The Fokker-Planck equation with BV drift coefficients

Master's Thesis

for the degree of

Master of Science (M.Sc.)

Mathematics

at the Faculty of Sciences of
Friedrich-Alexander-Universität Erlangen-Nürnberg

submitted on 13. 9. 2023

by **Christian Lange**

Supervisor: Prof. Dr. Hannes Meinlschmidt

Friedrich-Alexander-Universität
Naturwissenschaftliche Fakultät

Contents

1 Introduction

One of the easiest partial differential equations is the linear transport equation

$$\partial_t p + b_i \partial_i p = 0$$
$$p(0,\cdot) = p_0$$

where we have $p : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}$ (one may think for example of a distribution of particles at time t and at space x) and the vector field $b : \mathbb{R}_0^+ \times \mathbb{R}^n \to \mathbb{R}^n$, which is called the *drift vector field*, so it can be thought as the transport vector of the particles.
If b is of C^1-regularity, it is very easy to solve this equation by the method of characteristics: We define $\xi : \mathbb{R} \to \mathbb{R}^n$ by the following ODE:

$$\xi'(t) = b(t, \xi(t))$$

Then one can check, that solutions of the transport equations fulfill

$$\frac{\mathrm{d}}{\mathrm{d}t}[p(t, \xi(t))] = 0$$

So solutions are constant on the curves in space-time given by ξ (called *characteristics*). With this in hand it is possible to derive an explicit formula for solutions the linear transport equation.
Here one can make an interesting observation: The fact, that the solution of the transport equation is constant along the characteristics, does not change by replacing p by $\beta \circ p$ with a C^1-function $\beta : \mathbb{R} \to \mathbb{R}$. So, for a solution $p(x,t)$, also $\beta(p(x,t))$ is a solution (for a moment ignoring the initial data p_0). We call a solution a *renormalized solution*, if this concatination is also a solution for any C^1-function β.
This concept can be generalized very far, it is also well defined for drift fields b that have Sobolev-regularity or even BV-regularity (instead of C^1-regularity, which is necessary for the use of the Picard-Lindelöf-theorem in the method of characteristics). The idea is the following: We define weak solutions of the transport equation. Then we show that every weak solution also fulfills this renormalization property. After this, it is possible to show uniqueness of solutions with this renormalization property.
To show the renormalization property, we will use approximation by convolutions and have to deal with so called *commutators*, defined as following for a differential operator or a function c

$$[\rho_\varepsilon, c](f) = \rho_\varepsilon * (cf) - c(\rho_\varepsilon * f)$$

So a commutator marks the difference between convoluting first and applying c then and the other way around. It will be important to show that these commutators converge to 0 as $\varepsilon \to 0$. Therefore we will have some *commutator estimates*.

All these steps do not only work for a transport equation, but also for a fokker-planck-equation:

$$\partial_t p + \partial_i(p b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) = 0$$

Compared to the transport equation, we see that it is in divergence form (which is equivalent, if we have some regularity conditions on $b^\sigma = b - \frac{1}{2}\operatorname{div}(\sigma\sigma^t)$, see [5], Section 7) and that we have a diffusion Matrix σ. For intuition, if $b = 0$ and $\sigma = \mathbb{1}_{n\times n}$ we get a heat equation, so one can think of σ as a matrix which is describing the diffusion of the particle distribution p.
After adapting the steps above a little it, it is possible to define also renormalized solutions for fokker-planck-equations and use them to show uniqueness of solutions. The main source is [14], which is the first one showing uniqueness for a fokker-planck-equation with a drift with only BV-regularity in space.
Another source is [5]. There also a Fokker-planck-equation is considered, but with drift coefficients in a Sobolev-space.
The theory of a BV-drift was first solved in [2], but only for a transport equation. This is also the source of one of the two big commutator estimates. The other one, used mainly for the terms from the diffusion term is from [7].

2 Analytic preparations

In this chapter we prove some analytic lemmata, which we will need later.

2.1 Functions of bounded variation

So called *functions of bounded variation* will be very important in the following chapter, especially in the proof of Ambrosios commutator estimate Theorem 3.18. The definitions and the statements are from [2] and [3].

Definition 2.1. Let $b \in L^1(U)$ for $U \subset \mathbb{R}^n$ open. b is of **bounded Variation** or a **BV-function**, if its distributional derivative is given by a vector-valued finite Radon measure, so if there is a finite Radon measure $Db = (D_1 b, ..., D_n b)$ such that

$$\int_U b \frac{\partial \varphi}{\partial x_i} \, \mathrm{d}x = -\int_U \varphi \, \mathrm{d}D_i b$$

for all $\varphi \in C_c^\infty(U)$ and $i = 1, ..., n$.
The space of functions of bounded Variation is called $BV(U)$.
$BV_{\mathrm{loc}}(U)$ is the usual local version, so the space of all functions which are of bounded variation on every compact subset of R^n.

For a $\mathbb{R}^{m\times n}$-valued measure λ we have the total variation $|\lambda|$ given by

$$|\lambda|(C) := \sup \left\{ \sum_{i=1}^{\infty} |\lambda(C_i)| : C_i \in \mathscr{B}(\Omega) \text{ pairwise disjoint}, C_i \subset C \right\}$$

with the Hilbert-Schmidt-norm in the sum. As usual we decompose Db in its singular and absolute continoous part with respect to the Lebesgue-measure by the Radon-Nikodym theorem, so lets set $Db = D^a b + D^s b$ with $|D^a b| \ll \mathscr{L}^n$ and $|D^s b| \perp \mathscr{L}^n$. $\nabla b = \frac{\partial b}{\partial x_i}$ is the density of $D^a b$ with respect to $\mathscr{L}^n$.

Lemma 2.2. *Let there be $b \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ and $z \in \mathbb{R}^n$. Then there holds*

$$\int_K |b(x+z) - b(x)| \, \mathrm{d}x \le \left| \sum_{i=1}^{n} z_i D_i b \right| \left(K_{|z|} \right)$$

for a compact $K \subset \mathbb{R}^n$ and $K_{|z|} = \left\{ x \in \mathbb{R}^n | \operatorname{dist}(x,K) < |z| \right\}$ the $|z|$-neighborhood of K.

Proof. (see also [3], Lemma 3.24 and Remark 3.25) First we take a sequence $b_k \in C^\infty(K_{|z|}) \cap BV(K_{|z|})$ approximating b in the following sense (according to Theorem 5.3 in [10])

- $b_k \to b$ in $L^1(K_{|z|})$

- $\|Db_k\|(K_{|z|}) \to \|Db\|(K_{|z|})$

Then we have (by adding $-b_k(x+z)+b_k(x+z)-b_k(x)+b_k(x)$ and using the L^1-approximation, Fubinis theorem and the approximation of the derivative):

$$\begin{aligned}
\int_K |b(x+z)-b(x)|\,\mathrm{d}x &\leq \lim_{k\to\infty}\int_K |b_k(x+z)-b_k(x)|\,\mathrm{d}x \\
&= \lim_{k\to\infty}\int_K \left|\int_0^1 \sum_{i=1}^n D_i b_k(x+tz)z_i\,\mathrm{d}t\right|\mathrm{d}x \\
&\leq \lim_{k\to\infty}\int_0^1\int_K \left|\sum_{i=1}^n D_i b_k(x+tz)z_i\right|\mathrm{d}x\,\mathrm{d}t \\
&\leq \lim_{k\to\infty}\int_0^1 \left|\sum_{i=1}^n z_i D_i b_k\right|(K_{|z|})\,\mathrm{d}t \\
&= \left|\sum_{i=1}^n z_i D_i b\right|(K_{|z|})
\end{aligned}$$

□

Lemma 2.3. *Let there be μ a locally finite measure on $\mathbb{R}$. Then for $\varepsilon > 0$ we define the following functions:*

$$\hat{\mu}_\varepsilon(t) := \frac{\mu([t,t+\varepsilon])}{\varepsilon}$$

Then for a compact set $K \subset \mathbb{R}$ we have

$$\int_K \hat{\mu}_\varepsilon(t)\,\mathrm{d}t \leq \mu(K_\varepsilon) \tag{1}$$

with $K_\varepsilon = \{x \in \mathbb{R} \mid \operatorname{dist}(x,K) < \varepsilon\}$ the ε-neighborhood of K.
Additionally, if $\mu \ll \mathscr{L}^1$, $\hat{\mu}_\varepsilon$ converges in $L^1_{\mathrm{loc}}(\mathbb{R})$ to the density of μ with respect to $\mathscr{L}^1$ for $\varepsilon \to 0$

Proof. We prove (1) first. We have

$$\hat{\mu}_\varepsilon(t) = \int_{\mathbb{R}} \frac{\mathbb{1}_{[-\varepsilon,0]}}{\varepsilon}(t-s)\,\mathrm{d}\mu(s)$$

Thus we get using Fubinis theorem

$$\begin{aligned}
\int_K \hat{\mu}_\varepsilon(t)\,\mathrm{d}t &= \int_K\int_{\mathbb{R}} \frac{\mathbb{1}_{[-\varepsilon,0]}}{\varepsilon}(t-s)\,\mathrm{d}\mu(s)\,\mathrm{d}t = \int_{\mathbb{R}}\int_K \frac{\mathbb{1}_{[-\varepsilon,0]}}{\varepsilon}(t-s)\,\mathrm{d}t\,\mathrm{d}\mu(s) \\
&\leq \int_{K_\varepsilon}\int_K \frac{\mathbb{1}_{[-\varepsilon,0]}}{\varepsilon}(t-s)\,\mathrm{d}t\,\mathrm{d}\mu(s) \\
&\leq \int_{K_\varepsilon} 1\,\mathrm{d}\mu(s) \\
&= \mu(K_\varepsilon)
\end{aligned}$$

This shows (1).
For the second property, let there be f the density of μ with respect to $\mathscr{L}^1$. So for any compact set, we have to show $\|\hat{\mu}_\varepsilon - f\|_{L^1(K)} \to 0$ as $\varepsilon \to 0$:

$$\begin{aligned}
\|\hat{\mu}_\varepsilon - f\|_{L^1(K)} &= \int_K |\hat{\mu}_\varepsilon(t) - f(t)|\,\mathrm{d}t = \int_K \left|\frac{1}{\varepsilon}\int_t^{t+\varepsilon} f(s)\,\mathrm{d}s - f(t)\right| \mathrm{d}t \\
&\leq \frac{1}{\varepsilon}\int_K \int_t^{t+\varepsilon} |f(s) - f(t)|\,\mathrm{d}s\,\mathrm{d}t = \frac{1}{\varepsilon}\int_K \int_0^{\varepsilon} |f(s+t) - f(t)|\,\mathrm{d}s\,\mathrm{d}t \\
&= \frac{1}{\varepsilon}\int_0^{\varepsilon} \int_K |f(s+t) - f(t)|\,\mathrm{d}t\,\mathrm{d}s
\end{aligned}$$

Now we take a function $\tilde{f} \in C_c^\infty(\mathbb{R})$ (which will approximate f as $C_c^\infty(\mathbb{R})$ is dense in $L^1(\mathbb{R})$). Then we have for fixed t:

$$\begin{aligned}
&\int_K |f(s+t) - f(t)|\,\mathrm{d}s \\
&\quad \leq \int_K |f(s+t) - \tilde{f}(s+t)|\,\mathrm{d}t + \int_K |\tilde{f}(s+t) - \tilde{f}(t)|\,\mathrm{d}t + \int_K |\tilde{f}(t) - f(t)|\,\mathrm{d}t \\
&\quad \leq 2\|f - \tilde{f}\|_{L^1(K_\varepsilon)} + \int_K |\tilde{f}(s+t) - \tilde{f}(t)|\,\mathrm{d}t
\end{aligned}$$

By choosing $\tilde{f}$ we can get $\|f - \tilde{f}\|_{L^1(K_\varepsilon)}$ arbitrarily small, so it remains to show

$$\frac{1}{\varepsilon}\int_0^{\varepsilon} \int_K |f(s+t) - f(t)|\,\mathrm{d}t\,\mathrm{d}s \to 0$$

for a smooth f:

$$\begin{aligned}
\frac{1}{\varepsilon}\int_0^{\varepsilon} \int_K |f(s+t) - f(t)|\,\mathrm{d}t\,\mathrm{d}s &\leq \frac{1}{\varepsilon}\int_0^{\varepsilon} \int_K \|f'\|_\infty s\,\mathrm{d}t\,\mathrm{d}s \\
&\leq \frac{|K|\|f'\|_\infty}{\varepsilon}\int_0^{\varepsilon} s\,\mathrm{d}s \\
&= \frac{|K|\|f'\|_\infty}{\varepsilon}\frac{\varepsilon^2}{2} \to 0
\end{aligned}$$

as $\varepsilon \to 0$. This was to show. □

The following lemma is about splitting BV-functions into components. In one variable, we also write also b' for the density of the absolutely continuous part of the derivative of b:

Lemma 2.4. *Let there be $b \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ and $x' \in R^{n-1}$. Then we define $b_{x'}(s) = b(x', s)$ for $s \in \mathbb{R}$. For $\mathscr{L}^{n-1}$-almost every x' we have the following:*

- $b_{x'} \in BV_{\mathrm{loc}}(\mathbb{R})$
- $b'_{x'}(s) = \frac{\partial b}{\partial x_n}(x', s)$ *for* $\mathscr{L}^1$*-almost every* $s \in \mathbb{R}$
- *for any* $\varepsilon > 0$ *we have* $b_{x'}(s+\varepsilon) - b_{x'}(s) = Db_{x'}([s, s+\varepsilon])$ *for* $\mathscr{L}^1$*-almost every* $s \in \mathbb{R}$
- $\int_{\mathbb{R}^{n-1}} |D^s b'_x| \, \mathrm{d}x' \le |D^s b|$

Proof. see Theorem 3.103, Theorem 3.107 and (3.108) in [3] □

We will need the following Lemma of Ambrosio on difference quotients of BV-functions in the proof of the commutator estimate Theorem 3.18. It states, loosely spoken, that also the difference quotients of a BV-function can be decomposed in a singular and a absolutely continuous part:

Lemma 2.5. *Let* $b \in BV_{\mathrm{loc}}(\mathbb{R}^n)$ *and* $z \in R^n$*. Then for* $\varepsilon > 0$ *the difference quotient in direction* z *can be decomposed in functions* $b^1_\varepsilon(z)$ *(the "absolutely continuous" part) and* $b^2_\varepsilon(z)$ *(the "singular" part) both in* $L^1_{\mathrm{loc}}(\mathbb{R}^n)$*:*

$$\frac{b(x+\varepsilon z) - b(x)}{\varepsilon} = b^1_\varepsilon(z)(x) + b^1_\varepsilon(z)(x)$$

$b^1_\varepsilon(z)$ *and* $b^2_\varepsilon(z)$ *can be chosen with the following properties:*

- $b^1_\varepsilon(z)$ *converges strongly in* $L^1_{\mathrm{loc}}(\mathbb{R}^n)$ *to* $\sum_{i=1}^{\infty} \frac{\partial b}{\partial x_i}(x) z_i$ *as functions of* x *as* $\varepsilon \to 0$
- *For any compact* $K \subset \mathbb{R}^n$ *we have*

$$\limsup_{\varepsilon \to 0} \int_K |b^2_\varepsilon(z)(x)| \, \mathrm{d}x \le |z| |D^s b|(K)$$

- *For compact* $K, K' \subset \mathbb{R}^n$ *and* $\delta > 0$ *we have the uniform bound*

$$\sup_{z \in K'} \sup_{\varepsilon \in (0,\delta)} \int_K |b^1_\varepsilon(z)(x)| + |b^2_\varepsilon(z)(x)| \, \mathrm{d}x \le \sup_{z \in K'} |z| |Db| \big(\{x : \mathrm{dist}(x, K) \le \delta\} \big)$$

Proof. ([3], Theorem 2.4 and [6], Proposition 3.2) Lets assume $z = e_n$ first, we discuss scaling and rotation-invariance of the theorem later. Additionally let there be $x = (x', x_n)$ with $x' \in \mathbb{R}^{n-1}$ and $x_n \in \mathbb{R}$.
Using the definition of $\hat{\mu}_\varepsilon$ and the statement of Lemma 2.3 first we define

$$b^1_\varepsilon(x', x_n) := \widehat{\frac{\partial b}{\partial x_n}(x', \cdot) \mathscr{L}^1}_\varepsilon (x_n) = \frac{1}{\varepsilon} \int_{x_n}^{x_n+\varepsilon} \frac{\partial b}{\partial x_n}(x', s) \, \mathrm{d}s$$

Then, by Lemma 2.3 we have the convergence of b_ε^1 to $\frac{\partial b}{\partial x_n}$, what was the first thing to show.
Next we define $b_{x'}(s) = b(x', s)$ and use Lemma 2.4 to calculate

$$\begin{aligned}\frac{b(x', x_n+\varepsilon) - b(x', x_n)}{\varepsilon} = \frac{b_{x'}(x_n+\varepsilon) - b_{x'}(x_n)}{\varepsilon} &= \frac{Db_{x'}([x_n, x_n+\varepsilon])}{\varepsilon} \\ &= \frac{Db^a_{x'}([x_n, x_n+\varepsilon])}{\varepsilon} + \frac{Db^s_{x'}([x_n, x_n+\varepsilon])}{\varepsilon} \\ &= \widehat{Db^a_{x'}}_\varepsilon(x_n) + \widehat{Db^s_{x'}}_\varepsilon(x_n) \\ &= b^1_\varepsilon(x', x_n) + \widehat{Db^s_{x'}}_\varepsilon(x_n)\end{aligned}$$

for almost every x_n. So we have $b^2_\varepsilon(x', x_n) = \widehat{Db^s_{x'}}_\varepsilon(x_n)$. Thus we have using (1) and Lemma 2.4:

$$\begin{aligned}\int_K |b^2_\varepsilon|(x', x_n)\,\mathrm{d}x_n\,\mathrm{d}x' &\le \int_{\mathbb{R}^{n-1}} \int_{\{x_n : (x', x_n) \in K\}} \left|\widehat{D^s b_{x'}}_\varepsilon(x_n)\right| \mathrm{d}x_n\,\mathrm{d}x' \\ &\le \int_{\mathbb{R}^{n-1}} |D^s b_{x'}|(\{x_n : (x', x_n) \in K_\varepsilon\})\,\mathrm{d}x' \le |D^s b|(K_\varepsilon)\end{aligned}$$

This was the second thing to show. With the exactly same argument we get $\int_K |b^1_\varepsilon|\,\mathrm{d}x \le |D^a b|(K_\varepsilon)$. Thus we have

$$\int_K |b^1_\varepsilon(e_n)(x)| + |b^2_\varepsilon(e_n)(x)|\,\mathrm{d}x \le |D^s b|(K_\varepsilon) + |D^a b|(K_\varepsilon) = |Db|(K_\varepsilon)$$

The last equality is from the fact, that for singular measures there holds the triangle inequality in the variation norm also reverse and is hence an equality. This shows the last property.
The case for general z is just carefully reproducing the proof by setting $b^i_\varepsilon(z) = b^i_{\varepsilon|z|}\left(\frac{z}{|z|}\right)$ for the scaling invariance. Then we often relabel $\varepsilon|z| \to \varepsilon$. The rotation invariance is obvious as the integrals do not change under rotation. □

Next we will need Albertis rank-one-theorem:

Theorem 2.6. *Let there be $b \in BV(\Omega, \mathbb{R}^m)$ for $\Omega \subset \mathbb{R}^n$ open. Let $D^s b = M|D^s b|$ be the singular part of the distributional derivative. Then for $M(x)$ has rank one $D^s b$-almost everywhere, i. e. $M(x) = \eta(x) \otimes \xi(x)$ with $|\xi(x)| = |\eta(x)| = 1$ for $D^s b$ a. e. $x \in \Omega$.*

Proof. [1] □

2.2 Convergence Lemmata

We will need the following technical lemma in the existence proof.

