Stochastic scalar first-order conservation laws

Mini-course given at TIFR Bangalore, May 2018, 21-25

J. Vovelle

Contents

1 Homogeneous first-order conservation laws

1.1 Introduction

Let $d \geq 1$ be the space dimension. Let $A \in \mathrm{Lip}(\mathbb{R};\mathbb{R}^d)$ be the flux. Consider the PDE

$$\partial_t u(x,t) + \mathrm{div}_x(A(u(x,t)) = 0, \quad x \in \mathbb{T}^d, \quad t > 0, \tag{1.1}$$

where $\mathbb{T}^d$ is the d-dimensional torus. Eq. (1.1) is a non-linear first-order equation in conservative form. The corresponding non-conservative-form is

$$\partial_t u(x,t) + a(u(x,t)) \cdot \nabla_x u(x,t) = 0, \quad x \in \mathbb{T}^d, \quad t > 0, \tag{1.2}$$

where $a(\xi) := A'(\xi)$.

Transport equation. Consider the simple case $a = \mathrm{Cst}$. The solution to (1.2) with initial datum v is

$$(x,t) \mapsto v(x - ta).$$

The graph of $x \mapsto u(x,t)$ is transported at speed a.

Non-linear case. In the non-linear case, one can solve the equation for characteristics to solve (1.1). This works as long as the solution remains Lipschitz in the space variable. Graphically, on the plot of $x \mapsto v(x)$, this amounts to transport each v-slice at speed $a(v)$. Some simple examples, for example the non-viscous Burgers' equation $a(\xi) = \xi$, with a bump function as initial datum, show that shocks will appear at some time.

Kinetic unknown. Let us emphasize this idea of transport of the graph for solving (1.1). Introduce the characteristic function of the sub-graph of $x \mapsto u(x,t)$: this is the function

$$\mathtt{f}(t,x,\xi) := \mathbf{1}_{u(x,t)>\xi}. \tag{1.3}$$

Solve the free transport equation

$$\partial_t \mathtt{f} + a(\xi) \cdot \nabla_x \mathtt{f} = 0. \tag{1.4}$$

Again, this works until shocks appear again. The kinetic formulation will incorporate an additional term to (1.4) to take into account the formation of shocks and the loss of regularity of solutions.

1.2 Kinetic formulation

Definition 1.1 (Solution)**.** Let $u_0 \in L^\infty(\mathbb{T}^d)$, let $T > 0$. A function

$$u \in L^\infty(\mathbb{T}^d \times [0,T]) \cap C([0,T]; L^1(\mathbb{T}^d))$$

is said to be a solution to (1.1) on $[0,T]$ with initial datum u_0 if u and $\mathtt{f} := \mathbf{1}_{u>\xi}$ have the following properties: there exists a finite non-negative measure m on $\mathbb{T}^d \times \mathbb{R}$ such that, for all $\varphi \in C^1_c(\mathbb{T}^d \times \mathbb{R})$, for all $t \in [0,T]$,

$$\langle \mathtt{f}(t), \varphi\rangle = \langle \mathtt{f}_0, \varphi\rangle + \int_0^t \langle \mathtt{f}(s), a(\xi)\cdot\nabla\varphi\rangle ds - m_\varphi([0,t]), \tag{1.5}$$

where $\mathtt{f}_0(x,\xi) = \mathbf{1}_{u_0(x)>\xi}$ and the measure m_φ is

$$m_\varphi(A) = \iiint_{A\times\mathbb{T}^d\times\mathbb{R}} \partial_\xi\varphi(x,\xi) dm(t,x,\xi), \tag{1.6}$$

for all Borel set $A \subset [0,T]$.

One can give a formulation of solutions that is weak in time also: $\mathtt{f}$ should satisfy

$$\int_0^T \langle \mathtt{f}(t), \partial_t\psi(t)\rangle + \int_0^T \langle \mathtt{f}(t), a(\xi)\cdot\nabla\psi(t)\rangle dt + \langle m, \partial_\xi\psi\rangle + \langle \mathtt{f}_0, \psi(0)\rangle = 0, \tag{1.7}$$

for all $\psi \in C^1_c(\mathbb{T}^d \times [0,T) \times \mathbb{R})$. The formulation (1.7) follows from (1.5) and the Fubini Theorem (consider tensor test functions $\psi\colon (x,t,\xi) \mapsto \varphi(x,\xi)\theta(t)$ first). The converse is true in the context of generalized solutions (see Definition 1.6 below). It is more delicate when considering mere solutions. Indeed, one can deduce from (1.7) that $t \mapsto \langle \mathtt{f}(t), \varphi\rangle$ has right- and left- traces at every point t, but then one has to show that these traces have the representation $\langle \mathtt{f}(t), \varphi\rangle$. Thus, either the continuity in time of the solution, or a result of uniqueness is required to complete the arguments. We will make some specific efforts to work with the formulation (1.5), given at fixed t, because it is better adapted to the study of the stochastic perturbation of (1.1). Let us state the fundamental result of Lions, Perthame, Tadmor 1994, [13] (in which solutions are defined according to (1.7) actually).

Theorem 1.1 (Lions-Perthame-Tadmor 1994, [13])**.** *Let $u_0 \in L^\infty(\mathbb{T}^d)$, let $T > 0$. There exists a unique solution*

$$u \in L^\infty(\mathbb{T}^d \times [0,T]) \cap C([0,T]; L^1(\mathbb{T}^d))$$

to (1.1) on $[0,T]$ with initial datum u_0.

1.2.1 Entropy formulation - kinetic formulation

We have the fundamental identity

$$\int_{\mathbb{R}} (\mathbf{1}_{u>\xi} - \mathbf{1}_{0>\xi})\phi'(\xi) d\xi = \phi(u) - \phi(0), \tag{1.8}$$

for all $\varphi \in C^1(\mathbb{R})$, which establishes a relation between non-linear expressions of u and the integral of $\mathtt{f} := \mathbf{1}_{u>\xi}$ against a test-function in ξ. Using (1.8) in (1.5) with a test function

$$\varphi(x,\xi) = \psi(x)\eta'(\xi),$$

where $\eta \in C^2(\mathbb{R})$ is a convex function and $\psi \in C^1(\mathbb{T}^d)$ is non-negative, one obtains the *entropy inequality*

$$\begin{aligned}\langle \eta(u)(t), \psi \rangle &= \langle \eta(u)(0), \psi \rangle + \int_0^t \langle q(u)(s), \nabla \psi \rangle ds - m(\psi \eta'')([0,t]) \\ &\leq \langle \eta(u)(0), \psi \rangle + \int_0^t \langle q(u)(s), \nabla \psi \rangle ds, \end{aligned} \tag{1.9}$$

where

$$q'(\xi) := \eta'(\xi) a(\xi). \tag{1.10}$$

Note that (1.9) implies the distributional inequality

$$\partial_t \eta(u) + \mathrm{div}_x(q(u)) \leq 0. \tag{1.11}$$

Conversely, one can deduce (1.5) from (1.9) by setting

$$m(\cdot, \xi) = -(\partial_t \eta^+(u;\xi) + \mathrm{div}_x(q^+(u;\xi))),$$

where $\eta^+(u;\xi) = (u-\xi)^+$ is the semi Kruzhkov entropy and

$$q^+(u;\xi) = \mathrm{sgn}_+(u-\xi)(A(u) - A(\xi))$$

the corresponding flux.

1.2.2 Some facts on the defect measure

The measure m is a defect measure regarding the convergence of the parabolic approximation

$$\partial_t u^\varepsilon(x,t) + \mathrm{div}_x(A(u^\varepsilon(x,t)) - \varepsilon \Delta u^\varepsilon(x,t) = 0, \quad x \in \mathbb{T}^d, \quad t > 0, \tag{1.12}$$

to (1.12). Indeed, using the usual chain-rule and (1.8), we infer the kinetic formulation

$$\langle \mathtt{f}^\varepsilon(t), \varphi \rangle = \langle \mathtt{f}_0, \varphi \rangle + \int_0^t \langle \mathtt{f}^\varepsilon(s), a(\xi) \cdot \nabla \varphi - \varepsilon \Delta \varphi \rangle ds - m^\varepsilon_\varphi([0,t]), \tag{1.13}$$

where $\mathtt{f}^\varepsilon(t) = \mathbf{1}_{u^\varepsilon(t) > \xi}$ and

$$\langle m^\varepsilon, \phi \rangle := \iint_{\mathbb{T}^d \times [0,T]} \phi(x,t,u^\varepsilon(x,t)) \varepsilon |\nabla u^\varepsilon(x,t)|^2 dx.$$

With a slight abuse of notation, one writes $m^\varepsilon = \varepsilon |\nabla u^\varepsilon|^2 \delta_{u^\varepsilon = \xi}$. By the energy estimate (this amounts to take $\varphi(x,\xi) = \xi$ in (1.13)), one obtains the bound

$$m^\varepsilon(\mathbb{T}^d \times [0,T] \times \mathbb{R}) \lesssim 1, \tag{1.14}$$

where the notation $A_\varepsilon \lesssim B_\varepsilon$ means that $A_\varepsilon \leq C B_\varepsilon$ for a constant C independent on ε. Using in (1.13) test functions with a higher power, like

$$\varphi(x,\xi) = \int_0^\xi |\zeta| d\zeta,$$

one can also show the tightness condition

$$\iiint_{\mathbb{T}^d \times [0,T] \times \mathbb{R}} |\xi| dm(x,t,\xi) \lesssim 1. \tag{1.15}$$

It follows from (1.14) and (1.15) that, up to a subsequence, $\langle m^\varepsilon, \phi\rangle \to \langle m, \phi\rangle$ for all continuous bounded $\phi\colon \mathbb{T}^d \times [0,T]\times\mathbb{R} \to \mathbb{R}$, where m is a finite non-negative measure on $\mathbb{T}^d \times [0,T]\times\mathbb{R}$. This is what we call the weak convergence of measure (sometimes called narrow convergence of measures). Let us take the limit $\varepsilon \to 0$ in (1.13). We obtain

$$\langle f(t), \varphi\rangle = \langle \mathtt{f}_0, \varphi\rangle + \int_0^t \langle f(s), a(\xi)\cdot\nabla\varphi\rangle ds - m_\varphi([0,t]), \tag{1.16}$$

where $f(t)$ is the "limit" of $\mathtt{f}^\varepsilon(t)$, which has to be specified. We have also to specify the set of times t for which (1.16) is satisfied. For the moment, let us simply remark that, if we assume that (1.16) is true for all $t\in[0,T]$, and if we assume the strong convergence

$$u^\varepsilon \to u \text{ in } C([0,T];L^1(\mathbb{T}^d)),$$

which ensures that $f(t) = \mathtt{f}(t) = \mathbf{1}_{u(t)>\xi}$, then the limit u is a solution to (1.1).

1.3 Kinetic functions

Let us come back to the problem of taking the limit in (1.13) for $\varepsilon \in \varepsilon_\mathbb{N}$, where

$$\varepsilon_\mathbb{N} = \{\varepsilon_n; n\in\mathbb{N}\}, \quad (\varepsilon_n)\downarrow 0.$$

We know that $0 \le \mathtt{f}^\varepsilon \le 1$, therefore, up to subsequence, $\mathtt{f}^\varepsilon \rightharpoonup f$ in $L^\infty(\mathbb{T}^d\times[0,T]\times\mathbb{R})$ weak-$*$, where $0\le f\le 1$ a.e. (note however that nothing guarantees that f has the structure $f = \mathtt{f} = \mathbf{1}_{u>\xi}$). We can say more about $\mathtt{f}^\varepsilon$. Let us introduce the Young measure

$$\nu^\varepsilon_{x,t}(\xi) = -\partial_\xi \mathtt{f}^\varepsilon(x,t,\xi) = \delta_{u^\varepsilon(x,t)=\xi}, \tag{1.17}$$

or, more precisely, for all $\psi\in L^1(\mathbb{T}^d\times[0,T])$, for all $\phi\in C_b(\mathbb{R})$,

$$\iint_{\mathbb{T}^d\times[0,T]} \psi(x,t)\langle\phi,\nu^\varepsilon_{x,t}\rangle dxdt = \iint_{\mathbb{T}^d\times[0,T]} \psi(x,t)\phi(u^\varepsilon(x,t))dxdt. \tag{1.18}$$

We have the tightness estimate (p is any exponent ≥ 1)

$$\sup_{t\in[0,T]}\int_{\mathbb{T}^d}\langle|\xi|^p,\nu^\varepsilon_{x,t}\rangle dx = \sup_{t\in[0,T]}\iint_{\mathbb{T}^d}|u^\varepsilon(x,t)|^p dx \le \int_{\mathbb{T}^d}|u_0(x)|^p dx \lesssim 1. \tag{1.19}$$

This implies in particular that

$$\iint_{\mathbb{T}^d\times[0,T]}\langle|\xi|^p,\nu^\varepsilon_{x,t}\rangle dxdt \lesssim 1. \tag{1.20}$$

By the usual theory of Young Measures, [2], this shows that there exists a Young measure ν such that, up to a given subsequence, for all $\psi\in L^1(\mathbb{T}^d\times[0,T])$, for all $\phi\in C_b(\mathbb{R})$,

$$\iint_{\mathbb{T}^d\times[0,T]} \psi(x,t)\langle\phi,\nu^\varepsilon_{x,t}\rangle dxdt \to \iint_{\mathbb{T}^d\times[0,T]} \psi(x,t)\langle\phi,\nu_{x,t}\rangle dxdt. \tag{1.21}$$

We also know, by lower semi-continuity, that the following slightly weaker form of the estimate (1.19) holds true in the limit:

$$\sup_J \frac{1}{|J|}\iint_{\mathbb{T}^d\times J}\langle|\xi|^p,\nu_{x,t}\rangle dxdt < +\infty, \tag{1.22}$$

where the sup in (1.22) is over open intervals $J \subset [0,T]$. Using (1.21), the estimates (1.20), (1.22), and some approximation arguments, one can show that $\mathtt{f}^\varepsilon \rightharpoonup f$ in $L^\infty(\mathbb{T}^d \times [0,T] \times \mathbb{R})$ weak-$*$, where f is defined by

$$f(x,t,\xi) := \nu_{x,t}(\xi,+\infty). \tag{1.23}$$

What we have gained now is that we know that f has a special structure. We introduce some definitions related to this.

Definition 1.2 (Young measure)**.** Let $(X,\mathcal{A},\lambda)$ be a finite measure space. Let $\mathcal{P}_1(\mathbb{R})$ denote the set of probability measures on $\mathbb{R}$. We say that a map $\nu\colon X \to \mathcal{P}_1(\mathbb{R})$ is a *Young measure* on X if, for all $\phi \in C_b(\mathbb{R})$, the map $z \mapsto \langle \nu_z, \phi\rangle$ from X to $\mathbb{R}$ is measurable. We say that a Young measure ν vanishes at infinity if, for every $p \geq 1$,

$$\int_X \langle |\xi|^p, \nu_z\rangle d\lambda(z) = \int_X \int_{\mathbb{R}} |\xi|^p d\nu_z(\xi) d\lambda(z) < +\infty. \tag{1.24}$$

Definition 1.3 (Kinetic function)**.** Let $(X,\mathcal{A},\lambda)$ be a finite measure space. A measurable function $f\colon X \times \mathbb{R} \to [0,1]$ is said to be a *kinetic function* if there exists a Young measure ν on X that vanishes at infinity such that, for λ-a.e. $z \in X$, for all $\xi \in \mathbb{R}$,

$$f(z,\xi) = \nu_z(\xi,+\infty). \tag{1.25}$$

We say that f is an *equilibrium* if there exists a measurable function $u\colon X \to \mathbb{R}$ with $u \in L^p(X)$ for all finite p, such that $f(z,\xi) = \mathtt{f}(z,\xi) = \mathbf{1}_{u(z)>\xi}$ a.e., or, equivalently, $\nu_z = \delta_{\xi=u(z)}$ for a.e. $z \in X$.

Definition 1.4 (Conjugate function)**.** If $f\colon X \times \mathbb{R} \to [0,1]$ is a kinetic function, we denote by $\bar f$ the conjugate function $\bar f := 1 - f$.

We also denote by χ_f the function defined by $\chi_f(z,\xi) = f(z,\xi) - \mathbf{1}_{0>\xi}$. This correction to f is integrable on $\mathbb{R}$. Actually, it is decreasing faster than any power of $|\xi|$ at infinity. Indeed, we have $\chi_f(z,\xi) = -\nu_z(-\infty,\xi)$ when $\xi < 0$ and $\chi_f(z,\xi) = \nu_z(\xi,+\infty)$ when $\xi > 0$. Therefore

$$|\xi|^p \int_X |\chi_f(z,\xi)| d\lambda(z) \leq \int_X \int_{\mathbb{R}} |\zeta|^p d\nu_z(\zeta) d\lambda(z) < \infty, \tag{1.26}$$

for all $\xi \in \mathbb{R}$, $1 \leq p < +\infty$.
We will use the following compactness result on Young measures (see Proposition 2.3.1 and Corollary 4.3.7 in [2]).

Theorem 1.2 (Compactness of Young measures)**.** *Let $(X,\mathcal{A},\lambda)$ be a finite measure space such that $\mathcal{A}$ is countably generated. Let (ν^n) be a sequence of Young measures on X satisfying the tightness condition*

$$\sup_n \int_X \int_{\mathbb{R}} |\xi|^p d\nu^n_z(\xi) d\lambda(z) < +\infty, \tag{1.27}$$

for all $1 \leq p < +\infty$. Then there exists a Young measure ν on X and a subsequence still denoted (ν^n) such that, for all $h \in L^1(X)$, for all $\phi \in C_b(\mathbb{R})$,

$$\lim_{n\to+\infty} \int_X h(z) \int_{\mathbb{R}} \phi(\xi) d\nu^n_z(\xi) d\lambda(z) = \int_X h(z) \int_{\mathbb{R}} \phi(\xi) d\nu_z(\xi) d\lambda(z). \tag{1.28}$$

For kinetic functions, Theorem 1.2 gives the following corollary (see [5, Corollary 2.5]).

Corollary 1.3 (Compactness of kinetic functions)**.** *Let $(X, \mathcal{A}, \lambda)$ be a finite measure space such that $\mathcal{A}$ is countably generated. Let (f_n) be a sequence of kinetic functions on $X \times \mathbb{R}$, $f_n(z,\xi) = \nu_z^n(\xi,+\infty)$, where the Young measures ν^n are assumed to satisfy (1.27). Then there exists a kinetic function f on $X \times \mathbb{R}$ (related to the Young measure ν in Theorem 1.2 by the formula $f(z,\xi) = \nu_z(\xi,+\infty)$) such that, up to a subsequence, $f_n \rightharpoonup f$ in $L^\infty(X \times \mathbb{R})$ weak-*.*

At last, related to these convergence results, we give the following strong convergence criterion (see [5, Lemma 2.6]).

Lemma 1.4 (Convergence to an equilibrium)**.** *Let $(X, \mathcal{A}, \lambda)$ be a finite measure space. Let $p > 1$. Let (f_n) be a sequence of kinetic functions on $X \times \mathbb{R}$: $f_n(z,\xi) = \nu_z^n(\xi,+\infty)$ where ν^n are Young measures on X satisfying (1.27). Let f be a kinetic function on $X \times \mathbb{R}$ such that $f_n \rightharpoonup f$ in $L^\infty(X \times \mathbb{R})$ weak-*. Assume that f is an equilibrium: $f(z,\xi) = \mathtt{f}(z,\xi) = \mathbf{1}_{u(z)>\xi}$ and let*

$$u_n(z) = \int_{\mathbb{R}} \xi d\nu_z^n(\xi).$$

Then, for all $1 \le q < p$, $u_n \to u$ in $L^q(X)$ strong.

