problem of the zero-noise limit. However, the proof of (2) is usually very hard. Let us isolate an example in which it is quite easy.

2.3.2 Example

This example is taken from [27], where local times are used in elegant way to solve the problem. We have

$$dX_t^\varepsilon = sign\left(X_t^\varepsilon\right)dt + \varepsilon dW_t, \qquad X_0^\varepsilon = 0$$

$$\begin{aligned}
|X_t^\varepsilon| &= \int_0^t sign\left(X_s^\varepsilon\right)^2 ds + \varepsilon\int_0^t sign\left(X_s^\varepsilon\right)dW_s + L_{X^\varepsilon}^0(t) \\
&\geq \int_0^t sign\left(X_s^\varepsilon\right)^2 ds + \varepsilon\int_0^t sign\left(X_s^\varepsilon\right)dW_s \\
&= t + \varepsilon\int_0^t sign\left(X_s^\varepsilon\right)dW_s
\end{aligned}$$

(by Girsanov, $X_s^\varepsilon \neq 0$ for a.e. s, with probability one). Hence

$$P\left(|X_t^\varepsilon| < \frac{t}{2}\right) \leq P\left(\varepsilon\left|\int_0^t sign\left(X_s^\varepsilon\right)dW_s\right| < \frac{t}{2}\right) \leq \frac{4\varepsilon^2}{t}.$$

We may apply the separating class argument.

We deduce that any P is concentrated on $\pm t$. By symmetry, it is unique and symmetric and the full family P_ε converges.

2.4 The classical 1D example

Let us solve the zero-noise problem for the classical example

$$b(x) = \begin{cases} |x|^{\alpha} & \text{for} \quad x \geq 0 \\ -|x|^{\alpha} & \text{for} \quad x < 0 \end{cases}$$

for $\alpha \in (0,1)$ (we have taken $A^{+} = A^{-} = 1$ above). Recall that

$$x_t^{\pm} = \pm C_{\alpha} t^{\frac{1}{1-\alpha}}$$

where $C_{\alpha,A} = (1-\alpha)^{1/(1-\alpha)}$.

Theorem 22 *In the weak sense, as $\varepsilon \to 0$,*

$$P_{\varepsilon}^{x} \to \frac{1}{2}\delta_{x_{\cdot}^{+}} + \frac{1}{2}\delta_{x_{\cdot}^{-}}.$$

Remark 23 *For general A^{+}, A^{-} one can prove that*

$$P_{\varepsilon}^{x} \to p^{+}\delta_{x_{\cdot}^{+}} + p^{-}\delta_{x_{\cdot}^{-}},$$

where

$$p^{+} = \frac{(A^{-})^{-\frac{1}{1+\alpha}}}{(A^{+})^{-\frac{1}{1+\alpha}} + (A^{-})^{-\frac{1}{1+\alpha}}}, \quad p^{-} = \frac{(A^{+})^{-\frac{1}{1+\alpha}}}{(A^{+})^{-\frac{1}{1+\alpha}} + (A^{-})^{-\frac{1}{1+\alpha}}},$$

and

$$x_t^{+} = C_{\alpha,A^{+}} t^{\frac{1}{1-\alpha}}, \quad x_t^{-} = -C_{\alpha,A^{-}} t^{\frac{1}{1-\alpha}},$$

with $C_{\alpha,A^{\pm}} = \left(A^{\pm}(1-\alpha)\right)^{1/(1-\alpha)}$.

Remark 24 *When the exponents are different*

$$b(x) = \begin{cases} A|x|^{\alpha} & \text{for} \quad x \geq 0 \\ -B|x|^{\beta} & \text{for} \quad x < 0 \end{cases}$$

with $\alpha, \beta \in (0,1)$ and $A, B > 0$, if $\alpha < \beta$ then "α wins": $P_{\varepsilon}^{x} \to \delta_{x_t^{+}}$.

The proof of Theorem 22, and also of the more general cases mentioned in the remarks, has been given for the first time by [5], which remains the key reference on the subject. Let us also mention [18] for a very interesting LDP result, for the symmetric example.

Recently, two new proofs have been given, by [27] and [9]. We have shown the idea of [27] above for the simpler case of the *sign* function. Let us describe here the scheme of proof of [9] because it identifies the transition point and it has an interesting dynamical character.