Lemma 2.7. *Let f_k, g_k be sequences in $L^2(\Omega)$ (with Ω an open subset of $\mathbb{R}^n$). Let $f_k \to f$ and $g_k \rightharpoonup g$ with f, g in L^2 and the weak convergence of g_k in L^2. Then $\int_\Omega f_k g_k \to \int_\Omega f g$.*

Proof. We need to show $\int_\Omega f_k g_k - f g \to 0$:

$$\int_\Omega f_k g_k - f g = \int_\Omega (f_k - f)(g_k + g) + f g_k - g f_k = \int_\Omega (f_k - f)(g_k + g) + \int_\Omega f g_k - \int_\Omega g f_k$$

as all integrals exist (this will be seen in the proof). So we need to check the convergence of these three integrals:

- for the first one, we have by the Hölder-inequality

  $$\int_\Omega (f_k - f)(g_k + g) \leq \|f_k - f\|_2 \|g_k + g\|_2 \to 0$$

  as $\|f_k - f\|_2 \to 0$ by definition and $\|g_k + g\|$ is bounded, because weak convergent sequences are bounded.
- The second integral $\int_\Omega f g_k$ converges to $\int_\Omega f g$ by the the definition of weak convergence of g_k
- The third integral also converges to $\int_\Omega f g$, because strong convergence implies weak convergence

So, summed up we get $\lim_{k\to\infty} \int_\Omega f_k g_k - f g = 0$. This was to show. □

Next we define convergence in measure and prove two useful lemmata:

Definition 2.8. Let there be $f_n, f : \Omega \to \mathbb{R}$ measurable with Ω a measurable subset of $\mathbb{R}^N$. We say f_n converges in measure to f, if

$$\lim_{n\to\infty} \mathscr{L}^N\Big(\big\{|f_n - f| \geq \varepsilon\big\}\Big) = 0$$

for any $\varepsilon > 0$.

According to [8][p. 257], convergence in L^p as well as convergence almost everywhere implies convergence in measure locally. In this sense, Pratts theorem[8][p. 260] is a generalization of the dominated convergence theorem.
The first lemma states, that convergence in measure is stable under (uniformly) continuous functions:

Lemma 2.9. *Let there be $f_n \to f$ in measure and $\alpha : \mathbb{R} \to \mathbb{R}$ uniformly continuous (or let the f_n be uniformly essentially bounded and α only continuous, so α uniformly continuous on the image of the f_n). Then $\alpha \circ f_n \to \alpha \circ f$ in measure.*

Proof. For any $\varepsilon > 0$ we have a $\delta > 0$, such that

$$|f_n(x) - f(x)| < \delta \Rightarrow |\alpha(f_n(x)) - \alpha(f(x))| < \varepsilon$$

So the contraposition is

$$\Big|\alpha(f_n(x)) - \alpha(f(x))\Big| \geq \varepsilon \Rightarrow |f_n(x) - f(x)| \geq \delta$$

So $\big\{|\alpha \circ f_n - \alpha \circ f| \geq \varepsilon\big\} \subset \big\{|f_n - f| \geq \delta\big\}$, so

$$\lim_{n\to\infty} \mathscr{L}^N\Big(\big\{|\alpha \circ f_n - \alpha \circ f| \geq \varepsilon\big\}\Big) \leq \lim_{n\to\infty} \mathscr{L}^N\Big(\big\{|f_n - f| \geq \delta\big\}\Big) = 0$$

□

Next we have two variants of a convergence theorem:

Lemma 2.10. *Let there be $f_n \in L^\infty(\Omega)$ with a uniform bound, so $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ and $g_n \in L^1(\Omega)$ again with $\sup_{n\in\mathbb{N}} \|g_n\|_1 < \infty$ and a dominating function $g \in L^1(\Omega)$. Additionally let $f_n \to 0$ in measure. Then*

$$\lim_{n\to\infty} \int_\Omega |f_n g_n| = 0$$

Proof. For $\varepsilon > 0$ we have

$$\begin{aligned}\int_\Omega |f_n g_n| &= \int_{\{|f_n|>\varepsilon\}} |f_n g_n| + \int_{\{|f_n|\leq\varepsilon\}} |f_n g_n| \\ &\leq \int_\Omega |f_n||g_n| \mathbb{1}_{\{|f_n|>\varepsilon\}} + \int_{\{|f_n|\leq\varepsilon\}} \varepsilon |g_n| \\ &\leq \sup_{n\in\mathbb{N}} \|f_n\|_\infty \int_\Omega |g_n| \mathbb{1}_{\{|f_n|>\varepsilon\}} + \varepsilon \sup_{n\in\mathbb{N}} \|g_n\|_1\end{aligned}$$

Now we take the limit $n \to \infty$. According to the dominated convergence theorem, the first integral converges to 0, because $\mathbb{1}_{\{|f_n|>\varepsilon\}}$ converges pointwise almost everywhere to zero (because $f_n \to 0$ in measure) and g_n is a dominating function. So we have:

$$\lim_{n\to\mathbb{N}} \int_\Omega |f_n g_n| \leq \varepsilon \sup_{n\in\mathbb{N}} \|g_n\|_1$$

Now $\varepsilon \to 0$ proves the lemma. □

Lemma 2.11. *Let there be $f_n \in L^\infty(\Omega)$ with a uniform bound and Ω bounded, so $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$ and $g_n \in L^p(\Omega)$ again with $\sup_{n\in\mathbb{N}} \|g_n\|_p < \infty$ with $p > 1$. Additionally let $f_n \to 0$ in measure. Then*

$$\lim_{n\to\mathbb{N}} \int_\Omega |f_n g_n| = 0$$

Proof. For $\varepsilon > 0$ we have (with $q = \frac{p}{p-1}$ Hölder-conjugate to p)

$$\begin{aligned}\int_\Omega |f_n g_n| &= \int_{\{|f_n|>\varepsilon\}} |f_n g_n| + \int_{\{|f_n|\le\varepsilon\}} |f_n g_n| \\ &\le \int_\Omega |f_n||g_n|\mathbb{1}_{\{f_n>\varepsilon\}} + \int_{\{|f_n|\le\varepsilon\}} \varepsilon |g_n| \\ &\le \sup_{n\in\mathbb{N}} \|f_n\|_\infty \int_\Omega |g_n|\mathbb{1}_{\{|f_n|>\varepsilon\}} + \varepsilon \sup_{n\in\mathbb{N}} \|g_n\|_p \left|\{|f_n| \le \varepsilon\}\right|^{1/q} \\ &\le \sup_{n\in\mathbb{N}} \|f_n\|_\infty \sup_{n\in\mathbb{N}} \|g_n\|_p |\{|f_n| > \varepsilon\}|^{1/q} + \varepsilon \sup_{n\in\mathbb{N}} \|g_n\|_p \left|\{|f_n| \le \varepsilon\}\right|^{1/q}\end{aligned}$$

Now we take the limit $n \to \infty$. The first term goes to 0 by the convergence in measure of the f_n, in the second one $|\{|f_n| \le \varepsilon\}|^{1/q}$ can be estimated by the measure of Ω:

$$\lim_{n\to\mathbb{N}} \int_\Omega |f_n g_n| \le \varepsilon \sup_{n\in\mathbb{N}} \|g_n\|_p |\Omega|^{1/q}$$

Now $\varepsilon \to 0$ proves the lemma. □

2.3 Mollification of distributions

We also will need the mollification of distributions on $\mathbb{R}^n$. For more details see [12], Chapter 11. In the whole chapter we only use even convolution kernels, so our definition does not need the reflection used in the definition in [12]

Definition 2.12. Let there be an even convolution kernel ρ_ε and a distribution u, both on $\mathbb{R}^n$. Then there is also a distribution $\rho_\varepsilon * u$ on $\mathbb{R}^n$. We define it for $\varphi \in C_c^\infty(\mathbb{R}^n)$ by

$$\rho_\varepsilon * u(\varphi) := u(\rho_\varepsilon * \varphi)$$

Remark 2.13. This definition generalizes the convolution of a function with an even convolution kernel in the following sense: For a L^1—function f and the associated distribution $\operatorname{test} f$ (defined by $\operatorname{test} f(\varphi) = \int_{\mathbb{R}^n} \varphi(x) f(x)\,\mathrm{d}x$ for a test function φ) there holds $\operatorname{test}(\rho_\varepsilon * f) = \rho_\varepsilon * \operatorname{test} f$ for an even convolution kernel. We insert a test function φ:

$$\begin{aligned}\rho_\varepsilon * \operatorname{test} f(\varphi) = \operatorname{test} f(\rho_\varepsilon * \varphi) &= \int_{\mathbb{R}^n} \rho_\varepsilon * \varphi(x) f(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\varphi(y)\,\mathrm{d}y f(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \rho_\varepsilon(y-x)\varphi(y) f(x)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \rho_\varepsilon * f(y)\varphi(y)\,\mathrm{d}y \\ &= \operatorname{test}(\rho_\varepsilon * f)(\varphi)\end{aligned}$$

We used Fubini, this is justified as the integrals are all finite.

As we consider a parabolic equation, we have not distributions on $\mathbb{R}^n$ but on $[0,T)\times\mathbb{R}^n$. We also want to mollify them, but only in space:

Definition 2.14. Let there be an even convolution kernel ρ_ε on R^n and a distribution u on $[0,T)\times\mathbb{R}^n$. Then there is also a distribution $\rho_\varepsilon * u$ on $[0,T)\times\mathbb{R}^n$. We define it for $\varphi\in C_c^\infty([0,T)\times\mathbb{R}^n)$ by

$$\rho_\varepsilon * u(\varphi) := u(\rho_\varepsilon * \varphi)$$

Here $\rho_\varepsilon * \varphi(x,t) := \rho_\varepsilon * \varphi(\cdot,t)(x)$

We will have only distributions of order zero and one in t, so we have the following two lemmata:

Lemma 2.15. *Let u be a distribution on $[0,T)\times\mathbb{R}^n$ of order zero in t, so there are distributions u_t on $\mathbb{R}^n$ such that*

$$u(\varphi) = \int_0^T u_t(\varphi(\cdot,t))\,\mathrm{d}t$$

Then $\rho_\varepsilon * u$ is also of order zero and given by $\int_0^T \rho_\varepsilon * u_t\,\mathrm{d}t$:

Proof. This is easily proven by inserting $\varphi\in C_c^\infty([0,T)\times\mathbb{R}^n)$:

$$\rho_\varepsilon * u(\varphi) = u(\rho_\varepsilon * \varphi) = \int_0^T u_t(\rho_\varepsilon * \varphi(\cdot,t))\,\mathrm{d}t = \int_0^T \rho_\varepsilon * u_t(\varphi(\cdot,t))\,\mathrm{d}t$$

□

The other situation is a distribution of order one in t:

Lemma 2.16. *Let there be $u\in L^1_{\mathrm{loc}}([0,T)\times\mathbb{R}^n)$ with boundary data u_0 at 0. We consider the distribution $\partial_t u$(given by $\partial_t u(\varphi) = -\int_0^T\int_{\mathbb{R}^n} u\partial_t\varphi + \int_{\mathbb{R}^n} u_0\varphi(0,\cdot)$).*
Then we have $\rho_\varepsilon * \partial_t u = \partial_t(\rho_\varepsilon * u)$, with the second distribution seen with boundary data $\rho_\varepsilon * u_0$ for an even convolution kernel ρ_ε.

Proof.

$$\begin{aligned}
\rho_\varepsilon * \partial_t u(\varphi) &= \partial_t u(\rho_\varepsilon * \varphi) \\
&= -\int_0^T \int_{\mathbb{R}^n} u(x,t)\partial_t(\rho_\varepsilon * \varphi)(x,t) + \int_{\mathbb{R}^n} u_0(x)\rho_\varepsilon * \varphi(0,x)\,\mathrm{d}x \\
&= -\int_0^T \int_{\mathbb{R}^n} u(x,t)\rho_\varepsilon * \partial_t\varphi(x,t) + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u_0(x)\rho_\varepsilon(x-y)\varphi(0,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= -\int_0^T \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x,t)\rho_\varepsilon(x-y)\partial_t\varphi(y,t)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t \\
&\quad + \int_{\mathbb{R}^n} \varphi(0,y)\int_{\mathbb{R}^n} u_0(x)\rho_\varepsilon(y-x)\,\mathrm{d}x\,\mathrm{d}y \\
&= -\int_0^T \int_{\mathbb{R}^n} \partial_t\varphi(y,t)\int_{\mathbb{R}^n} u(x,t)\rho_\varepsilon(y-x)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_{\mathbb{R}^n} \varphi(0,y)\rho_\varepsilon * u_0(y)\,\mathrm{d}y \\
&= -\int_0^T \int_{\mathbb{R}^n} \partial_t\varphi(y,t)\rho_\varepsilon * u(y,t) + \int_{\mathbb{R}^n} \varphi(0,y)\rho_\varepsilon * u_0(y)\,\mathrm{d}y \\
&= \partial_t(\rho_\varepsilon * u)(\varphi)
\end{aligned}$$

□

2.4 A distributional Gronwall inequality

Next we have the following distributional version of Gronwalls inequality:

Lemma 2.17. *Let there be a function $f \in C([0,T])$ with $f(0)=0$ and $g,h \in L^1([0,T])$. f, g and h are assumed to be nonnegative. Additionally $f' \leq fg + h$ distributionally, so for every nonnegative test function $\varphi \in C_c^\infty((0,T))$ we have*

$$-\int_0^T \varphi'(t)f(t)\,\mathrm{d}t \leq \int_0^T f(t)g(t)\varphi(t)\,\mathrm{d}t + \int_0^T h(t)\varphi(t)\,\mathrm{d}t$$

Then

$$f(t) \leq e^{\int_0^t g(r)\,\mathrm{d}r}\int_0^t h(s)\,\mathrm{d}s$$

almost everywhere in $[0,T]$.

Proof. In the distributional formulation we test with $\widetilde{\varphi}(s) = e^{-\int_0^s g(r)\,\mathrm{d}r}\varphi$ for a nonnegative test function φ. This is not an element of $C_c^\infty((0,T))$, but monotone, bounded and weakly differentiable, as $s \mapsto \int_0^s g(r)\,\mathrm{d}r$ is weakly differentiable with derivative g, so we can approximate it by test functions and the equation holds.

Testing with $\widetilde{\varphi}$ leads to

$$
\begin{aligned}
-\int_0^T \varphi'(s)e^{-\int_0^s g(r)\,\mathrm{d}r} f(s)\,\mathrm{d}s + \int_0^T \varphi(s)e^{-\int_0^s g(r)\,\mathrm{d}r} g(s)f(s)\,\mathrm{d}s \\
\leq \int_0^T \varphi(s)g(s)f(s)e^{-\int_0^s g(r)\,\mathrm{d}r}\,\mathrm{d}s + \int_0^T \varphi(s)h(s)e^{-\int_0^s g(r)\,\mathrm{d}r}\,\mathrm{d}s
\end{aligned}
$$

$$
\Longrightarrow -\int_0^T \varphi'(s)e^{-\int_0^s g(r)\,\mathrm{d}r} f(s)\,\mathrm{d}s \leq \int_0^T \varphi(s)h(s)e^{-\int_0^s g(r)\,\mathrm{d}r}\,\mathrm{d}s
$$

Defining $R(s) = e^{-\int_0^s g(r)\,\mathrm{d}r} f(s)$ and estimating $e^{-\int_0^s g(r)\,\mathrm{d}r} \leq 1$ we have $R \in C([0,T])$, $R(0) = 0$ and

$$
-\int_0^T R(s)\varphi'(s)\,\mathrm{d}s \leq \int_0^T \varphi(s)h(s)\,\mathrm{d}s
$$

so $R' \leq h$ distributionally. Lets define $\overline{R}(t) = \int_0^t h(s)\,\mathrm{d}s$, so $\overline{R}'(t) = h(t)$ almost everywhere. So we have

$$
\begin{aligned}
-\int_0^T (R(s) - \overline{R}(s))\varphi'(s)\,\mathrm{d}s &= -\int_0^T R(s)\varphi'(s) + \int_0^T \overline{R}(s)\varphi'(s) \\
&\leq \int_0^T \varphi(s)h(s)\,\mathrm{d}s - \int_0^T \overline{R}'(s)\varphi(s)\,\mathrm{d}s \\
&= \int_0^T \varphi(s)h(s)\,\mathrm{d}s - \int_0^T \varphi(s)h(s)\,\mathrm{d}s \\
&= 0
\end{aligned}
$$

According to the following lemma, this leads to $R - \overline{R} \leq 0$, so $R(t) \leq \int_0^t h(s)\,\mathrm{d}s$ on $[0,T]$, so

$$
f(t) \leq e^{\int_0^t g(r)\,\mathrm{d}r} \int_0^t h(s)\,\mathrm{d}s
$$

almost everywhere. This was to show. □

Lemma 2.18. *Let there be $R \in C([0,T]), R(0) = 0$ and $R' \leq 0$ distributionally:*

$$
\int_0^T R(s)\varphi'(s)\,\mathrm{d}s \geq 0
$$

for all nonnegative test functions φ. Then $R \leq 0$ in $[0,T]$.