1.4 Generalized solutions

1.4.1 Limit kinetic equation, up to a negligible set

Again, we come back to the problem of taking the limit in (1.13) for $\varepsilon \in \varepsilon_{\mathbb{N}}$. Recall the following

Definition 1.5 (Weak convergence of measures)**.** Let E be a metric space. A sequence of finite Borel measures (μ_n) on E is said to converge weakly to a finite Borel measure μ (denoted $\mu_n \rightharpoonup \mu$) if

$$\langle \mu_n, \phi \rangle \to \langle \mu, \phi \rangle,$$

for all $\phi \in C_b(E)$.

Recall also (this is one of the assertions of the Portmanteau theorem, [1, Theorem 2.1]) that $\mu_n \rightharpoonup \mu$ if, and only if, $\mu_n(A) \to \mu(A)$ for all Borel set A such that $\mu(\partial A) = 0$. Consequently, in (1.13), and by considering the measures on $E = \mathbb{R}_+$, we have

$$m_\varphi^\varepsilon([0,t]) \to m_\varphi([0,t]), \ \forall t \notin B_{\mathrm{at}}, \tag{1.29}$$

where

$$B_{\mathrm{at}} = \{t \in [0,T]; |m_\varphi|(\{t\}) > 0\}. \tag{1.30}$$

The measure $|m_\varphi|$ is the total variation of m_φ. For each $k \in \mathbb{N}^*$, the set $\{t \in [0,T]; |m_\varphi|(\{t\}) \ge k^{-1}\}$ is finite since $|m_\varphi|$ is finite. Therefore B_{at} is at most countable. By the dominated convergence theorem, it follows that the sequence of element $t \mapsto m_\varphi^\varepsilon([0,t])$ is converging to $t \mapsto m_\varphi([0,t])$ in $L^\infty(0,T)$ weak-$*$. Note that we can also simply use the the Fubini theorem to show this result. Indeed, if $\theta \in L^1([0,T])$ is given, we have

$$\int_0^T \theta(t) m_\varphi^\varepsilon([0,t]) dt = \int_{[0,T]} \Theta(t) dm_\varphi^\varepsilon(t) = \langle \Theta, m_\varphi^\varepsilon \rangle, \quad \Theta(t) := \int_t^T \theta(s) ds.$$

Consequently, using the Fubini theorem again, we obtain the convergence

$$\int_0^T \theta(t) m_\varphi^\varepsilon([0,t]) dt \to \langle \Theta, m_\varphi \rangle = \int_0^T \theta(t) m_\varphi([0,t]) dt.$$

We also have the convergence

$$\int_0^t \langle \mathtt{f}^\varepsilon(s), a(\xi)\cdot\nabla\varphi - \varepsilon\Delta\varphi\rangle ds \to \int_0^t \langle f(s), a(\xi)\cdot\nabla\varphi\rangle ds, \tag{1.31}$$

for all $t\in[0,T]$, and thus in $L^\infty(0,T)$ weak-$*$. This shows that $\langle \mathtt{f}^\varepsilon(t),\varphi\rangle$ is converging in $L^\infty(0,T)$ weak-$*$ to a certain quantity

$$F_\varphi(t) := \langle \mathtt{f}_0,\varphi\rangle + \int_0^t \langle f(s), a(\xi)\cdot\nabla\varphi\rangle ds - m_\varphi([0,t]). \tag{1.32}$$

We also know that

$$\int_0^T \langle \mathtt{f}^\varepsilon(t),\varphi\rangle\theta(t)dt \to \int_0^T \langle f(t),\varphi\rangle\theta(t)dt,$$

for all $\theta\in L^1([0,T])$. Consequently, $F_\varphi(t)$ and $\langle f(t),\varphi\rangle$ coincide for a.e. $t\in[0,T]$:

$$\langle f(t),\varphi\rangle = \langle \mathtt{f}_0,\varphi\rangle + \int_0^t \langle f(s), a(\xi)\cdot\nabla\varphi\rangle ds - m_\varphi([0,t]),\ \forall t\in N_0, \tag{1.33}$$

where N_0 has measure zero in $[0,T]$.

1.4.2 Modification as a càdlàg function

Proposition 1.5. *There exists a kinetic function $f^+\colon \mathbb{T}^d\times[0,T]\times\mathbb{R}\to[0,1]$ such that*

1. *$f^+ = f$ a.e. on $\mathbb{T}^d\times[0,T]\times\mathbb{R}$,*
2. *for all $\varphi\in C_c^1(\mathbb{T}^d\times\mathbb{R})$, $t\mapsto\langle f^+(t),\varphi\rangle$ is a càdlàg function,*
3. *the identity*

$$\langle f^+(t),\varphi\rangle = \langle \mathtt{f}_0,\varphi\rangle + \int_0^t \langle f^+(s), a(\xi)\cdot\nabla\varphi\rangle ds - m_\varphi([0,t]),\ \forall t\in[0,T]. \tag{1.34}$$

is satisfied for all $\varphi\in C_c^1(\mathbb{T}^d\times\mathbb{R})$.

Proof. Recall the definition (1.32) of F_φ. Note first that everything reduces to find a kinetic function $f^+\colon \mathbb{T}^d\times[0,T]\times\mathbb{R}$ satisfying the identity $\langle f^+(t),\varphi\rangle = F_\varphi(t)$ for all $\varphi\in C_c^1(\mathbb{T}^d\times\mathbb{R})$, for all $t\in[0,T]$. Indeed, item 1 then follows from (1.33). This in turn implies (1.34) since we can replace $f(s)$ by $f^+(s)$ in the transport term. Item 2 is obvious, since $t\mapsto F_\varphi(t)$ is a càdlàg function. For $t_*\in[0,T)$ fixed, we set

$$\nu^+_{x,t_*} = \lim_{\delta\to0}\frac{1}{\delta}\int_{t_*}^{t_*+\delta}\nu_{x,t}dt,\quad f^+(x,t_*,\xi) = \nu^+_{x,t_*}(\xi,+\infty). \tag{1.35}$$

The limit in (1.35) is in the sense of Young measures on $\mathbb{T}^d$:

$$\int_{\mathbb{T}^d}\psi(x)\langle\phi,\nu^+_{x,t_*}\rangle dx = \lim_{\delta\to0}\int_{\mathbb{T}^d}\psi(x)\langle\phi,\nu^\delta_{x,t_*}\rangle dx,\quad \nu^\delta_{x,t_*} := \frac{1}{\delta}\int_{t_*}^{t_*+\delta}\nu_{x,t}dt, \tag{1.36}$$

for all $\psi\in L^1(\mathbb{T}^d)$, for all $\phi\in C_b(\mathbb{R})$. Let us justify the existence of the limit in (1.35). If $(\delta_n)\downarrow0$, then (1.22) shows that the sequence $(\nu^{\delta_n}_{x,t_*})$ is compact in the sense of Young measures.

It has therefore an adherence value ν^+_{x,t_*} in the sense of (1.36). For f^+ defined as in (1.35), we deduce, for $\varphi \in C^1_c(\mathbb{T}^d \times \mathbb{R})$, that

$$\langle f^+(t_*), \varphi\rangle = \lim_{k\to+\infty} \frac{1}{\delta_{n_k}} \int_{t_*}^{t_*+\delta_{n_k}} \langle f^+(t), \varphi\rangle dt. \tag{1.37}$$

Since $\langle f^+(t), \varphi\rangle = F_\varphi(t)$ for almost all $t \in [0, T]$, (1.37) gives

$$\langle f^+(t_*), \varphi\rangle = \lim_{k\to+\infty} \frac{1}{\delta_{n_k}} \int_{t_*}^{t_*+\delta_{n_k}} F_\varphi(t) dt = F_\varphi(t_*). \tag{1.38}$$

The last identity in (1.38) is due to the fact that F_φ is càdlàg. The relation 1.38 shows that $f^+(t_*)$, and thus $\nu^+_{x,t_*} = -\partial_\xi f^+(t_*)$ are uniquely defined. Consequently, the convergence (1.36) is indeed true for the whole sequence. In this way, we have defined a kinetic function f^+. The identity (1.38) being satisfied at every point t_*, the result follows. □

Eventually, we have shown the convergence of $\mathtt{f}^\varepsilon$ to a *generalized solution* f with initial datum $\mathtt{f}_0$, according to the following definition.

Definition 1.6 (Generalized solution)**.** Let $f_0 \colon \mathbb{T}^d \times \mathbb{R} \to [0, 1]$ be a kinetic function. A kinetic function $f \colon \mathbb{T}^d \times [0, T] \times \mathbb{R} \to [0, 1]$ is said to be a *generalized* solution to (1.1) on $[0, T]$ with initial datum f_0 if

1. for all $\varphi \in C^1_c(\mathbb{T}^d \times \mathbb{R})$, $t \mapsto \langle f(t), \varphi\rangle$ is a càdlàg function,
2. there exists a finite non-negative measure m on $\mathbb{T}^d \times \mathbb{R}$ such that

$$\langle f(t), \varphi\rangle = \langle f_0, \varphi\rangle + \int_0^t \langle f(s), a(\xi) \cdot \nabla\varphi\rangle ds - m_\varphi([0, t]), \tag{1.39}$$

   for all $\varphi \in C^1_c(\mathbb{T}^d \times \mathbb{R})$, for all $t \in [0, T]$.

Remark 1.1 (Measure-valued solutions). One can use the relation (1.25) to express the identity (1.39) in terms of the Young measure $\nu_{x,t}$ only. This relates our notion of generalized solution to the notion of measure-valued solution as developed by Di Perna for systems of first-order conservation laws, [4].

The next steps then are the following ones:

1. prove a result of *reduction*, that states that every generalized solution starting from an initial datum at equilibrium remains an equilibrium for all time,
2. deduce the strong convergence of u^ε in $L^p(\mathbb{T}^d \times [0, T])$ to the unique solution of (1.1).

This will be established in Section 4.3, in the stochastic framework. In the next section we complete the analysis of generalized solutions with some results that will be useful later.

1.4.3 Behaviour of the defect measure at a given time

Let f be a generalized solution. By considering the averages

$$\frac{1}{\delta} \int_{t_*-\delta}^{t_*} \nu_{x,t} dt,$$

in a manner similar to the proof of Proposition 1.5, one can show that the limit from the left $\langle f(t-),\varphi\rangle$ of $t\mapsto\langle f(t),\varphi\rangle$ is represented by a kinetic function f^-, in the sense that

$$\lim_{\delta\to 0+}\langle f(t-\delta),\varphi\rangle=\langle f^-(t),\varphi\rangle.$$

By (1.39), we have then, if $t\in(0,T)$, the relation

$$\langle f(t),\varphi\rangle=\langle f^-(t),\varphi\rangle-m_\varphi(\{t\}). \tag{1.40}$$

For $t=0$, we obtain

$$\langle f(0),\varphi\rangle=\langle f_0,\varphi\rangle-m_\varphi(\{0\}). \tag{1.41}$$

We would like to deduce from (1.41) that $f(0)=f_0$. This is an expected identity by consistency. This is true indeed *if f_0 is at equilibrium*, according to the following proposition.

Proposition 1.6 (The case of equilibrium)**.** *Suppose that f_0 is at equilibrium, $f_0=\mathtt{f}_0$, in (1.41). Then $f(0)=f_0$ and $m(\mathbb{T}^d\times\{0\}\times\mathbb{R})=0$.*

(Sketch of the proof). Taking $\varphi(x,\xi)=\varphi(x)$ (this has to be justified), we deduce from (1.41) that

$$\int_\mathbb{R}\chi_f(x,0,\xi)d\xi=\int_\mathbb{R}\chi_{f_0}(x,\xi)d\xi,\quad \chi_f(x,t,\xi)=f(x,t,\xi)-\mathbf{1}_{0>\xi},$$

for a.e. $x\in\mathbb{T}^d$. If $f_0(x,\xi)=\mathbf{1}_{u_0>\xi}$, this shows that

$$\int_\mathbb{R}\xi d\nu_{x,0}(\xi)=\int_\mathbb{R}\chi_f(x,0,\xi)d\xi=u_0(x) \tag{1.42}$$

for a.e. $x\in\mathbb{T}^d$. Subtracting $\mathbf{1}_{0>\xi}$ to both sides of (1.41) and taking $\varphi(x,\xi)=\psi(x)\eta'(\xi)$ with η convex and $\psi\geq 0$, we obtain then

$$\int_{\mathbb{T}^d}\psi(x)\left[\int_\mathbb{R}\eta(\xi)d\nu_{x,0}(\xi)-\eta(u_0(x))\right]dx+m_\varphi(\{0\})=0. \tag{1.43}$$

In (1.43), we have $m_\varphi(\{0\})\geq 0$ since η is convex and $\psi\geq 0$. By the Jensen inequality and (1.42), we also have

$$\int_\mathbb{R}\eta(\xi)d\nu_{x,0}(\xi)-\eta(u_0(x))\geq 0.$$

Consequently, all the terms in (1.43) are trivial. □

2 Some basic facts on stochastic processes

2.1 Stochastic processes

Definition 2.1 (Stochastic process)**.** Let E be a metric space, I a subset of $\mathbb{R}$ and $(\Omega,\mathcal{F},\mathbb{P})$ a probability space. An E-valued *stochastic process* $(X_t)_{t\in I}$ is a collection of random variables $X_t\colon\Omega\to E$ indexed by I.

Definition 2.2 (Processes with independent increments)**.** Let E be a metric space. A process $(X_t)_{t\in[0,T]}$ with values in E is said to have independent increments if, for all $n\in\mathbb{N}^*$, for all $0\leq t_1<\ldots<t_n\leq T$, the family $\{X_{t_{i+1}}-X_{t_i};i=1,\ldots,n-1\}$ of E-valued random variables is independent.

Definition 2.3 (Processes with continuous trajectories)**.** Let E be a metric space. A process $(X_t)_{t\in[0,T]}$ with values in E is said to have continuous trajectories, if for all $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is continuous from $[0,T]$ to E. If this is realized only almost surely (for ω in a set of full measure), then we say that (X_t) is almost surely continuous, or has almost surely continuous trajectories.

Similarly, one defines processes that are *càdlàg*: for all $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is continuous from the right and has limit from the left (continue à droite, limite à gauche, *i.e.* càdlàg in french). We also speak of process with almost sure càdlàg trajectories. An important class of càdlàg processes are the jump processes. The trajectories of a process $(X_t)_{t\in[0,T]}$ may have more regularity than the C^0-regularity. Consider for example a process satisfying: there exists $\alpha \in (0,1)$ such that, for $\mathbb{P}$-almost all $\omega \in \Omega$, there exists a constant $C(\omega) \geq 0$ such that

$$d_E(X_t(\omega), X_s(\omega)) \leq C(\omega)|t-s|^\alpha, \tag{2.1}$$

for all $t, s \in [0,T]$. Then we say that $(X_t)_{t\in[0,T]}$ has almost surely α-Hölder trajectories, or is almost-surely C^α.

2.2 Law of a process

2.2.1 Cylindrical sets

Let E be a metric space. A process $(X_t)_{t\in[0,T]}$ with values in E can be seen as a function

$$X\colon \Omega \to E^{[0,T]}, \tag{2.2}$$

where $E^{[0,T]}$ is the set of the applications $[0,T] \to E$. Let $\mathcal{F}_{\text{cyl}}$ denote the cylindrical σ-algebra on $E^{[0,T]}$. This is the coarsest (minimal) σ-algebra that makes the projections

$$\pi_t\colon E^{[0,T]} \to E, \quad Y \mapsto Y_t$$

measurable. It is called cylindrical because it is generated by the *cylindrical sets*, which are subsets of $E^{[0,T]}$ of the form

$$D = \pi_{t_1}^{-1}(B_1) \bigcap \cdots \bigcap \pi_{t_n}^{-1}(B_n) = \left\{ Y \in E^{[0,T]}; Y_{t_1} \in B_1, \ldots, Y_{t_n} \in B_n \right\}, \tag{2.3}$$

where $t_1, \ldots, t_n \in [0,T]$ for a given $n \in \mathbb{N}^*$, and $B_1, \ldots, B_n$ are Borel subsets of E. Roughly speaking, in (2.3), D is the product of $B_1 \times \cdots \times B_n$ with the whole space $\prod_{t\neq t_j} E$. This is why we speak of cylinder set. We have

$$X^{-1}(D) = \bigcap_{j=1}^n X_{t_j}^{-1}(B_j) \in \mathcal{F},$$

hence $X\colon (\Omega, \mathcal{F}) \to (E^{[0,T]}, \mathcal{F}_{\text{cyl}})$ is a *random variable*.

Definition 2.4 (Law of a stochastic process)**.** Let E be a metric space. The law of an E-valued stochastic process $(X_t)_{t\in[0,T]}$ is the probability measure μ_X on $(E^{[0,T]}, \mathcal{F}_{\text{cyl}})$ induced by the map X in (2.2).

Remark 2.1. The σ-algebra $\mathcal{F}_{\text{cyl}}$ being generated by the cylindrical sets, the law of X is characterized by the data

$$\mathbb{P}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n),$$

which are called the *finite-dimensional distributions* of X.

We can be more specific on $\mathcal{F}_{\rm cyl}$. Each cylindrical set in (2.3) is of the form

$$\left\{ Y \in E^{[0,T]}; (Y_t)_{t\in J} \in B \right\}, \tag{2.4}$$

where J is a *countable* (since finite) subset of $[0,T]$ and B an element of the product σ-algebra $\Pi_{t\in J}\mathcal{B}(E_t)$, where $E_t = E$ for all t (this latter is the cylindrical σ-algebra for E^J). The collection of sets of the form (2.4) is precisely $\mathcal{F}_{\rm cyl}$.

Lemma 2.1 (Countably generated sets)**.** *The cylindrical σ-algebra $\mathcal{F}_{\rm cyl}$ is the collections of sets of the form (2.4), for $J \subset [0,T]$ countable and B in the cylindrical σ-algebra of E^J.*

Proof of Lemma 2.1. Let us call $\mathcal{F}_\circ$ the collection of sets of the form (2.4), for $J\subset[0,T]$ countable and B in the cylindrical σ-algebra of E^J. The countable union of countable sets being countable, $\mathcal{F}_\circ$ is stable by countable union. Clearly it contains the empty set and is stable when taking the complementary since

$$\left\{ Y \in E^{[0,T]}; (Y_t)_{t\in J} \in B \right\}^c = \bigcup_{t\in J} \pi_t^{-1}(C_t), \quad C_t = (\pi_t(B))^c \in \mathcal{B}(E).$$

Therefore, $\mathcal{F}_\circ$ is a σ-algebra. Since $\mathcal{F}_\circ$ contains cylindrical sets (case J finite in (2.4)), $\mathcal{F}_\circ = \mathcal{F}_{\rm cyl}$. □

A corollary of this characterization of $\mathcal{F}_{\rm cyl}$ is that a lot of sets described in terms of an uncountable set of values X_t of the process $(X_t)_{t\in[0,T]}$ are *not measurable*, *i.e* not in $\mathcal{F}_{\rm cyl}$. This is due to the fact that $[0,T]$ is uncountable. For processes indexed by countable sets (discrete time processes), these problems of non-measurable sets do not appear.

Exercise 2.5. Show that the following sets are not in $\mathcal{F}_{\rm cyl}$:

1. $A_1 = \{X \equiv 0\} = \bigcap_{t\in[0,T]} \pi_t^{-1}(\{0\})$,
2. $A_2 = \{t \mapsto X_t \text{ is continuous}\}$.

2.2.2 Continuous processes

Now, assume that E is a Banach space and $(X_t)_{t\in[0,T]}$ is a process with almost-sure continuous trajectories. Then we would like to say that, instead of (2.2), we have

$$X\colon \Omega \to C([0,T];E), \tag{2.5}$$

In that case, the sets A_1 and A_2 in Exercise 2.5 *are* measurable.