We may identify the existence of two regimes. At the beginning of time, the solution which starts from $x = 0$ behaves like the Brownian motion $(\varepsilon W_t)_{t \geq 0}$, although ε is very small, because the drift is much smaller (this happens also for a Lipschitz drift). But close to the time-space points

$$(t_{\varepsilon}, x_{\varepsilon}) := \left(\varepsilon^{\frac{2(1-\alpha)}{1+\alpha}}, x_{t_{\varepsilon}}^{\pm}\right) = \left(\varepsilon^{\frac{2(1-\alpha)}{1+\alpha}}, \pm C_{\alpha}\varepsilon^{\frac{2}{1+\alpha}}\right) \quad (3)$$

a transition occurs: the drift becomes much stronger than the noise (the usual fluctuations of the noise do not contrast the drift anymore) and pushes the trajectories far away from the neighborhood of $x = 0$, roughly along one of the trajectories $(x_t^{\pm})_{t\geq 0}$.

The easiest way to identify heuristically the transition time-space point (3) is to compare x_t^+ (or x_t^-) with εW_t. Forgetting the scale constants C_α in the definition of x_t^+ and x_t^-, we thus claim that the typical time t at which transition occurs is given by the solution of the equation

$$t^{\frac{1}{1-\alpha}} = \varepsilon t^{\frac{1}{2}}. \tag{4}$$

We then get $t_\varepsilon := \varepsilon^{\frac{2(1-\alpha)}{1+\alpha}}$ as typical transition time. Accordingly, we define

$$x_\varepsilon := t_\varepsilon^{\frac{1}{1-\alpha}} = \varepsilon^{\frac{2}{1+\alpha}}$$

as the typical space scale for observing the transition between the two regimes. This simple intuition is confirmed by the proofs below.

For every $r > 0$, let us denote by τ_r the exit time from $(-r, r)$, defined on the canonical space $C([0, +\infty); \mathbb{R})$, as

$$\tau_r(\xi) = \inf\{t > 0 : |\xi_t| \geq r\}, \qquad \xi \in C([0, +\infty); \mathbb{R}),$$

when this set is not empty, $\tau_r(\xi) = +\infty$ otherwise.

Proposition 25 *Let t_ε, x_ε as above. For every function $\widetilde{t}_\varepsilon > t_\varepsilon$ such that $\lim_{\varepsilon\to 0} \widetilde{t}_\varepsilon / t_\varepsilon = +\infty$, we have*

$$\lim_{\varepsilon\to 0} \sup_{x\in(-x_\varepsilon, x_\varepsilon)} P_\varepsilon^x\left(\tau_{x_\varepsilon} > \widetilde{t}_\varepsilon\right) = 0.$$

The proposition states that with high probability the system reaches $\pm x_\varepsilon$ in a time just a little greater than t_ε, when it starts inside $(-x_\varepsilon, x_\varepsilon)$. This fact is mainly due to the fluctuations of the noise, the drift playing a negligible role on a time interval of length of the same order as t_ε. We prove this result in any dimension, under quite general conditions.

Then we show that, starting from x_ε or above, the solution remains above $(1-\gamma)\, x_t^+$, for any arbitrarily prescribed $\gamma \in (0, 1)$, with probability larger than some $\lambda > 0$ (similarly if it starts from $-x_\varepsilon$). The result is complemented by the following fact: for every function $\widetilde{x}_\varepsilon > x_\varepsilon$ such that $\lim_{\varepsilon\to 0} \widetilde{x}_\varepsilon / x_\varepsilon = +\infty$, if we start above $\widetilde{x}_\varepsilon$ then the solution remains above x_t^+ forever, with probability close to one. The formal statement is:

Theorem 26 *Let $\gamma \in (0, 1)$ be given. Then there exists a constant $\lambda_\gamma > 0$, independent of $\varepsilon \in (0, 1)$, such that*

$$\inf_{x\geq x_\varepsilon} P\left(X_t^{x,\varepsilon} \geq (1-\gamma)\, x_t^+, \quad \forall t \geq 0\right) \geq \lambda_\gamma > 0.$$

Moreover, for every function $\widetilde{x}_\varepsilon > x_\varepsilon$ such that $\lim_{\varepsilon\to 0} \widetilde{x}_\varepsilon / x_\varepsilon = +\infty$, we have

$$\lim_{\varepsilon\to 0} \inf_{x\geq \widetilde{x}_\varepsilon} P\left(X_t^{x,\varepsilon} \geq (1-\gamma)\, x_t^+, \quad \forall t \geq 0\right) = 1.$$

Similar results hold for $(x_t^-)_{t\geq 0}$.