Proof. We argue by contradiction, so lets assume the existence of a $t \in [0,T]$ with $R(t) > 0$. We take mollified functions R_ε of R (to achieve this, extend R on some interval $[-\tau, T+\tau]$ constant outside of $[0,T]$ and continuous in 0 and T), so R_ε is well defined. As R is continuous, we have $R_\varepsilon \to R$ uniformly and for a $\delta > 0$, we have an $\varepsilon > 0$ such that $\|R_\varepsilon - R\|_\infty \leq \delta$.
So, for any positive test function φ, we have

$$\begin{aligned} 0 \leq \int_0^T R(s)\varphi'(s)\,\mathrm{d}s &= \int_0^T R_\varepsilon(s)\varphi'(s)\,\mathrm{d}s + \int_0^T (R(s)-R_\varepsilon(s))\varphi'(s)\,\mathrm{d}s \\ &\leq \int_0^T R_\varepsilon(s)\varphi'(s)\,\mathrm{d}s + \|\varphi'\|_1\|R-R_\varepsilon\|_\infty \\ &\leq \int_0^T R_\varepsilon(s)\varphi'(s)\,\mathrm{d}s + \|\varphi'\|_1\delta \end{aligned}$$

So, by a partial integration, we get, that for every $\delta > 0$ there exists a $\varepsilon > 0$ such that

$$\int_0^T R'_\varepsilon(s)\varphi(s) \leq \|\varphi'\|_1\delta \tag{2}$$

Now we take nonnegative test functions φ_k for $k \in \mathbb{N}$ with the following properties, let there be

- $\varphi_k(0) = 0$
- $\varphi_k = 1$ on $[\frac{1}{k}, t-\frac{1}{k}]$
- $\varphi_k = 0$ on $[t, T]$
- $|\varphi'_k| \leq 2K$, so especially $\|\varphi'_k\|_1 \leq 4$ independent of k as $\varphi'_k \neq 0$ only in $[0, \frac{1}{k}]$ and $[t-\frac{1}{k}, t]$

So we have $\varphi_k \to \mathbb{1}_{[0,t]}$ pointwise almost everywhere.
Inserting φ_k in (2) leads to

$$\int_0^T R'_\varepsilon(s)\varphi_k(s)\,\mathrm{d}s \leq \|\varphi'_k\|_1\delta \leq 4\delta$$

Now lets take $k \to \infty$. As $|R_\varepsilon|$ is a continuous function on $[0,T]$, it is integrable and hence a suitable dominating function for the left-hand side (because $|\varphi_k| \leq 1$). So we can apply the dominated convergence theorem and get for any $\delta > 0$ an $\varepsilon > 0$ such that

$$\int_0^t R'_\varepsilon(s)\,\mathrm{d}s = R_\varepsilon(t) - R_\varepsilon(0) \leq 4\delta$$

Now let $\delta \to 0$, so also $\varepsilon \to 0$. Then $R_\varepsilon(0) \to 0$ and $R_\varepsilon(t) \to R(t) > 0$ by assumption. This is a contradiction as the right hand side goes to 0. □

2.5 Youngs inequality for integral operators

From harmonic analysis we use the following result (also known as *Schur's test*), which is also valid in more general versions, see [15], Theorem 0.3.1:

Theorem 2.19. *Let* $K : \mathbb{R}^n \times \mathbb{R}^n \to [0, \infty)$ *and* $f \in L^p(\mathbb{R}^n)$ *for* $p \geq 1$. *Let there be*

$$\int_{\mathbb{R}^n} K(x, y)\,\mathrm{d}y \leq 1 \textit{ for almost every } x \in \mathbb{R}^n$$

and

$$\int_{\mathbb{R}^n} K(x, y)\,\mathrm{d}x \leq 1 \textit{ for almost every } y \in \mathbb{R}^n$$

Then we have

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x, y)|f(y)|^p\,\mathrm{d}y\,\mathrm{d}x \leq \int_{\mathbb{R}^n} |f(y)|^p\,\mathrm{d}y$$

Proof. [15], Theorem 0.3.1 □

This leads to the following lemma

Lemma 2.20. *Let there be* $f \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ *for* $p \geq 1$. *Then we have for* $R > 1$ *and* $0 < \varepsilon < 1$

$$\int_{B_R} \int_{B_1(0)} |f(x + \varepsilon y)|^p\,\mathrm{d}y\,\mathrm{d}x \leq \|f\|^p_{L^p(B_{R+1})}$$

Proof. Use theorem 2.19 (after setting f to 0 outside of B_{R+1}) and

$$K(x, y) = \begin{cases} \frac{1}{|B_\varepsilon|} & |x - y| \leq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

□

2.6 Bouchuts Lemma

In the proof of the renormalization theorem we will define $\Lambda(M, \rho) = \int_{\mathbb{R}^n} \left|\langle Mz, \nabla\rho(z)\rangle\right| \mathrm{d}z$ for a $n \times n$-Matrix and $\rho \in C_c^\infty(\mathbb{R}^n)$. We will try to get $\Lambda(M, \rho)$ as small as possible by choosing the convolution kernel ρ. The following lemma of Bouchut gives an answer to this question if M has rank one (this will be satisfied by Albertis rank one theorem):

Lemma 2.21. *Let there be* $\xi, \eta \in \mathbb{R}^n$ *with* $\xi \perp \eta$ *and with* $\xi = \eta = 1$. *Then, for any given* ε *we find a even* $\rho \in C_c^\infty(\mathbb{R}^n)$*, such that* $\Lambda(\eta \otimes \xi, \rho) < \varepsilon$*, this means*

$$\int_{\mathbb{R}^n} \left|\langle z, \xi\rangle\right| \left|\langle \nabla\rho(z), \eta\rangle\right| \mathrm{d}z < \varepsilon$$

Proof. ([2], Lemma 3.3) We first prove the Lemma for $n = 2$. Without loss of generality we can assume $\xi = e_1$ and $\eta = e_2$. Lets define the rectangle $R_\varepsilon = [\frac{\varepsilon}{2}, \frac{\varepsilon}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$. Then we take

$$\rho = \frac{\mathbb{1}_{R_\varepsilon}}{\varepsilon} * h_\delta$$

for a convolution kernel h_δ.
So we have $\int_{\mathbb{R}^n} \rho = 1$ and

$$\left| \frac{\partial(\rho * h_\delta)}{\partial z_2} \right| \to \frac{|\nu_2|}{\varepsilon} \mathscr{H}^1 \llcorner \partial R_\varepsilon$$

as $\delta \to 0$ in the sense of measure with $\nu = (\nu_1, \nu_2)$ the inner unit normal to R_ε.
Then we have

$$\lim_{\delta \to 0} \Lambda(\eta \otimes \xi, \rho) = \lim_{\delta \to 0} \int_{\mathbb{R}^2} |z_1| \left| \frac{\partial(\rho * h_\delta)}{\partial z_2} \right| \mathrm{d}z = \frac{2}{\varepsilon} \int_{-\frac{\varepsilon}{2}}^{\frac{\varepsilon}{2}} |z_1| \, \mathrm{d}z_1 = \frac{\varepsilon}{2}$$

So we can choose δ small enough to get a suitable ρ. If $n > 2$ we just multiply this 2-dimensional kernel with a fixed kernel in the other dimensions (orthogonal to ξ and η). □

3 The Fokker-Planck equation

We are going to consider a Fokker-Planck-equation of the following form:

$$\partial_t p + \partial_i(p b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) = 0 \tag{3}$$

This is a time-dependent equation, so we consider it on a time interval $[0, T]$. We have problem data

- A drift field $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$
- A diffusion term $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$

and a solution

- $p : [0, T] \times \mathbb{R}^n \to \mathbb{R}$

For $\sigma = 0$, (3) becomes the standard transport equation. The main difficulty is that we consider a drift-field b which has only BV-regularity in the spatial variables.
We will consider weak solutions in the following sense:

Definition 3.1. Let there be an initial condition $p_0 \in L^2 \cap L^\infty$ (in the formal sense that $p|_{t=0} = p_0$). Then a function $p \in L^\infty([0, T], L^2 \cap L^\infty)$ satisfying $\sigma^*\nabla p \in L^2([0, T], L^2)$ (where σ^* is the transpose of σ) is called a *weak solution* to (3) if

$$\int_0^T \int_{\mathbb{R}^n} p\partial_t\varphi \,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^n} p_0\varphi(0,\cdot)\,\mathrm{d}x = -\int_0^T \int_{\mathbb{R}^n} p\langle b, \nabla\varphi\rangle \,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_0^T \int_{\mathbb{R}^n} \langle\sigma^*\nabla p, \sigma^*\nabla\varphi\rangle \,\mathrm{d}x\,\mathrm{d}p$$

for all test functions $\varphi \in C_c^\infty([0, T) \times \mathbb{R}^n)$

The main theorem we prove is the following (Theorem 1.1 in [14]):

Theorem 3.2. *Let b and σ be as above, satisfying the following regularity assumptions:*

- $b \in \Big(L^1\big([0, T], BV_{\mathrm{loc}}(\mathbb{R}^n)\big)\Big)^n$
- $\frac{b}{1+|x|} \in \Big(L^1\big([0, T], L^1 + L^\infty(\mathbb{R}^n)\big)\Big)^n$
- $\mathrm{div}(b) \in L^1\big([0, T], L^1_{\mathrm{loc}}(\mathbb{R}^n)\big)$
- $[\mathrm{div}(b)]^- \in L^1\big([0, T], L^\infty(\mathbb{R}^n)\big)$

and

- $\sigma \in \left(L^2\left([0,T], W^{1,2}_{\text{loc}}(\mathbb{R}^n)\right)\right)^{n\times m}$
- $\frac{\sigma}{1+|x|} \in \left(L^2\left([0,T], L^2 + L^\infty(\mathbb{R}^n)\right)\right)^{n\times m}$

Then, for any initial condition $p|_{t=0} = p_0$ with $p_0 \in L^2 \cap L^\infty$, (3) has a unique weak solution in the space

$$X_2 = \left\{p \in L^\infty\left([0,T], L^2 \cap L^\infty\right), \sigma^* \nabla p \in L^2\left([0,T], L^2\right)\right\}$$

3.1 Definition and properties of $\sigma^*\nabla p$

In the definition of the Fokker-Planck equation (3) (respectively in the weak formalization of Definition 3.1) we need to define the meaning of $\sigma^*\nabla p$, as p is only assumed to be in $L^2 \cap L^\infty$:

Remark 3.3. For the existence of a weak solution of (3) we state the existence of $\sigma^*\nabla p(t)$ for almost every t in the following distributional sense: There exists an $u \in L^2(\mathbb{R}^n)^n$, such that, for all test functions $\varphi \in H^1_0(\mathbb{R}^n)^n$:

$$\int_{\mathbb{R}^n} u(t,x)\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t = -\int_{\mathbb{R}^n} p \cdot \operatorname{div}(\sigma\varphi)\,\mathrm{d}x\,\mathrm{d}t$$

In this sense we will also write $u = \sigma^*\nabla p$

Next, let there be p_ε mollified versions (only in space) of p, then we show the following convergence:

Lemma 3.4. *We have $\sigma^*\nabla p_\varepsilon \to \sigma^*\nabla p$ in $L^2\left([0,T], L^2_{\text{loc}}\right)$.*

Proof. For a compact $K \subset \mathbb{R}^n$ we estimate by adding the zero $(\sigma^*\nabla p)_\varepsilon - (\sigma^*\nabla p)_\varepsilon$ with $(\sigma^*\nabla p)_\varepsilon$ being the mollified version of $\sigma^*\nabla p$

$$\begin{aligned}
&\int_0^T \|\sigma^*\nabla p_\varepsilon(t) - \sigma^*\nabla p(t)\|_{L^2(K)}\,\mathrm{d}t \\
\leq &\int_0^T \|\sigma^*\nabla p_\varepsilon(t) - (\sigma^*\nabla p)_\varepsilon(t)\|_{L^2(K)}\,\mathrm{d}t + \int_0^T \|(\sigma^*\nabla p)_\varepsilon(t) - \sigma^*\nabla p(t)\|_{L^2(K)}\,\mathrm{d}t
\end{aligned}$$

The first integral goes to zero by Lemma 3.16 as it is exactly R_ε. (Note that this is not a circular reasoning argument, as Lemma 3.16 is proven without using this lemma or anything else depending on this lemma).
The second integral also converges to zero, as $\|(\sigma^*\nabla p)_\varepsilon(t) - \sigma^*\nabla p(t)\|_{L^2(K)} \to 0$ for every t pointwise by the properties of the convolution and as we have the dominating function $t \mapsto 2\|\sigma^*\nabla p(t)\|_{L^2(K)}$. □

For the definition of *renormalized solution* we will have to give also a meaning to the expression $\sigma^*\nabla\beta(p)$ for a $\beta \in C^2(\mathbb{R})$:

Remark 3.5. Let there be a function $u \in L^2(\mathbb{R}^n)^n$, such that, for all test functions $\varphi \in H^1_0(\mathbb{R}^n)^n$:

$$\int_{\mathbb{R}^n} u(t,x)\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t = -\int_{\mathbb{R}^n} \beta(p)\cdot \operatorname{div}(\sigma\varphi)\,\mathrm{d}x\,\mathrm{d}t$$

In this sense we will also write $u = \sigma^*\nabla\beta(p)$.

Also in this distributional sense we have the following chain-rule:

Lemma 3.6. *In the sense of the last remark we have* $\sigma^*\nabla\beta(p) = \beta'(p)\cdot\sigma^*\nabla p$

Proof. We approximate $\sigma \in W^{1,2}_{\text{loc}}(\mathbb{R}^n)$ and $p \in L^2 \cap L^\infty$ by smooth functions $\sigma_n \in C^\infty \cap W^{1,2}_{\text{loc}}$ and $p_n \in C^\infty_c \cap L^2 \cap L^\infty$ (note that the p_n should have compact support), such that $\sigma_n \to \sigma$ in $W^{1,2}_{\text{loc}}(\mathbb{R}^n)$ and $p_n \to p$ only in L^2 (not in L^∞, because this is not possible in general). Additionally, let $\|\sigma_n\|_2 \leq \|\sigma\|_2$, $\|\nabla\sigma_n\|_2 \leq \|\nabla\sigma\|_2$ and $\|p_n\|_2 \leq \|p\|_2$, $\|p_n\|_\infty \leq \|p\|_\infty$. This can be achieved for example by mollifying σ and p.
At first we show that $\sigma_n^*\nabla p_n \rightharpoonup \sigma^*\nabla p$ weak in L^2_{loc}, possibly after picking a subsequence (especially $\sigma_n^*\nabla p_n \in L^2_{\text{loc}}$ as ∇p_n is bounded). We have for a test function φ

$$\int_{\mathbb{R}^n} \sigma_n^*\nabla p_n\varphi = -\int_{\mathbb{R}^n} p_n\nabla(\sigma_n\varphi) \to -\int_{\mathbb{R}^n} p\nabla(\sigma\varphi) = \int_{\mathbb{R}^n} \sigma^*\nabla p\varphi$$

The convergence is because $\sigma_n \to \sigma$ in $W^{1,2}_{\text{loc}}(\mathbb{R}^n)$ and $p_n \to p$ in L^2 both strongly.
We want to show $\beta'(p)\cdot\sigma^*\nabla p = \sigma^*\nabla\beta(p)$ in the following big frame, for any test function φ:

$$\begin{aligned}\int_{\mathbb{R}^n} \beta'(p)\cdot\sigma^*\nabla p\cdot\varphi &\overset{A}{=} \lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p_n)\cdot\sigma_n^*\nabla p_n\cdot\varphi = \lim_{n\to\infty}\int_{\mathbb{R}^n} \nabla(\beta(p_n))\cdot(\sigma_n\varphi)\\ &= \lim_{n\to\infty} -\int_{\mathbb{R}^n} \beta(p_n)\operatorname{div}(\sigma_n\varphi) \overset{B}{=} -\int_{\mathbb{R}^n} \beta(p)\cdot\nabla(\sigma\varphi)\end{aligned}$$

By the last remark 3.6 this would show the Lemma, but of course we need to justify A and B. For both we will need, that $\beta'\circ p_n \to \beta'\circ p$ in measure. This holds because $p_n \to p$ in measure (L^p-convergence implies convergence in measure) and Lemma 2.9, as p_n are uniformly bounded and $\beta \in C^2(\mathbb{R})$.

Step A: We have

$$\lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p)\cdot\sigma^*\nabla p\cdot\varphi - \beta'(p_n)\cdot\sigma_n^*\nabla p_n\cdot\varphi$$
$$= \lim_{n\to\infty}\int_{\mathbb{R}^n} \big(\beta'(p)-\beta'(p_n)\big)\cdot\big(\sigma^*\nabla p+\sigma_n^*\nabla p_n\big)\cdot\varphi$$
$$+ \lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p_n)\cdot\sigma^*\nabla p\cdot\varphi$$
$$- \lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p)\cdot\sigma_n^*\nabla p_n\cdot\varphi$$

So we need to check the behavior of these three limits:

- The first integral $\int_{\mathbb{R}^n}\big(\beta'(p)-\beta'(p_n)\big)\cdot\big(\sigma^*\nabla p+\sigma_n^*\nabla p_n\big)\cdot\varphi$ converges to zero by Lemma 2.11 as $\beta'(p)-\beta'(p_n)$ goes to zero in measure and as $\sigma_n^*\nabla p_n$ is bounded in L^2 (because its weakly convergent in L^2_{loc} and we integrate on the compact support of φ
- The second integral $\int_{\mathbb{R}^n}\beta'(p_n)\cdot\sigma^*\nabla p\cdot\varphi$ goes to $\int_{\mathbb{R}^n}\beta'(p)\cdot\sigma^*\nabla p\cdot\varphi$ by Lemma 2.10
- The third integral $\int_{\mathbb{R}^n}\beta'(p)\cdot\sigma_n^*\nabla p_n\cdot\varphi$ goes to $\int_{\mathbb{R}^n}\beta'(p)\cdot\sigma^*\nabla p\cdot\varphi$ by the weak convergence of the $\sigma_n^*\nabla p_n$

So the sum converges to 0, this was to show.