Exercise 2.6. Let

$$\mathcal{F}_{\rm cts} = \mathcal{F}_{\rm cyl} \cap C([0,T];E).$$

Show that the σ-algebra $\mathcal{F}_{\rm cts}$ coincides with the Borel σ-algebra on $C([0,T];E)$, the topology on $C([0,T];E)$ being the topology of Banach space with norm

$$X \mapsto \sup_{t\in[0,T]} \|X(t)\|_E.$$

Then show that the sets A_1 and A_2 in Exercise 2.5 are measurable.

Actually, starting from (2.2), we have (2.5) indeed only if we first redefine X on $\Omega\backslash\Omega_{\text{cts}}$ where Ω_{cts} is the set of ω such that $t \mapsto X_t(\omega)$ is continuous. However, it is not ensured that $\Omega_{\text{cts}}(=X^{-1}(A_2)$ with the notation of Exercise 2.5) is measurable. A correct procedure is the following one (we modify not only Ω, but $\mathbb{P}$ also, [16]). Define the probability measure Q on $\mathcal{F}_{\text{cts}}$ by

$$Q(A) = \mathbb{P}(X \in \tilde{A}), \quad A = \tilde{A} \cap C([0,T];E), \quad \tilde{A} \in \mathcal{F}_{\text{cyl}}. \tag{2.6}$$

for all $A \in \mathcal{F}_{\text{cts}}$. By definition, each $A \in \mathcal{F}_{\text{cts}}$ can be written as in (2.6). If two decompositions

$$A = \tilde{A}_1 \cap C([0,T];E) = \tilde{A}_2 \cap C([0,T];E)$$

are possible, then the definition of $Q(A)$ is unambiguous since $\mathbb{P}(X \in \tilde{A}_1) = \mathbb{P}(X \in \tilde{A}_2)$. Indeed, by hypothesis, there exists a measurable subset G of Ω of full measure such that: $\omega \in G$ implies that $t \mapsto X_t(\omega)$ is continuous (*i.e.* $G \subset \Omega_{\text{cts}}$). If $\omega \in X^{-1}(\tilde{A}_1) \cap G$, then

$$X(\omega) \in \tilde{A}_1 \cap C([0,T];E) = \tilde{A}_2 \cap C([0,T];E),$$

hence $X^{-1}(\tilde{A}_1) \cap G \subset X^{-1}(\tilde{A}_2) \cap G$. It follows that

$$\mathbb{P}(X \in \tilde{A}_1) = \mathbb{P}(X^{-1}(\tilde{A}_1) \cap G) \leq \mathbb{P}(X^{-1}(\tilde{A}_2) \cap G) = \mathbb{P}(X \in \tilde{A}_2).$$

By symmetry of $\tilde{A}_1$ and $\tilde{A}_2$, we obtain the result. We consider then the canonical process

$$Y_t \colon C([0,T];E) \to \mathbb{R}, \quad Y_t(\omega) = \omega(t).$$

The law of Y on $(C([0,T];E), \mathcal{F}_{\text{cts}}, Q)$ is the same as X (*cf.* Remark 2.1), thus considering X or Y is equivalent, and Y has the desired path-space $C([0,T];E)$.

Definition 2.7 (Modification)**.** Let E be a metric space and let $(X_t)_{t\in[0,T]}$, $(Y_t)_{t\in[0,T]}$ be two stochastic processes on E. If $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ have the same law, they are said to be *equivalent*. One say that $(Y_t)_{t\in[0,T]}$ is a *modification* of $(X_t)_{t\in[0,T]}$ if

$$\forall t \in [0,T],\ \mathbb{P}(X_t \neq Y_t) = 0.$$

Exercise 2.8. Show that *modification* implies *equivalent*.

2.3 The Wiener process

Definition 2.9 (Wiener process)**.** A d-dimensional Wiener process is a process $(B_t)_{t\geq 0}$ with values in $\mathbb{R}^d$ such that: $B_0 = 0$ almost-surely, $(B_t)_{t\geq 0}$ has independent increments, and, for all $0 \leq s < t$, the increment $B_t - B_s$ follows the normal law $\mathcal{N}(0, (t-s)\mathrm{I}_d)$.

Exercise 2.10. Show that the properties above depend only on the *law* of the process, *i.e.* if $(B_t)_{t\geq 0}$ and $(\tilde{B}_t)_{t\geq 0}$ are some equivalent processes on $\mathbb{R}^d$ and $(B_t)_{t\geq 0}$ is a d-dimensional Wiener process, then $(\tilde{B}_t)_{t\geq 0}$ is a d-dimensional Wiener process as well.

A consequence of the criterion of continuity of Kolmogorov (which we do not state), is the following continuity result.

Proposition 2.2 (Continuity of the Wiener process)**.** *If $(B_t)_{t\geq 0}$ is a d-dimensional Wiener process is a process, then there is a modification $(\tilde{B}_t)_{t\geq 0}$ of $(\tilde{B}_t)_{t\geq 0}$ that has C^α trajectories for all $\alpha < 1/2$.*

A corollary of the following result on the quadratic variation of the Wiener process is that Proposition 2.2 cannot be true if $\alpha > 1/2$.

Proposition 2.3 (Quadratic variation)**.** *Let $(B_t)_{t\geq 0}$ be a d-dimensional Wiener process is a process. For $\sigma = (t_i)_{0,n}$ a subdivision*

$$0 = t_0 < \cdots < t_n = t$$

of the interval $[0,t]$ of step $|\sigma| = \sup_{0\leq i<n}(t_{i+1} - t_i)$, define

$$V_2^\sigma(t) = \sum_{i=0}^{n-1} |B_{t_{i+1}} - B_{t_i}|^2.$$

Then $V_2^\sigma(t) \to t$ in $L^2(\Omega)$ when $|\sigma| \to 0$.

Proof. Let $\xi_i = |B_{t_{i+1}} - B_{t_i}|^2 - (t_{i+1} - t_i)$.

$$\mathbb{E}\,|V_2^\sigma(t) - t|^2 = \mathbb{E}\left|\sum_{i=0}^{n} \xi_i\right|^2 = \sum_{0\leq i,j<n} \mathbb{E}[\xi_i\xi_j]. \tag{2.7}$$

The random variables $\xi_0, \ldots, \xi_{n-1}$ are centred, $\mathbb{E}[\xi_i] = 0$, and independent. Therefore in the sum over i, j in (2.7), only the perfect squares (case $i = j$) are contributing. Since $\mathbb{E}|\xi|^2$, the variance of $|B_{t_{i+1}} - B_{t_i}|^2$, is of order $(t_{i+1} - t_i)^2$, the result follows. □

2.4 Filtration, stochastic basis

Definition 2.11 (Filtration)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathcal{F}$ is said to be a *filtration* if the family is increasing with respect to t: $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s \leq t$. The space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is called a filtered space. If $(\mathcal{F}_t)_{t\geq 0}$ we set $\mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s$. We say that $(\mathcal{F}_t)_{t\geq 0}$ is *continuous from the right* if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all t. We say that $(\mathcal{F}_t)_{t\geq 0}$ is *complete* if $\mathcal{F}_t$ is complete: it contains all $\mathbb{P}$-negligible sets. We say that $(\mathcal{F}_t)_{t\geq 0}$ *satisfies the usual condition* if $(\mathcal{F}_t)_{t\geq 0}$ is continuous from the right and complete.

Definition 2.12 (Adapted process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and E a metric space. An E-valued process $(X_t)_{t\geq 0}$ is said to be *adapted* if, for all $t \geq 0$, X_t is $\mathcal{F}_t$-measurable.

Note that this means $\sigma(X_t) \subset \mathcal{F}_t$ for all $t \geq 0$.

Example 2.2. If $(X_t)_{t\geq 0}$ is a process over $(\Omega, \mathcal{F}, \mathbb{P})$, we introduce

$$\mathcal{F}_t^X = \sigma(\{X_s; 0 \leq s \leq t\}) \tag{2.8}$$

the σ-algebra generated by all random variables $(X_{s_1}, \ldots, X_{s_N})$ for $N \in \mathbb{N}^*$, $s_1, \ldots, s_N \in [0,t]$. Then $(\mathcal{F}_t^X)_{t\geq 0}$ is a filtration and $(X_t)_{t\geq 0}$ is adapted to this filtration: $(\mathcal{F}_t^X)_{t\geq 0}$ is called the natural filtration of the process, or the filtration generated by $(X_t)_{t\geq 0}$.

Exercise 2.13. Let $(X_t)_{t\geq 0}$ be a continuous process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Show that $(\mathcal{F}_t^X)_{t\geq 0}$ is not necessarily continuous from the right. *Hint:* you may consider $X_t = tY$, Y being given.

Proposition 2.4. *We assume that $(\mathcal{F}_t)$ is complete and that E is complete. Then any limit (a.s., or in probability, or in $L^p(\Omega)$) of adapted processes is adapted.*

Proof of Proposition 2.4. Note that requiring $\mathcal{F}_0$ to be complete is equivalent to require all the σ-algebras $\mathcal{F}_t$ to be complete. Let X^n and X be some E-valued random variables such that $(X^n)_{n\in\mathbb{N}}$ is converging to X for one of the modes of convergence that we are considering. We just have to consider convergence almost-sure since convergence in probability or in $L^p(\Omega)$ implies convergence a.s. of a subsequence. If all the X^n are $\mathcal{G}$-measurable, where $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$, then the set of points where (X_n) is converging is in $\mathcal{G}$ (we use the Cauchy criterion to characterize the convergence). Consequently, X is equal $\mathbb{P}$-a.e. to a $\mathcal{G}$-measurable function. If $\mathcal{G}$ is complete, we deduce that X is $\mathcal{G}$-measurable. □

Definition 2.14 (Progressively measurable process)**.** Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. An E-valued process $(X_t)_{t\in[0,T]}$ is said to be *progressively measurable* (with respect to $(\mathcal{F}_t)_{t\in[0,T]}$) if, for all $t\in[0,T]$, the map $(s,\omega)\mapsto X_s(\omega)$ from $[0,t]\times\Omega$ to E is $\mathcal{B}([0,t])\times\mathcal{F}_t$-measurable.

Definition 2.15 (Stochastic basis)**.** Let $(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\geq 0})$ be a filtered space. Let $m\geq 1$ and let $(B(t))_{t\geq 0}$ be an m-dimensional Wiener process such that $(B(t))_{t\geq 0}$ is $(\mathcal{F}_t)$-adapted and, for all $0\leq s<t$, $B(t)-B(s)$ is independent on $\mathcal{F}_s$. Then one says that

$$(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\geq 0},(B(t))_{t\geq 0})$$

is a stochastic basis.

3 Stochastic integration

Let $(\beta(t))$ be a one dimensional Wiener process over $(\Omega,\mathcal{F},\mathbb{P})$. Let K be a separable Hilbert space and let $(g(t))$ be a K-valued stochastic process. The first obstacle to the definition of the stochastic integral

$$I(g)=\int_0^T g(t)d\beta(t) \tag{3.1}$$

is the lack of regularity of $t\mapsto\beta(t)$, which has almost-surely a regularity $1/2-$: for all $\alpha\in[0,1/2)$, almost-surely, β is in $C^\alpha([0,T])$ and not in $C^{1/2}([0,T])$. Young's integration theory can be used to give a meaning to (3.1) for integrands $g\in C^\gamma([0,T])$ when $\gamma>1/2$, but this not applicable here, since the resolution of stochastic differential equation requires a definition of $I(\beta)$. In that context, one has to expand the theory of Young's or Riemann − Stieltjes' Integral, this is one of the purpose of rough paths' theory, *cf.* [8]. Below, it is the martingale properties of the Wiener process which are used to define the stochastic integral (3.1).

3.1 Stochastic integration of elementary processes

Let $(\mathcal{F}_t)_{t\geq 0}$ be a given filtration, such that $(\beta(t))$ is $(\mathcal{F}_t)$-adapted, and the increment $\beta(t)-\beta(s)$ is independent on $\mathcal{F}_s$ for all $0\leq s\leq t$. Let $(g(t))_{t\in[0,T]}$ be a K-valued stochastic process which is adapted, simple and L^2, in the sense that

$$g(\omega,t)=g_{-1}(\omega)\mathbf{1}_{\{0\}}(t)+\sum_{i=0}^{n-1}g_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(t), \tag{3.2}$$

where $0\leq t_0\leq\cdots\leq t_n\leq T$, g_{-1} is $\mathcal{F}_0$-measurable, each g_i, $i\in\{0,\dots,n-1\}$ is $\mathcal{F}_{t_i}$-measurable and in $L^2(\Omega;K)$. For such an integrand g, we define $I(g)$ as the following Riemann sum

$$I(g)=\sum_{i=0}^{n-1}(\beta(t_{i+1})-\beta(t_i))g_i. \tag{3.3}$$

Remark 3.1. Let λ denote the Lebesgue measure on $[0,T]$. For g as in (3.2), we have

$$g(\omega,t)=\sum_{i=0}^{n-1} g_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(t),$$

for $\mathbb{P}\times\lambda$-almost all $(\omega,t)\in\Omega\times[0,T]$ since the singleton $\{0\}$ has λ-measure 0. We include the term $g_{-1}(\omega)\mathbf{1}_{\{0\}}(t)$ in (3.2) to be consistent with the definition of the predictable σ-algebra in the next section 3.2. Consistency here is in the sense that the predictible σ-algebra $\mathcal{P}_T$ as defined in Section 3.2 is precisely the σ-algebra generated by the elementary processes.

Note that g as in (3.2) belongs to $L^2(\Omega\times[0,T],\mathbb{P}\times\lambda)$ and that

$$\int_0^T \mathbb{E}\|g(t)\|_K^2 dt=\sum_{i=0}^{n-1}(t_{i+1}-t_i)\mathbb{E}\left[\|g_i\|_K^2\right]. \tag{3.4}$$

In (3.3), g_i and the increment $\beta(t_{i+1})-\beta(t_i)$ are independent. Using this fact, we can prove the following proposition.

Proposition 3.1 (Itō's isometry). *We have $I(g)\in L^2(\Omega;K)$ and*

$$\mathbb{E}\left[I(g)\right]=0,\quad \mathbb{E}\left[\|I(g)\|_K^2\right]=\int_0^T \mathbb{E}\|g(t)\|_K^2 dt. \tag{3.5}$$

Proof of Proposition 3.1. We develop the square of the norm of $I(g)$:

$$\begin{aligned}\|I(g)\|_K^2 &= \sum_{i=0}^{n-1}|\beta(t_{i+1})-\beta(t_i)|^2\|g_i\|_K^2\\ &\qquad\qquad + 2\sum_{0\leq i<j\leq n-1}(\beta(t_{i+1})-\beta(t_i))(\beta(t_{j+1})-\beta(t_j))\langle g_i,g_j\rangle_K. \end{aligned}\tag{3.6}$$

By independence, the expectancy of the second term (cross-products) in (3.6) vanishes, while the expectancy of the first term gives

$$\sum_{i=0}^{n-1}(t_{i+1}-t_i)\mathbb{E}\left[\|g_i\|_E^2\right]=\int_0^T \mathbb{E}\|g(t)\|_E^2 dt$$

since $\mathbb{E}\left[|\beta(t_{i+1})-\beta(t_i)|^2\right]=(t_{i+1}-t_i)$. This shows that $I(g)\in L^2(\Omega;K)$ and the second equality in (3.5). The first equality follows from the identity

$$\mathbb{E}\left[(\beta(t_{i+1})-\beta(t_i))g_i\right]=\mathbb{E}\left[(\beta(t_{i+1})-\beta(t_i))\right]\mathbb{E}\left[g_i\right]=0,$$

for all $i\in\{0,\dots,n-1\}$. □

3.2 Extension

Let $\mathcal{E}_T$ denote the set of L^2-elementary predictable functions in the form (3.2). This is a subset of $L^2(\Omega\times[0,T];K)$ (the measure on $\Omega\times[0,T]$ being the product measure $\mathbb{P}\times\lambda$). The second identity in (3.5) shows that

$$I\colon \mathcal{E}_T\subset L^2(\Omega\times[0,T];K)\to L^2(\Omega;K) \tag{3.7}$$

is a linear isometry. The stochastic integral $I(g)$ is the extension of this isometry to the closure $\overline{\mathcal{E}_T}$ of $\mathcal{E}_T$ in $L^2(\Omega\times[0,T];K)$. It is clear that (3.5) (Itō's isometry) is preserved in this extension operation. To understand what is $I(g)$ exactly, we have to identify the closure $\overline{\mathcal{E}_T}$, or, at least certain sub-classes of $\overline{\mathcal{E}_T}$. For this purpose, we introduce $\mathcal{P}_T$, the predictable sub-σ-algebra of $\mathcal{F}\times\mathcal{B}([0,T])$ generated by the sets $F_0\times\{0\}$, $F_s\times(s,t]$, where F_0 is $\mathcal{F}_0$-measurable, $0\leq s<t\leq T$ and F_s is $\mathcal{F}_s$-measurable. We have denoted by $\mathcal{B}([0,T])$ the Borel σ-algebra on $[0,T]$. It is clear that each element in $\mathcal{E}_T$ is $\mathcal{P}_T$ measurable. We will admit without proof the following propositions (Proposition 3.2 and Proposition 3.3).

Proposition 3.2. *Assume that the filtration $(\mathcal{F}_t)$ is complete and continuous from the right. Then the σ-algebra generated on $\Omega\times[0,T]$ by adapted left-continuous (respectively, adapted continuous processes) coincides with the predictable σ-algebra $\mathcal{P}_T$.*

Proof of Proposition 3.2. Exercise, or see [16, Proposition 5.1, p. 171]. □

A $\mathcal{P}_T$-measurable process is called a predictable process. Denote by $\mathcal{P}_T^*$ the completion of $\mathcal{P}_T$. By Proposition 3.2, any adapted a.s. left-continuous or continuous process is $\mathcal{P}_T^*$-measurable.

Proposition 3.3. *Assume that the filtration $(\mathcal{F}_t)$ is complete and continuous from the right. Define*

1. *the optional σ-algebra to be the σ-algebra $\mathcal{O}$ generated by adapted càdlàg processes,*
2. *the progressive σ-algebra to be the σ-algebra* Prog *generated by the progressively measurable processes (Definition 2.14).*

Then we have the inclusion

$$\mathcal{P}_T\subset\mathcal{O}\subset\mathrm{Prog}\subset\mathcal{P}_T^*,\tag{3.8}$$

and the identity

$$\overline{\mathcal{E}_T}=L^2(\Omega\times[0,T],\mathcal{P}_T^*;K).\tag{3.9}$$

Proof of Proposition 3.3. See [3, Lemma 2.4] and [3, Chapter 3]. □

In what follows we will always assume that the filtration $(\mathcal{F}_t)$ is complete and continuous from the right.

Note that a function is in $L^2(\Omega\times[0,T],\mathcal{P}_T^*;K)$ if it is equal $\mathbb{P}\times\lambda$-a.e. to a function of $L^2(\Omega\times[0,T];K)$ which is $\mathcal{P}_T$-measurable.

A consequence of Proposition 3.2 and Proposition 3.3 is that we can define the stochastic integral $I(g)$ of processes $(g(t))$ which are either adapted and left-continuous or continuous or càdlàg or progressively measurable. We will use the notation $\int_0^T g(t)d\beta(t)$ for $I(g)$.

Exercise 3.1. Show that (in the case $K=\mathbb{R}$)

1. if $(g(t))$ is an adapted process such $g\in C([0,T];L^2(\Omega))$, then

$$\int_0^T g(t)d\beta(t)=\lim_{|\sigma|\to 0}\sum_{i=0}^{n-1}g(t_i)(\beta(t_{i+1})-\beta(t_i)),\tag{3.10}$$

where $\sigma=\{0=t_0\leq\cdots\leq t_n=T\}$ and $\sigma=\sup_{0\leq i<n}(t_{i+1}-t_i)$.