These two steps lead to the solution of our problem. Indeed, with large probability, we reach $\pm x_\varepsilon$ in a very short time of order $\widetilde{t}_\varepsilon$ (Proposition 25). Then, iterating the argument of escape with probability larger than λ_γ, we reach a prescribed $\pm\widetilde{x}_\varepsilon$ with probability close to one, again in a short time. Finally, restarting from $\pm\widetilde{x}_\varepsilon$, we escape above or below $(1-\gamma)\,x_t^\pm$ forever (Theorem 26). Using the separating class $C(t) = x_t^+(-\delta,\delta)$ for small $\delta > 0$, and Lemma 17, we get the result.

We omit the details of the argument based on strong Markov property that leads to the final result starting from Proposition 25 and Theorem 26 and limit ourselves to the proof of these two statements.

2.5 Proof of Proposition 25

The proof of this proposition can be made in every dimension. Let $b:\mathbb{R}^d \to \mathbb{R}^d$ be a function such that there exist $M > 0$ and $\alpha \in (0,1)$ such that

$$\sup_{|x|\le r} |b(x)| \le M r^\alpha \qquad \forall r > 0.$$

We set $r_\varepsilon := \varepsilon^{\frac{2}{1+\alpha}}$.

Proposition 27 *Under the previous conditions we have*

$$\sup_{|x|\le r_\varepsilon} P_\varepsilon^x(\tau_{r_\varepsilon} > t_\varepsilon) \le \theta := P(|Z| \le 2 + M)$$

where $Z \sim N(0, Id_d)$ is the standard Gaussian vector in $\mathbb{R}^d$.

Proof. If $\tau_{r_\varepsilon}(X^{x,\varepsilon}) > t_\varepsilon$ then $|X_t^{x,\varepsilon}| \le r_\varepsilon$ for $t \in [0, t_\varepsilon]$. Hence, from

$$X_t^{x,\varepsilon} = x + \int_0^t b(X_s^{x,\varepsilon})\,ds + \varepsilon W_t,$$

with $|x| \le r_\varepsilon$, we get

$$\begin{aligned}\varepsilon|W_t| &\le |X_t^{x,\varepsilon}| + |x| + \int_0^t |b(X_s^{x,\varepsilon})|\,ds \\ &\le 2r_\varepsilon + tMr_\varepsilon^\alpha \le 2r_\varepsilon + t_\varepsilon M r_\varepsilon^\alpha\end{aligned}$$

for $t \in [0, t_\varepsilon]$. Moreover,

$$t_\varepsilon r_\varepsilon^\alpha = r_\varepsilon.$$

Hence

$$\varepsilon|W_t| \le (2+M)\,r_\varepsilon \qquad \text{for } t \in [0, t_\varepsilon].$$

In particular, this implies $\varepsilon|W_{t_\varepsilon}| \le (2+M)\,r_\varepsilon$. Since the random vector $Z := t_\varepsilon^{-1/2} W_{t_\varepsilon}$ is $N(0, Id)$, we deduce that

$$\begin{aligned}P\big(\varepsilon|W_{t_\varepsilon}| \le (2+M)\,r_\varepsilon\big) &= P\left(\left|t_\varepsilon^{-1/2} W_{t_\varepsilon}\right| \le \frac{(2+M)\,r_\varepsilon}{\varepsilon t_\varepsilon^{1/2}}\right) \\ &= P(|Z| \le 2+M) =: \theta < 1.\end{aligned}$$

We have proved

$$P_\varepsilon^x(\tau_{r_\varepsilon} > t_\varepsilon) \le \theta.$$

■

Corollary 28 *Assume that, for every $\varepsilon \in (0,1)$, the solutions $(X_t^{x,\varepsilon})_{t\geq 0}$ are a strong Markov family w.r.t. the initial condition. Then, for any integer $n \geq 1$,*

$$\sup_{|x|\leq r_\varepsilon} P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon\right) \leq \theta^n.$$

In particular, for every function $\widetilde{t}_\varepsilon > t_\varepsilon$ such that $\lim_{\varepsilon\to 0} \widetilde{t}_\varepsilon / t_\varepsilon = +\infty$, we have

$$\lim_{\varepsilon\to 0} \sup_{|x|\leq r_\varepsilon} P_\varepsilon^x \left(\tau_{r_\varepsilon} > \widetilde{t}_\varepsilon\right) = 0.$$

Proof. We know that

$$P_\varepsilon^x \left(\tau_{r_\varepsilon} > t_\varepsilon\right) \leq \theta = P\left(|Z| \leq 2 + M\right).$$