Step B: We have

$$\lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p)\cdot\operatorname{div}(\sigma\varphi)-\beta'(p_n)\cdot\operatorname{div}(\sigma_n\varphi)$$
$$= \lim_{n\to\infty}\int_{\mathbb{R}^n} \big(\beta'(p)-\beta'(p_n)\big)\cdot(\operatorname{div}(\sigma_n\varphi)+\operatorname{div}(\sigma\varphi))$$
$$- \lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p)\operatorname{div}(\sigma_n\varphi)$$
$$+ \lim_{n\to\infty}\int_{\mathbb{R}^n} \beta'(p_n)\operatorname{div}(\sigma\varphi)$$

Again, we check the three integrals:

- The first integral $\int_{\mathbb{R}^n}(\beta'(p)-\beta'(p_n))\cdot(\operatorname{div}(\sigma_n\varphi)+\operatorname{div}(\sigma\varphi))$ goes to zero by Lemma 2.10
- The second integral $\int_{\mathbb{R}^n}\beta'(p)\operatorname{div}(\sigma_n\varphi)$ goes to $\int_{\mathbb{R}^n}\beta'(p)\operatorname{div}(\sigma\varphi)$ by the strong convergence of $\nabla\sigma_n$ to $\nabla\sigma$ in L^2_{loc}

- The third integral $\int_{\mathbb{R}^n} \beta'(p_n)\,\mathrm{div}(\sigma\varphi)$ goes to $\int_{\mathbb{R}^n} \beta'(p)\,\mathrm{div}(\sigma\varphi)$ by Lemma 2.10

So also here the sum goes to 0 and the proof is finished. □

3.2 Existence of solutions

Theorem 3.7. *Let there be b, σ as in Theorem 3.2. Then, for any $p_0 \in L^2 \cap L^\infty$, there exists a solution of*

$$\partial_t p + \partial_i(p b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) = 0 \tag{4}$$

in the space

$$X_2 = \Big\{p \in L^\infty\big([0,T], L^2 \cap L^\infty\big), \sigma^* \nabla p \in L^2\big([0,T], L^2\big)\Big\}$$

(see [5] p. 20 and p. 24 for the idea of the proof). The strategy is, to consider a regularized version of (4) (by approximating b and σ by smooth functions) and to show that the solutions of the smoothed PDE converge weakly. Then we show that the weak limit solves the actual PDE.
At first, we need to derive the a-priori-estimates of the following lemma. We assume b, σ and p_0 to be smooth. Then, according according to [11], Theorem 2 a solution exists, even smooth in space and continuously differentiable in time. For this solution we get some estimates:

Lemma 3.8. *Let there be b, σ and p_0 as in 3.2 and additionally smooth and let p be the solution of (4). Then we have constants C_1, C_2 and C_3, depending only on b and p_0, such that for all t (uniformly)*

$$\|p(t)\|_\infty \le C_1 \tag{5}$$

$$\|p(t)\|_2 \le C_2 \tag{6}$$

$$\|\sigma^* \nabla p\|_{L^2([0,T],L^2)} \le C_3 \tag{7}$$

Proof. So assume p to be a solution (continously differentiable in time and smooth in space) of (4) with $p|_{t=0} = p_0$ and $p_0 \in L^2 \cap L^\infty$ also smooth by [11].
We start with the L^∞-bound of p. Lets define $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\partial_t \varphi(t) = \sup_{x \in \mathbb{R}} \big[\mathrm{div} b(t,x)\big]^- \cdot \varphi(t) + \varepsilon$$

with $\varepsilon > 0$ and $\varphi(0) = \|p_0\|_\infty + \varepsilon$. Then, by [4, p. 53], φ exists and is bounded on $[0,T]$.
We show that $p(t,x) < \varphi(t)$ for all t, x by contradiction. Then, this is enough to show (5), as φ is bounded on $[0,T]$. So lets assume the existence of x_0, t_0 such that $p(t_0, x_0) \ge \varphi(t_0)$. Let t_0 be the minimal t under this assumption (φ is a monotonously

increasing continuous function and at $t = 0$ we have $p(x,0) < \varphi(0)$, so such a minimal t_0 is well defined), so, as p is continuous we especially get

$$p(t_0, x_0) = \varphi(t_0)$$

Now we claim that x_0 is a global maximum in space of p: If there was a x_1 with $p(t_0, x_1) > p(t_0, x_0)$, we would take the function $t \to p(t, x_1) - \varphi(t)$. This function is negative in 0 and strictly positive in t_0, so it has to have a zero-value less than t_0. This is a contradiction to the minimality of t_0, so x_0 is a maximum. We have, that

$$\partial_t p(t_0, x_0) - \partial_t \varphi(t_0) \geq 0 \tag{8}$$

because t_0 is by definition the smallest zero of $p - \varphi$, and as $p(0) - \varphi(0) < 0$ by definition, the derivative has to by nonnegative in t_0.
Now we go back to the differential equation (4) and split up the derivatives with the product rule:

$$\partial_t p + (\nabla p) \cdot b + p \operatorname{div} b - \frac{1}{2} \partial_i (\sigma_{ik} \sigma_{jk}) \partial_j p) - \frac{1}{2} \sigma_{ik} \sigma_{jk} \partial_i \partial_j p = 0$$

Lets consider this equation in (t_0, x_0). As x_0 is a maximum in space, we have, that $\nabla p(t_0, x_0) = 0$, and, that the Hessematrix is negative definit in x_0, so we have

$$\sigma_{ik} \sigma_{jk} \partial_i \partial_j p(x_0, t_0) \leq 0$$

So we have

$$\partial_t p(t_0, x_0) + \operatorname{div}(b(t_0, x_0)) \cdot p(t_0, x_0) = \frac{1}{2} \sigma_{ik} \sigma_{jk} \partial_i \partial_j p(t_0, x_0) \leq 0$$

This leads to:

$$\partial_t p(t_0, x_0) \leq -\operatorname{div}(b(t_0, x_0)) \cdot p(t_0, x_0)$$

Additionally φ is always nonnegative (by the defining ODE, the derivative and the initial value is nonnegative), so by $p(t_0, x_0) = \varphi(t_0)$ also $p(t_0, x_0)$ is nonnegative. This is important in the following estimate:

$$\begin{aligned} \partial_t p(t_0, x_0) &\leq -\operatorname{div}(b(t_0, x_0)) \cdot p(t_0, x_0) \leq \sup_{x \in \mathbb{R}} \big[\operatorname{div} b(t_0, x)\big]^{-} p(t_0, x_0) \\ &= \sup_{x \in \mathbb{R}} \big[\operatorname{div} b(t_0, x)\big]^{-} \varphi(t_0) = \partial_t \varphi(t_0) - \varepsilon \end{aligned}$$

This is a contradiction to (8). For a lower bound lets take $-p$, which also solves (4) (with the sign-flipped initial data) and for which we have established an upper bound.

For the L^2-bound we may multiply (4) with p and integrate in space over some $\mathbb{R}^n$. We assume that all integrals exist and are finite, which will be justified later:

$$\int_{\mathbb{R}^n} p \partial_t p + \int_{\mathbb{R}^n} p \partial_i (p b_i) - \int_{\mathbb{R}^n} \frac{1}{2} \big(\partial_i (\sigma_{ik} \sigma_{jk} \partial_j p)\big) p = 0$$

Now $p\partial_t p = \partial_t \frac{p^2}{2}$, and p is continously differentiable in time, so we may change integration and differentiation in the first integral:

$$\frac{1}{2}\partial_t \int_{\mathbb{R}^n} p^2 + \int_{\mathbb{R}^n} p\partial_i(pb_i) - \frac{1}{2}\int_{\mathbb{R}^n} \big(\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p)\big)p = 0 \tag{9}$$

In the second integral we integrate by parts two times:

$$\int_{\mathbb{R}^n} p\partial_i(pb_i) = -\int_{\mathbb{R}^n} \partial_i p p b_i = -\int_{\mathbb{R}^n} \partial_i \frac{p^2}{2} b_i = \int_{\mathbb{R}^n} \frac{p^2}{2} \operatorname{div} b$$

Multiplying with 2 and another integration by parts in the last term in (9) leads to

$$\partial_t \int_{\mathbb{R}^n} p^2 + \int_{\mathbb{R}^n} p^2 \operatorname{div} b + \int_{\mathbb{R}^n} (\sigma_{ik}\sigma_{jk}\partial_j p)\partial_i p = 0$$

In the last term we have $(\sigma_{ik}\sigma_{jk}\partial_j p)\partial_i p = |\sigma^*\nabla p|^2$:

$$\partial_t \int_{\mathbb{R}^n} p^2 + \int_{\mathbb{R}^n} p^2 \operatorname{div} b + \int_{\mathbb{R}^n} |\sigma^*\nabla p|^2 = 0 \tag{10}$$

One of the regularity assumptions of 3.2 is, that $[\operatorname{div} b]^- \in L^1([0,T], L^1_{\text{loc}}(\mathbb{R}^n))$, so let there be a function $C \in L^1([0,T])$ such that $[\operatorname{div} b]^- < C(t)$. Since $\int_{\mathbb{R}^n} |\sigma^*\nabla p|^2 \geq 0$, (10) leads to

$$\partial_t \int_{\mathbb{R}^n} p^2 \leq C(t) \int_{\mathbb{R}^n} p^2$$

Now we can apply Gronwalls inequality in differential form [9, p. 711], which leads to

$$||p(t)||_2^2 \leq e^{\|C(\cdot)\|_1 t}||p_0||_2^2 = C_1||p_0||_2^2 \tag{11}$$

with $C_1 = e^{\|C(\cdot)\|_1 T}$, so (6) is proven.
We want to get an analog bound for the L^2-Norm of $t \to \|\sigma^*(\cdot,t)\nabla p(\cdot,t)\|_2$ (in fact the squared L^2-Norm) using (10) :

$$\begin{aligned}
\int_0^T \|\sigma^*\nabla p\|_2^2 \,\mathrm{d}t &= \int_0^T \int_{\mathbb{R}^n} |\sigma^*(x,t)\nabla p(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t \\
&= \int_0^T \left(-\partial_t \int_{\mathbb{R}^n} p^2 \,\mathrm{d}x - \int_{\mathbb{R}^n} p^2 \operatorname{div} b \,\mathrm{d}x\right)\mathrm{d}t \\
&= -\int_{\mathbb{R}^n} p(x,T)^2 \,\mathrm{d}x + \|p_0\|_2^2 - \int_0^T \int_{\mathbb{R}^n} p^2 \operatorname{div} b \,\mathrm{d}x\,\mathrm{d}t \\
&\leq \int_0^T \|[\operatorname{div} b(t)]^-\|_\infty \int_{\mathbb{R}^n} p^2 \,\mathrm{d}x\,\mathrm{d}t + \|p_0\|_2^2
\end{aligned}$$

We can use the L^2-estimate (11) and get

$$\begin{aligned}\int_0^T \|\sigma^*\nabla p\|_2^2\,\mathrm{d}t &\le \int_0^T \|[\operatorname{div} b(t)]^-\|_\infty C_1\|p_0\|_2^2\,\mathrm{d}t + \|p_0\|_2^2\\ &\le C_1\|p_0\|_2^2 \int_0^T \|[\operatorname{div} b(t)]^-\|_\infty\,\mathrm{d}t + \|p_0\|_2^2\\ &\le C_1\|p_0\|_2^2 \cdot \|\|[\operatorname{div} b(\cdot)]^-\|_\infty\|_1 + \|p_0\|_2^2\\ &\le C_2\|p_0\|_2^2 \qquad (12)\end{aligned}$$

with $C_2 = C_1\left\|\|[\operatorname{div} b(\cdot)]^-\|_\infty\right\|_1 + 1$. Here $\left\|\|[\operatorname{div} b(\cdot)]^-\|_\infty\right\|_1$ is finite because the components of b are assumed to be in $L^1\big([0,T], L^\infty(\mathbb{R}^n)\big)$. So also (7) is proven and the proof is finished. □

Now we are able to prove Theorem 3.7:

Proof of Theorem 3.7. Lets take convolution kernels in space and time ρ_ε and define $b_\varepsilon = \rho_\varepsilon * b$, $\sigma_\varepsilon = \rho_\varepsilon * \sigma$ and $p_{0\varepsilon} = \rho_\varepsilon * p_0$. To convolute in time, we need to define b and σ on $[-\varepsilon, T+\varepsilon]$ by reflecting on the interval bound, so for $0 < \tau < \epsilon$ we extend b by $b(-\tau) = b(\tau)$ and $b(T+\tau) = b(T-\tau)$. σ is extended analogously.
Then a solution to the according smooth problem exists according to [11], Theorem 2, let it be assigned with p_ε. This solution exists even smooth in space and continuously differentiable in time, so the a-priori-estimates (5), (6) and (7) hold for p_ε. Additionally, by the properties of the convolution, we know that the norms of $\operatorname{div} b_\varepsilon$ and $p_{0\varepsilon}$ are lower or equal then the correspondent non-smoothed ones, so there are uniform constants in the a-priori-estimates (this can be easily checked by going through the proof of Lemma 3.8 and verifying, that the constants only depend on the norms of $\operatorname{div} b$ and p_0).
So we have a bounded sequence p_ε in $L^2\big([0,T]\times\mathbb{R}^n\big)$. By picking a subsequence we get a weak convergent subsequence, and, as $\|\sigma_\varepsilon^t \nabla p_\varepsilon\|$ is also bounded, we can take a weak convergent subsubsequence (but a priori it is not clear that it converges to $\sigma^*\nabla p$!), which we call without loss of generality again p_ε, so we have functions p and u such that

$$\begin{aligned} p_\varepsilon &\rightharpoonup p\\ \sigma_\varepsilon^t \nabla p_\varepsilon &\rightharpoonup u\end{aligned}$$

For a test function $\varphi \in C_c^\infty([0,T)\times\mathbb{R}^n)$, we have

$$\int_0^T\int_{\mathbb{R}^n} u\varphi = \lim_{\varepsilon\to 0}\int_0^T\int_{\mathbb{R}^n} \sigma_\varepsilon^t \nabla p_\varepsilon \varphi = -\lim_{\varepsilon\to 0}\int_0^T\int_{\mathbb{R}^n} p_\varepsilon \cdot \nabla(\sigma_\varepsilon\varphi) = -\int_0^T\int_{\mathbb{R}^n} p\cdot\nabla(\sigma\varphi)$$

with the last equality by the weak convergence of p_ε and, the strong convergence of $\nabla(\sigma_\varepsilon\varphi)$ and Lemma 2.7. So $u = \sigma^t\nabla p$ in the distributional sense of Remark 3.3.
So we need to show that p is a weak solution of (4). This is done by taking the

weak formulation with the smoothed terms and checking the convergence to the corresponding terms. We have

$$\int_0^T \int_{\mathbb{R}^n} p_\varepsilon \partial_t \varphi \,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^n} p_{0\varepsilon}\varphi(0,\cdot)\,\mathrm{d}x = \\ -\int_0^T \int_{\mathbb{R}^n} p_\varepsilon \langle b_\varepsilon, \nabla\varphi\rangle \,\mathrm{d}x\,\mathrm{d}t + \frac{1}{2}\int_0^T \int_{\mathbb{R}^n} \langle \sigma_\varepsilon^* \nabla p_\varepsilon, \sigma_\varepsilon^* \nabla \varphi\rangle \,\mathrm{d}x\,\mathrm{d}p$$

for any test function $\varphi \in C_c^\infty([0,T) \times \mathbb{R}^n)$. Now we need to check the convergence of these integrals:

- The first integral $\int_0^T \int_{\mathbb{R}^n} p_\varepsilon \partial_t \varphi \,\mathrm{d}x\,\mathrm{d}t$ goes to $\int_0^T \int_{\mathbb{R}^n} p \partial_t \varphi \,\mathrm{d}x\,\mathrm{d}t$ by the weak convergence of p_ε
- The second integral $\int_{\mathbb{R}^n} p_{0\varepsilon}\varphi(0,\cdot)\,\mathrm{d}x$ goes to $\int_{\mathbb{R}^n} p_0\varphi(0,\cdot)\,\mathrm{d}x$ by the strong convergence of $p_{0\varepsilon}$ to p_0 (which implies weak convergence).
- The third integral $\int_0^T \int_{\mathbb{R}^n} p_\varepsilon \langle b_\varepsilon, \nabla\varphi\rangle \,\mathrm{d}x\,\mathrm{d}t$ goes to $\int_0^T \int_{\mathbb{R}^n} p \langle b, \nabla\varphi\rangle \,\mathrm{d}x\,\mathrm{d}t$ by Lemma 2.7 with $\langle b_\varepsilon, \nabla\varphi\rangle$ strong convergent and p_ε weak convergent
- The last integral $\int_0^T \int_{\mathbb{R}^n} \langle \sigma_\varepsilon^* \nabla p_\varepsilon, \sigma_\varepsilon^* \nabla \varphi\rangle \,\mathrm{d}x\,\mathrm{d}p$ goes to $\int_0^T \int_{\mathbb{R}^n} \langle \sigma^* \nabla p, \sigma^* \nabla \varphi\rangle \,\mathrm{d}x\,\mathrm{d}p$ also by Lemma 2.7, because $\sigma_\varepsilon^* \nabla p_\varepsilon$ converges weakly to $\sigma^* p$ (recalling the definition of $\sigma^* p$ above in the distributional sense) and $\sigma_\varepsilon^* \nabla \varphi$ is strong convergent in L^2 to $\sigma^* \nabla\varphi$

So all integrals converge to the corresponding terms in p, p_0, b and σ, so p is in fact a solution of (4). □

3.3 Uniqueness of solutions

3.3.1 The technique of renormalized solutions

To get a suitable definition of a renormalized solution, we need to calculate at first only formally, taking no account of any regularity or the difference between weak and strong solutions. We start with (3), but of course inserting $\beta(p)$ instead of p with

$\beta : \mathbb{R} \to \mathbb{R}$ and assuming that p solves (3)

$$\begin{aligned}
&\partial_t \beta(p) + \partial_i(\beta(p)b_i) - \frac{1}{2}\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p)\big) \\
&= \beta'(p)\cdot\partial_t p + \beta'(p)\partial_i p b_i + \beta(p)\operatorname{div}(b) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\beta'(p)\partial_j p) \\
&= \beta'(p)\cdot\partial_t p + \beta'(p)\partial_i p b_i + \beta(p)\operatorname{div}(b) - \frac{1}{2}\beta'(p)\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_j p\beta''(p)\partial_i p \\
&= \beta'(p)\Big(\partial_t p + \partial_i p b_i - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p)\Big) + \beta(p)\operatorname{div}(b) - \frac{1}{2}\beta''(p)\sigma_{ik}\sigma_{jk}\partial_j p\partial_i p \\
&= \beta'(p)\underbrace{\Big(\partial_t p + \partial_i p b_i + p\operatorname{div} b - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p)\Big)}_{=0 \text{ by } (3)} - \beta'(p)p\operatorname{div}(b) + \beta(p)\operatorname{div}(b) \\
&\quad - \frac{1}{2}\beta''(p)|\sigma^*\nabla p|^2 \\
&= \big(\beta(p) - p\beta'(p)\big)\operatorname{div}(b) - \frac{1}{2}\beta''(p)|\sigma^*\nabla p|^2
\end{aligned}$$

So the following is a natural definition of a renormalized solution, of course now again understood distributionally:

Definition 3.9. A solution of (3) is called a *renormalized solution*, if, for all $\beta \in C^2(\mathbb{R})$, the following equation holds (in the distributional sense as in definition 3.1)

$$\partial_t\beta(p) + \partial_i(\beta(p)b_i) - \frac{1}{2}\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p)\big) - \big(\beta(p) - p\beta'(p)\big)\operatorname{div}(b) + \frac{1}{2}\beta''(p)|\sigma^*\nabla p|^2 = 0$$

Remark 3.10. The distribution in definition 3.9 (and of course also the distribution in definition 3.1) can also be tested with a test function only in space $\varphi_2 \in C_c^\infty(\mathbb{R}^n)$ to get a distribution u_{φ_2} on $[0,T]$, seen as a functional on $C_0^\infty([0,T])$. Formally, we take $\varphi_1 \in C_0^\infty([0,T])$ and $\varphi_2 \in C_c^\infty(\mathbb{R}^n)$ and test definition 3.9 with $\varphi(x,t) = \varphi_1(t)\varphi_2(x)$:

$$\begin{aligned}
u_{\varphi_2}(\varphi_1) = &-\int_0^T\int_{\mathbb{R}^n}\beta(p)\varphi_2\partial_t\varphi_1 - \int_0^T\int_{\mathbb{R}^n}\beta(p)b_i\partial_i\varphi_2\varphi_1 \\
&+ \frac{1}{2}\int_0^T\int_{\mathbb{R}^n}\varphi_1\big\langle\sigma^*\nabla\varphi_2, \sigma^*\nabla\beta(p)\big\rangle \\
&+ \int_0^T\int_{\mathbb{R}^n}\big(-\beta(p)\operatorname{div}(b) + p\beta'(p)\operatorname{div}(b) + \frac{1}{2}\beta''(p)|\sigma^*\nabla p|^2\big)\varphi_1\varphi_2 \\
= 0&
\end{aligned}$$

In this distributional sense, we also write $\big(\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^n}\beta(p)\varphi_2\big)(\varphi_1) = -\int_0^T\int_{\mathbb{R}^n}\beta(p)\varphi_2\partial_t\varphi_1$ for the first integral, as a distribution on $[0,T]$.