2. Show that the result (3.10) holds true if $(g(t))$ is a continuous adapted process such that $\sup_{t\in[0,T]} \mathbb{E}|g(t)|^q$ is finite for a $q > 2$.

3. If $g \in L^2(0,T)$ is deterministic, then $\int_0^T g(t)d\beta(t)$ is a gaussian random variable $\mathcal{N}(0,\sigma^2)$ of variance
$$\sigma^2 = \int_0^t |g(t)|^2 dt.$$

3.3 Itô's Formula

3.3.1 Dimension one

Proposition 3.4 (Itô's Formula)**.** *Assume that the filtration $(\mathcal{F}_t)$ is complete and continuous from the right. Let $g \in L^2(\Omega\times[0,T],\mathcal{P}_T^*;\mathbb{R})$, $f \in L^1(\Omega\times[0,T],\mathcal{P}_T^*;\mathbb{R})$, let $x\in\mathbb{R}$ and let*
$$X_t = x + \int_0^t f(s)ds + \int_0^t g(s)d\beta(s).$$
Let $u\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ be a function of class $C_b^{1,2}$. Then
$$u(t,X_t) = u(0,x) + \int_0^t \left[\frac{\partial u}{\partial s}(s,X_s) + \frac{\partial u}{\partial x}(s,X_s)f(s) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(s,X_s)|g(s)|^2\right] ds + \int_0^t \frac{\partial u}{\partial x}(s,X_s)g(s)d\beta(s), \quad (3.11)$$
for all $t\in[0,T]$.

Proof of Proposition 3.4. We do the proof in the case where u is independent on t and $f\equiv 0$ since the more delicate (and remarkable) term in (3.11) is the Itô's correction involving the second derivative of u. By approximation, it is also sufficient to consider the case where u is in C_b^3 and g is the elementary process
$$g = \sum_{l=0}^{m-1} g_l \mathbf{1}_{(s_l,s_{l+1}]},$$
where $(s_l)_{0,m}$ is a subdivision of $[0,T]$ and g_l is a.s. bounded: $|g_l|\leq M$ a.s. Let $\sigma=(t_i)_{0,n}$ be a subdivision of $[0,T]$ which is a refinement of (s_l). Let us consider the case $t=T$ only (for general times t, replace t_i by $t_i\wedge t$ in the formulas below). We decompose
$$u(X_T)-u(x) = \sum_{i=0}^{n-1} u(X_{t_{i+1}}) - u(X_{t_i}),$$
and use the Taylor formula to get
$$u(X_T)-u(x) = \sum_{i=0}^{n-1} u'(X_{t_i})(X_{t_{i+1}}-X_{t_i}) + \frac{1}{2}u''(X_{t_i})(X_{t_{i+1}}-X_{t_i})^2 + r_\sigma^1, \quad (3.12)$$
where
$$|r_\sigma^1| \leq \frac{1}{6}\|u^{(3)}\|_{C_b(\mathbb{R})}\sum_{i=0}^{n-1}|X_{t_{i+1}}-X_{t_i}|^3. \quad (3.13)$$

Since $X_{t_{i+1}} - X_{t_i} = g(t_i)\delta\beta(t_i)$, $\delta\beta(t_i) := \beta(t_{i+1}) - \beta(t_i)$, we deduce from (3.12)-(3.13) that

$$u(X_T) - u(x) = \sum_{i=0}^{n-1} u'(X_{t_i})g(t_i)\delta\beta(t_i) + \frac{1}{2}u''(X_{t_i})|g(t_i)|^2|\delta\beta(t_i)|^2 + r_\sigma^1, \tag{3.14}$$

and that

$$\mathbb{E}|r_\sigma^1| \leq \frac{1}{6}\|u^{(3)}\|_{C(\mathbb{R})} M \sum_{i=0}^{n-1} \mathbb{E}|\delta\beta(t_i)|^3 = \mathcal{O}\left(\sum_{i=0}^{n-1}(t_{i+1}-t_i)^{3/2}\right) = \mathcal{O}(|\sigma|^{1/2}). \tag{3.15}$$

By (3.14), we get

$$u(X_T) - u(x) = \int_0^T u'(X_t)g(t)d\beta(t) + \int_0^T \frac{1}{2}u''(X_t)|g(t)|^2 dt + r_\sigma^3 + r_\sigma^2 + r_\sigma^1, \tag{3.16}$$

where the remainder r_σ^3 and r_σ^2 are such that

$$\sum_{i=0}^{n-1} u'(X_{t_i})g(t_i)\delta\beta(t_i) = \int_0^T u'(X_t)g(t)d\beta(t) + r_\sigma^3,$$

and

$$\sum_{i=0}^{n-1} \frac{1}{2}u''(X_{t_i})|g(t_i)|^2|\delta\beta(t_i)|^2 = \int_0^T \frac{1}{2}u''(X_t)|g(t)|^2 dt + r_\sigma^2. \tag{3.17}$$

By Itô's Isometry, we have the estimate

$$\begin{aligned}\mathbb{E}|r_\sigma^3|^2 &= \sum_{i=0}^{n-1} \mathbb{E}\int_{t_i}^{t_{i+1}} |u'(X_t) - u'(X_{t_i})|^2|g(t_i)|^2 dt \\ &\leq M^2\|u''\|_{C_b(\mathbb{R})}^2 \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}|X_t - X_{t_i}|^2 dt.\end{aligned}$$

Since $\mathbb{E}|X_t - X_{t_i}|^2 = \mathbb{E}|g(t_i)|^2(t-t_i) \leq M^2(t-t_i)$, we deduce that

$$\mathbb{E}|r_\sigma^3|^2 \leq M^4\|u''\|_{C_b(\mathbb{R})}^2 \sum_{i=0}^{n-1}(t_{i+1}-t_i)^2 = \mathcal{O}(|\sigma|). \tag{3.18}$$

Some similar estimates show that we can replace X_t by the step function equal to X_{t_i} on $(t_i, t_{i+1}]$ in the right-hand side of (3.17) and that this contributes to an error of order $|\sigma|$: $r_\sigma^2 = r_\sigma^4 + r_\sigma^5$, where $\mathbb{E}|r_\sigma^4|^2 = \mathcal{O}(|\sigma|)$, where the remainder term r_σ^5 is defined by

$$r_\sigma^5 = \sum_{i=0}^{n-1} \frac{1}{2}u''(X_{t_i})|g(t_i)|^2[|\delta\beta(t_i)|^2 - (t_{i+1}-t_i)].$$

Since $(t_{i+1} - t_i) = \mathbb{E}[|\delta\beta(t_i)|^2|\mathcal{F}_{t_i}]$, cancellations occur when we develop the square of r_σ^5 and take the expectation: only the pure squares remain, and we get

$$\mathbb{E}|r_\sigma^5|^2 \leq \|u''\|_{C_b(\mathbb{R})}^2 M^4 \sum_{i=0}^{n-1} \mathbb{E}||\delta\beta(t_i)|^2 - (t_{i+1}-t_i)]|^2 = \mathcal{O}(|\sigma|). \tag{3.19}$$

Using (3.15), (3.18), (3.19), we can pass to the limit $|\sigma| \to 0$ in (3.16) to get (3.11) in our simplified case. □

3.3.2 Higher dimensions

We explain briefly what is the Itô's formula for the stochastic integral against a d-dimensional Wiener process, for integrand taking values in a given separable Hilbert space K. A d-dimensional Wiener process $(B(t))_{t\geq 0}$ admits the decomposition

$$B(t)=\sum_{k=1}^{d}\beta_k(t)e_k, \tag{3.20}$$

where (e_k) is the canonical basis of $\mathbb{R}^d$ and $\beta_1(t),\dots,\beta_d(t)$ are independent one-dimensional processes. Let $(\mathcal{F}_t)_{t\geq 0}$ be a given filtration, such that, for all k, $(\beta_k(t))$ is $(\mathcal{F}_t)$-adapted, and the increment $\beta_k(t)-\beta_k(s)$ is independent on $\mathcal{F}_s$ for all $0\leq s\leq t$. Let K be a separable Hilbert space. Let $(g(t))$ be a process with values in $\mathcal{L}(\mathbb{R}^d;K)$ such that

$$g\in L^2(\Omega\times[0,T],\mathcal{P}_T^*;\mathcal{L}(\mathbb{R}^d;K)). \tag{3.21}$$

We set

$$\int_0^T g(t)dB(t)=\sum_{k=1}^{d}\int_0^T g(t)e_k d\beta_k(t). \tag{3.22}$$

This defines an element of $L^2(\Omega;K)$ and, using the independence of $\beta_1(t),\dots,\beta_d(t)$, we have the Itô isometry

$$\mathbb{E}\left\|\int_0^T g(t)dB(t)\right\|_K^2=\sum_{k=1}^{d}\int_0^T \mathbb{E}\|g(t)e_k\|_K^2 dt. \tag{3.23}$$

Let us examine the generalization of the Itô Formula. We refer to the proof of Proposition 3.4. If $u\in C_b^3(K;\mathbb{R})$, we have the Taylor expansion (which generalizes (3.12))

$$u(X_{t_{i+1}})-u(X_{t_i})=Du(X_{t_i})\cdot(X_{t_{i+1}}-X_{t_i})+\frac{1}{2}D^2u(X_{t_i})\cdot(X_{t_{i+1}}-X_{t_i})^{\otimes 2}+\mathcal{O}(|X_{t_{i+1}}-X_{t_i}|^3).$$

The increment being here $X_{t_{i+1}}-X_{t_i}=\sum_{1\leq k\leq d}g(t_i)e_k\delta\beta_k(t_i)$, we have to examine in particular the term

$$\sum_{1\leq k,l\leq d}D^2u(X_{t_i})\cdot(g(t_i)e_k,g(t_i)e_l)\delta\beta_k(t_i)\delta\beta_l(t_i). \tag{3.24}$$

Ii is treated like the left-hand side of (3.17), with the additional fact that the independence of $\beta_1(t),\dots,\beta_d(t)$ comes into play and that the off-diagonal terms in (3.19), the sum over $k\neq l$, is negligible when $|\sigma|\to 0$. We obtain the Itô Formula

$$\begin{aligned}u(t,X_t)=u(0,x)&+\int_0^t\left[\frac{\partial u}{\partial s}(s,X_s)+Du(s,X_s)\cdot f(s)\right]ds\\ &\sum_{k=1}^{d}\frac{1}{2}\int_0^t D^2u(s,X_s)\cdot(g(s)e_k,g(s)e_k)ds+\int_0^t Du(s,X_s)\cdot g(s)dB(s),\end{aligned} \tag{3.25}$$

for

$$X_t=x+\int_0^t f(s)ds+\int_0^t g(s)dB(s), \tag{3.26}$$

where D in (3.25) means D_x. In (3.25), $u\colon[0,T]\times K\to\mathbb{R}$ is of class $C_b^{1,2}$. In (3.25) and (3.26), the integrands are in the following classes:

$$f\in L^1(\Omega\times[0,T],\mathcal{P}_T^*;K),\quad g\in L^2(\Omega\times[0,T],\mathcal{P}_T^*;\mathcal{L}(\mathbb{R}^d;K))).$$

A standard instance of (3.25) and (3.26) is when K is finite dimensional, $K=\mathbb{R}^m$ (often with $m=d$). Then $g(t)\in\mathcal{L}(\mathbb{R}^d;\mathbb{R}^m)$ is assimilated with its matrix representation ($d\times m$ matrix) in the canonical bases of $\mathbb{R}^d$ and $\mathbb{R}^m$, $D^2u(t,x)$, which is a bilinear form on $\mathbb{R}^m$ is assimilated to a $m\times m$ matrix, and the Itô correction term rewritten

$$\sum_{k=1}^{d}\frac{1}{2}D^2u(s,X_s)\cdot(g(s)e_k,g(s)e_k)=\frac{1}{2}\mathrm{Trace}(g(s)^*D^2u(s,X_s)g(s)). \tag{3.27}$$

3.4 Martingales and martingale characterization of the stochastic integral

Definition 3.2 (Martingale)**.** Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq0},\mathbb{P})$ be a filtered space and E a separable Banach space. Let $(X_t)_{t\geq0}$ be a L^1, E-valued process: for all $t\geq0$, $X_t\in L^1(\Omega;E)$. The process $(X_t)_{t\geq0}$ is said to be a *martingale* if, for all $0\leq s\leq t$, $X_s=\mathbb{E}(X_t|\mathcal{F}_s)$.

Proposition 3.5 (Martingale characterization of the stochastic integral)**.** *Let $(B(t))_{t\geq0}$ be a d-dimensional Wiener process given by (3.20). Let $(\mathcal{F}_t)_{t\geq0}$ be a given filtration, such that, for all k, $(\beta_k(t))$ is $(\mathcal{F}_t)$-adapted, and the increment $\beta_k(t)-\beta_k(s)$ is independent on $\mathcal{F}_s$ for all $0\leq s\leq t$. Let K be a separable Hilbert space. Let g be an integrand as in (3.21) and let $(X_t)_{t\in[0,T]}$ be an adapted, continuous, L^2-process such that $X_0=0$. Then*

$$X_t=\int_0^t g(s)dB(s),\ \forall t\in[0,T], \tag{3.28}$$

if, and only if, the processes

$$X_t,\quad \beta_k(t)X_t-\int_0^t g(s)e_kds,\quad \|X_t\|_K^2-\sum_{k=1}^{d}\int_0^t\|g(s)e_k\|_K^2ds \tag{3.29}$$

are $(\mathcal{F}_t)$-martingales.

To illustrate the interest of this proposition, consider the case where $(\mathcal{F}_t)$ is the filtration generated by $Z_t:=(X_t,B_t,Y_t)$, where (Y_t) is an other process, on a Polish space F. To test that the process (X_t) is an $(\mathcal{F}_t)$-martingale, one has to prove, for fixed $0\leq s<t\leq T$, that

$$\mathbb{E}[h_sX_t]=\mathbb{E}[h_sX_s] \tag{3.30}$$

for all h_s that is $\mathcal{F}_s$-measurable. Since $(\mathcal{F}_s)$ is generated by $(X_r,B_r,Y_r)_{0\leq r\leq s}$, this amounts to prove that for all $0\leq t_1<\cdots<t_n\leq s$, for all continuous function φ on G^n, where $G=E\times\mathbb{R}^d\times F$, one has

$$\mathbb{E}[\varphi(Z_{t_1},\dots,Z_{t_n})X_t]=\mathbb{E}[\varphi(Z_{t_1},\dots,Z_{t_n})X_s]. \tag{3.31}$$

The strength of (3.31) is that it involves the *law* of (Z_t) only. Similarly, the test of the martingale character of the two other processes in (3.29) involves the law of (Z_t) only.

Proof of Proposition 3.5. If (3.28) is realized, then (X_t) is a martingale. Indeed, every stochastic integral is a martingale. This can be seen by approximation of the integrand by elementary functions. By the Itô formula, we have

$$\|X_t\|_K^2-\sum_{k=1}^{d}\int_0^t\|g(s)e_k\|_K^2ds=2\sum_{k=1}^{d}\int_0^t\langle X_s,g(s)e_k\rangle_Kd\beta_k(s).$$

By the Itô formula and polarization, we also have

$$\beta_k(t)X_t - \int_0^t g(s)e_k ds = \int_0^t X(s)d\beta_k(s) + \sum_{l=1}^d \int_0^t \beta_k(s)g(s)e_l d\beta_l(s).$$

Once again we conclude by the fact that a stochastic integral is a martingale. The proof of the converse statement can be found in [11, Proposition A.1] for instance. Let us give some details about it. We first claim that the following identity is satisfied:

$$\mathbb{E}\left[\int_s^t \langle (X_t - X_s), \theta(\sigma)\rangle_K d\beta_k(\sigma) - \int_s^t \langle g(\sigma)e_k, \theta(\sigma)\rangle_K d\sigma \middle| \mathcal{F}_s\right] = 0 \tag{3.32}$$

for all $0 \leq s \leq t \leq T$, all $k \geq 1$ and all $\theta \in L^2_{\mathcal{P}}([0,T]\times\Omega; K)$. The proof consists in approximating θ on the interval $[s,t]$ by predictable simple functions. Note that (3.32) uses only the fact that

$$X_t, \quad X_t\beta_k(t) - \int_0^t g(s)e_k ds$$

are $(\mathcal{F}_t)$-martingales. We apply (3.32) with $s = 0$ and $\theta(\sigma) = g(\sigma)e_k$ and sum over k to obtain

$$\mathbb{E}\langle X_t, \bar{X}_t\rangle_K = \mathbb{E}\sum_{k=1}^d \int_0^t \|g(s)e_k\|_K^2 ds, \quad \bar{X}_t := \int_0^t g(s)dB(s). \tag{3.33}$$

This gives the expression of the cross-product when we expand the term $\mathbb{E}\|X_t - \bar{X}_t\|_K^2$. Using the fact that

$$\|X_t\|_K^2 - \sum_{k=1}^d \int_0^t \|g(s)e_k\|_K^2 ds$$

is a $(\mathcal{F}_t)$-martingale and applying Itō's Isometry to $\mathbb{E}\|\bar{X}(t)\|_K^2$ shows that the square terms are also given by

$$\mathbb{E}\|X_t\|_K^2 = \mathbb{E}\|\bar{X}_t\|_K^2 = \sum_{k=1}^d \int_0^t \|g(s)e_k\|_K^2 ds.$$

It follows that $X_t = \bar{X}_t$. □

4 Stochastic scalar conservation laws

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}, (B(t))_{t\geq 0})$ be a stochastic basis ($B(t)$ is an m-dimensional Wiener process). In this section we study now the Cauchy problem for the stochastic scalar conservation law

$$du(x,t) + \operatorname{div}(A(u(x,t)))dt = \sum_{k=1}^m g_k(x, u(x,t))d\beta_k(t), \quad x \in \mathbb{T}^d, t > 0, \tag{4.1}$$

where $A \in \operatorname{Lip}(\mathbb{R}; \mathbb{R}^d)$ is such that $a\colon \xi \mapsto A'(\xi)$ has at most a polynomial growth and is continuous, in the following sense: there exists $p_* \geq 1$, $C \geq 0$, such that

$$|a(\xi)| \leq C(1 + |\xi|^{p_*-1}), \tag{4.2}$$

$$\sup_{|\zeta|<\delta} |a(\xi) - a(\xi+\zeta)| \leq C(1 + |\xi|^{p_*-1})\omega(\delta), \tag{4.3}$$

for all $\xi \in \mathbb{R}$ and $\delta > 0$, where $\lim_{\delta\to 0} \omega(\delta) = 0$. In the noise term, we assume that the functions g_k satisfy, for some given constants $D_0, D_1 \geq 0$, and some exponent $\vartheta \in (1, 2]$,

$$\mathbf{G}^2(x,u) = \sum_{k=1}^{m} |g_k(x,u)|^2 \leq D_0(1 + |u|^2), \tag{4.4}$$

$$\sum_{k=1}^{m} |g_k(x,u) - g_k(y,v)|^2 \leq D_1(|x-y|^2 + |u-v|^\vartheta), \tag{4.5}$$

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

4.1 Solutions, generalized and approximate solutions

Remember that, due to the identity (1.8), the kinetic formulation can be found by computing $d\eta(u)$, for a given function η. More rigorously, this should be done on the parabolic approximation, for instance, to (4.1). In this way, one would get a term that vanishes when $\varepsilon \to 0$ and a term giving the defect measure. There are some additional terms due to the noise that are given by the Itô formula:

$$\sum_{k=1}^{m} g_k(x,u(x,t))\eta'(u(x,t))d\beta_k(t) \text{ and } \frac{1}{2}\sum_{k=1}^{m} |g_k(x,u(x,t))|^2\eta''(u(x,t))dt.$$

This yields the following definitions (compare to Definition 1.1).
Let $\mathcal{M}_b(\mathbb{T}^d \times [0,T] \times \mathbb{R})$ be the set of finite Borel signed measures on $\mathbb{T}^d \times [0,T] \times \mathbb{R}$. We denote by $\mathcal{M}_b^+(\mathbb{T}^d \times [0,T] \times \mathbb{R})$ the subset of non-negative measures.