We consider the sequence of stopping times

$$\tau_{r_\varepsilon}^{(n)} := (nt_\varepsilon) \wedge \tau_{r_\varepsilon}$$

for $n = 0, 1, ...$ and apply strong Markov property to get

$$\begin{aligned} P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon\right) &= P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon \middle| |\xi_{\tau_{r_\varepsilon}^{(n-1)}}| < r_\varepsilon\right) P_\varepsilon^x \left(|\xi_{\tau_{r_\varepsilon}^{(n-1)}}| < r_\varepsilon\right) \\ &\quad + P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon \middle| |\xi_{\tau_{r_\varepsilon}^{(n-1)}}| = r_\varepsilon\right) P_\varepsilon^x \left(|\xi_{\tau_{r_\varepsilon}^{(n-1)}}| = r_\varepsilon\right) \\ &\leq \theta P_\varepsilon^x \left(\tau_{r_\varepsilon} > (n-1)\, t_\varepsilon\right), \end{aligned}$$

because $P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon \middle| |\xi_{\tau_{r_\varepsilon}^{(n-1)}}| = r_\varepsilon\right) = 0$ and $P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon \middle| |\xi_{\tau_{r_\varepsilon}^{(n-1)}}| < r_\varepsilon\right) \leq \theta$. Therefore

$$P_\varepsilon^x \left(\tau_{r_\varepsilon} > nt_\varepsilon\right) \leq \theta^n.$$

■

2.6 Proof of Theorem 26

Let τ_{γ,x^+} be the random time, defined on the canonical space $C([0,+\infty);\mathbb{R})$:

$$\tau_{\gamma,x^+}(\xi) = \inf\left\{t > 0 : \xi(t) < (1-\gamma)\, x_t^+\right\} \tag{5}$$

(equal to $+\infty$ if this event never happens).

For an initial condition $x \geq x_\varepsilon$, we have $X_0^{x,\varepsilon} > (1-\gamma)\, x_0^+ = 0$ at time zero. Since the processes are continuous, $P_x^\varepsilon\left(\tau_{\gamma,x^+} > 0\right) = 1$. For $s \in \left[0, \tau_{\gamma,x^+}(X^{x,\varepsilon})\right]$ (all $s \geq 0$ if $\tau_{\gamma,x^+}(X^{x,\varepsilon}) = +\infty$), we have $X_s^{x,\varepsilon} \geq (1-\gamma)\, x_s^+$, so that (by definition of b)

$$b\left(X_s^{x,\varepsilon}\right) \geq (1-\gamma)^\alpha\, b\left(x_s^+\right).$$

Therefore, for every $t \in \left[0, \tau_{\gamma,x^+}(X^{x,\varepsilon})\right]$,

$$\begin{aligned} X_t^{x,\varepsilon} &= x + \int_0^t b\left(X_s^{x,\varepsilon}\right) ds + \varepsilon W_t \\ &\geq x + (1-\gamma)^\alpha \int_0^t b\left(x_s^+\right) ds + \varepsilon W_t = x + (1-\gamma)^\alpha\, x_t^+ + \varepsilon W_t, \end{aligned}$$

because $x_t^+ = \int_0^t b(x_s^+)\,ds$.

Now consider $\eta \in (0,1)$ such that $1-\eta$ is the mid point between $(1-\gamma)$ and $(1-\gamma)^\alpha$. We have (with equal distance)

$$(1-\gamma) < 1-\eta < (1-\gamma)^\alpha .$$

We rewrite the inequality above in the form

$$\begin{aligned} X_t^{x,\varepsilon} &\geq (1-\eta)\, x_t^+ + R_t^{x,\varepsilon,\gamma}, \\ R_t^{x,\varepsilon,\gamma} &:= x + (\gamma-\eta)\, x_t^+ + \varepsilon W_t, \end{aligned}$$

which, we recall, holds for every $t \in \left[0, \tau_{\gamma,x^+}(X^{x,\varepsilon})\right]$. Letting $A(x,\varepsilon,\gamma)$ be the event

$$A(x,\varepsilon,\gamma) = \left\{R_t^{x,\varepsilon,\gamma} \geq 0, \quad \forall t \geq 0\right\},$$

we deduce from next lemma that

$$\begin{aligned} &\inf_{x \geq x_\varepsilon} P(A(x,\varepsilon,\gamma)) \geq \lambda_\gamma > 0 \\ &\lim_{\varepsilon\to 0} \inf_{x \geq g(\varepsilon)x_\varepsilon} P(A(x,\varepsilon,\gamma)) = 1, \end{aligned}$$