3.3.2 Proof of uniqueness under the renormalization assumption

Theorem 3.11. *If any weak solution of the Fokker-Planck-equation (3) is also a renormalized solution, the uniqueness in Theorem 3.2 holds.*

Proof. ([14], p. 8): As we consider a linear equation, its enough to prove, that $p_0 \equiv 0$ implies $p(t) \equiv 0$ for almost every t.
We choose $\beta(s) = s^2$, so we have $\beta'' = 2$ and $\beta(p) - p\beta'(p) = -p^2$, so definition 3.9 together with Lemma 3.6 leads to

$$\partial_t(p^2) + \partial_i(p^2 b_i) - \partial_i(\sigma_{ik}\sigma_{jk} p \partial_j p) + p^2 \operatorname{div}(b) = -|\sigma^* \nabla p|^2 \leq 0 \tag{13}$$

Of course, the $\leq$ is meant in the sense, that testing with a nonnegative function leads to a nonnegative result.
As mentioned in 3.10 it is possible to test with a test function in $C_c^\infty(\mathbb{R}^n)$ to get a distribution on $[0,T]$. We choose especially a function φ_R with $R > 0$, defined by a nonnegative function $\varphi \in C_c(\mathbb{R}^n, [0,1])$ satisfying $\varphi|_B(1) \equiv 1$ and $\operatorname{spt}(\varphi) \subset B(2)$. Then let φ_R be the stretched version of φ by the parameter R, so

$$\varphi_R(x) = \varphi\left(\frac{x}{R}\right)$$

So we have $\nabla\varphi_R(x) = \frac{1}{R}\varphi(\frac{x}{R})$.
Testing with this function in (13) leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^n} p^2\varphi_R - \int_{\mathbb{R}^n} p^2\langle b, \nabla\varphi_R\rangle + \int_{\mathbb{R}^n} p\langle\sigma^*\nabla\varphi_R, \sigma^*\nabla p\rangle + \int_{\mathbb{R}^n} p^2\varphi_R \operatorname{div}(b) \leq 0$$

So:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^n} p^2\varphi_R \leq \int_{\mathbb{R}^n} p^2\langle b, \nabla\varphi_R\rangle - \int_{\mathbb{R}^n} p\langle\sigma^*\nabla\varphi_R, \sigma^*\nabla p\rangle - \int_{\mathbb{R}^n} p^2\varphi_R \operatorname{div}(b) \tag{14}$$

(14) is an inequality of distributions on $[0,T]$, but the three integrals on the right hand side exist also in a classical sense, so we can try to estimate them. In the first one we have:

$$\left|\int_{\mathbb{R}^n} p^2\langle b, \nabla\varphi_R\rangle\right| = \left|\int_{\{R\leq|x|\leq 2R\}} p^2\langle b, \frac{1}{R}\nabla\varphi\left(\frac{\cdot}{R}\right)\rangle\right| \leq \|\nabla\varphi\|_\infty \int_{\{R\leq|x|\leq 2R\}} p^2|b|\frac{1}{R}$$

Now we have $|x| \leq 2R$, so we can estimate with a generic constant C for R big enough:

$$\left|\int_{\mathbb{R}^n} p^2\langle b, \nabla\varphi_R\rangle\right| \leq C\|\nabla\varphi\|_\infty \int_{\{x\geq R\}} p^2\frac{|b|}{1+|x|}$$

Next we use, that $\frac{b}{1+|x|} \in \Big(L^1\big([0,T], L^1 + L^\infty(\mathbb{R}^n)\big)\Big)^n$ by assumption, so let there be vector fields b_1, b_2 such that $b = b_1 + b_2$ and $\frac{|b_1|}{1+|x|} \in L^1\big([0,T], L^1(\mathbb{R}^n)\big)$, $\frac{|b_2|}{1+|x|} \in L^1\big([0,T], L^\infty(\mathbb{R}^n)\big)$. So, taking $\|\nabla\varphi\|_\infty$ into the generic constant, we get

$$\left|\int_{\mathbb{R}^n} p^2\langle b, \nabla\varphi_R\rangle\right| \leq C\int_{\{x\geq R\}} p(t,\cdot)^2\frac{|b_1(t)|}{1+|x|} + C\int_{\{x\geq R\}} p(t,\cdot)^2\frac{|b_2(t)|}{1+|x|} \tag{15}$$

We want to use dominated convergence for $R \to \infty$, so we need dominating integrable functions for both integrals, seen as functions of t:

$$C \int_{\{x \geq R\}} p^2 \frac{|b_1(t)|}{1+|x|} \leq C \|p^2(t,\cdot)\|_\infty \left\| \frac{b_1(t)}{1+x} \right\|_1$$

As $p \in L^\infty([0,T], L^\infty)$, the first norm is finite and the last one is an integrable function of t by definition of b_1.
Next we have

$$C \int_{\{x \geq R\}} p(t,\cdot)^2 \frac{|b_2(t)|}{1+|x|} \leq C \left\| \frac{b_2(t)}{1+|x|} \right\|_\infty \int_{\{x \geq R\}} p(t,\cdot)^2 \leq C \left\| \frac{b_2(t)}{1+|x|} \right\|_\infty \|p(t,\cdot)\|_2^2$$

As $p \in L^\infty\big([0,T], L^2\big)$, we also have a dominating function. So we can use dominated convergence in (15), and as all terms on the right hand side go to 0 for almost every fixed t at $R \to \infty$, we conclude

$$\lim_{R \to \infty} \int_0^T \left| \int_{R^n} p^2 \langle b, \nabla \varphi_R \rangle \right| = 0 \tag{16}$$

Next, we estimate $\int_{\mathbb{R}^n} p \langle \sigma^* \nabla \varphi_R, \sigma^* \nabla p \rangle$ with the same estimate for $\frac{1}{R}$:

$$\begin{aligned} \left| \int_{\mathbb{R}^n} p \langle \sigma^* \nabla \varphi_R, \sigma^* \nabla p \rangle \right| &\leq \int_{\{R \leq |x| \leq 2R\}} |p| \cdot |\sigma^* \nabla p| \cdot \frac{1}{R} \|\nabla \varphi\|_\infty |\sigma| \\ &\leq C \|\nabla \varphi\|_\infty \int_{\{|x| \geq R\}} |p| \cdot |\sigma^* \nabla p| \cdot \frac{|\sigma|}{1+|x|} \end{aligned}$$

Again by assumption, we split σ in $\sigma = \sigma_1 + \sigma_2$, with

$$\begin{aligned} \frac{|\sigma_1|}{1+|x|} &\in L^2([0,T], L^2(\mathbb{R}^n)) \\ \frac{|\sigma_2|}{1+|x|} &\in L^2([0,T], L^\infty(\mathbb{R}^n)) \end{aligned}$$

So we have:

$$\left| \int_{\mathbb{R}^n} p \langle \sigma^* \nabla \varphi_R, \sigma^* \nabla p \rangle \right| \leq C \int_{\{|x| \geq R\}} |p| \cdot |\sigma^* \nabla p| \cdot \frac{|\sigma_1|}{1+|x|} + C \int_{\{|x| \geq R\}} |p| \cdot |\sigma^* \nabla p| \cdot \frac{|\sigma_2|}{1+|x|} \tag{17}$$

Now, as above, we estimate both terms:

$$\begin{aligned} &C \int_{\{|x| \geq R\}} |p| \cdot |\sigma^* \nabla p| \cdot \frac{|\sigma_1|}{1+|x|} \\ &\qquad \leq C \|p\|_\infty \int_{\{|x| \geq R\}} |\sigma^* \nabla p| \cdot \frac{|\sigma_1|}{1+|x|} \leq C \|p\|_\infty \|\sigma^*(t) \nabla p(t)\|_2 \left\| \frac{|\sigma_1(t)|}{1+|x|} \right\|_2 \end{aligned}$$

As functions of t, both $t \to \|\sigma^*(t)\nabla p(t)\|_2$ and $t \to \left\|\frac{|\sigma_1(t)|}{1+|x|}\right\|_2$ are L^2-functions by assumption, so the product is a L^1-function and thus we have found a dominating function.
Next we estimate the second term:

$$\begin{aligned}
C\int_{\{|x|\geq R\}} |p|\cdot|\sigma^*\nabla p|\cdot\frac{|\sigma_2|}{1+|x|} &\leq C\left\|\frac{|\sigma_2(t)|}{1+|x|}\right\|_\infty \int_{\{|x|\geq R\}} |p|\cdot|\sigma^*\nabla p| \\
&\leq C\left\|\frac{|\sigma_2(t)|}{1+|x|}\right\|_\infty \|p(t)\|_2\|\sigma^*(t)\nabla p(t)\|_2 \\
&\leq C\left\|\frac{|\sigma_2(t)|}{1+|x|}\right\|_\infty \|p\|_{L^\infty([0,T],L^2(\mathbb{R}^n))}\|\sigma^*(t)\nabla p(t)\|_2
\end{aligned}$$

Again, as functions of t, both $t \to \|\sigma^*(t)\nabla p(t)\|_2$ and $t \to \left\|\frac{|\sigma_2(t)|}{1+|x|}\right\|_\infty$ are L^2-functions by assumption and $\|p\|_{L^\infty([0,T],L^2(\mathbb{R}^n))}$ is a constant, so we have found a dominating function.
So we can use the dominated convergence theorem in (17) and again we have pointwise convergence to 0, so we have:

$$\lim_{R\to\infty}\int_0^T \left|\int_{\mathbb{R}^n} p\langle\sigma^*\nabla\varphi_R, \sigma^*\nabla p\rangle\right| = 0 \tag{18}$$

At last we have to estimate the last integral of the right hand side of (14):

$$-\int_{\mathbb{R}^n} p^2\varphi_R \operatorname{div}(b) \leq \|[\operatorname{div}(b)]^-\|_\infty \int_{\mathbb{R}^n} p^2\varphi_R \tag{19}$$

Now we are able to estimate (14). To simplify notation, we define:

$$A_R(t) = \int_{\mathbb{R}^n} p^2(t)\langle b(t), \nabla\varphi_R\rangle - \int_{\mathbb{R}^n} p(t)\big\langle\sigma^*(t)\nabla\varphi_R, \sigma^*(t)\nabla p(t)\big\rangle$$

Thus, (16) and (18) lead to

$$\lim_{R\to\infty}\int_0^T |A_R(t)|dt = 0 \tag{20}$$

So, testing (14) with a nonnegative test function ψ in time, and using (19) we get

$$\begin{aligned}
&-\int_0^T\int_{\mathbb{R}^n} p(x,t)^2\varphi_R(x)\psi'(t)\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad\leq \int_0^T A_R(t)\psi(t)\,\mathrm{d}t + \int_0^T \|[\operatorname{div}(b(t))]^-\|_\infty \int_{\mathbb{R}^n} p^2(x,t)\varphi_R(x)\,\mathrm{d}x\,\psi(t)\,\mathrm{d}t
\end{aligned}$$

Now we use the distributional form of Gronwalls Lemma (Lemma 2.17) with

- $f(t) = \int_{\mathbb{R}^n} p(x,t)^2 \varphi_R(x)\,\mathrm{d}x$
- $g(t) = \|[\mathrm{div}(b(t))]^-\|_\infty$
- $h(t) = A_R(t)$

The assumptions of Lemma 2.17 are all obviously fulfilled except the continuity of f. But also this holds as f is a Sobolev-function in one dimension and hence continuous. Thus we have

$$\int_{\mathbb{R}^n} p(x,t)^2 \varphi_R(x)\,\mathrm{d}x \le \exp\left(\int_0^t \|[\mathrm{div}(b(r))]^-\|_\infty\,\mathrm{d}r\right)\int_0^t |A_R(s)|\,\mathrm{d}s$$

Here we take the limit $R \to \infty$. On the right hand side we use (20) and that

$$\exp\left(\int_0^t \left\|\left[\mathrm{div}(b(r))\right]^-\right\|_\infty \mathrm{d}r\right) \le \exp\left(\left\|[\mathrm{div}(b)]^-\right\|_1\right)$$

So the integral exists as $[\mathrm{div}(b)]^- \in L^1\big([0,T], L^\infty(\mathbb{R}^n)\big)$ and the right hand side goes to 0.
On the left hand side we use the monotone convergence theorem as φ_R is converging monotone against the constant function with value 1 and p^2 is nonnegative, so we get:

$$\int_{\mathbb{R}^n} p(x,t)^2\,\mathrm{d}x \le 0$$

for almost every t, so $p \equiv 0$ almost everywhere. This was to show. □

3.3.3 Commutators and the commutator estimate of DiPerna and Lions

The proof of the renormalization assumption will regularize the renormalized Fokker-Planck equation to approximate the renormalized Fokker-Planck equation up to some error terms which will converge to zero. For this aim, we have to regularize the standard-Fokker-Planck equation at first, as the error terms arising here will also arise in the regularized renormalized Fokker-Planck equation.
So lets take even convolution kernels ρ with support in $B_1(0)$ and define $p_\varepsilon = \rho_\varepsilon * p$. Then, we are interested in the term

$$\partial_t p_\varepsilon + \partial_i(p_\varepsilon b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon) \tag{21}$$

Therefore we regularize (3) in the spatial variables, as explained in Section 2.3, especially in the lemmata 2.15 and 2.16, so we get:

$$\partial_t(\rho_\varepsilon * p) + \rho_\varepsilon * \partial_i(p b_i) - \frac{1}{2}\rho_\varepsilon * \partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) = 0 \tag{22}$$

For the following calculations we define commutators:

Definition 3.12. For a differential operator or a function c we define

$$[\rho_\varepsilon, c](f) = \rho_\varepsilon * (cf) - c(\rho_\varepsilon * f)$$

Remark 3.13. Note that these commutators are in general only distributions, as cf does not need to be defined as a function if f is for example only in a L^p-space. See section 2.3 for the details of the mollifications of distributions.

So we analyse the terms of (22) trying to get terms of (21):

$$\begin{aligned}\rho_\varepsilon * \partial_i(pb_i) &= \rho_\varepsilon * (\operatorname{div}(b)p) + \rho_\varepsilon * (b_i\partial_i p)\\ &= \rho_\varepsilon * (\operatorname{div}(b)p) - \operatorname{div}(b)p_\varepsilon + \rho_\varepsilon * (b_i\partial_i p) - b_i\partial_i p_\varepsilon + \partial_i(b_i p_\varepsilon)\\ &= [\rho_\varepsilon, \operatorname{div}(b)](p) + [\rho_\varepsilon, b_i\partial_i](p) + \partial_i(b_i p_\varepsilon)\\ &=: Q_{1,\varepsilon} + Q_{2,\varepsilon} + \partial_i(b_i p_\varepsilon)\end{aligned}$$

Next we have

$$\begin{aligned}&\rho_\varepsilon * \partial_i(\sigma_{ik}\sigma_{jk}\partial_j p) = \rho_\varepsilon * \big((\partial_i\sigma_{ik})\sigma_{jk}\partial_j p + \sigma_{ik}\partial_i(\sigma_{jk}\partial_j p)\big)\\ &= [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j p) + (\partial_i\sigma_{ik})\rho_\varepsilon * (\sigma_{jk}\partial_j p) + [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j p) + \sigma_{ik}\partial_i(\rho_\varepsilon * (\sigma_{jk}\partial_j p))\\ &= [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j p) + [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j p) + \partial_i\left(\sigma_{ik}\cdot\rho_\varepsilon * (\sigma_{jk}\partial_j p)\right)\\ &=: S_\varepsilon + T_\varepsilon + \partial_i\left(\sigma_{ik}\rho_\varepsilon * (\sigma_{jk}\partial_j p)\right)\end{aligned}$$

We further analyse the term $\sigma_{ik}\cdot\rho_\varepsilon * (\sigma_{jk}\partial_j p)$ by defining $R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jk}\partial_j](p)$:

$$\sigma_{ik}\rho_\varepsilon * (\sigma_{jk}\partial_j p) = \sigma_{ik}R_{k,\varepsilon} + \sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon$$

So, (22) leads to:

$$\partial_t p_\varepsilon + \partial_i(p_\varepsilon b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon) = -Q_{1,\varepsilon} - Q_{2,\varepsilon} + \frac{1}{2}(\partial_i(\sigma_{ik}R_{k,\varepsilon}) + S_\varepsilon + T_\varepsilon) \tag{23}$$

with the above defined error terms:

- $Q_{1,\varepsilon} = [\rho_\varepsilon, \operatorname{div}(b)](p) = \rho_\varepsilon * (\operatorname{div}(b)p) - \operatorname{div}(b)p_\varepsilon$
- $Q_{2,\varepsilon} = [\rho_\varepsilon, b_i\partial_i](p) = \rho_\varepsilon * (b_i\partial_i p) - b_i\partial_i p_\varepsilon$
- $R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jk}\partial_j](p) = \rho_\varepsilon * (\sigma_{jk}\partial_j p) - \sigma_{jk}\partial_j p_\varepsilon$
- $S_\varepsilon = [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j p) = \rho_\varepsilon * ((\partial_i\sigma_{ik})\sigma_{jk}\partial_j p) - (\partial_i\sigma_{ik})\cdot\rho_\varepsilon * \sigma_{jk}\partial_j p$
- $T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j p) = \rho_\varepsilon * ((\sigma_{ik}\partial_i)\sigma_{jk}\partial_j p) - \sigma_{ik}\partial_i(\rho_\varepsilon * \sigma_{jk}\partial_j p)$

Later, in the proof of the renormalization assumption (Theorem 3.20), we will need the behavior of the error terms at $\varepsilon \searrow 0$. The term $Q_{2,\varepsilon}$ is difficult in the BV setting, but the other terms can be dealt with a commutator estimate from the Di-Perna-Lions Theory:

Lemma 3.14. *For* $1 \leq r, \alpha, r_1, \alpha_1 \leq \infty$ *we define* β *and* β_1 *by* $\frac{1}{\beta} = \frac{1}{r} + \frac{1}{\alpha}$ *and* $\frac{1}{\beta_1} = \frac{1}{r_1} + \frac{1}{\alpha_1}$. *Let* c, f *and* g *be in the following spaces:*

- $c \in \left(L^{\alpha_1}([0,T], W^{1,\alpha}_{\text{loc}}(\mathbb{R}^n))\right)^n$
- $f \in L^{r_1}\left([0,T], L^r_{\text{loc}}(\mathbb{R}^n)\right)$
- $g \in L^{\alpha_1}\left([0,T], L^\alpha_{\text{loc}}(\mathbb{R}^n)\right)$

Then the following commutators converge at $\varepsilon \to 0$

$$[\rho_\varepsilon, c_i \partial_i](f) \to 0 \text{ in } L^{\beta_1}\left([0,T], L^\beta_{\text{loc}}(\mathbb{R}^n)\right) \tag{24}$$

$$[\rho_\varepsilon, g](f) \to 0 \text{ in } L^{\beta_1}\left([0,T], L^\beta_{\text{loc}}(\mathbb{R}^n)\right) \tag{25}$$

Proof. (see [7][Theorem II.1]) The proof is done in five steps. In the first three steps we prove (24), but only time independent, in the fourth step we include the time dependency and in the last step we show how to prove (25). Also mind the following Remark 3.15 for the case when $r = r_1 = \infty$.