Definition 4.1 (Random measure)**.** A map m from Ω to $\mathcal{M}_b(\mathbb{T}^d \times [0,T] \times \mathbb{R})$ is said to be a random signed measure (on $\mathbb{T}^d \times [0,T] \times \mathbb{R}$) if, for each $\phi \in C_b(\mathbb{T}^d \times [0,T] \times \mathbb{R})$, $\langle m, \phi\rangle \colon \Omega \to \mathbb{R}$ is a random variable. If almost surely $m \in \mathcal{M}_b^+(\mathbb{T}^d \times [0,T] \times \mathbb{R})$, we simply speak of random measure.

Notation: if $1 \leq p < +\infty$, and U is a given open set in $\mathbb{R}^m$, we denote by $L^p_{\mathcal{P}}(U \times [0,T] \times \Omega)$ the set $L^p(\Omega \times [0,T], \mathcal{P}_T^*; L^p(U))$. The index $\mathcal{P}$ therefore stand for "predictable".

Definition 4.2 (Solution)**.** Let $u_0 \in L^\infty(\mathbb{T}^d)$. An $L^1(\mathbb{T}^d)$-valued stochastic process $(u(t))_{t\in[0,T]}$ is said to be a solution to (4.1) with initial datum u_0 if u and $\mathtt{f} := \mathbf{1}_{u>\xi}$ have the following properties:

1. $u \in L^1_{\mathcal{P}}(\mathbb{T}^d \times [0,T] \times \Omega)$,
2. for all $\varphi \in C^1_c(\mathbb{T}^d \times \mathbb{R})$, almost surely, $t \mapsto \langle \mathtt{f}(t), \varphi\rangle$ is càdlàg,
3. for all $p \in [1, +\infty)$, there exists $C_p \geq 0$ such that

$$\mathbb{E}\left(\sup_{t\in[0,T]} \|u(t)\|^p_{L^p(\mathbb{T}^d)}\right) \leq C_p, \tag{4.6}$$

4. there exists a random measure m with first moment,

$$\mathbb{E}\|m\|_{\mathrm{TV}} = \mathbb{E}m(\mathbb{T}^d \times [0,T] \times \mathbb{R}) < +\infty, \tag{4.7}$$

such that for all $\varphi\in C^1_c(\mathbb{T}^d\times\mathbb{R})$, for all $t\in[0,T]$,

$$\begin{aligned}\langle \mathtt{f}(t),\varphi\rangle=\langle \mathtt{f}_0,\varphi\rangle&+\int_0^t\langle \mathtt{f}(s),a(\xi)\cdot\nabla\varphi\rangle ds\\&+\sum_{k=1}^m\int_0^t\int_{\mathbb{T}^d}g_k(x,u(x,s))\varphi(x,u(x,s))dxd\beta_k(s)\\&+\frac{1}{2}\int_0^t\int_{\mathbb{T}^d}\partial_\xi\varphi(x,u(x,s))\mathbf{G}^2(x,u(x,s))dxds-m_\varphi([0,t]),\end{aligned}\tag{4.8}$$

a.s., where $\mathtt{f}_0(x,\xi)=\mathbf{1}_{u_0(x)>\xi}$ and m_φ is defined by (1.6).

Remark 4.1. Note that we need to prove the measurability of the function $\sup_{t\in[0,T]}\|u(t)\|_{L^p(\mathbb{T}^d)}$ to give a sense to (4.6). Let us denote by $\bar{\mathtt{f}}=1-\mathtt{f}=\mathbf{1}_{u\leq\xi}$ the conjugate function of $\mathtt{f}$. By the identity

$$|u|^p=\int_{\mathbb{R}}\left[\mathtt{f}\mathbf{1}_{\xi>0}+\bar{\mathtt{f}}\mathbf{1}_{\xi<0}\right]p|\xi|^{p-1}d\xi,\tag{4.9}$$

we have, for $p\in[1,+\infty)$,

$$\|u(t)\|^p_{L^p(\mathbb{T}^d)}=\sup_{\psi_+\in F_+,\psi_-\in F_-}\langle \mathtt{f}(t),\psi_+\rangle+\langle\bar{\mathtt{f}}(t),\psi_-\rangle,\tag{4.10}$$

where the sup is taken over some countable sets F_+ and F_- of functions ψ chosen as follows: $F_\pm=\{\psi_n;n\geq1\}$, where (ψ_n) is a sequence of non-negative functions in $C^\infty_c(\mathbb{R})$ which converges point-wise monotonically to $\xi\mapsto p|\xi^\pm|^{p-1}$ if $p>1$ and to $\xi\mapsto\mathrm{sgn}_\pm(\xi)$ if $p=1$. By (4.10), we have

$$\sup_{t\in[0,T]}\|u(t)\|^p_{L^p(\mathbb{T}^d)}=\sup_{\psi_\pm\in F_\pm}\sup_{t\in[0,T]}\langle \mathtt{f}(t),\psi_+\rangle+\langle\bar{\mathtt{f}}(t),\psi_-\rangle.\tag{4.11}$$

By Item (2) in Definition 4.2, we know that the function

$$\sup_{t\in[0,T]}\langle \mathtt{f}(t),\psi_+\rangle+\langle\bar{\mathtt{f}}(t),\psi_-\rangle$$

is $\mathcal{F}$-measurable for all $\psi_\pm\in F_\pm$. Indeed, the sup over $[0,T]$ of a càdlàg function is the sup of the function on any dense countable subset of $[0,T]$ containing the terminal point T. By (4.11), the function $\sup_{t\in[0,T]}\|u(t)\|_{L^p(\mathbb{T}^d)}$ is measurable.

We have seen in the homogeneous case, that approximation procedures of (4.1) lead naturally to a certain notion of *generalized solution*. We give the corresponding definition in Definition 4.4 below. First we give a still more general notion of *approximate generalized solution*. This latter notion will be flexible enough to be used to show the convergence of various approximations, like parabolic, BGK, numerical, to (4.1). See Section 4.2.

Definition 4.3 (Approximate generalized solution)**.** Let $f_0\colon\mathbb{T}^d\times\mathbb{R}\to[0,1]$ be a kinetic function. An $L^\infty(\mathbb{T}^d\times\mathbb{R};[0,1])$-valued process $(f(t))_{t\in[0,T]}$ is said to be an approximate generalized solution to (4.1) of order N, of error term η, and of initial datum f_0 if $f(t)$ and $\nu_t:=-\partial_\xi f(t)$ have the following properties:

1. for all $t\in[0,T]$, almost surely, $f(t)$ is a kinetic function, and, for all $R>0$, $f\in L^1_{\mathcal{P}}(\mathbb{T}^d\times(0,T)\times(-R,R)\times\Omega)$,
2. for all $\varphi\in C^N_c(\mathbb{T}^d\times\mathbb{R})$, almost surely, the map $t\mapsto\langle f(t),\varphi\rangle$ is càdlàg,

3. for all $p \in [1,+\infty)$, there exists $C_p \geq 0$ such that

$$\mathbb{E}\left(\sup_{t\in[0,T]}\int_{\mathbb{T}^d}\int_{\mathbb{R}}|\xi|^p d\nu_{x,t}(\xi)dx\right) \leq C_p, \tag{4.12}$$

4. there exists a random measure m with first moment (4.7), there exists some adapted continuous stochastic processes $(\eta(t,\varphi))_{t\in[0,T]}$ defined for all $\varphi \in C_c^N(\mathbb{T}^d \times \mathbb{R})$ such that, for all $\varphi \in C_c^N(\mathbb{T}^d \times \mathbb{R})$, for all $t \in [0,T]$, almost surely,

$$\begin{aligned}\langle f(t),\varphi\rangle =&\langle f_0,\varphi\rangle + \int_0^t \langle f(s), a(\xi)\cdot\nabla_x\varphi\rangle ds\\ &+\int_0^t\int_{\mathbb{T}^d}\int_{\mathbb{R}} g_k(x,\xi)\varphi(x,\xi)d\nu_{x,s}(\xi)dxd\beta_k(s) + \eta(t,\varphi)\\ &+\frac{1}{2}\int_0^t\int_{\mathbb{T}^d}\int_{\mathbb{R}}\mathbf{G}^2(x,\xi)\partial_\xi\varphi(x,\xi)d\nu_{x,s}(\xi)dxds - m(\partial_\xi\varphi)([0,t]),\end{aligned} \tag{4.13}$$

Remark 4.2. We prove that

$$\sup_{t\in[0,T]}\int_{\mathbb{T}^d}\int_{\mathbb{R}}|\xi|^p d\nu_{x,t}(\xi)dx$$

is measurable, which ensures that one can take the expectation in (4.12), as in Remark 4.1.

Definition 4.4 (Generalized solution)**.** A generalized solution to (4.1) of initial datum f_0 is an approximate generalized solution to (4.1) of initial datum f_0 and error $\eta \equiv 0$.

4.2 Examples

4.2.1 Vanishing viscosity method

Consider the following parabolic approximation to (4.1):

$$du^\varepsilon + \operatorname{div}(A(u^\varepsilon))dt - \varepsilon\Delta u^\varepsilon dt = \sum_{k=1}^m g_k(x,u^\varepsilon(x,t))d\beta_k(t) \tag{4.14}$$

It defines an approximate generalized solution (f^ε) of order 2, with random measure m^ε and error η^ε defined as follows:

$$\begin{aligned} f^\varepsilon &= \mathtt{f}^\varepsilon = \mathbf{1}_{u^\varepsilon>\xi},\\ \langle m^\varepsilon,\phi\rangle &= \iint_{\mathbb{T}^d\times(0,T)}\phi(x,t,u^\varepsilon(x,t))\varepsilon|\nabla_x u^\varepsilon(x,t)|^2dxdt,\\ \eta^\varepsilon(t,\varphi) &= \varepsilon\int_0^t\iint_{\mathbb{T}^d\times\mathbb{R}}\mathtt{f}^\varepsilon(x,s,\xi)\Delta\varphi(x,\xi)d\xi dxds.\end{aligned}$$

4.2.2 Approximation by the Finite Volume method

The approximation of (4.1) by the Finite Volume method is considered in [6].

4.3 Main results

4.3.1 Uniqueness, reduction

Theorem 4.1 (Uniqueness, Reduction)**.** *Let $u_0 \in L^\infty(\mathbb{T}^d)$. Assume (4.4)-(4.5). Assume $\omega(\delta) = o(\delta^{1/2})$ in (4.3). Then we have the following results:*

1. *there is at most one solution with initial datum u_0 to (4.1).*
2. *If f is a generalized solution to (4.1) with initial datum f_0* at equilibrium: *$f_0 = \mathbf{1}_{u_0>\xi}$, then there exists a solution u to (4.1) with initial datum u_0 such that $f(x,t,\xi) = \mathbf{1}_{u(x,t)>\xi}$ a.s., for a.e. (x,t,ξ).*
3. *if u_1, u_2 are two solutions to (4.1) associated with the initial data $u_{1,0}$, $u_{2,0} \in L^\infty(\mathbb{T}^d)$ respectively, then*
$$\mathbb{E}\|(u_1(t)-u_2(t))^+\|_{L^1(\mathbb{T}^d)} \leq \|(u_{1,0}-u_{2,0})^+\|_{L^1(\mathbb{T}^d)}, \tag{4.15}$$
for all $t \in [0,T]$. This implies the L^1-contraction property, and the comparison principle for solutions.

Corollary 4.2 (Continuity in time)**.** *Let $u_0 \in L^\infty(\mathbb{T}^d)$. Assume (4.4)-(4.5). Assume $\omega(\delta) = o(\delta^{1/2})$ in (4.3). Then, for every $p \in [1,+\infty)$, the solution u to (4.1) with initial datum u_0 has a representative in the space $L^p(\Omega; L^\infty(0,T; L^p(\mathbb{T}^d)))$ with almost sure continuous trajectories in $L^p(\mathbb{T}^d)$.*

4.3.2 Convergence in law

Let us first recall the Prohorov's theorem and Skorohod's representation theorem, [1].

Theorem 4.3 (Prohorov's theorem)**.** *Let E be a Polish*[1] *space endowed with the Borel σ-algebra. Let μ_n, $n = 1,2,\dots$ be some Borel probability measures over E. Then there is equivalence between:*

1. *each subsequence of (μ_n) admits a subsequence converging weakly: for all $\phi \in C_b(E)$,*
$$\langle \mu_{n_k}, \phi\rangle \to \langle \mu, \phi\rangle,$$
2. *the family $\{\mu_n; n \in \mathbb{N}\}$ is tight: for all $\varepsilon > 0$, there is a compact $K_\varepsilon \subset E$ such that $\mu_n(K_\varepsilon) > 1-\varepsilon$*

Let $X_1, X_2, \dots$ be some E-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Assume that (X_n) is tight: for all $\varepsilon > 0$, there is a compact $K_\varepsilon \subset E$ such that $\mathbb{P}(X_n \in K_\varepsilon) > 1-\varepsilon$. The Prohorov's theorem says that, up to subsequence, (X_n) is converging in law: $\mathbb{E}\phi(X_n) \to \langle \mu, \phi\rangle$, for all $\phi \in C_b(E)$. By changing the probability space and the random variables, one can exhibit a limit random variable corresponding to the limit law. Consider for example the probability space and the random variables

$$(\Omega \times E, \mathcal{F} \times \mathcal{B}(E), \mathbb{P} \times \mu), \quad \tilde{X}_n(\omega, x) = X_n(\omega), \quad \tilde{X}(\omega, x) = x.$$

Then $\mathrm{Law}(\tilde{X}_n) = \mathrm{Law}(X_n)$, $\mathrm{Law}(\tilde{X}) = \mu$ and $(\tilde{X}_n)$ is converging in law to $\tilde{X}$. By modification of the probability space and of the random variables, one can even get almost sure convergence.

[1] separable, complete, metric space

Theorem 4.4 (Skorohod's representation theorem)**.** *Let E be a Polish space endowed with the Borel σ-algebra. Let (X_n) be a sequence of random variables which converges in law to a random variable X. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and some random variables $\tilde{X}_n$, $\tilde{X}$ on $\tilde{\Omega}$ such that*

1. *for all $n \in \mathbb{N}^*$, the random variables $\tilde{X}_n$ and X_n have the same law; $\tilde{X}$ and X have the same law,*
2. *$(\tilde{X}_n)$ is converging to $\tilde{X}$, $\tilde{\mathbb{P}}$-almost-surely.*

We use this point of view (independence regarding the probabilistic data) to define a notion of solution in law.

Definition 4.5 (Solution in law)**.** Let $u_0 \in L^\infty(\mathbb{T}^d)$. We say that $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, (\tilde{\mathcal{F}}_t)_{t\geq 0}, (\tilde{B}(t))_{t\geq 0}, \tilde{u})$ is solution in law to (4.1) with initial datum u_0 if, $\tilde{u}$ is solution to (4.1) with initial datum u_0 in the sense of Definition 4.2, where every instance of the $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0}, (B(t))_{t\geq 0})$ has been replaced by the equivalent notion with tildas (this involves the σ-algebra of predictable sets also). We have a similar definition of generalized solution in law.

We will now give the main result, about the convergence in law of a sequence of approximate solutions. The following result, Theorem 4.5 states that a sequence of approximate solutions is converging in law to a generalized solution. What is converging actually? We are manipulating different objects: f, $\nu = -\partial_\xi f$, m... We choose to put the emphasis on ν. See [5, Theorem 4.6] for a more complete result.

Let us denote by $\mathcal{P}_1(E)$ the set of probability measures on a Polish space E and by $\mathcal{Y}_1$ the set of Young measures $\mathbb{T}^d \times [0,T] \to \mathcal{P}_1(\mathbb{R})$. The set $\mathcal{Y}_1$ is itself a Polish space ([5, Proposition 4.3]).

Theorem 4.5 (Convergence in law of a sequence of approximate solutions)**.** *Let (f^n) be a sequence of approximate generalized solutions with order N, error η^n and initial datum f_0^n. Assume that*

- *for all $\varphi \in C_c^N(\mathbb{T}^d \times \mathbb{R})$, $(\eta^n(t,\varphi))_{t\in[0,T]}$ tends to 0 in probability on $C([0,T])$,*
- *the sequence (f_0^n) is converging to f_0 in $L^\infty(\mathbb{T}^d \times \mathbb{R})$ weak-$*$, where f_0 is a kinetic function,*
- *there exists $C_p \geq 0$ independent on n such that $\nu^n := -\partial_\xi f^n$ satisfies the uniform bound*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \int_{\mathbb{T}^d} \int_{\mathbb{R}} |\xi|^p d\nu_{x,t}^n(\xi) dx\right] \leq C_p, \tag{4.16}$$

- *the defect measure m^n has a first moment uniformly bounded:*

$$\mathbb{E} m^n(\mathbb{T}^d \times [0,T] \times \mathbb{R}) \text{ is uniformly bounded,} \tag{4.17}$$

and m^n vanishes for large ξ uniformly in n: if $B_R^c = \{\xi \in \mathbb{R}, |\xi| \geq R\}$, then

$$\lim_{R\to+\infty} \sup_n \mathbb{E} m^n(\mathbb{T}^d \times [0,T] \times B_R^c) = 0. \tag{4.18}$$

Then there is a generalized solution in law *$(f(t))_{t\in[0,T]}$ to (4.1) with initial datum f_0 such that, up to subsequence, the sequence of Young measures (ν^n) associated to f^n is converging in law on $\mathcal{Y}_1$ to the Young measure ν associated to f.*

4.3.3 Stability and reduction

Let us consider Theorem 4.5 in the special case where f_0 is at equilibrium (in practice, this amounts to require the strong convergence of the initial datum, which is very natural). By the reduction result in Theorem 4.1, we know that $f(t)$ is at equilibrium for all $t \in [0,T]$: $f(t) = \mathbf{1}_{u(t)>\xi}$, or, equivalently, $\nu_{x,t} = \delta_{u(x,t)}$, where u is the solution (in law) of (4.1) with initial datum u_0. Let

$$u^n(x,t) = \int_{\mathbb{R}} \xi d\nu^n_{x,t}(\xi). \tag{4.19}$$

The particular structure (Dirac mass) of the limit Young measure implies that there is strong convergence of (u^n) to u, in the sense that (u^n) is converging in law to u on $L^p(\mathbb{T}^d \times [0,T])$ for every $1 \leq p < +\infty$. To explain this strong convergence result, we refer to Lemma 1.4, or [5, Lemma 2.6]. Let us also give a brief proof in the case $p = 2$ by forgetting the random character of the objects (assume that everything is deterministic). We know that

$$\begin{aligned}\iint_{\mathbb{T}^d\times[0,T]} \psi(x,t) \int_{\mathbb{R}} \phi(\xi) d\nu^n_{x,t}(\xi) dxdt \to \iint_{\mathbb{T}^d\times[0,T]} \psi(x,t) \int_{\mathbb{R}} \phi(\xi) d\nu_{x,t}(\xi) dxdt \\ = \iint_{\mathbb{T}^d\times[0,T]} \psi(x,t)\phi(u(x,t)) dxdt,\end{aligned} \tag{4.20}$$

for all $\psi \in L^1(\mathbb{T}^d \times [0,T])$ and for all $\phi \in C(\mathbb{R})$ such that $|\phi(\xi)| \leq C(1 + |\xi|^p)$ for a given $p \geq 1$. Taking $\psi \in L^2(\mathbb{T}^d \times [0,T])$ and $\phi(\xi) = \xi$, we obtain $u^n \rightharpoonup u$ in $L^2(\mathbb{T}^d \times [0,T])$-weak. Taking $\psi = 1$ and $\phi(\xi) = |\xi|^2$, we obtain

$$\iint_{\mathbb{T}^d\times[0,T]} \psi(x,t) \int_{\mathbb{R}} |\xi|^2 d\nu^n_{x,t}(\xi) dxdt \to= \iint_{\mathbb{T}^d\times[0,T]} |u(x,t)|^2 dxdt. \tag{4.21}$$

By the Jensen inequality, it follows from (4.21) that

$$\limsup_{n\to+\infty} \|u^n\|^2_{L^2(\mathbb{T}^d\times[0,T])} \leq \|u^n\|^2_{L^2(\mathbb{T}^d\times[0,T])}.$$

Since

$$\|u - u^n\|^2_{L^2(\mathbb{T}^d\times[0,T])} = \|u\|^2_{L^2(\mathbb{T}^d\times[0,T])} + \|u^n\|^2_{L^2(\mathbb{T}^d\times[0,T])} - 2\langle u, u^n\rangle_{L^2(\mathbb{T}^d\times[0,T])},$$

we deduce that $\limsup_{n\to+\infty} \|u - u^n\|^2_{L^2(\mathbb{T}^d\times[0,T])} \leq 0$, which gives the desired result. At that stage, there remains to show that the weak mode of convergence used for random variables (convergence in law) is actually strong (convergence $L^p(\Omega)$). To do that (see the details in [5, Theorem 4.15]), we use the Gyöngy-Krylov argument, [10, Lemma 1.1]. The basis of the Gyöngy-Krylov argument is this simple fact: if a couple (X_n, Y_n) of random variables converges in law to a random variable written (Z, Z), *i.e.* concentrated on the diagonal, then $(X_n - Y_n)$ converges to 0 in probability. See Section 5.3 for more details on the Gyöngy-Krylov argument. In conclusion, we obtain the following final result.