whenever $g(\varepsilon) \to +\infty$ as $\varepsilon \to 0$. This implies the two claims of the theorem. Indeed, on the event $A(x,\varepsilon,\gamma)$, it holds

$$X_t^{x,\varepsilon} \geq (1-\eta)\, x_t^+ > (1-\gamma)\, x_t^+$$

for every $t \in \left[0, \tau_{\gamma,x^+}(X^{x,\varepsilon})\right]$. But this is compatible only with the statement $\tau_{\gamma,x^+}(X^{x,\varepsilon}) = +\infty$. Hence $A(x,\varepsilon,\gamma) \subset \left\{\tau_{\gamma,x^+}(X^{x,\varepsilon}) = +\infty\right\}$. The proof of Theorem 26 is complete.

We now prove

Lemma 29 *Given $A > 0$, there is a constant λ_A, independent of ε, such that*

$$P\left(\varepsilon^{\frac{2}{1+\alpha}} + At^{\frac{1}{1-\alpha}} + \varepsilon W_t > 0, \text{ for all } t \geq 0\right) \geq \lambda_A > 0.$$

Moreover, given $g(\varepsilon)$ such that $\lim_{\varepsilon\to 0} g(\varepsilon) = +\infty$,

$$\lim_{\varepsilon\to 0} P\left(\varepsilon^{\frac{2}{1+\alpha}} g(\varepsilon) + At^{\frac{1}{1-\alpha}} + \varepsilon W_t > 0, \text{ for all } t \geq 0\right) = 1.$$

Proof. Since the process $(W_t)_{t\geq 0}$ has the same law as $(\beta^{-1/2} W_{\beta t})_{t\geq 0}$ for every $\beta > 0$, we have

$$\begin{aligned} &P\left(\varepsilon^{\frac{2}{1+\alpha}} g(\varepsilon) + At^{\frac{1}{1-\alpha}} + \varepsilon W_t > 0, \text{ for all } t \geq 0\right) \\ &= P\left(\beta^{\frac{1}{2}} \varepsilon^{\frac{2}{1+\alpha}-1} g(\varepsilon) + \beta^{\frac{1}{2}} A t^{\frac{1}{1-\alpha}} \varepsilon^{-1} + W_{\beta t} > 0, \text{ for all } t \geq 0\right). \end{aligned}$$

Choose $\beta = \beta_\varepsilon$ so that $\beta_\varepsilon^{\frac{1}{2}} \varepsilon^{\frac{2}{1+\alpha}-1} = 1$, namely $\beta_\varepsilon^{\frac{1}{2}} = \varepsilon^{-\frac{1-\alpha}{1+\alpha}}$. Then, since $\varepsilon = \beta_\varepsilon^{-\frac{1}{2}\frac{1+\alpha}{1-\alpha}}$,

$$\beta_\varepsilon^{\frac{1}{2}} t^{\frac{1}{1-\alpha}} \varepsilon^{-1} = \beta_\varepsilon^{\frac{1}{2}\frac{1+\alpha}{1-\alpha}} \beta_\varepsilon^{\frac{1}{2}} t^{\frac{1}{1-\alpha}} = \beta_\varepsilon^{\frac{1}{1-\alpha}} t^{\frac{1}{1-\alpha}} = (\beta_\varepsilon t)^{\frac{1}{1-\alpha}} .$$

Therefore

$$\begin{aligned}
&P\left(\beta_\varepsilon^{\frac{1}{2}}\varepsilon^{\frac{2}{1+\alpha}-1}g\left(\varepsilon\right)+\beta_\varepsilon^{\frac{1}{2}}At^{\frac{1}{1-\alpha}}\varepsilon^{-1}+W_{\beta_\varepsilon t}>0,\ \text{for all } t\geq 0\right)\\
&=P\left(g\left(\varepsilon\right)+A\left(\beta_\varepsilon t\right)^{\frac{1}{1-\alpha}}+W_{\beta_\varepsilon t}>0,\ \text{for all } t\geq 0\right)\\
&=P\left(g\left(\varepsilon\right)+As^{\frac{1}{1-\alpha}}+W_s>0,\ \text{for all } s\geq 0\right).
\end{aligned}$$