Step 1: An estimate

As said, we fix the time dependency in the fourth step, so we consider $c \in W^{1,\alpha}_{\text{loc}}(\mathbb{R}^n)^n$ and $f \in L^r_{\text{loc}}(\mathbb{R}^n)$ and show $[\rho_\varepsilon, c_i\partial_i](f) \to 0$ in $L^\beta_{\text{loc}}(\mathbb{R}^n)$

$$\begin{aligned}
[\rho_\varepsilon, c_i\partial_i](f)(x) &= (c_i\partial_i f) * \rho_\varepsilon(x) - c_i(x)\partial_i(f * \rho_\varepsilon)(x) \\
&= (c_i\partial_i f) * \rho_\varepsilon(x) - c_i(x)(f * \partial_i\rho_\varepsilon)(x) \\
&= \int_{\mathbb{R}^n} c_i(y)\partial_i f(y)\rho_\varepsilon(x-y)\,\mathrm{d}y - \int_{\mathbb{R}^n} c_i(x)f(y)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}^n} -\partial_i c_i(y)f(y)\rho_\varepsilon(x-y) + c_i(y)f(y)\partial_i\rho_\varepsilon(x-y) \\
&\quad - c_i(x)f(y)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y \\
&= -(f \operatorname{div} c) * \rho_\varepsilon(x) + \int_{\mathbb{R}^n} f(y)\big(c_i(y) - c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y
\end{aligned}$$

The first term goes to $-f \operatorname{div} c$ in $L^\beta_{\text{loc}}(\mathbb{R}^n)$, so we consider the second one. In the following step, we will take the L^β-norm on a ball B_R, as we want to show convergence in $L^\beta_{\text{loc}}(\mathbb{R}^n)$. C will denote various constants, changing from line to line and independent of ε, f, c, only depending on the convolution kernel ρ and R:

$$\begin{aligned}
&\left\|\int_{\mathbb{R}^n} f(y)\big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y\right\|_{L^\beta(B_R)}^\beta \\
&= \int_{B_R}\left|\int_{\mathbb{R}^n} f(y)\big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y\right|^\beta \mathrm{d}x \\
&= \int_{B_R}\left|\int_{B_\varepsilon(x)} f(y)\big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y\right|^\beta \mathrm{d}x \\
&= \int_{B_R}\left|\int_{B_\varepsilon(x)} f(y)\big(c_i(y)-c_i(x)\big)\left(\frac{1}{\varepsilon^{N+1}}\partial_i\rho\left(\frac{x-y}{\varepsilon}\right)\right)\mathrm{d}y\right|^\beta \mathrm{d}x \\
&= \int_{B_R}\left|\int_{B_\varepsilon(x)} f(y)\left(\frac{c_i(y)-c_i(x)}{\varepsilon}\right)\partial_i\rho\left(\frac{x-y}{\varepsilon}\right)\frac{1}{\varepsilon^N}\,\mathrm{d}y\right|^\beta \mathrm{d}x \\
&= \int_{B_R}\left|\int_{B_1(0)} f(x+\varepsilon z)\left(\frac{c_i(x+\varepsilon z)-c_i(x)}{\varepsilon}\right)\partial_i\rho\,(z)\,\mathrm{d}z\right|^\beta \mathrm{d}x
\end{aligned}$$

Next we use that $\partial_i\rho$ is bounded, so we can estimate it by C. In the following estimate we use Hölder with the pair of Hölder-conjugate exponents $\frac{r}{\beta}$ and $\frac{\alpha}{\beta}$ and Jensens inequality to pull a $^\beta$ into the integrals, maybe by changing the constant C:

$$\begin{aligned}
&\leq C\int_{B_R}\left|\int_{B_1(0)} f(x+\varepsilon z)\left(\frac{c_i(x+\varepsilon z)-c_i(x)}{\varepsilon}\right)\mathrm{d}z\right|^\beta \mathrm{d}x \qquad (26)\\
&\leq C\int_{B_R}\left(\int_{B_1(0)}|f(x+\varepsilon z)|^{\frac{r}{\beta}}\,\mathrm{d}z\right)^{\frac{\beta^2}{r}}\left(\int_{B_1(0)}\left(\frac{|c_i(x+\varepsilon z)-c_i(x)|}{\varepsilon}\right)^{\frac{\alpha}{\beta}}\mathrm{d}z\right)^{\frac{\beta^2}{\alpha}}\mathrm{d}x \\
&\leq C\int_{B_R}\left(\int_{B_1(0)}|f(x+\varepsilon z)|^{r}\,\mathrm{d}z\right)^{\frac{\beta}{r}}\left(\int_{B_1(0)}\left(\frac{|c_i(x+\varepsilon z)-c_i(x)|}{\varepsilon}\right)^{\alpha}\mathrm{d}z\right)^{\frac{\beta}{\alpha}}\mathrm{d}x
\end{aligned}$$

Next we use again Hölder with the pair of Hölder-conjugate exponents $\frac{r}{\beta}$ and $\frac{\alpha}{\beta}$, but this time in the x-integral:

$$\leq C\left(\int_{B_R}\int_{B_1(0)}|f(x+\varepsilon z)|^r\,\mathrm{d}z\,\mathrm{d}x\right)^{\frac{\beta}{r}}\left(\int_{B_R}\int_{B_1(0)}\left(\frac{|c_i(x+\varepsilon z)-c_i(x)|}{\varepsilon}\right)^{\alpha}\mathrm{d}z\,\mathrm{d}x\right)^{\frac{\beta}{\alpha}}$$

By Lemma 2.20 we know, that the first integral is estimated by $\|f\|_{L^r(B_{R+1})}^r$, thus we can take both sides $\frac{1}{\beta}$ to get:

$$\left\| \int_{\mathbb{R}^n} f(y)\,(c_i(y) - c_i(x))\,\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y \right\|_{L^\beta(B_R)}$$
$$\leq C \|f\|_{L^r(B_{R+1})} \left(\int_{B_R} \int_{B_1(0)} \left(\frac{|c_i(x+\varepsilon z) - c_i(x)|}{\varepsilon} \right)^\alpha \mathrm{d}z\,\mathrm{d}x \right)^{\frac{1}{\alpha}}$$

So we estimate the integral using Fubini and the estimate of the L^α-norm of difference quotients against the L^α-norm of the gradient ([9], p. 277):

$$\begin{aligned} \int_{B_R} \int_{B_1(0)} \left(\frac{|c_i(x+\varepsilon z) - c_i(x)|}{\varepsilon} \right)^\alpha \mathrm{d}z\,\mathrm{d}x &= \int_{B_1(0)} |z|^\alpha \int_{B_R} \left(\frac{|c_i(x+\varepsilon z) - c_i(x)|}{\varepsilon |z|} \right)^\alpha \mathrm{d}x\,\mathrm{d}z \\ &\leq C \int_{B_1(0)} |z|^\alpha \|\nabla c\|^\alpha_{L^\alpha(B_{R+1})}\,\mathrm{d}z \\ &\leq C \|\nabla c\|^\alpha_{L^\alpha(B_{R+1})} \end{aligned}$$

So we get

$$\left\| \int_{\mathbb{R}^n} f(y)\big(c_i(y) - c_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y \right\|_{L^\beta(B_R)} \leq C \|f\|_{L^r(B_{R+1})} \|\nabla c\|_{L^\alpha(B_{R+1+C})} \tag{27}$$

Step 2: Reducing the problem to smooth f and c
We want to show

$$\int_{\mathbb{R}^n} f(y)\big(c_i(y) - c_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y \to f \operatorname{div} c$$

in $L^\beta(B_R)$ as functions of x. Here we show that it is enough to show this for smooth f and c. So lets take $\tilde{f}$ and $\tilde{c}$ smooth with $\|\tilde{f} - f\|_r \leq \varepsilon_1$ and $\|\tilde{c} - c\|_{W^{1,\alpha}_{\mathrm{loc}}} \leq \varepsilon_2$. Then we have

$$\begin{aligned} &\left\| \int_{\mathbb{R}^n} f(y)\big(c_i(y) - c_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y - f \operatorname{div} c \right\|_{L^\beta(B_R)} \\ &\leq \left\| \int_{\mathbb{R}^n} \tilde{f}(y)\big(\tilde{c}_i(y) - \tilde{c}_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y - \tilde{f} \operatorname{div} \tilde{c} \right\|_{L^\beta(B_R)} \\ &\quad + \left\| \int_{\mathbb{R}^n} \tilde{f}(y)\big(c_i(y) - \tilde{c}_i(y) + c_i(x) - \tilde{c}_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y \right\|_{L^\beta(B_R)} \\ &\quad + \left\| \int_{\mathbb{R}^n} \big(f(y) - \tilde{f}(y)\big)\big(c_i(y) - c_i(x)\big)\partial_i \rho_\varepsilon(x-y)\,\mathrm{d}y \right\|_{L^\beta(B_R)} \\ &\quad + \left\| \int_{\mathbb{R}^n} f \operatorname{div} c - \tilde{f} \operatorname{div} \tilde{c} \right\|_{L^\beta(B_R)} \end{aligned}$$

But the error terms (the last three integrals) can be estimated by (27), in the order of the terms by

- $C\|f\|_{L^r(B_{R+1})}\|\nabla(c-\tilde{c})\|_{L^\alpha(B_{R+1+C})} \leq C\|f\|_{L^r(B_{R+1})}\varepsilon_2$
- $C\|f-\tilde{f}\|_{L^r(B_{R+1})}\|\nabla c\|_{L^\alpha(B_{R+1+C})} \leq C\varepsilon_1\|\nabla c\|_{L^\alpha(B_{R+1+C})}$
- and the last one, by easy arguments by $C\max(\|f\|_{L^r(B_{R+1})}, \|\nabla c\|_{L^\alpha(B_{R+1+C})})\max(\varepsilon_1, \varepsilon_2)$.

So all terms are controlled by ε_1 or ε_2, so it is enough to show the convergence of the first integral to 0, which is the same as assuming c and f to be smooth for the rest of the proof.

Step 3: Solving the time-independent problem
So now we only have to show

$$\int_{\mathbb{R}^n} f(y)\big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y \to f\,\mathrm{div}\,c$$

in $L^\beta(B_R)$ as functions of x for smooth f and c. At first we want to replace $f(y)$ by $f(x)$, so we use the smoothness of f and c to control

$$\int_{\mathbb{R}^n} \big(f(y)-f(x)\big)\big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y$$
$$\leq \|\nabla f\|_\infty\|\nabla c\|_\infty \int_{\mathbb{R}^n} |x-y|^2\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y \leq C\|\nabla f\|_\infty\|\nabla c\|_\infty \int_{\mathbb{R}^n} \varepsilon^2\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y$$

As $\int_{\mathbb{R}^n} \partial_i\rho_\varepsilon \in \mathcal{O}(\frac{1}{\varepsilon})$ the term converges to 0, so we can ignore it.
So now we have to estimate

$$f(x)\int_{\mathbb{R}^n} \big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y$$

We have:

$$c_i(y)-c_i(x) = \frac{\partial c_i(x)}{\partial x_j}(y_j-x_j) + \mathcal{O}((y_j-x_j)^2) = \frac{\partial c_i(x)}{\partial x_j}(y_j-x_j) + \mathcal{O}(\varepsilon^2)$$

As $\int_{\mathbb{R}}^n \partial_i\rho_\varepsilon \in \mathcal{O}(\frac{1}{\varepsilon})$ we can again ignore the ε^2-term in the following calculation:

$$f(x)\int_{\mathbb{R}^n} \big(c_i(y)-c_i(x)\big)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y = f(x)\int_{\mathbb{R}^n} \frac{\partial c_i(x)}{\partial x_j}(y_j-x_j)\partial_i\rho_\varepsilon(x-y)\,\mathrm{d}y$$

Substituting $z = x-y$ leads to

$$-f(x)\frac{\partial c_i(x)}{\partial x_j}\int_{\mathbb{R}^n} z_j\partial_i\rho_\varepsilon(z)\,\mathrm{d}z$$

A integration by parts leads to

$$\int_{\mathbb{R}^n} z_j\partial_i\rho_\varepsilon(z)\,\mathrm{d}y = -\int_{\mathbb{R}^n} \partial_i(z_j)\rho_\varepsilon(z)\,\mathrm{d}z = -\delta_{ij}$$

So we get

$$f(x)\frac{\partial c_i(x)}{\partial x_j}\delta_{ij} = f(x)\operatorname{div} c(x)$$

Step 4: Including the time-dependency
In the first three steps we have shown $[\rho_\varepsilon, c_i\partial_i](f)(t) \to 0$ in $L^\beta_{\text{loc}}(\mathbb{R}^n)$ for almost every t. Now we want to show

$$\int_0^T \left\|[\rho_\varepsilon, c_i\partial_i](f)(t)\right\|^{\beta_1}_{L^\beta_{\text{loc}}(\mathbb{R}^n)} \to 0$$

As we have convergence against 0 pointwise almost everywhere, we want to use the dominated convergence theorem, so we need a dominating function. But the estimate (27) of Step 1 leads exactly to

$$\left\|[\rho_\varepsilon, c_i\partial_i](f)\right\|^{\beta_1}_{L^\beta_{\text{loc}}(\mathbb{R}^n)}(t) \le C\|f\|^{\beta_1}_{L^r_{\text{loc}}(\mathbb{R}^n)}(t)\|\nabla c\|^{\beta_1}_{L^\alpha_{\text{loc}}(\mathbb{R}^n)}(t)$$

As $\|f\|_{L^r} \in L^{r_1}([0,T])$ and $\|\nabla c\|_{L^\alpha} \in L^{\alpha_1}([0,T])$ and $\frac{1}{\beta_1} = \frac{1}{r_1} + \frac{1}{\alpha_1}$, this is in fact an integrable function.

Step 5: The second commutator (25)
Here we can argue more directly:

$$[\rho_\varepsilon, g](f) = (gf) * \rho_\varepsilon - g(f_\varepsilon)$$

As gf is in $L^\beta_{\text{loc}}(\mathbb{R}^n)$, by the properties of the mollification we know, that the first term goes to gf in $L^\beta_{\text{loc}}(\mathbb{R}^n)$. So does the second term, as the following calculation leads to (using Hölder with the pair $\frac{\alpha}{\beta}$ and $\frac{r}{\beta}$):

$$\|gf - gf_\varepsilon\|_{L^\beta(B_R)} \le \|g\|_{L^\alpha(B_R)}\|f - f_\varepsilon\|_{L^r(B_R)}$$

which goes also to 0 by basic properties of the mollification. The time dependency is included analog with the dominating function $2\|f\|_{L^r_{\text{loc}}(\mathbb{R}^n)}\|g\|_{L^\alpha_{\text{loc}}(\mathbb{R}^n)}$ □

Remark 3.15. We proved Lemma 3.14 only for all exponents $< \infty$. The lemma is also valid for exponents being ∞, and the proof can be easily adapted. As we only use it for $r = r_1 = \infty$, we only consider this case in detail (so $\alpha = \beta$ and $\alpha_1 = \beta_1$). The proof for $\alpha = \beta = 1$ can also be found in a slightly different form in [6], Lemma 2.2.
Again we do at first the time-independent case and include the time-dependency later. With the same steps as in Step 1 of the Lemma 3.14 we need to show

$$\int_{B_R}\left|\int_{B_1(0)} f(x+\varepsilon z)\left(\frac{c_i(x+\varepsilon z) - c_i(x)}{\varepsilon}\right)\partial_i\rho(z)\,\mathrm{d}z + f(x)\operatorname{div} c(x)\right|^\beta \mathrm{d}x \to 0$$

The different sign compared to the calculations above is from the transformation from y to z in the derivative of ρ. Above there was no difference because we only considered absolute values.
Now we use dominated convergence: f and ρ are bounded, and the sequence of difference quotients of c converges in $L^\beta(B_R)$ to $\partial_z c_i$ (see [13], p.182, theorem 9.1.1 for the whole space, this can be easily adapted to B_R as the proof works just by an approximation with smooth functions), thus we can find an almost everywhere convergent subsequence. $f(x+\varepsilon z)$ is the translation, and it is well known, that the translation converges in L^p, so we can extract a subsequence converging pointwise almost everywhere to f (so formally we use the fact, that a sequence converges if we can extract a subsequence of every given subsequence converging to the same limit). So we use dominated convergence and get

$$\int_{B_R}\left|\int_{B_1(0)} f(x)\partial_z c_i(x)\partial_i\rho(z)\,\mathrm{d}z + f(x)\operatorname{div} c(x)\right|^\beta \mathrm{d}x$$
$$= \int_{B_R}\left|f(x)\partial_j c_i(x)\int_{B_1(0)} z_j\partial_i\rho(z)\,\mathrm{d}z + f(x)\operatorname{div} c(x)\right|^\beta \mathrm{d}x$$

As before we have $\int_{B_1(0)} z_j\partial_i\rho(z)\,\mathrm{d}z = -\delta_{ij}$, so the above term is 0.
Including the time-dependency can be done with exactly the same argument as in Step 4.
So only the second commutator (25) has to be done in the case $r_1 = r = \infty$:

$$\begin{aligned}\int_{B_R}\left|[\rho_\varepsilon, g](f)(x)\right|^\beta \mathrm{d}x &= \int_{B_R}\left|\int_{B_\varepsilon(x)}\big(g(x)-g(y)\big)f(y)\rho_\varepsilon(x-y)\,\mathrm{d}y\right|^\beta \mathrm{d}x\\ &\le \|f\|^\beta_{L^\infty(B_{R+1})}\int_{B_R}\left|\int_{B_\varepsilon(x)}\big(g(x)-g(y)\big)\rho_\varepsilon(x-y)\,\mathrm{d}y\right|^\beta \mathrm{d}x\\ &= \|f\|^\beta_{L^\infty(B_{R+1})}\int_{B_R}\left|g(x)-g_\varepsilon(x)\right|^\beta \mathrm{d}x\end{aligned}$$

This converges to zero as $\varepsilon \to 0$ by basic properties of the mollification. The time dependency is again included analogously.