Theorem 4.6 (Strong convergence of a sequence of approximate solutions)**.** *Assume* (4.4)-(4.5). *Assume* $\omega(\delta) = o(\delta^{1/2})$ *in* (4.3)*. Let* (f^n) *be a sequence of approximate generalized solutions as in Theorem 4.5. Assume that the limit* f_0 *is at equilibrium,* $f_0 = \mathbf{1}_{u_0>\xi}$*. Then the sequence* (u^n) *given by* (4.19) *is converging in* $L^p(\mathbb{T}^d \times [0,T] \times \Omega)$ *for all finite* $p \geq 1$ *to the unique solution* u *of* (4.1) *with initial datum* u_0*.*

4.4 Some elements of proof

4.4.1 Uniqueness, reduction

See [5, Section 3] for a complete proof. Here, we will simply explain the main ideas. If f_1 and f_2 are two generalized solution, we have, in a weak sense, using the variables (x,ξ) for f_1 and (y,ζ) for f_2,

$$f_1(x,t,\xi) = f_1(x,0,\xi) - \int_0^t a(\xi)\cdot\nabla_x f_1(x,s,\xi)ds - \sum_{k=1}^m \int_0^t \mathbf{G}^2(x,\xi)\partial_\xi \nu^1_{x,s}(\xi)ds + \partial_\xi m^1([0,t]) + \sum_{k=1}^m \int_0^t g_k(x,\xi)\nu^1_{x,s}(\xi)d\beta_k(s), \quad (4.22)$$

and

$$f_2(y,t,\zeta) = f_2(y,0,\zeta) - \int_0^t a(\zeta)\cdot\nabla_y f_2(y,s,\zeta)ds - \sum_{k=1}^m \int_0^t \mathbf{G}^2(y,\zeta)\partial_\zeta \nu^2_{y,s}(\zeta)ds + \partial_\zeta m^2([0,t]) + \sum_{k=1}^m \int_0^t g_k(y,\zeta)\nu^2_{y,s}(\zeta)d\beta_k(s). \quad (4.23)$$

Introduce the conjugate function $\bar{f}_2 = 1 - f_2$, which satisfies

$$\bar{f}_2(y,t,\zeta) = \bar{f}_2(y,0,\zeta) - \int_0^t a(\zeta)\cdot\nabla_y \bar{f}_2(y,s,\zeta)ds + \sum_{k=1}^m \int_0^t \mathbf{G}^2(y,\zeta)\partial_\zeta \nu^2_{y,s}(\zeta)ds - \partial_\zeta m^2([0,t]) - \sum_{k=1}^m \int_0^t g_k(y,\zeta)\nu^2_{y,s}(\zeta)d\beta_k(s). \quad (4.24)$$

Using the Itô formula, we deduce that, in the weak sense,

$$\begin{aligned} f_1(x,t,\xi)\bar{f}_2(y,t,\zeta) &= f_1(x,0,\xi)\bar{f}_2(y,0,\zeta) \\ &- \int_0^t (a(\xi)\cdot\nabla_x f_1(x,s,\xi))\bar{f}_2(y,s,\zeta) + (a(\zeta)\cdot\nabla_y \bar{f}_2(y,s,\zeta))f_1(x,t,\xi)ds \\ &- \sum_{k=1}^m \int_0^t \mathbf{G}^2(x,\xi)\partial_\xi \nu^1_{x,s}(\xi)\bar{f}_2(y,s,\zeta) - \mathbf{G}^2(y,\zeta)\partial_\zeta \nu^2_{y,s}(\zeta)f_1(x,s,\xi)ds \\ &+ \int_{[0,t]} (\partial_\xi m^1(x,s,\xi)\bar{f}_2(y,s,\zeta) - \partial_\zeta m^2(y,s,\zeta)f_1(x,s,\xi)) \\ &+ \sum_{k=1}^m \int_0^t g_k(x,\xi)(\nu^1_{x,s}(\xi)\bar{f}_2(y,s,\zeta) - \nu^2_{y,s}(\zeta)f_1(x,s,\xi))d\beta_k(s), \\ &- \sum_{k=1}^m \int_0^t g_k(x,\xi)\nu^1_{x,s}(\xi)g_k(y,\zeta)\nu^2_{y,s}(\zeta)ds. \end{aligned} \quad (4.25)$$

We test (4.25) against a test function of the form $\varphi(x,y,\xi,\zeta) = \rho(x-y)\psi(\xi-\zeta)$ with $\rho,\psi \geq 0$. Using the identities

$$(\nabla_x + \nabla_y)\varphi = 0, \quad (\partial_\xi + \partial_\zeta)\varphi = 0,$$

and taking expectancy, we obtain

$$\mathbb{E}\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\rho(x-y)\psi(\xi-\zeta)f_1(x,t,\xi)\bar{f}_2(y,t,\zeta)d\xi d\zeta dxdy$$
$$=\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\rho(x-y)\psi(\xi-\zeta)f_{1,0}(x,\xi)\bar{f}_{2,0}(y,\zeta)d\xi d\zeta dxdy+\mathrm{K}+\mathrm{I}_\rho+\mathrm{I}_\psi, \quad (4.26)$$

where

$$\mathrm{K}=\mathbb{E}\int_{[0,t]}\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}f_1(x,s,\xi)\partial_\zeta\varphi dm^2(y,s,\zeta)-\bar{f}_2(y,s,\zeta)\partial_\xi\varphi dm^1(x,s,\xi),$$

and

$$\mathrm{I}_\rho=\mathbb{E}\int_0^t\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}f_1(x,s,\xi)\bar{f}_2(y,s,\zeta)(a(\xi)-a(\zeta))\psi(\xi-\zeta)d\xi d\zeta\cdot\nabla_x\rho(x-y)dxdyds$$

and

$$\mathrm{I}_\psi=\frac{1}{2}\int_{(\mathbb{T}^d)^2}\rho(x-y)\mathbb{E}\int_0^t\int_{\mathbb{R}^2}\psi(\xi-\zeta)\sum_{k=1}^m|g_k(x,\xi)-g_k(y,\zeta)|^2d(\nu^1_{x,s}\otimes\nu^2_{y,s})(\xi,\zeta)dxdyds. \quad (4.27)$$

To obtain the expression (4.27), we have used the following kind of transformations:

$$\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\mathbf{G}^2(x,\xi)\partial_\xi\nu^1_{x,s}(\xi)\bar{f}_2(y,s,\zeta)\varphi(x,y,\xi,\zeta)$$
$$=-\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\mathbf{G}^2(x,\xi)\nu^1_{x,s}(\xi)\bar{f}_2(y,s,\zeta)\partial_\xi\varphi(x,y,\xi,\zeta)$$
$$=\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\mathbf{G}^2(x,\xi)\nu^1_{x,s}(\xi)\bar{f}_2(y,s,\zeta)\partial_\zeta\varphi(x,y,\xi,\zeta)$$
$$=-\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\mathbf{G}^2(x,\xi)\nu^1_{x,s}(\xi)\partial_\zeta\bar{f}_2(y,s,\zeta)\varphi(x,y,\xi,\zeta)$$
$$=-\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\mathbf{G}^2(x,\xi)\nu^1_{x,s}(\xi)\nu^2_{y,s}(\zeta)\varphi(x,y,\xi,\zeta).$$

The same kind of computations give

$$\mathrm{K}=-\mathbb{E}\int_{[0,t]}\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\varphi\nu^1_{x,s}(\xi)dm^2(y,s,\zeta)+\varphi\nu^2_{y,s}(\zeta)dm^1(x,s,\xi),$$

and thus $\mathrm{K}\leq 0$. Consequently, we have the estimate

$$\mathbb{E}\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\rho(x-y)\psi(\xi-\zeta)f_1(x,t,\xi)\bar{f}_2(y,t,\zeta)d\xi d\zeta dxdy$$
$$\leq\int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2}\rho(x-y)\psi(\xi-\zeta)f_{1,0}(x,\xi)\bar{f}_{2,0}(y,\zeta)d\xi d\zeta dxdy+\mathrm{I}_\rho+\mathrm{I}_\psi, \quad (4.28)$$

We then take $\psi:=\psi_\delta$ and $\rho=\rho_\varepsilon$ where (ψ_δ) and (ρ_ε) are approximations of the identity on $\mathbb{R}$ and $\mathbb{T}^d$ respectively, *i.e.*

$$\psi_\delta(\xi)=\frac{1}{\delta}\psi\left(\frac{\xi}{\delta}\right), \quad \rho_\varepsilon(x)=\frac{1}{\varepsilon^d}\rho\left(\frac{x}{\varepsilon}\right),$$

where ψ and ρ are some given smooth probability densities on $\mathbb{R}$ and $\mathbb{T}^d$ respectively. Using the estimate (4.5), it is easy to get the bound

$$\mathrm{I}_\psi \leq \frac{tD_1}{2}(\varepsilon^2\delta^{-1} + \delta^{\vartheta-1}). \tag{4.29}$$

Integration by parts also gives

$$\mathrm{I}_\rho = \mathbb{E}\int_0^t \int_{(\mathbb{T}^d)^2}\int_{\mathbb{R}^2} \Upsilon_\delta(\xi,\zeta) d(\nu^1_{x,s}\otimes\nu^2_{y,s})(\xi,\zeta)\cdot\nabla_x\rho(x-y)dxdyds,$$
$$\Upsilon_\delta(\xi,\zeta) := \int_{\xi'=-\infty}^{\xi}\int_{\zeta'=\zeta}^{+\infty}(a(\xi')-a(\zeta'))\psi_\delta(\xi'-\zeta')d\xi'd\zeta'.$$

We write

$$\Upsilon_\delta(\xi,\zeta) := \int_{\xi'=-\infty}^{\xi}\int_{\zeta'=\zeta-\xi'}^{+\infty}(a(\xi')-a(\zeta'+\xi'))\psi_\delta(\zeta')d\xi'd\zeta'$$

and use (4.3) to get the estimate

$$\Upsilon_\delta(\xi,\zeta)\leq C\omega(\delta)\int_{\zeta+\delta}^{\xi}(1+|\xi'|^{p-1})d\xi \lesssim \omega(\delta)(1+|\xi|^p+|\zeta|^p).$$

It follows that

$$|\mathrm{I}_\rho| \lesssim t\varepsilon^{-1}\omega(\delta). \tag{4.30}$$

If $\omega(\delta)=o(\delta^{1/2})$, say $\omega(\delta)=\delta^{1/2}\omega'(\delta)$, we take $\varepsilon=(\delta\omega'(\delta))^{1/2}$ and obtain

$$|\mathrm{I}_\rho|+|\mathrm{I}_\psi| = \mathcal{O}((\omega'(\delta))^{1/2}+\delta^{\vartheta/2}) = o(1).$$

Taking the limit $\delta\to 0$ in (4.28) gives us

$$\mathbb{E}\int_{\mathbb{T}^d}\int_{\mathbb{R}} f_1(t)\bar{f}_2(t)dxd\xi \leq \int_{\mathbb{T}^d}\int_{\mathbb{R}} f_{1,0}\bar{f}_{2,0}dxd\xi. \tag{4.31}$$

Take $f_1=f_2=0$. If f_0 is at equilibrium, then $f_0\bar{f}_0\equiv 0$. Therefore $f(t)\bar{f}(t)\equiv 0$, almost surely, by (4.31). This implies that $f(t)$ is at equilibrium. If $f_i=\mathtt{f}_i=\mathbf{1}_{u_i>\xi}$, then

$$\int_{\mathbb{R}} f_1\bar{f}_2 d\xi = (u_1-u_2)^+.$$

The estimate (4.15) is therefore a consequence of (4.31).

4.4.2 Convergence of approximations

A consequence of (4.16) is the bound

$$\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}\int_{\mathbb{R}}|\xi|^p d\nu^n_{x,t}(\xi)dxdt\right]\leq C_pT, \tag{4.32}$$

Let us first explain how one can use the bound (4.32). Recall that $\mathcal{Y}_1$ is the set of Young measures $\mathbb{T}^d\times[0,T]\to\mathcal{P}_1(\mathbb{R})$. It is a standard result from the theory of Young measures that, for all $R>0$, the set

$$K_R=\left\{\nu\in\mathcal{Y}_1;\int_0^T\int_{\mathbb{T}^d}\int_{\mathbb{R}}|\xi|^p d\nu^n_{x,t}(\xi)dxdt\leq R\right\}$$

is compact. We consider here the topology $\tau_{\mathcal{Y}^1}^W$ defined in [2, p.21], which admits a compatible, complete, metric, [2, Proposition 2.3.1]. A sequence $(\nu^\varepsilon)_{\varepsilon\in\varepsilon_\mathbb{N}}$ of Young measures is converging to a Young measure ν if (1.21) is satisfied when $\varepsilon\to 0$ in $\varepsilon_\mathbb{N}$ for all $\psi\in L^1(\mathbb{T}^d\times[0,T])$ and $\phi\in C_b(\mathbb{R})$. We exploit (4.32) in a very simple way: by the Markov inequality, we have

$$\mathbb{P}(\nu^n\notin K_R)\leq\frac{1}{R}\mathbb{E}\left[\int_0^T\int_{\mathbb{T}^d}\int_\mathbb{R}|\xi|^p d\nu^n_{x,t}(\xi)dxdt\right]\leq\frac{C_pT}{R}.$$

Let $\varepsilon>0$. For $R>C_pT\varepsilon^{-1}$, the probability $\mathbb{P}(\nu^n\notin K_R)$ is $<\varepsilon$, uniformly in n. This shows that (ν^n) is tight in $\mathcal{Y}_1$. By the Skorokhod theorem, up to a change of probability space, ν^n is converging almost surely to a limiting Young measure ν. The same kind of arguments, based on the bounds (4.17) and (4.18), provides a limit in law for the sequence (m^n). Taking the limit in the stochastic integral is done using the martingale characterization of the stochastic integral, Proposition 3.5, which is adapted to *convergence in law* (see (3.31) and the related comments). The remaining arguments are quite similar to those developed in Section 1, when considering the convergence of the parabolic approximation. Let us explain more precisely how we take the limit in the stochastic integral. To simplify the presentation, we will give the main arguments in the next section in finite dimension.

4.4.3 Convergence in law in SDEs

Let $X^n_t\in\mathbb{R}$ solve the SDE

$$dX^n_t=b(X^n_t)dt+\sigma(X^n_t)dB_t,$$

in the sense that

$$X^n_t=x_0+\int_0^t b(X^n_s)ds+\int_0^t\sigma(X^n_s)dB_s,\tag{4.33}$$

for all $t\geq 0$, where $x_0\in\mathbb{R}$ is given. Suppose that the following uniform bound is satisfied: there exists $\alpha\in(0,1)$ and $C\geq 0$ such that

$$\sup_n\mathbb{E}\|X^n\|_{C^\alpha([0,T])}\leq C.\tag{4.34}$$

Assume also b,σ continuous, with sub-linear growth. We may assume that b and σ are Lipschitz continuous, in which case one can prove (4.34) using some relatively standard arguments (Kolmogorov's continuity criterion, martingale estimates...).

Tightness. Let $R>0$ and

$$K_R=\left\{X\in C([0,T]);\|X\|_{C^\alpha([0,T])}\leq R\right\}.$$

By the Ascoli theorem, K_R is a compact subset of $C([0,T])$. By (4.34), and the Markov inequality, we have

$$\mathbb{P}(X^n\notin K_R)=\mathbb{E}\mathbf{1}_{R<\|X^n\|_{C^\alpha([0,T])}}\leq\frac{\mathbb{E}\|X^n\|_{C^\alpha([0,T])}}{R}\leq\frac{C}{R}.$$

Therefore (X^n) is tight in $C([0,T])$. By Prokhorov's Theorem, up to a subsequence, (X^n) is converging in law to a process X. Two related questions arise then:

1. what is the equation satisfied by X?
2. How to pass to the limit in Equation (4.33)?