Whenever $g(\varepsilon)=1$, the latter probability is positive and independent of ε; we call it λ_A and the first claim of the lemma is proved; whenever $\lim_{\varepsilon\to 0} g(\varepsilon)=+\infty$, the latter probability tends to one (see [9] for details). . ■

2.7 Scaling and canonical equation

Some problems with special scaling properties allow to convert the zero-noise problem into an asymptotic problem ($t\to\infty$) for an equation independent of ε, that we call the *canonical equation.*

Consider

$$dX_t^\varepsilon = sign\left(X_t^\varepsilon\right)\left|X_t^\varepsilon\right|^\alpha dt+\varepsilon dW_t,\qquad X_0^\varepsilon=0.$$

Set

$$Y_t^\varepsilon = hX_{tk}^\varepsilon.$$

Then ($W'_{tk}\sim\frac{W_{tk+\epsilon k}-W_{tk}}{\epsilon k}=k^{1/2}\frac{k^{-1/2}W_{(t+\epsilon)k}-k^{-1/2}W_{tk}}{\epsilon k}\overset{\mathcal{L}}{=}k^{-1/2}\frac{W_{t+\epsilon}-W_t}{\epsilon}=k^{-1/2}W'_t$)

$$\begin{aligned}
dY_t^\varepsilon &= hk\left(d_sX_s^\varepsilon\right)_{s=tk}\\
&= hksign\left(X_{tk}^\varepsilon\right)\left|X_{tk}^\varepsilon\right|^\alpha dt+hk\varepsilon\left(d_sW_s\right)_{s=tk}\\
&\overset{\mathcal{L}}{=} h^{1-\alpha}ksign\left(Y_t^\varepsilon\right)\left|Y_t^\varepsilon\right|^\alpha dt+hk^{1/2}\varepsilon dW_t
\end{aligned}$$

and now we try to set

$$\begin{aligned}
hk^{1/2}\varepsilon &= 1\\
h^{1-\alpha}k &= 1.
\end{aligned}$$

This has the solution $k=h^{\alpha-1}$, $hh^{\alpha/2-1/2}\varepsilon=1$, $h^{(\alpha+1)/2}\varepsilon=1$

$$\begin{aligned}
h &= \varepsilon^{-2/(1+\alpha)}\\
k &= \varepsilon^{2(1-\alpha)/(1+\alpha)}.
\end{aligned}$$

Proposition 30 *If X_t^ε solves*

$$dX_t^\varepsilon = sign\left(X_t^\varepsilon\right)\left|X_t^\varepsilon\right|^\alpha dt+\varepsilon dW_t,\qquad X_0^\varepsilon=0$$

then $Y_t^\varepsilon=h_\varepsilon X_{tk_\varepsilon}^\varepsilon$ with

$$\begin{aligned}
h_\varepsilon &= \varepsilon^{-2/(1+\alpha)}\\
k_\varepsilon &= \varepsilon^{2(1-\alpha)/(1+\alpha)}
\end{aligned}$$

solves (weakly)

$$dY_t = sign\left(Y_t\right)\left|Y_t\right|^\alpha dt+dW_t,\qquad Y_0=0$$

This equation is independent of ε and thus we call it the *canonical equation* (corresponding to the original SDE). It allows us to translate the problem into a $t \to \infty$ one. First, notice that

$$\lim_{\varepsilon \to 0} P\left(|X_t^\varepsilon| < \gamma t^{\frac{1}{1-\alpha}}\right) = 0$$

is a criterion of separation, for sufficiently small $\gamma > 0$ (precisely any $\gamma < (1-\alpha)^{\frac{1}{1-\alpha}}$).

Proposition 31 *If*

$$\lim_{s \to \infty} P\left(|Y_s| < \gamma s^{\frac{1}{1-\alpha}}\right) = 0$$

then, for every $t > 0$,

$$\lim_{\varepsilon \to 0} P\left(|X_t^\varepsilon| < \gamma t^{\frac{1}{1-\alpha}}\right) = 0$$

and this (for small γ) implies that P_ε converges to $\frac{1}{2}\delta_{x^+} + \frac{1}{2}\delta_{x^-}$, $x^\pm(t) = (1-\alpha)^{\frac{1}{1-\alpha}} t^{\frac{1}{1-\alpha}}$.