With Lemma 3.14 we can control the error terms in (23) except $Q_{2,\varepsilon}$:

Lemma 3.16. *As $\varepsilon \to 0$ the terms $Q_{1,\varepsilon}, S_\varepsilon, T_\varepsilon$ converge strongly to 0 in $L^1\big([0,T], L^1_{\text{loc}}(\mathbb{R}^n)\big)$, the term $R_{k,\varepsilon}$ converges to 0 even in $L^2\big([0,T], L^2_{\text{loc}}(\mathbb{R}^n)\big)$.*

Proof. This is just a trivial consequence of Lemma 3.14 with the regularity assumptions:

- For $Q_{1,\varepsilon} = [\rho_\varepsilon, \operatorname{div}(b)](p)$ we use the second limit (25) with $\alpha_1 = \alpha = \beta_1 = \beta = 1$ and $r_1 = r = \infty$

- For $S_\varepsilon = [\rho_\varepsilon, \partial_i \sigma_{ik}](\sigma_{jk}\partial_j p)$ we use the second limit (25) with $\alpha_1 = \alpha = r_1 = r = 2$ and $\beta_1 = \beta = 1$
- For $T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j p)$ we use the first limit (24) with $\alpha_1 = \alpha = r_1 = r = 2$ and $\beta_1 = \beta = 1$
- For $R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jk}\partial_j](p)$ we use the first limit (24) with $\alpha_1 = \alpha = \beta_1 = \beta = 2$ and $r_1 = r = \infty$

□

Lemma 3.17. *We have $R_{k,\varepsilon} \in L^2([0,T], W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n))$, so it has first order Sobolev regularity with respect to the spatial variables.*

Proof. (see [14], Remark 2.2 for a more general version) We have

$$\begin{aligned} R_{k,\varepsilon}(x) &= \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\sigma_{jk}(y)\partial_j p(y) - \sigma_{jk}(x)\partial_j\rho_\varepsilon(x-y)p(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \partial_j\rho_\varepsilon(x-y)\sigma_{jk}(y)p(y) - \rho_\varepsilon(x-y)\partial_j\sigma_{jk}(y)p(y) \\ &\quad - \sigma_{jk}(x)\partial_j\rho_\varepsilon(x-y)p(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \partial_j\rho_\varepsilon(x-y)\big(\sigma_{jk}(y)-\sigma_{jk}(x)\big)p(y)\,\mathrm{d}y - \big((p\cdot\partial_j\sigma_{jk})*\rho_\varepsilon\big)(x) \end{aligned}$$

These expressions are all well defined functions, so it is in fact well defined to consider $R_{k,\varepsilon}$ as a function and not only as a distribution. The term $(p\cdot\partial_j\sigma_{jk})*\rho_\varepsilon$ is smooth, so we check the regularity of the integral:

$$\int_{\mathbb{R}^n} \partial_j\rho_\varepsilon(x-y)\big(\sigma_{jk}(y)-\sigma_{jk}(x)\big)p(y)\,\mathrm{d}y = \big((p\sigma_{jk})*\partial_j\rho_\varepsilon\big)(x) - \sigma_{jk}(x)\cdot p*\partial_j\rho_\varepsilon(x)$$

$(p\sigma_{jk})*\partial_j\rho_\varepsilon$ and $p*\partial_j\rho_\varepsilon$ are again smooth, so as $\sigma \in \Big(L^2\big([0,T], W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)\big)\Big)^{n\times m}$ has Sobolev regularity in space also $R_{k,\varepsilon} \in L^2\big([0,T], W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)\big)$. □

For simplification of notation we denote

$$U_\varepsilon = -Q_{1,\varepsilon} + \frac{1}{2}(S_\varepsilon + T_\varepsilon)$$

So we have $U_\varepsilon \to 0$ in $L^1\left([0,T], L^1_{\mathrm{loc}}(\mathbb{R}^n)\right)$ and (23) becomes

$$\partial_t p_\varepsilon + \partial_i(p_\varepsilon b_i) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon) = U_\varepsilon - Q_{2,\varepsilon} + \frac{1}{2}\partial_i(\sigma_{ik}R_{k,\varepsilon}) \tag{28}$$

3.3.4 The commutator estimate of Ambrosio

For this proof we will need another commutator estimate to deal with $Q_{2,\varepsilon}$. Therefore we define M_t by $D^s b_t = M_t |D^s b_t|$

Theorem 3.18. *For a $n \times n$-Matrix M and $\rho \in C_c^\infty(\mathbb{R}^n)$ we define*

$$\Lambda(M,\rho) = \int_{\mathbb{R}^n} \left| \left\langle Mz, \nabla\rho(z) \right\rangle \right| \mathrm{d}z$$
$$I(\rho) = \int_{\mathbb{R}^n} |z| \cdot |\nabla\rho(z)| \, \mathrm{d}z$$

Then for any compact $K \subset (0,T) \times \mathbb{R}^n$ we get

$$\limsup_{\varepsilon\to 0} \int_K |Q_{2,\varepsilon}| \, \mathrm{d}x \, \mathrm{d}t \le \|p\|_\infty \int_K \Lambda(M_t(x),\rho) \, \mathrm{d}|D^s b|(t,x) + \|p\|_\infty (n + I(\rho)) |D^a b|(K) \tag{29}$$

and

$$\limsup_{\varepsilon\to 0} \int_K |Q_{2,\varepsilon}| \, \mathrm{d}x \, \mathrm{d}t \le \|p\|_\infty I(\rho) |D^s b|(K) \tag{30}$$

Remark 3.19. $D^s b$ is a measure on K in the following way: Defining $A_t = \{x \in \mathbb{R}^n : (x,t) \in A\}$ for $A \subset K$ we have

$$D^s b(A) = \int_0^T D^s b_t(A_t) \, \mathrm{d}t$$

This is in fact well defined as $b \in \left(L^1\left([0,T], BV_{\mathrm{loc}}(\mathbb{R}^n)\right)\right)^n$. $D^a b$ is defined analogously.

Proof of Theorem 3.18. ([2], Theorem 3.2) At first we derive an identity for $Q_{2,\varepsilon}$. Mind, that the distribution $b_i \partial_i p$ is defined as following (for a test function φ and as usual ignoring the time-dependency):

$$b_i \partial_i p(\varphi) = -\int_{\mathbb{R}^n} p b_i \partial_i \varphi - \int_{\mathbb{R}^n} p \operatorname{div} b \varphi$$

With this and the definitions of the mollification of distributions of section 2.3 and the evenness of ρ in mind we do the following calculation:

$$\begin{aligned}
Q_{2,\varepsilon}(\varphi) &= [\rho_\varepsilon, b_i\partial_i](p)(\varphi) \\
&= \rho_\varepsilon * (b_i\partial_i p)(\varphi) - b_i\partial_i p_\varepsilon(\varphi) = b_i\partial_i p(\rho_\varepsilon * \varphi) - b_i\partial_i p_\varepsilon(\varphi) \\
&= -\int_{\mathbb{R}^n} b_i(x)p(x)\partial_i(\rho_\varepsilon * \varphi)(x) + p(x)\operatorname{div} b(x)\rho_\varepsilon * \varphi(x) \\
&\quad + b_i(x)\partial_i p_\varepsilon(x)\varphi(x)\,\mathrm{d}x \\
&= -\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} b_i(x)p(x)\partial_i\rho_\varepsilon(x-y)\varphi(y) + p(x)\operatorname{div} b(x)\rho_\varepsilon(x-y)\varphi(y) \\
&\quad + b_i(x)\partial_i\rho_\varepsilon(x-y)p(y)\varphi(x)\,\mathrm{d}y\,\mathrm{d}x \\
&= -\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} -b_i(y)p(y)\partial_i\rho_\varepsilon(x-y)\varphi(x) + p(y)\operatorname{div} b(y)\rho_\varepsilon(x-y)\varphi(x) \\
&\quad + b_i(x)\partial_i\rho_\varepsilon(x-y)p(y)\varphi(x)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} p(y)\big(b_i(y)-b_i(x)\big)\partial_i\rho_\varepsilon(x-y)\varphi(x)\,\mathrm{d}y\,\mathrm{d}x \\
&\quad - \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} p(y)\operatorname{div} b(y)\rho_\varepsilon(x-y)\varphi(x)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} p(y)\big(b_i(y)-b_i(x)\big)\partial_i\rho_\varepsilon(x-y)\varphi(x)\,\mathrm{d}y\,\mathrm{d}x - p\operatorname{div} b * \rho_\varepsilon(\varphi)
\end{aligned}$$

In the fifth $=$ we switched x and y in the first two integrals and used the evenness of ρ (and hence the in-evenness of $\partial_i\rho$). So $Q_{2,\varepsilon}$ is in fact represented by a function. Changing variables $y = x - \varepsilon z$ leads to:

$$Q_{2,\varepsilon}(x) = \int_{\mathbb{R}^n} p(x-\varepsilon z)\left(\frac{b(x-\varepsilon z)-b(x)}{\varepsilon}\right)\nabla\rho(z)\,\mathrm{d}z - p\operatorname{div} b * \rho_\varepsilon \tag{31}$$

So now we consider $\limsup_{\varepsilon\to 0}\int_K |Q_{2,\varepsilon}|\,\mathrm{d}x\,\mathrm{d}t$ using the decomposition of the difference quotient of the BV-function b of Lemma 2.5 into b^1_ε and b^2_ε (by slightly relabeling $b^1_\varepsilon(t,x,z) = (b_t)^1_\varepsilon(-z)(x)$ and analog with $b^2_\varepsilon(t,x,z)$) and defining $\tilde{p}_\varepsilon(t,x,z) = p(t,x-$

εz) (from now on we also keep again track of the time-dependency)

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\int_K |Q_{2,\varepsilon}|\,\mathrm{d}x\,\mathrm{d}t\\
&\qquad = \limsup_{\varepsilon\to 0}\int_K \left|\int_{\mathbb{R}^n} \tilde p_\varepsilon (b^1_\varepsilon + b^2_\varepsilon)\nabla\rho(z)\,\mathrm{d}z - p(t,x)\operatorname{div} b(t,x) * \rho_\varepsilon\right|\mathrm{d}x\,\mathrm{d}t\\
&\qquad \le \limsup_{\varepsilon\to 0}\int_K \left|\int_{\mathbb{R}^n} \tilde p_\varepsilon b^1_\varepsilon \nabla\rho(z)\,\mathrm{d}z - p(t,x)\operatorname{div} b(t,x)\right|\mathrm{d}x\,\mathrm{d}t\\
&\qquad\quad + \limsup_{\varepsilon\to 0}\int_K |p(t,x)\operatorname{div} b(t,x) * \rho_\varepsilon - p(t,x)\operatorname{div} b(t,x)|\,\mathrm{d}t\,\mathrm{d}x\\
&\qquad\quad + \limsup_{\varepsilon\to 0}\|p\|_\infty \int_{\mathbb{R}^n} |\nabla\rho(z)| \int_K |b^2_\varepsilon(t,x,z)|\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}z
\end{aligned}$$

We estimate these three limes superior to prove (30) first.
The second lim sup is 0 as convolutions are converging in L^1 in space and then dominated convergence in t with dominating function $2\|p\|_\infty \|\operatorname{div}(x)\|_1(\cdot)$ in t.
The third lim sup can be estimated as following using Lemma 2.5:

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\|p\|_\infty \int_{\mathbb{R}^n} |\nabla\rho(z)| \int_K |b^2_\varepsilon(t,x,z)|\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}z\\
&\qquad\qquad\qquad \le \|p\|_\infty \int_{\mathbb{R}^n} |\nabla\rho(z)| \int_0^T |z||D^s b_t(K_t)|\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}z
\end{aligned}$$

With Remark 3.19 and the definition of I this is estimated by $\|p\|_\infty I(\rho)|D^s b|(K)$, so only the first integral remains and it remains to show

$$\limsup_{\varepsilon\to 0}\int_K \left|\int_{\mathbb{R}^n} \tilde p_\varepsilon(t,x,z) b^1_\varepsilon(t,x,z)\nabla\rho(z)\,\mathrm{d}z - p(t,x)\operatorname{div} b(t,x)\right|\mathrm{d}x\,\mathrm{d}t = 0$$

At first we leave t fixed and consider the functions $\tilde p_\varepsilon(t,\cdot,\cdot)b^1_\varepsilon(t,\cdot,\cdot)$ as functions in $L^1_{\text{loc}}(\mathbb{R}^n\times\mathbb{R}^n)$. We have convergence in $L^1_{\text{loc}}(\mathbb{R}^n\times\mathbb{R}^n)$ to $p(x)\nabla b_t(x)(-z)$, because (again leaving away the time dependency as t is fixed for the moment)

$$\begin{aligned}
&\left\|\tilde p_\varepsilon(x,z)b^1_\varepsilon(x,z) - p(x)\nabla b_t(x)(-z)\right\|_1\\
&\qquad \le \left\|\big(\tilde p_\varepsilon(x,z) - p(x)\big)\nabla b_t(x)(-z)\right\|_1 + \left\|\tilde p_\varepsilon(x,z)\big(b^1_\varepsilon(x,z) - \nabla b_t(x)(-z)\big)\right\|_1
\end{aligned}$$

In the first norm we apply dominated convergence: We have convergence pointwise to 0 almost everywhere and, as p and $\tilde p_\varepsilon$ are both bounded by $\|p\|_\infty$ (mind that $\tilde p_\varepsilon$ is only a translation of p), we have the dominating function $2\|p\|_\infty \nabla b_t(x)(-z)$.
The second norm also converges to 0 as $\tilde p_\varepsilon$ is bounded again and
$\|(b^1_\varepsilon(x,z) - \nabla b_t(x)(-z))\|_1 \to 0$ by Lemma 2.5.
Of course multiplying with the in ε constant, bounded function $\nabla\rho$ does not change

this convergence, and the integration in t is included analogously with dominated convergence as Lemma 2.5 gives a uniform bound, which is also bounded by an L^1-function in t. So it just remains to show:

$$\int_K \left| p(t,x)\left(\left(\int_{\mathbb{R}^n} \frac{\partial b_t^j}{\partial x_i}(x) z_i \frac{\partial \rho}{\partial z_j}(z)\,\mathrm{d}z + \operatorname{div} b(t,x)\right)\right| \mathrm{d}x\,\mathrm{d}t = 0$$

But this is clear by $\int z_i \frac{\partial \rho}{\partial z_j}\,\mathrm{d}z = -\delta_{ij}$ by a partial integration. So (30) is shown.

Next we show (29): We start with (31), so we can estimate:

$$\|Q_{2,\varepsilon}\|_{L^1(K)} \leq \|p\|_\infty \int_{\mathbb{R}^n}\int_K \left|\frac{b(t,x-\varepsilon z)-b(t,x)}{\varepsilon}\right| \nabla\rho(z)\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}z + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K_\varepsilon)} \tag{32}$$

with $K_\varepsilon = \{x \in \mathbb{R}^n | \operatorname{dist}(x,K) < \varepsilon\}$ the ε-neighborhood of K.
We fix $z \in \mathbb{R}^n$ and define $w_{tz}(x) = b(t,x)\nabla\rho(z)$. We estimate the first integral of (32) using Lemma 2.2 and $K_t = \{x \in \mathbb{R}^n | (x,t) \in K\}$ and $K_{t\,\varepsilon|z|}$ the $\varepsilon|z|$ neighborhood of K_t:

$$\begin{aligned}\int_K \frac{1}{\varepsilon}|w_{tz}(x-\varepsilon z) - w_{tz}(x)|\,\mathrm{d}x\,\mathrm{d}t &= \int_0^T \int_{K_t} \frac{1}{\varepsilon}|w_{tz}(x-\varepsilon z) - w_{tz}(x)|\,\mathrm{d}x\,\mathrm{d}t \\ &\leq \int_0^T \frac{1}{\varepsilon}\left|\sum_{i=1}^n \varepsilon z_i D_i w_{tz}\right|(K_{t\,\varepsilon|z|})\,\mathrm{d}t \\ &\leq \int_0^T \left|\sum_{i=1}^n z_i D_i w_{tz}\right|(K_{t\,\varepsilon|z|})\,\mathrm{d}t\end{aligned}$$

So we are interested in the distributional derivative of w_{tz}, which is given by:

$$D_i w_{tz} = \frac{\partial b_t^j}{\partial x_i}\frac{\partial \rho}{\partial z_j}(z)\mathscr{L}^n + (M_t)_{ji}\frac{\partial \rho}{\partial z_j}|D^s b_t| \tag{33}$$

So (32) leads to:

$$\begin{aligned}&\limsup_{\varepsilon\to 0} \|Q_{2,\varepsilon}\|_{L^1(K)} \\ &\quad\leq \limsup_{\varepsilon\to 0} \|p\|_\infty \int_{\mathbb{R}^n}\int_0^T \left|\sum_{i=1}^n z_i D_i w_{tz}\right|(K_{t\,\varepsilon|z|})\,\mathrm{d}t\,\mathrm{d}z + \limsup_{\varepsilon\to 0}\|p\|_\infty \|\operatorname{div} b\|_{L^1(K_\varepsilon)}\end{aligned}$$

By basic properties of measure theory $\limsup_{\varepsilon\to 0}\|p\|_\infty\|\operatorname{div} b\|_{L^1(K_\varepsilon)} = \|p\|_\infty\|\operatorname{div} b\|_{L^1(K)}$.
Also in the first integral we can apply dominated convergence, as the integral in z is in fact only on the support of ρ and not on the whole $\mathbb{R}^n$, thus we have an integrable

dominating function as $w_t z \in BV_{\text{loc}}$ with a norm which is bounded in z. Thus we can replace $K_{t\varepsilon|z|}$ by K_t. So we have, using (33)

$$
\begin{aligned}
&\limsup_{\varepsilon\to 0} \|Q_{2,\varepsilon}\|_{L^1(K)} \\
&\leq \|p\|_\infty \int_{\mathbb{R}^n}\int_0^T \left| z_i \left(\frac{\partial b_t^j}{\partial x_i}\frac{\partial \rho}{\partial z_j}(z)\mathscr{L}^n + (M_t)_{ji}\frac{\partial \rho}{\partial z_j}|D^s b_t| \right) \right| (K_t)\,\mathrm{d}t\,\mathrm{d}z + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K)} \\
&\leq \|p\|_\infty \int_{\mathbb{R}^n}\int_0^T\int_{K_t} \left| z_i \frac{\partial b_t^j}{\partial x_i}\frac{\partial \rho}{\partial z_j}(z) \right| \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}z \\
&\quad + \|p\|_\infty \int_{\mathbb{R}^n}\int_0^T\int_{K_t} \left| z_i (M_t)_{ji}\frac{\partial \rho}{\partial z_j} \right| \mathrm{d}|D^s b_t|(x)\,\mathrm{d}t\,\mathrm{d}z + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K)} \\
&\leq \|p\|_\infty \int_{\mathbb{R}^n}\int_K |z||\nabla b_t||\nabla \rho(z)|\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}z + \|p\|_\infty \int_{\mathbb{R}^n}\int_K \left|\langle M_t(x) z, \nabla\rho(z)\rangle\right| \mathrm{d}|D^s b|(t,x)\,\mathrm{d}z \\
&\quad + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K)} \\
&= \|p\|_\infty I(\rho)\int_K |\nabla b_t(x)|\,\mathrm{d}x\,\mathrm{d}t + \|p\|_\infty \int_K \Lambda(M_t(x),\rho)\,\mathrm{d}|D^s b|(t,x) + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K)} \\
&= \|p\|_\infty \int_K \Lambda(M_t(x),\rho)\,\mathrm{d}|D^s b|(t,x) + \|p\|_\infty I(\rho)|D^a b|(K) + \|p\|_\infty \|\operatorname{div} b\|_{L^1(K)}
\end{aligned}
$$

As $\|\operatorname{div} b\|_{L^1(K)} \leq n|D^a b|(K)$ (because $\frac{\partial b_j}{x_i}$ is the density of $D^a b$ with respect to the Lebesgue-measure), the proof is finished. □

3.3.5 Proof of the renormalization assumption

Now we have all the tools to prove the renormalization property:
p_ε is smooth in the spatial variables and by (28) (using the regularity assumptions on b and σ and of Lemma 3.16) we know that $\partial_t p_\varepsilon \in L^1_{\text{loc}}\big([0,T]\times\mathbb{R}^n\big)$, as $R_{k,\varepsilon} \in L^2\big([0,T], W^{1,2}_{\text{loc}}(\mathbb{R}^n)\big)$ by Lemma 3.17. So $p_\varepsilon \in W^{1,1}_{\text{loc}}([0,T]\times\mathbb{R}^n)$. Hence we can do the calculations at the beginning of section 3.3.1 now rigorously for p_ε instead of p using

the chain rule for Sobolev-functions:

$$
\begin{aligned}
&\partial_t \beta(p_\varepsilon) + \partial_i(\beta(p_\varepsilon) b_i) - \frac{1}{2}\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j \beta(p_\varepsilon)\big) \\
&= \beta'(p_\varepsilon)\cdot \partial_t p_\varepsilon + \beta'(p_\varepsilon)\partial_i p_\varepsilon b_i + \beta(p_\varepsilon)\operatorname{div}(b) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\beta'(p_\varepsilon)\partial_j p_\varepsilon) \\
&= \beta'(p_\varepsilon)\cdot \partial_t p_\varepsilon + \beta'(p_\varepsilon)\partial_i p_\varepsilon b_i + \beta(p_\varepsilon)\operatorname{div}(b) - \frac{1}{2}\beta'(p_\varepsilon)\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon) \\
&\quad - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon \beta''(p_\varepsilon)\partial_i p_\varepsilon \\
&= \beta'(p_\varepsilon)\left(\partial_t p_\varepsilon + \partial_i p_\varepsilon b_i - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon)\right) + \beta(p_\varepsilon)\operatorname{div}(b) - \frac{1}{2}\beta''(p_\varepsilon)\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon \partial_i p_\varepsilon \\
&= \beta'(p_\varepsilon)\underbrace{\left(\partial_t p_\varepsilon + \partial_i p_\varepsilon b_i + p_\varepsilon \operatorname{div} b - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j p_\varepsilon)\right)}_{=U_\varepsilon - Q_{2,\varepsilon} + \frac{1}{2}\partial_i(\sigma_{ik}R_{k,\varepsilon}) \text{ by (28)}} - \beta'(p_\varepsilon)p_\varepsilon \operatorname{div}(b) + \beta(p_\varepsilon)\operatorname{div}(b) \\
&\quad - \frac{1}{2}\beta''(p_\varepsilon)|\sigma^*\nabla p_\varepsilon|^2
\end{aligned}
$$