Skorokhod's theorem. Let us apply the Skorokhod theorem. There exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and some random variables $\tilde{X}^n$, $\tilde{X}$ such that

$$\mathrm{Law}(\tilde{X^n}) = \mathrm{Law}(X^n), \quad \mathrm{Law}(\tilde{X}) = \mathrm{Law}(X),$$

and $\tilde{X}^n \to \tilde{X}$ a.s. in $C([0,T])$. By the dominated convergence theorem, we have, $\tilde{\mathbb{P}}$-almost surely,

$$\int_0^t b(\tilde{X}_s^n)ds \to \int_0^t b(\tilde{X}_s)ds, \forall t \in [0,T].$$

What can we say about the stochastic integral

$$\int_0^t \sigma(X_s^n)dB_s \ ? \tag{4.35}$$

A preliminary problem is to determine what is the right version of (4.35) on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Skorokhod's theorem for (X^n, B). The couple $(X_t^n, B_t)_{t\in[0,T]}$ is tight in $C([0,T])\times C([0,T])$ (this is obvious since (B_t) is a stationary sequence). Consequently, there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ such that, up to subsequence, there exists some random variables $(\tilde{X}^n, \tilde{B}^n)$, $(\tilde{X}, \tilde{B})$ such that

$$\mathrm{Law}(\tilde{X}^n, \tilde{B}^n) = \mathrm{Law}(X^n, B),$$

and $(\tilde{X}^n, \tilde{B}^n) \to (\tilde{X}, \tilde{B})$ a.s. In particular, we have $\tilde{B}^n \to \tilde{B}$ in law and thus $\mathrm{Law}(B) = \mathrm{Law}(\tilde{B})$. By Exercise 2.10, we know then that $(\tilde{B}_t)_{t\in[0,T]}$ is a Wiener process. Let $(\tilde{\mathcal{F}}_t^0)_{t\in[0,T]}$ be the filtration generated by $(\tilde{X}_t, \tilde{B}_t)_{t\in[0,T]}$. Let $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$ be the natural augmentation of $(\tilde{\mathcal{F}}_t^0)_{t\in[0,T]}$, which is defined as follows: one first take the completion of $(\tilde{\mathcal{F}}_t^0)_{t\in[0,T]}$:

$$\tilde{\mathcal{F}}_t^1 = \left(\tilde{\mathcal{F}}_t^0\right)^* .$$

Recall (see for instance [17], Section *The role played by the sets of measure zero*) that $\tilde{\mathcal{F}}_t^1$ is the collections of sets E, for which there exists $A, B \in \tilde{\mathcal{F}}_t^0$ such that

$$A \subset E \subset B, \quad \tilde{\mathbb{P}}(B \setminus A) = 0. \tag{4.36}$$

Recall also that the probability measure is also extended in this procedure, into a probability measure $\tilde{\mathbb{P}}_t^*$ defined by $\tilde{\mathbb{P}}_t^*(E) = \tilde{\mathbb{P}}(B) = \tilde{\mathbb{P}}(A)$. We have indicated by a subscript t the dependence of $\tilde{\mathbb{P}}_t^*$ on t: this is the case since the sets A, B such that (4.36) occurs are taken in $\tilde{\mathcal{F}}_t^0$. However, if

$$A' \subset E \subset B', \quad \tilde{\mathbb{P}}(B' \setminus A') = 0,$$

with $A', B' \in \tilde{\mathcal{F}}_s^0$ for a $s \neq t$, then $A \setminus A' \subset B' \setminus B$, so $\tilde{\mathbb{P}}(A \setminus A') = 0$. Similarly, we have $\tilde{\mathbb{P}}(A' \setminus A) = 0$, hence $\tilde{\mathbb{P}}(A) = \tilde{\mathbb{P}}(A')$. Therefore the extension $\tilde{\mathbb{P}}_t^*$ does not depend on t. For simplicity, we will also drop the star, and simply denote it by $\tilde{\mathbb{P}}$. Once $\tilde{\mathcal{F}}_s^1$ has been defined, one sets

$$\tilde{\mathcal{F}}_t = \tilde{\mathcal{F}}_{t+}^1 = \bigcap_{s>t} \tilde{\mathcal{F}}_s^1.$$

We have then the following propositions.

Proposition 4.7. *The set $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, (\tilde{\mathcal{F}}_t)_{t\in[0,T]}, (\tilde{B}_t)_{t\in[0,T]})$ is a stochastic basis.*

Proposition 4.8. *Let*

$$\tilde{Y}_t = \tilde{X}_t - x_0 - \int_0^t b(\tilde{X}_s)ds.$$

Then the processes

$$\tilde{Y}_t, \quad \tilde{Y}_t\tilde{B}_t - \int_0^t \sigma(\tilde{X}_s)ds, \quad |\tilde{Y}_t|^2 - \int_0^t |\sigma(\tilde{X}_s)|^2 ds$$

are $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$-martingales.

We use the martingale characterization of the stochastic integral, Proposition 3.5, to deduce that

$$\tilde{Y}_t = \int_0^t \sigma(\tilde{X}_s)dB_s.$$

This shows that $(\tilde{X}_t)$ satisfies the limit equation

$$\tilde{X}(t) = x_0 + \int_0^t b(\tilde{X}(s))ds + \int_0^t \sigma(\tilde{X}(s))d\tilde{\beta}(s). \tag{4.37}$$

Note that the stochastic integral in (4.37) is well defined because $(\tilde{X}_t)$ is a continuous process and is $\tilde{F}_t$-measurable.

Proof of Proposition 4.7. It is clear that $\tilde{B}_t$ is $\tilde{\mathcal{F}}_t$-measurable. Let $0 \leq s < t \leq T$. To show that the increment $\tilde{B}_t - \tilde{B}_s$ is independent on $\tilde{\mathcal{F}}_t$, we proceed as follows. Let $s < \tau < \sigma < t$. Let us first prove that $\tilde{B}_t - \tilde{B}_\sigma$ is independent on $\tilde{\mathcal{F}}^0_\tau$. Since $(\tilde{\mathcal{F}}^0_t)_{t\in[0,T]}$ is the filtration generated by $(\tilde{X}_t, \tilde{B}_t)_{t\in[0,T]}$, this amounts to prove that, for every $k \geq 1$, for every $0 \leq \tau_1 \leq \cdots \leq \tau_k \leq \tau$, for all bounded continuous function Φ on $\mathbb{R}^{2k}$, and all bounded continuous function ψ on $\mathbb{R}$, one has

$$\tilde{\mathbb{E}}\left[\Phi(\tilde{X}_{\tau_1}, \tilde{B}_{\tau_1}, \ldots, \tilde{X}_{\tau_k}, \tilde{B}_{\tau_k})\psi(\tilde{B}_t - \tilde{B}_\sigma)\right] = \tilde{\mathbb{E}}\left[\Phi(\tilde{X}_{\tau_1}, \tilde{B}_{\tau_1}, \ldots, \tilde{X}_{\tau_k}, \tilde{B}_{\tau_k})\right]\tilde{\mathbb{E}}\left[\psi(\tilde{B}_t - \tilde{B}_\sigma)\right]. \tag{4.38}$$

We will only consider the case $\psi(x) = x$ for simplicity, since in that case the right-hand side of (4.38) is zero (the case $\psi(x) = x$ also happens to be the model case for the adaptation of the present proof to the proof of Proposition 4.8). By identity of the laws, we know that

$$\tilde{\mathbb{E}}\left[\Phi(\tilde{X}^n_{\tau_1}, \tilde{B}^n_{\tau_1}, \ldots, \tilde{X}^n_{\tau_k}, \tilde{B}^n_{\tau_k})(\tilde{B}^n_t - \tilde{B}^n_\sigma)\right] = \mathbb{E}\left[\Phi(X^n_{\tau_1}, B_{\tau_1}, \ldots, X^n_{\tau_k}, B_{\tau_k})(B_t - B_\sigma)\right], \tag{4.39}$$

for every n. Since (X^n_t) is adapted and $(B_t - B_\sigma)$ is independent on $\mathcal{F}_r$ for all $r \leq \sigma$, the right-hand side of (4.39) is zero. Using the Vitali convergence theorem, we also know that the limit of the left-hand side of (4.39) is the left-hand side of (4.38). Indeed, let $\tilde{\zeta}_n$ denote the argument of $\tilde{\mathbb{E}}$ in the left-hand side of (4.39). It satisfies

$$\tilde{\zeta}_n = \Phi(\tilde{X}^n_{\tau_1}, \tilde{B}^n_{\tau_1}, \ldots, \tilde{X}^n_{\tau_k}, \tilde{B}^n_{\tau_k})(\tilde{B}^n_t - \tilde{B}^n_\sigma) \to \tilde{\zeta} := \Phi(\tilde{X}_{\tau_1}, \tilde{B}_{\tau_1}, \ldots, \tilde{X}_{\tau_k}, \tilde{B}_{\tau_k})(\tilde{B}_t - \tilde{B}_\sigma)$$

almost surely. Furthermore, the sequence $(\tilde{\zeta}_n)$ is equi-integrable on $\tilde{\Omega}$ because

$$\tilde{\mathbb{E}}|\tilde{\zeta}_n|^2 \leq \|\Phi\|^2_{C_b(\mathbb{R}^{2k})}\tilde{\mathbb{E}}|\tilde{B}^n_t - \tilde{B}^n_\sigma|^2 = \|\Phi\|^2_{C_b(\mathbb{R}^{2k})}(t - \sigma).$$

Therefore we obtain (4.38), which gives

$$\tilde{\mathbb{E}}\left[\mathbf{1}_A(\tilde{B}_t - \tilde{B}_\sigma)\right] = 0, \tag{4.40}$$

for any set $A \in \tilde{\mathcal{F}}^0_\tau$. It is clear that this remains true if we replace A by E where E is such that (4.36) is satisfied for $A, B \in \tilde{\mathcal{F}}^0_\tau$. Therefore (4.40) holds true when $A \in \tilde{\mathcal{F}}^1_\tau$. Since $\tau \in (s, \sigma)$ is arbitrary, we obtain (4.40) for $A \in \tilde{\mathcal{F}}^1_{s+} = \tilde{\mathcal{F}}_s$. Then we can take the limit $\sigma \to s$. Here we use the L^2-continuity of the Wiener process, for example. Indeeed, $\mathbb{E}|\tilde{B}_\sigma - \tilde{B}_s|^2 = (\sigma - s) \to 0$ when $\sigma \to s$. Eventually, we obtain

$$\tilde{\mathbb{E}}\left[\mathbf{1}_A(\tilde{B}_t - \tilde{B}_s)\right] = 0,$$

for all $A \in \tilde{\mathcal{F}}_s$. □

Proof of Proposition 4.8. Let us show for example that

$$\tilde{Z}_t := \tilde{Y}_t\tilde{B}_t - \int_0^t \sigma(\tilde{X}_s)ds$$

is a $(\tilde{\mathcal{F}}_t)$-martingale. Let $0 \leq s < t \leq T$. Assume we have proved that, for every $s < \tau < \sigma < t$, we have

$$\mathbb{E}\left[\mathbf{1}_A\tilde{Z}_t\right] = \mathbb{E}\left[\mathbf{1}_A\tilde{Z}_\sigma\right], \tag{4.41}$$

for every A in $\tilde{\mathcal{F}}^0_\tau$. Then, as in the proof of Proposition 4.7, we deduce that (4.41) holds true for $A \in \tilde{\mathcal{F}}_s$, and then we use an argument of continuity to take the limit $\sigma \to s$, to get

$$\mathbb{E}\left[\mathbf{1}_A\tilde{Z}_t\right] = \mathbb{E}\left[\mathbf{1}_A\tilde{Z}_s\right], \tag{4.42}$$

for all A in $\tilde{\mathcal{F}}_s$. Since $\tilde{Z}_s$ is $\tilde{\mathcal{F}}_s$-measurable (it is $\tilde{\mathcal{F}}^0_s$-measurable), the right-hand side of (4.42) is $\tilde{Z}_s$ almost-surely. This gives us the desired result. To establish (4.41), we proceed as in the proof of Proposition 4.7, we just need to replace $\tilde{B}_t - \tilde{B}_\sigma$ by $\tilde{Z}_t - \tilde{Z}_\sigma$ and to do the minor adaptations that are required. □

5 Compensated compactness

In this section, we consider the case $d = 1$ (spatial dimension 1) and assume that the flux A satisfies (4.2) and the following non-degeneracy hypothesis

$$\xi \mapsto a(\xi) \text{ is not constant on any open interval.} \tag{5.1}$$

We will prove the convergence (convergence in law first, then convergence in probability) on $L^p(\mathbb{T} \times [0, T])$, $1 \leq p < +\infty$, of (u^ε), the solution to the parabolic approximation (4.14), to the solution u of (4.1). We will use a compensated compactness method based on the div-curl lemma. Extensions of this compensated compactness method to higher space-dimensions exist, at least in the deterministic case. They use microlocal defect measures, [9], or the equivalent concept of H-measures, [19], see for example the works of E. Yu. Panov and collaborators [14, 12, 15].

5.1 Estimates on the divergence

Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be a convex function of class C^2 with η' and η'' bounded. Let q be associated entropy flux defined by (1.10). By the Itô formula and (4.14), the (t, x)-divergence of the 2 dimensional vector $(\eta(u^\varepsilon), q(u^\varepsilon))$ is given by

$$\partial_t\eta(u^\varepsilon) + \partial_x q(u^\varepsilon) = \varepsilon\partial_x^2\eta(u^\varepsilon) - \varepsilon\eta''(u^\varepsilon)|\partial_x u^\varepsilon|^2 + \frac{1}{2}\mathbf{G}^2(x, u^\varepsilon)\eta''(u^\varepsilon) + \partial_t M^\varepsilon_t, \tag{5.2}$$

where

$$M_t^\varepsilon(x,t)=\int_0^t\sum_{k=1}^m g_k(x,u^\varepsilon(x,t))\eta'(u^\varepsilon(x,t))d\beta_k(t). \tag{5.3}$$

Let us assume without loss of generality that $T=1$. If $v\in L^1(\mathbb{T}\times(0,1))$ and $\theta\in C_c^\infty(0,1)$ is a cut off function, we can view the product θv as a 1-periodic function in x, but also in t. Therefore, we consider θv as a function of $L^1(\mathbb{T}^2)$. Recall that, for $s\in\mathbb{R}$, the Sobolev space $H^s(\mathbb{T}^2)$ is defined as the set of tempered distributions v whose Fourier coefficients $\hat{v}(n)=\langle v,e_n\rangle$, $e_n(x)=\exp(2\pi i x\cdot n)$ satisfy

$$\|v\|^2_{H^s(\mathbb{T}^2)}=\sum_{n\in\mathbb{Z}^2}\langle n\rangle^{2s}|\hat{v}(n)|^2<+\infty,\quad \langle n\rangle=(1+|n|^2)^{1/2}.$$

By analysing each the terms in the right-hand side of (5.2), we will show the following result.

Proposition 5.1. *Let u^ε be the solution to (4.14) with initial datum $u_0\in L^\infty(\mathbb{T})$. Let $\varepsilon_\mathbb{N}=\{\varepsilon_n;n\in\mathbb{N}\}$, where $(\varepsilon_n)\downarrow 0$. Let $\theta\in C_c^\infty(0,1)$. The sequence of general term*

$$\mathrm{div}_{t,x}(\theta\eta(u^\varepsilon),\theta q(u^\varepsilon))=\partial_t(\theta\eta(u^\varepsilon))+\partial_x(\theta q(u^\varepsilon)),\quad \varepsilon\in\varepsilon_\mathbb{N},$$

is tight in $H^{-1}(\mathbb{T}^2)$: for all $\delta>0$, there exists a compact set K_δ of $H^{-1}(\mathbb{T}^2)$ such that

$$\mathbb{P}(\mathrm{div}_{t,x}(\eta(u^\varepsilon),q(u^\varepsilon))\in K_\delta)>1-\delta,$$

for all $\varepsilon\in\varepsilon_\mathbb{N}$.

Proof. we know that u^ε satisfies the bounds

$$\mathbb{E}\|u^\varepsilon\|^p_{L^p(\mathbb{T}\times(0,1))}\lesssim_p 1,\quad \mathbb{E}\left[\varepsilon\|\partial_x u^\varepsilon\|^2_{L^2(\mathbb{T}\times(0,1))}\right]\lesssim 1. \tag{5.4}$$

Since η has at sublinear growth, the first estimate in (5.4) gives

$$\mathbb{E}\|\theta'\eta(u^\varepsilon)\|^2_{L^2(\mathbb{T}^2)}\lesssim 1. \tag{5.5}$$

For $R>0$. We introduce the set

$$A_{1,R}=\left\{v\in L^2(\mathbb{T}^2);\|v\|_{L^2(\mathbb{T}^2)}\leq R\right\},$$

which is relatively compact in $H^{-1}(\mathbb{T}^2)$. Using (5.4) and the Markov inequality, we obtain the estimate

$$\mathbb{P}(\eta(u^\varepsilon)\in A_{1,R})\leq\frac{1}{R}\mathbb{E}\|\eta(u^\varepsilon)\|_{L^2(\mathbb{T}\times(0,1))}\lesssim R^{-1}.$$

We have $\varepsilon\theta\partial_x^2\eta(u^\varepsilon)\to 0$ in $H^{-1}(\mathbb{T}^2)$ when $\varepsilon\to 0$ in $\varepsilon_\mathbb{N}$, conditionally to the event

$$B_{\varepsilon,R}=\left\{\varepsilon^{1/2}\partial_x u^\varepsilon\in A_{1,R}\right\}.$$

Due to (5.4) and the Markov inequality, we have $\mathbb{P}(B_{\varepsilon,R})\lesssim R^{-1}$. This implies that $\varepsilon\theta\partial_x^2\eta(u^\varepsilon)\to 0$ in $H^{-1}(\mathbb{T}^2)$ in probability when $\varepsilon\to 0$ in $\varepsilon_\mathbb{N}$, which in turn implies that the sequence $(\varepsilon\theta\partial_x^2\eta(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ is tight in $H^{-1}(\mathbb{T}^2)$. Since η'' is bounded, the second estimate in (5.4) gives

$$\mathbb{P}(\varepsilon\theta\eta''(u^\varepsilon)|\partial_x u^\varepsilon|^2\in A_{2,R})\lesssim R^{-1},\quad A_{2,R}=\left\{v\in L^1(\mathbb{T}^2);\|v\|_{L^1(\mathbb{T}^2)}\leq R\right\}.$$

Similarly, due to (4.4) and (5.4), we have

$$\mathbb{E}[\|\theta\mathbf{G}^2(x,u^\varepsilon)\eta''(u^\varepsilon)\|_{L^1(\mathbb{T}^2)}] \lesssim 1,$$

from which we deduce that

$$\mathbb{P}(\psi_\varepsilon \in A_{2,R}) \lesssim R^{-1}, \quad \psi\varepsilon = \theta(-\varepsilon\eta''(u^\varepsilon)|\partial_x u^\varepsilon|^2 + \frac{1}{2}\mathbf{G}^2(x,u^\varepsilon)\eta''(u^\varepsilon)).$$

Let $\theta\partial_t M_t^\varepsilon$ denote the distribution $\partial_t(\theta M_t^\varepsilon) - \theta' M_t^\varepsilon$. At that stage, we will admit (this is the object of Lemma 5.2 below), that $\mathbb{P}(\theta\partial_t M_t^\varepsilon \in A_{3,R}) \lesssim R^{-1}$, where $A_{3,R}$ is a set that is relatively compact in $H^{-1}(\mathbb{T}^2)$. We can thus write

$$\operatorname{div}_{t,x}(\theta\eta(u^\varepsilon), \theta q(u^\varepsilon)) = \chi^\varepsilon + \psi^\varepsilon,$$

where

$$\chi^\varepsilon := \theta'\eta(u^\varepsilon) + \varepsilon\theta\partial_x^2\eta(u^\varepsilon) + \theta\partial_t M_t^\varepsilon$$

defines a tight sequence in $H^{-1}(\mathbb{T}^2)$ and

$$\psi^\varepsilon := \theta(-\varepsilon|\partial_x u^\varepsilon|^2 + \frac{1}{2}\mathbf{G}^2(x,u^\varepsilon))\eta''(u^\varepsilon)$$

defines a sequence which is stochastically bounded in $L^1(\mathbb{T}^2)$. Since the sequences $(\eta(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ and $(q(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ are bounded, in expectancy, in L^p (the bound on $q(u^\varepsilon)$ uses Hypothesis (4.2)), they are stochastically bounded, and thus $\operatorname{div}_{t,x}(\theta\eta(u^\varepsilon), \theta q(u^\varepsilon))$ is stochastically bounded in $W^{-1,p}(\mathbb{T}^2)$ for a given $p > 2$. We can apply Lemma A.3 in [7], which is a probabilistic version of Murat's Lemma, to conclude. □

Lemma 5.2. *The sequence $(\partial_t(\theta M_t^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ is tight in $H^{-1}(\mathbb{T}^2)$.*

Proof. The proof is in essential the proof of Lemma 4.19 in [7]. However, we will proceed slightly differently (instead of using Marchaud fractional derivative we work directly with fractional Sobolev spaces and an Aubin-Simon compactness lemma). Let $0 \le s \le t \le T$. By the Burkholder-Davis-Gundy Inequality, we have

$$\mathbb{E}\|M^\varepsilon(t) - M^\varepsilon(s)\|_{L^4(\mathbb{T})}^4 \lesssim \int_\mathbb{T} \mathbb{E}\left|\int_s^t |\eta'(u^\varepsilon)|^2\mathbf{G}^2(x,u^\varepsilon)d\sigma\right|^2 dx,$$

and, using the Hölder Inequality, the bound (4.4) and the fact that η' is sub-linear,

$$\mathbb{E}\|M^\varepsilon(t) - M^\varepsilon(s)\|_{L^4(\mathbb{T})}^4 \lesssim |t-s| \int_s^t \mathbb{E}\int_\mathbb{T}(1 + |u^\varepsilon(x,\sigma)|^4)d\sigma dx.$$