Proof.

$$\begin{aligned}
\lim_{\varepsilon \to 0} P\left(|X_t^\varepsilon| < \gamma t^{\frac{1}{1-\alpha}}\right) &= \lim_{\varepsilon \to 0} P\left(h_\varepsilon \left|X^\varepsilon_{tk_\varepsilon k_\varepsilon^{-1}}\right| < h_\varepsilon \gamma t^{\frac{1}{1-\alpha}}\right) \\
&= \lim_{\varepsilon \to 0} P\left(\left|Y_{tk_\varepsilon^{-1}}\right| < h_\varepsilon k_\varepsilon^{\frac{1}{1-\alpha}} \gamma \left(tk_\varepsilon^{-1}\right)^{\frac{1}{1-\alpha}}\right) \\
&= \lim_{s \to \infty} P\left(|Y_s| < \gamma s^{\frac{1}{1-\alpha}}\right) = 0
\end{aligned}$$

because

$$h_\varepsilon k_\varepsilon^{\frac{1}{1-\alpha}} = 1.$$

■

2.7.1 The asymptotic problem for the canonical equation

We have reduced the problem to show that

$$\lim_{t \to \infty} P\left(|Y_t| < \gamma t^{\frac{1}{1-\alpha}}\right) = 0$$

for $\gamma > 0$ small enough, where

$$dY_t = sign\,(Y_t)\,|Y_t|^\alpha\, dt + dW_t, \qquad Y_0 = 0$$

The idea we are going to describe is similar to the one already used above: we reach (by noise) a certain threshold and then that from that threshold we go to infinity with the prescribed speed; all up to arbitrarily small probability events. The main difference with respect to the argument above is that we do not need to control constants so precisely.

Remark 32 *In a sense, the argument now looks more of ergodic type. It would be interesting to restate the problem as a true ergodic question and apply known techniques.*

Lemma 33 *For any $R > 0$,*

$$P\left(|Y_t| \le R \text{ for all } t \ge 0\right) = 0.$$

This lemma can be proved in several ways and it is essentially due to the fact that the noise is non-degenerate; think for instance to the support theorems of Stroock-Varadhan. Let us repeat that we do not need any control on the exit time from $[-R, R]$, opposite to the proof above.

When, at time $t_0(\omega)$, we are at level R, for a very large R, the solution is for a long time over level $R/2$. During this period we have

$$Y_t(\omega) \geq R + \left(\frac{R}{2}\right)^{\alpha} (t - t_0(\omega)) + W_t(\omega) - W_{t_0(\omega)}(\omega).$$

Since the growth of $W_t(\omega) - W_{t_0(\omega)}(\omega)$ is much slower than linear, it is intuitively clear that the term $\left(\frac{R}{2}\right)^{\alpha} (t - t_0(\omega))$ dominates and we get that:

Lemma 34 *For any $R > 0$,*

$$P\left(|Y_t| \geq R \text{ eventually}\right) = 1.$$

Let us complete the argument only in the simplest case $\alpha = 0$, otherwise we have to use a simplified version of the proof of Theorem 26.

Assume then $\alpha = 0$ and that after time $t_1(\omega)$ we have $|Y_t| \geq 1$. Then, for $t \geq t_1(\omega)$,

$$Y_t(\omega) \geq 1 + t - t_1(\omega) + W_t(\omega) - W_{t_1(\omega)}(\omega).$$

Since $\lim_{t\to\infty} \frac{W_t(\omega) - W_{t_1(\omega)}(\omega)}{t - t_1(\omega)} = 0$, we get

$$Y_t(\omega) \geq 1 + \frac{1}{2}(t - t_1(\omega))$$

for $t \geq t_2(\omega)$ which is stronger than

$$\lim_{t\to\infty} P\left(|Y_t| < \frac{1}{4}t\right) = 0.$$

Remark 35 *This scaling argument adapts very well to the Vlasov example, although it requires more work. One has to prove the separating condition*

$$\lim_{\varepsilon\to 0} P\left(2\left|X_t^{0,\varepsilon}\right| + \left|V_t^{0,\varepsilon}\right|^2 < t^2\right) = 0$$

which implies that any P^0 is concentrated on $\pm\left(\frac{t^2}{2}, t\right)$ and that by symmetry, P^0 is unique and symmetric and the full family P_ε^0 converges. To prove the separating condition one can apply a scaling argument and restate it as an asymptotic problem ($t \to \infty$) for the canonical equation

$$\begin{aligned} dX_t &= V_t dt \\ dV_t &= sign(X_t)\, dt + dW_t. \end{aligned}$$

Now the argument is more involved than in the 1D case since it is necessary to divide the space in more regions, but at the end the key properties are similar to the case just described in the 1D case.

References

[1] M. Aizenman, On vector fields as generators of flows: a counterexample to Nelson's conjecture, *Ann. Math.* (2) **107** (1978), no. 2, 287–296.