By (5.4), it follows that

$$\mathbb{E}\|M^\varepsilon(t) - M^\varepsilon(s)\|_{L^4(\mathbb{T})}^4 \lesssim |t-s|^2. \tag{5.6}$$

Integrating with respect to t and s, we find that

$$\mathbb{E}\int_0^1\int_0^1 \frac{\|(\theta M^\varepsilon)(t) - (\theta M^\varepsilon)(s)\|_{L^4(\mathbb{T})}^4}{|t-s|^{1+2\nu}} dtds \lesssim 1, \tag{5.7}$$

as soon as $\nu < 1/2$. The left-hand side in this inequality (5.7) is the norm of θM^ε in the space $L^4(\Omega; W^{\nu,4}(0,1;L^4(\mathbb{T})))$. Since $L^4(\mathbb{T}) \hookrightarrow H^{-1}(\mathbb{T})$, it follows that

$$\mathbb{E}\|\theta M^\varepsilon\|_{W^{\nu,4}(0,1;H^{-1}(\mathbb{T}))}^4 \lesssim 1.$$

We use the continuous injection

$$W^{\nu,4}(0,1;H^{-1}(\mathbb{T}))\hookrightarrow C^{0,\mu}([0,1];H^{-1}(\mathbb{T}))$$

for every $0<\mu<\nu-\frac{1}{4}$ to obtain

$$\mathbb{E}\|\theta M^\varepsilon\|^4_{C^{0,\mu}([0,1];H^{-1}(\mathbb{T}))}\lesssim 1. \tag{5.8}$$

Besides, taking $s=0$ in (5.6) and summing with respect to $t\in(0,T)$ gives also

$$\mathbb{E}\|\theta M^\varepsilon\|^4_{L^4(\mathbb{T}\times(0,1))}\lesssim 1. \tag{5.9}$$

By the Aubin-Simon compactness Lemma, [18], the set

$$A_R:=\left\{M\in L^2(\mathbb{T}\times(0,1));\|M\|_{C^{0,\mu}([0,1];H^{-1}(\mathbb{T}))}\leq R,\ \|M\|_{L^4(\mathbb{T}\times(0,1))}\leq R\right\}$$

is compact in $C([0,1];H^{-1}(\mathbb{T}))$, hence compact in $L^2(0,1;H^{-1}(\mathbb{T}))$. Consequently (5.8) and (5.9) show that (θM^ε) is tight as a $L^2(0,1;H^{-1}(\mathbb{T}))$-random variable, and we conclude that $(\partial_t(\theta M^\varepsilon))$ is tight as a $H^{-1}(\mathbb{T}^2)$-random variable. □

5.2 Application of the div-curl lemma

Let $\nu^\varepsilon_{x,t}(\xi)=\delta_{u^\varepsilon(x,t)}(\xi)$ be the Young measure associated to u^ε. By (5.4), we have the estimate

$$\mathbb{E}\iint_{\mathbb{T}\times[0,1]}\int_{\mathbb{R}}|\xi|^p d\nu^\varepsilon_{x,t}(\xi)dxdt\lesssim_p 1.$$

Like in Theorem 4.5, we deduce from this estimate and from the Prokhorov theorem that the sequence $(\nu^\varepsilon)_{\varepsilon\in\varepsilon_\mathbb{N}}$ has a subsequence converging in law in $\mathcal{Y}_1$ (*i.e.* in the sense of Young measures) to a Young measure ν. Let $\Upsilon=\{\eta_k;k\in\mathbb{N}\}$ be a countable collection sub-quadratic entropies, with associated fluxes q_k, and let $\Theta=\{\theta_j;j\in\mathbb{N}\}$ be a countable set of cut-off functions on the time interval $(0,1)$. By a diagonal process and the Prokhorov theorem, there is a subsequence (not relabeled) such that, for all $k,j\in\mathbb{N}$, the sequence $\mathrm{div}_{t,x}(\theta_j\eta_k(u^\varepsilon),\theta_j q_k(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ is converging in law in $H^{-1}(\mathbb{T}^2)$. The stationary sequence (B_t) is of course also tight in $C([0,T];\mathbb{R}^m)$. We use the Skorokhod theorem: by changing the probability space, we may assume a.e. convergence of all the random variables considered above. More precisely, we obtain simultaneous convergence in law of the set of unknowns

$$\left(\{\nu^\varepsilon,\mathrm{div}_{t,x}(\theta\eta(u^\varepsilon),\theta q(u^\varepsilon)),(B_t);\theta\in\Theta,\eta\in\Upsilon\}\right)_{\varepsilon\in\varepsilon_\mathbb{N}}. \tag{5.10}$$

We will keep the same notation $(\Omega,\mathcal{F},\mathbb{P})$ for the probability space and assume $\mathbb{P}$-a.e. convergence since it will not cause any confusion in the following arguments. Indeed, our aim is simply to explain how do we apply the div-curl lemma in our case. We proceed as follows: by (5.4), we know that, for each k, there is a set $N_k\subset\Omega$ of measure zero such that, on the complement on N_k, $(\eta_k(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ and $(q_k(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ are bounded in $L^2(\mathbb{T}\times(0,1))$ (the bound may depend on the $\omega\in N_k^c$). Let $N_{(1)}$ be the union over k of the sets N_k. We have $\mathbb{P}(N_{(1)})=0$. Similarly, there is a measurable set $N_{(2)}$ of probability zero, such that for every $\omega\notin N_{(2)}$, for all j,k, the sequence $(\mathrm{div}_{t,x}(\theta_j\eta_k(u^\varepsilon),\theta_j q_k(u^\varepsilon))_{\varepsilon\in\varepsilon_\mathbb{N}}$ is relatively compact in $H^{-1}(\mathbb{T}^2)$. Eventually, there is a set $N_{(3)}$ of measure zero, such that, for all $\omega\notin N_{(3)}$, for all $\phi\in C_b(\mathbb{R})$, and when $\varepsilon\to 0$ in $\varepsilon_\mathbb{N}$,

$$\phi(u^\varepsilon)\to\langle\nu,\phi\rangle \text{ in } L^\infty(\mathbb{T}\times(0,1)) \text{ weak-}*\,. \tag{5.11}$$

It is easy to prove, by means of a cut-off function and (5.4), that the convergence (5.11) holds true when ϕ is continuous and polynomially bounded, and that the limit occurs also in $L^2(\mathbb{T}\times(0,1))$-weak. Let $N = N_{(1)} \cup N_{(2)} \cup N_{(3)}$. We use the identity

$$\operatorname{curl}\begin{pmatrix} \theta_j q_k(u^\varepsilon) \\ -\theta_j \eta_k(u^\varepsilon)\end{pmatrix} = \begin{pmatrix} 0 & \operatorname{div}_{t,x}(\theta_j\eta_k(u^\varepsilon),\theta_j q_k(u^\varepsilon))_{\varepsilon\in\varepsilon_{\mathbb{N}}} \\ -\operatorname{div}_{t,x}(\theta_j\eta_k(u^\varepsilon),\theta_j q_k(u^\varepsilon))_{\varepsilon\in\varepsilon_{\mathbb{N}}} & 0\end{pmatrix},$$

and the div-curl lemma (on the complementary set of N) to infer that, for all j,k,k', we have

$$\theta_j(\eta_k(u^\varepsilon)q_{k'}(u^\varepsilon) - \eta_{k'}(u^\varepsilon)q_k(u^\varepsilon)) \to \theta_j(\langle\nu,\eta_k\rangle\langle\nu,q_{k'}\rangle - \langle\nu,\eta_{k'}\rangle\langle\nu,q_k\rangle) \tag{5.12}$$

in the sense of measures on $\mathbb{T}\times[0,1]$. We also know by (5.11) that

$$\theta_j(\eta_k(u^\varepsilon)q_{k'}(u^\varepsilon) - \eta_{k'}(u^\varepsilon)q_k(u^\varepsilon)) \to \theta_j\langle\nu,\eta_k q_{k'} - \eta_{k'}q_k\rangle.$$

Let $\{\theta_j; j\in\mathbb{N}\}$ be dense in $C_c(0,1)$ for the C^0 convergence. It follows then that

$$\langle\nu_{x,t},\eta_k q_{k'} - \eta_{k'}q_k\rangle = \langle\nu_{x,t},\eta_k\rangle\langle\nu_{x,t},q_{k'}\rangle - \langle\nu_{x,t},\eta_{k'}\rangle\langle\nu_{x,t},q_k\rangle, \tag{5.13}$$

for a.e. $(x,t)\in\mathbb{T}\times(0,1)$. There are various ways to conclude, once the identity (5.13) has been obtained. We will give a proof using the kinetic function $f(x,t,\xi) = \nu_{x,t}(\xi,+\infty)$. Take for η_k the entropy $u\mapsto(u-\xi_k)^+$ where $\{\xi_k; k\in\mathbb{N}\}$ is a dense subset of $\mathbb{R}$. It then follows from (5.13) that

$$\langle\nu_{x,t},\eta\tilde{q} - \tilde{\eta}q\rangle = \langle\nu_{x,t},\eta\rangle\langle\nu_{x,t},\tilde{q}\rangle - \langle\nu_{x,t},\tilde{\eta}\rangle\langle\nu_{x,t},q\rangle, \tag{5.14}$$

for each η, $\tilde{\eta}$ of the form $\eta(u) = (u-\xi)^+$ and $\tilde{\eta}(u) = (u-\zeta)^+$, where $\xi,\zeta\in\mathbb{R}$. We differentiate with respect to ξ and ζ in (5.14) to get, after some computations, the following identity:

$$(a(\xi)-a(\zeta))\left[\int_{\mathbb{R}}\mathbf{1}_{u>\xi\vee\zeta}d\nu_{x,t}(u) - \int_{\mathbb{R}}\mathbf{1}_{u>\xi}d\nu_{x,t}(u)\int_{\mathbb{R}}\mathbf{1}_{u>\zeta}d\nu_{x,t}(u)\right] = 0.$$

This can be written more simply as

$$(a(\xi)-a(\zeta))(f(x,t,\xi\vee\zeta) - f(x,t,\xi)f(x,t,\zeta)) = 0. \tag{5.15}$$

Let us fix $\xi\in\mathbb{R}$. Due to the non-degeneracy condition (5.1), there is a sequence $(\zeta_n)\uparrow\xi$ such that $a(\zeta_n)\neq a(\xi)$ for all n. From (5.15), we deduce that $f(x,t,\xi)(1-f(x,t,\zeta_n)) = 0$ for all n. We then take the limit $n\to+\infty$. Since $\zeta\mapsto f(x,t,\zeta)$ is continuous from the left, we deduce that $f(x,t,\xi)(1-f(x,t,\xi)) = 0$. We have seen in the end of Section 4.4.1 that this implies the fact that f is at equilibrium, *i.e.* $\nu_{x,t}$ is a Dirac mass. Note (*cf.* Lemma 1.4) that the fact that $\nu_{x,t}$ is a Dirac mass, say $\delta_{u(x,t)}$, implies the convergence in law of the considered sequence of $(u^\varepsilon)_{\varepsilon\in\varepsilon_{\mathbb{N}}}$ to u on the space $L^p(\mathbb{T}\times(0,1))$ with its strong topology ($p\geq 1$ is finite). The limit u is a martingale solution of (4.1).

In conclusion, we obtain, up to subsequence, the convergence in law to a martingale solution of (4.1). In the next section we explain how convergence in probability can be proved, using pathwise uniqueness for the limit problem.

Before we come to that point, note that one can wonder if there may be two subsequences of $(u^\varepsilon)_{\varepsilon\in\varepsilon_{\mathbb{N}}}$ having two different limits in law. This will not be the case if uniqueness of martingale solution has been proved. We do not know how to give a specific proof of uniqueness of martingale solutions. All that we know is how to prove the pathwise uniqueness result of Theorem 4.1 (*pathwise* uniqueness result because (4.15) gives, in the case $u_{0,1} = u_{0,2}$, the *almost-sure* identity $u_1 = u_2$ in $L^1(\mathbb{T}^d\times(0,T))$). However, using such a uniqueness result is quite in contradiction with showing compensated compactness, since, as we have seen, combining generalized solutions and a slightly reinforced uniqueness result (*i.e.* the reduction result of Theorem 4.1) we obtain the strong convergence of the parabolic approximation, without such restriction as $d = 1$ and (5.1). To prove the convergence of the parabolic approximation of the first-order stochastic scalar conservation law (4.1), compensated compactness may be not the most appropriate tool therefore. Note however that the restriction on the dimension, $d = 1$, can be relaxed by means of a different analysis using microlocal defect measures, [9], *cf.* [15]. Note also that this technique can be applied to mixed first-second order conservation laws, [12], while, for such equations, an approach using generalized kinetic solutions and a reduction result has not been developed yet.

5.3 Gyöngy-Krylov argument

Let $1 \leq p < +\infty$ be fixed. We want to prove that the sequence $(u^{\varepsilon_n})_{n\in\mathbb{N}}$ is converging in probability in $L^p(\mathbb{T}\times(0,T))$. Let $\delta > 0$. Assume, by contradiction, that the Cauchy condition

$$\lim_{n,m\to+\infty} \mathbb{P}(\|u^{\varepsilon_n} - u^{\varepsilon_m}\|_{L^p(\mathbb{T}\times(0,T))} > \delta) = 0$$

is not satisfied. This means that there exists $\eta > 0$, a subsequence $(u^{\varepsilon_{n_k}})$, a subsequence $(u^{\varepsilon_{m_l}})$ such that

$$\mathbb{P}(\|u^{\varepsilon_{n_k}} - u^{\varepsilon_{m_l}}\|_{L^p(\mathbb{T}\times(0,T))} > \delta) > \eta, \quad \forall k, l \in \mathbb{N}. \tag{5.16}$$

Let us now consider the following set of unknowns (except for the noise component, we duplicate the unknowns considered in (5.10)):

$$(\{\nu^{\varepsilon_{n_k}}, \mathrm{div}_{t,x}(\theta\eta(u^{\varepsilon_{n_k}}), \theta q(u^{\varepsilon_{n_k}})), \nu^{\varepsilon_{m_l}}, \mathrm{div}_{t,x}(\theta\eta(u^{\varepsilon_{m_l}}), \theta q(u^{\varepsilon_{m_l}})), (B_t); \theta \in \Theta, \eta \in \Upsilon\})_{k,l\in\mathbb{N}}.$$

We apply the analysis of Section 5.2 to deduce, up to a further subsequence which we will not indicate (this does not affect (5.16)), that there is a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and some random variables $(\tilde{u}^{\varepsilon_{n_k}})$, $(\tilde{u}^{\varepsilon_{m_l}})$ having the same laws as $(u^{\varepsilon_{n_k}})$ and $(u^{\varepsilon_{m_l}})$ respectively, such that

$$\tilde{u}^{\varepsilon_{n_k}} \to \tilde{u}, \quad \tilde{u}^{\varepsilon_{m_l}} \to \tilde{u}'$$

almost surely in $L^p(\mathbb{T}\times(0,T))$. We obtain two martingale solutions $\tilde{u}$ and $\tilde{u}'$ which are defined on the same probability space and satisfy (4.1) with the same driving noise term. At this stage, we should also discuss what filtration is considered, to justify that it is the same for both $\tilde{u}$ and $\tilde{u}'$, let us simply say that the filtration $(\tilde{\mathcal{F}}_t)$ that we consider is the "natural augmentation" (this corresponds to taking completion and then right-limit, see Section 4.4.3 for more details) of the natural filtration of the limit process $(\tilde{u}(t), \tilde{u}'(t), \tilde{B}_t)$. Since $\tilde{u}$ and $\tilde{u}'$ are solutions with the same set of probabilistic data, we can apply the pathwise result of Theorem 4.1 to deduce that $\tilde{u} = \tilde{u}'$ almost surely. Let $h_\delta \in C(\mathbb{R}_+)$ be the following regularization of the function $s \mapsto \mathbf{1}_{s>\delta}$: $h_\delta(s) = \min(1, \delta^{-1}s)$. We have then $\mathbf{1}_{s>\delta} \leq h_\delta(s)$, which gives the bound from above

$$\mathbf{1}_{\|u-v\|_{L^p(\mathbb{T}\times(0,T))}>\delta} \leq \Phi_\delta(u,v) := h_\delta(\|u-v\|_{L^p(\mathbb{T}\times(0,T))}),$$

where Φ_δ is continuous bounded function on $L^p(\mathbb{T}\times(0,T)) \times L^p(\mathbb{T}\times(0,T))$. By (5.16), we have

$$\eta < \mathbb{E}\Phi_\delta(u^{\varepsilon_{n_k}}, u^{\varepsilon_{m_l}}),$$

and by the result of convergence in law that we have established,

$$\mathbb{E}\Phi_\delta(u^{\varepsilon_{n_k}}, u^{\varepsilon_{m_l}}) = \tilde{\mathbb{E}}\Phi_\delta(\tilde{u}^{\varepsilon_{n_k}}, \tilde{u}^{\varepsilon_{m_l}}) \to \tilde{\mathbb{E}}\Phi_\delta(\tilde{u}, \tilde{u}) = 0,$$

and thus $\eta < 0$, a contradiction.

References

[1] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[2] C. Castaing, P. Raynaud de Fitte, and M. Valadier. *Young measures on topological spaces*, volume 571 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2004. With applications in control theory and probability theory.

[3] K. L. Chung and R. J. Williams. *Introduction to stochastic integration*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, second edition, 1990.

[4] R. DiPerna. Measure-valued solutions to conservation laws. *Arch. Rational Mech. Anal.*, 88(3):223–270, 1985.

[5] S. Dotti and J. Vovelle. Convergence of approximations to stochastic scalar conservation laws. Nov. 2016.

[6] S. Dotti and J. Vovelle. Convergence of the finite volume method for scalar conservation laws with multiplicative noise: an approach by kinetic formulation. *hal-01391073*, 2016.

[7] J. Feng and D. Nualart. Stochastic scalar conservation laws. *J. Funct. Anal.*, 255(2):313–373, 2008.

[8] P. K. Friz and M. Hairer. *A Course on Rough Paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[9] P. Gérard. Microlocal defect measures. *Comm. Partial Differential Equations*, 16(11):1761–1794, 1991.

[10] I. Gyöngy and N. Krylov. Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields*, 105(2):143–158, 1996.

[11] M. Hofmanová. Degenerate parabolic stochastic partial differential equations. *Stochastic Process. Appl.*, 123(12):4294–4336, 2013.

[12] H. Holden, K. H. Karlsen, D. Mitrovic, and E. Y. Panov. Strong compactness of approximate solutions to degenerate elliptic-hyperbolic equations with discontinuous flux function. *Acta Math. Sci. Ser. B (Engl. Ed.)*, 29(6):1573–1612, 2009.

[13] P.-L. Lions, B. Perthame, and E. Tadmor. A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Amer. Math. Soc.*, 7(1):169–191, 1994.

[14] E. Y. Panov. On existence and uniqueness of entropy solutions to the Cauchy problem for a conservation law with discontinuous flux. *J. Hyperbolic Differ. Equ.*, 6(3):525–548, 2009.

[15] E. Y. Panov. Existence and strong pre-compactness properties for entropy solutions of a first-order quasilinear equation with discontinuous flux. *Arch. Ration. Mech. Anal.*, 195(2):643–673, 2010.

[16] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[17] W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[18] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

[19] L. Tartar. H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 115(3-4):193–230, 1990.