[2] S. Attanasio, Stochastic flows of diffeomorphisms for one-dimensional SDE with discontinuous drift, *Electron. Commun. Probab.* **15** (2010), 213–226.

[3] S. Attanasio, Master thesis.

[4] S. Attanasio, F. Flandoli, Zero-noise solutions of linear transport equations without uniqueness: an example, *C. R. Math. Acad. Sci. Paris* **347** (2009), no. 13-14, 753–756.

[5] R. Bafico, P. Baldi, Small random perturbations of Peano phenomena, *Stochastics* **6** (1982), n. 3, 279-292.

[6] V.S. Borkar, K. Suresh Kumar, A New Markov Selection Procedure for Degenerate Diffusions, *J. Theor. Probab.* **23** (2010), 729–747.

[7] R. Buckdahn, M. Quincampoix, Y. Ouknine, On limiting values of stochastic differential equations with small noise intensity tending to zero, *Bull. Sci. Math.* **133** (2009), 229-237.

[8] G. Da Prato, F. Flandoli, E. Priola, M. Röckner, Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift, to appear on *Annals of Probab.*

[9] F. Delarue, F. Flandoli, The transition point in the zero noise limit for a 1D Peano example, to appear on *Discr. Cont. Dynam. Syst.*

[10] F. Delarue, F. Flandoli, D. Vincenzi, Noise prevents collapse of Vlasov-Poisson point charges, to appear on *Comm. Pure Appl. Math.*

[11] E. Fedrizzi, F. Flandoli, Noise prevents singularities in linear transport equations, *J. Funct. Anal.* **264** (2013), no. 6, 1329–1354.

[12] F. Flandoli, *Random Perturbation of PDEs and Fluid Dynamic Models*, Saint Flour summer school lectures 2010, Lecture Notes in Mathematics n. 2015, Springer, Berlin 2011.

[13] F. Flandoli, M. Gubinelli, E. Priola, Well posedness of the transport equation by stochastic perturbation, *Invent. Math.* **180** (2010), 1-53.

[14] F. Flandoli, M. Gubinelli, E. Priola, Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift, *Bull. Sci. Math.* **134** (2010), no. 4, 405–422.

[15] F. Flandoli, F. Russo, Generalized integration and stochastic ODEs, *Ann. Probab.* **30** (2002), no. 1, 270–292.

[16] F. Flandoli, T. Nilssen, F.N. Proske, in preparation

[17] M. I. Freidlin, A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, Berlin 1998.

[18] M. Gradinaru, S. Herrmann, B. Roynette, A singular large deviations phenomenon, *Ann. Inst. H. Poincaré Probab. Statist.* **37** (2001), no. 5, 555–580.

[19] S. Herrmann, Phénomène de Peano et grandes déviations, *C. R. Acad. Sci. Paris*, Sér. I, Math. **332** (11) (2001) 1019–1024.

[20] N. V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics, 12. American Mathematical Society, Providence, RI, 1996.

[21] N. V. Krylov, The heat equation in Lq((0;T);Lp)spaces with weights, SIAM J. on Math. Anal. 32, 1117-1141 (2001).

[22] N. V. Krylov, M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theory Related Fields* **131** (2005), 154-196.

[23] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, Ecole d'été de probabilités de Saint-Flour, XII—1982, 143-303, Lecture Notes in Math. **1097**, Springer, Berlin, 1984.

[24] S.-E. A. Mohammed, T. Nilssen, F.N. Proske, Sobolev Differentiable Stochastic Flows of SDE's with Measurable Drift and Applications, arXiv:1204.3867.

[25] T. Nilssen, preprint.

[26] D. W. Stroock, S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin 1979.

[27] D. Trevisan, Zero noise limits using local times, arXiv:1301.6510, to appear on *Electr. Comm. Probab.*

[28] Y. A. Veretennikov, On strong solution and explicit formulas for solutions of stochastic integral equations, *Math. USSR Sb.* **39**, 387-403 (1981).

[29] Xicheng Zhang, Stochastic flows of SDEs with irregular coefficients and stochastic transport equations, *Bull. Sci. Math.* **134** (2010), no. 4, 340–378.

[30] Xicheng Zhang, Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients, *Electron. J. Probab.* **16** (2011), no. 38, 1096–1116.

[31] Xicheng Zhang, Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients, *Rev. Mat. Iberoam.* **29** (2013), no. 1, 25–52.