So we have in a distributional sense:

$$
\begin{aligned}
\partial_t \beta(p_\varepsilon) + \partial_i(\beta(p_\varepsilon) b_i) - \frac{1}{2}\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j \beta(p_\varepsilon)\big) - \big(\beta(p_\varepsilon) - p_\varepsilon \beta'(p_\varepsilon)\big)\operatorname{div}(b) \\
+ \frac{1}{2}\beta''(p_\varepsilon)|\sigma^*\nabla p_\varepsilon|^2 = \beta'(p_\varepsilon)\left(U_\varepsilon - Q_{2,\varepsilon} + \frac{1}{2}\partial_i(\sigma_{ik}R_{k,\varepsilon})\right) \quad (34)
\end{aligned}
$$

Theorem 3.20. *Let b,σ be as in Theorem 3.2. Then a weak solution in the sense of definition 3.1 is always a renormalized solution.*

Proof. [14], Theorem 2.5: We take (34) and let $\varepsilon \to 0$, so we need to check the behavior of all terms in (34) as distributions. For the argumensts we will often need Lemma 2.10, and, to use this, that $\beta(p_\varepsilon) - \beta(p)$ converges to 0 and is uniformly bounded in ε. This holds because $p_\varepsilon \to p$ in measure (L^p-convergence implies convergence in measure) and Lemma 2.9, as p_ε are uniformly bounded by $\|p\|_\infty$ and $\beta \in C^2(\mathbb{R})$, so the analog statement also holds for $\beta'(p_\varepsilon)$ and $\beta''(p_\varepsilon)$:

Step 1: $\partial_t \beta(p_\varepsilon) \to \partial_t \beta(p)$
So we have to show

$$\int_0^T \int_{\mathbb{R}^n} \big(\beta(p_\varepsilon) - \beta(p)\big)\cdot \partial_t \varphi \to 0$$

and

$$\int_{\mathbb{R}^n} \big(\beta(p_\varepsilon)|_{t=0} - \beta(p)|_{t=0}\big)\varphi|_{t=0} \to 0$$

for a test function φ. This is both clear by Lemma 2.10, for the first integral we have the convergence at first pointwise for a fixed t and then by the dominated convergence theorem, as $\beta(p_\varepsilon)-\beta(p)$ is also bounded in t, so we have an integrable dominating function.

Step 2: $\partial_i(\beta(p_\varepsilon)b_i) \to \partial_i(\beta(p)b_i)$

So

$$\int_0^T \int_{\mathbb{R}^n} \big(\beta(p_\varepsilon)-\beta(p)\big)\cdot b_i \cdot \partial_i\varphi \to 0$$

which is clear again by Lemma 2.10 and dominated convergence in t, as $b \in L^1_{\text{loc}}$, so $b\cdot\nabla\varphi \in L^1$.

Step 3: $\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p_\varepsilon)\big) \to \partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p)\big)$
We have (by using $\sigma^*\nabla\beta(p)=\beta'(p)\cdot\sigma^*\nabla p$ by Lemma 3.6)

$$\begin{aligned}
-\int_0^T \int_{\mathbb{R}^n} &\varphi\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p_\varepsilon)\big)-\varphi\partial_i\big(\sigma_{ik}\sigma_{jk}\partial_j\beta(p_\varepsilon)\big)\\
&= \int_0^T \int_{\mathbb{R}^n} \big\langle \sigma^*\nabla\varphi, \sigma^*\nabla\beta(p_\varepsilon)-\sigma^*\nabla\beta(p)\big\rangle\\
&= \int_0^T \int_{\mathbb{R}^n} \big\langle \sigma^*\nabla\varphi, \beta'(p_\varepsilon)\sigma^*\nabla p_\varepsilon-\beta(p)\sigma^*\nabla p\big\rangle\\
&= \int_0^T \int_{\mathbb{R}^n} \big(\beta'(p_\varepsilon)+\beta'(p)\big)\cdot\langle \sigma^*\nabla\varphi, \sigma^*\nabla p_\varepsilon-\sigma^*\nabla p\rangle\\
&\quad+\int_0^T \int_{\mathbb{R}^n} \beta'(p_\varepsilon)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p\rangle\\
&\quad-\int_0^T \int_{\mathbb{R}^n} \beta'(p)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p_\varepsilon\rangle
\end{aligned}$$

So again we have to check the three integrals:

- $\int_0^T\int_{\mathbb{R}^n}\big(\beta'(p_\varepsilon)+\beta'(p)\big)\cdot\langle \sigma^*\nabla\varphi, \sigma^*\nabla p_\varepsilon-\sigma^*\nabla p\rangle \to 0$ by using the uniform boundedness of $\beta'(p_\varepsilon)+\beta'(p)$, the Cauchy-Schwarz inequality and Lemma 3.4, by which $\sigma^*\nabla p_\varepsilon-\sigma^*\nabla p\to 0$ in $L^2\big([0,T],L^2_{\text{loc}}\big)$.
- $\int_0^T\int_{\mathbb{R}^n}\beta'(p_\varepsilon)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p\rangle \to \int_0^T\int_{\mathbb{R}^n}\beta'(p)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p\rangle$ by Lemma 2.10 and dominated convergence in t
- $\int_0^T\int_{\mathbb{R}^n}\beta'(p)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p_\varepsilon\rangle \to \int_0^T\int_{\mathbb{R}^n}\beta'(p)\,\langle \sigma^*\nabla\varphi,\sigma^*\nabla p\rangle$ by the strong convergence of Lemma 3.4

Step 4: $\beta(p_\varepsilon)\operatorname{div}(b) \to \beta(p)\operatorname{div}(b)$
So

$$\int_0^T \int_{\mathbb{R}^n} \big(\beta(p_\varepsilon) - \beta(p)\big) \cdot \operatorname{div}(b)\varphi \to 0$$

which is again clear by Lemma 2.10 and dominated convergence in t, as $\operatorname{div}(b) \in L^1_{\text{loc}}$, so $\operatorname{div}(b)\varphi \in L^1$.

Step 5: $p_\varepsilon\beta'(p_\varepsilon)\operatorname{div}(b) \to p\beta'(p)\operatorname{div}(b)$

Here we have to show

$$\int_0^T \int_{\mathbb{R}^n} \big(p_\varepsilon\beta(p_\varepsilon) - p\beta(p)\big) \cdot \operatorname{div}(b)\varphi \to 0$$

Again, we reduce it to the spatial problem and use dominated convergence, so it suffices to show

$$\int_{\mathbb{R}^n} \big(p_\varepsilon\beta(p_\varepsilon) - p\beta(p)\big) \cdot \operatorname{div}(b)\varphi \to 0$$

for fixed t. We have

$$\begin{aligned}\int_{\mathbb{R}^n} \big(p_\varepsilon\beta(p_\varepsilon) - p\beta(p)\big) \cdot \operatorname{div}(b)\varphi = &\int_{\mathbb{R}^n} (p_\varepsilon - p)\big(\beta(p_\varepsilon) + \beta(p)\big) \cdot \operatorname{div}(b)\varphi \\ &+ \int_{\mathbb{R}^n} p\beta'(p_\varepsilon)\operatorname{div}(b)\varphi \\ &- \int_{\mathbb{R}^n} p_\varepsilon\beta'(p)\operatorname{div}(b)\cdot\varphi\end{aligned}$$

So we check the three integrals:

- $\int_{\mathbb{R}^n} (p_\varepsilon - p)\big(\beta(p_\varepsilon)+\beta(p)\big)\cdot\operatorname{div}(b)\varphi \le \|\beta'(p_\varepsilon)+\beta'(p)\|_\infty \int_{\operatorname{spt}\varphi} |p_\varepsilon - p|\operatorname{div}(b) \to 0$ according to the dominated convergence theorem, as $p_\varepsilon \to p$ pointwise almost everywhere and as p_ε is bounded
- $\int_{\mathbb{R}^n} p\beta'(p_\varepsilon)\operatorname{div}(b)\varphi \to \int_{\mathbb{R}^n} p\beta'(p)\operatorname{div}(b)\varphi$ as in Step 4 after taking $\|p\|_\infty$ out of the integral
- $\int_{\mathbb{R}^n} p_\varepsilon\beta'(p)\operatorname{div}(b)\cdot\varphi \to \int_{\mathbb{R}^n} p\beta'(p)\operatorname{div}(b)\cdot\varphi$ again with dominated convergence.

So the sum converges to 0, which was to show.

Step 6: $\beta''(p_\varepsilon)|\sigma^*\nabla p_\varepsilon|^2 \to \beta''(p)|\sigma^*\nabla p|^2$

So we have to show:

$$\int_0^T \int_{\mathbb{R}^n} \left(\beta''(p_\varepsilon)|\sigma^*\nabla p_\varepsilon|^2 - \beta''(p)|\sigma^*\nabla p|^2\right)\varphi \to 0$$

Again we reduce it to a spatial problem and estimate:

$$\begin{aligned}\int_{\mathbb{R}^n} \left(\beta''(p_\varepsilon)|\sigma^*\nabla p_\varepsilon|^2 - \beta''(p)|\sigma^*\nabla p|^2\right)\varphi &= \int_{\mathbb{R}^n} \left(\beta''(p_\varepsilon) + \beta''(p)\right)\left(|\sigma^*\nabla p_\varepsilon|^2 - |\sigma^*\nabla p|^2\right)\varphi \\ &+ \int_{\mathbb{R}^n} \beta''(p_\varepsilon)|\sigma^*\nabla p|^2\varphi \\ &- \int_{\mathbb{R}^n} \beta''(p)|\sigma^*\nabla p_\varepsilon|^2\varphi\end{aligned}$$

In the three integrals we have

- $\int_{\mathbb{R}^n}\left(\beta''(p_\varepsilon) + \beta''(p)\right)\left(|\sigma^*\nabla p_\varepsilon|^2 - |\sigma^*\nabla p|^2\right)\varphi \to 0$ by the strong convergence of $|\sigma^*\nabla p_\varepsilon|^2$ to $|\sigma^*\nabla p|^2$ by Lemma 3.4 and as $\beta''(p_\varepsilon) + \beta''(p)$ and φ are bounded.
- $\int_{\mathbb{R}^n} \beta''(p_\varepsilon)|\sigma^*\nabla p|^2\varphi \to \int_{\mathbb{R}^n} \beta''(p)|\sigma^*\nabla p|^2\varphi$ by Lemma 2.10
- $\int_{\mathbb{R}^n} \beta''(p)|\sigma^*\nabla p_\varepsilon|^2\varphi \to \int_{\mathbb{R}^n} \beta''(p)|\sigma^*\nabla p|^2\varphi$ by the strong convergence of $|\sigma^*\nabla p_\varepsilon|^2$ to $|\sigma^*\nabla p|^2$ by Lemma 3.4

Step 7: $\beta'(p_\varepsilon)U_\varepsilon \to 0$
We have

$$\int_0^T \int_{\mathbb{R}^n} \beta'(p_\varepsilon)U_\varepsilon\varphi \le \sup_{\varepsilon>0} \|\beta'(p_\varepsilon)\|_\infty \int_0^T \int_{\operatorname{spt}\varphi} |U_\varepsilon| \to 0$$

as $U_\varepsilon \to 0$ in $L^1\left([0,T], L^1_{\text{loc}}(\mathbb{R}^n)\right)$ according to Lemma 3.16.
Step 8: $\beta'(p_\varepsilon)Q_{2,\varepsilon} \to 0$
As this step is more complicated than the others, it is done in the following Lemma 3.21

Step 9: $\beta'(p_\varepsilon)\partial_i(\sigma_{ik}R_{k,\varepsilon}) \to 0$
By partial integration and the product rule $\nabla(\beta'(p_\varepsilon)\varphi) = \beta'(p_\varepsilon)\nabla\varphi + \beta''(p_\varepsilon)\nabla p_\varepsilon\varphi$ we have to consider at first the integral

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^n} \beta'(p_\varepsilon)\langle\sigma^*\nabla\varphi, R_\varepsilon\rangle &\le \|\beta'\|_\infty\|\nabla\varphi\|_\infty \int_0^T \int_{\operatorname{spt}\varphi} |\sigma||R_\varepsilon| \\ &\le \|\beta'\|_\infty\|\nabla\varphi\|_\infty \int_0^T \|\sigma|_{\operatorname{spt}\varphi}\|_2\|R_\varepsilon\|_2 \to 0\end{aligned}$$

by Lemma 3.16.
The other integral we have to estimate is the following:

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^n} \varphi \beta''(p_\varepsilon)\langle \sigma^* \nabla p_\varepsilon, R_\varepsilon\rangle &\leq \|\varphi\|_\infty \|\beta''(p_\varepsilon)\|_\infty \int_0^T \int_{\mathbb{R}^n} \langle \sigma^* \nabla p_\varepsilon, R_\varepsilon\rangle \\ &\leq \|\varphi\|_\infty \|\beta''(p_\varepsilon)\|_\infty \int_0^T \|\sigma^* \nabla p_\varepsilon\|_2 \|R_\varepsilon\|_2 \to 0\end{aligned}$$

again by Lemma 3.16 and Lemma 3.4. □

Lemma 3.21. *We have the convergence $\beta'(p_\varepsilon)Q_{2,\varepsilon} \to 0$ in the distributional sense.*

Proof. ([14][p.6] und [2][p. 241]) At first we fix the convolution kernel ρ, from which $Q_{2,\varepsilon}$ depends. By Theorem 3.18 we know, that $|Q_{2,\varepsilon}|$ as a function of ε is bounded in $L^1_{\text{loc}}((0,T)\times\mathbb{R}^n)$, thus also $|\beta'(p_\varepsilon)Q_{2,\varepsilon}|$ is bounded. We can consider them as Radon-measures. So, by the Riesz-Markov Representation theorem, we can also see them as elements of the dual space of $C_c((0,T)\times\mathbb{R}^n)$, which is a separable Banach space. So we have a bounded sequence in the dual space of a separable Banach space and can pick a weakly-∗-convergent subsequence to some measure Q_ρ. This measure is in fact indepent of ρ, because in the proof of Theorem 3.20 we have already seen, that all other terms except $\beta'(p_\varepsilon)Q_{2,\varepsilon}$ in (34) converge in a distributional sense to terms independent of ρ. So lets set $Q := Q_\rho$. Of course we want to show $Q = 0$.
For a test function $\varphi \in C_c((0,T)\times\mathbb{R}^n)$, which we insert in the measure Q and using Theorem 3.18 (extended to inserting test functions instead of compact subsets by monotone convergence), we have:

$$\begin{aligned}Q(\varphi) &= \lim_{\varepsilon\to 0}\int_{(0,T)\times\mathbb{R}^n} |\beta'(p_\varepsilon)Q_{2,\varepsilon}|\varphi \,\mathrm{d}x\,\mathrm{d}t \\ &\leq \|\beta'(p_\varepsilon)\|_\infty \limsup_{\varepsilon\to 0}\int_{(0,T)\times\mathbb{R}^n} |Q_{2,\varepsilon}|\varphi \,\mathrm{d}x\,\mathrm{d}t \\ &\leq \|\beta'(p_\varepsilon)\|_\infty \|p\|_\infty I(\rho)|D^s b|(\varphi)\end{aligned}$$

Thus Q is absolutely continuous with respect to $|D^s b|$ and we can define g as the Radon-Nikodym density of Q with respect to $|D^s b|$, so we have

$$Q(K) = \int_K g(t,x)\,\mathrm{d}|D^s b|(t,x)$$

for any compact K.
Thus we get, this time with the first estimate of Theorem 3.18:

$$\int_K g(t,x)\,\mathrm{d}|D^s b|(t,x) = Q(K) \leq \|\beta'(p_\varepsilon)\|_\infty \left(\limsup_{\varepsilon\to 0}\int_K |Q_{2,\varepsilon}|\,\mathrm{d}x\,\mathrm{d}t\right)$$

$$\leq \|\beta'(p_\varepsilon)\|_\infty \|p\|_\infty \int_K \Lambda(M_t(x),\rho)\,\mathrm{d}|D^s b|(t,x) + \|\beta'(p_\varepsilon)\|_\infty \|p\|_\infty (n + I(\rho))|D^a b|(K)$$

This holds for arbitrary compact sets K, so we can especially choose $|D^a b|$-null-sets, and, as $|D^s b|$ and $|D^a b|$ are singular, we hence get with $C = \|\beta'(p_\varepsilon)\|_\infty \|p\|_\infty$

$$g(t,x) \leq C\Lambda(M_t(x),\rho)$$

for $|D^s b|$-almost-every (t,x) and for every convolution kernel ρ.
Let D be a countable dense subset of the set of convolution kernels with respect to the $W^{1,1}$-norm (which is separable). So we have

$$g(t,x) \leq C \inf_{\rho\in D} \Lambda(M_t(x),\rho)$$

for $|D^s b|$-almost-every (t,x). From the definition of Λ we get that the mapping $\rho \to \Lambda(M_t(x),\rho)$ is continuous for fixed (x,t), thus we have

$$\inf_{\rho\in D} \Lambda(M_t(x),\rho) = \inf_{\rho} \Lambda(M_t(x),\rho)$$

where the right infimum is taken with respect to all convolution kernels. But this infimum is 0, as $M_t(x)$ has rank one $|D^s b|$ almost everywhere by Theorem 2.6 and for such a matrix the infimum of Λ is 0 by Lemma 2.21. So $g = 0$ $|D^s b|$-almost-everywhere, thus $Q = 0$ and $\beta'(p_\varepsilon)Q_{2,\varepsilon} \to 0$ in the distributional sense. □

Remark 3.22. Taking the countable dense subset in the proof of Lemma 3.21 was necessary because the uncountable infimum of measurable functions does need need to be measurable anymore, so speaking about inequalities $|D^s b|$-almost everywhere, which contain $\inf_{\rho} \Lambda(M_t(x),\rho)$ directly is not well defined (see also [6], Theorem 3.6 for this subtility from another point of view).

4 Summary

After the introduction in section 1 we collected some analytic tools used later in various proofs in section 3. At first some properties on BV-functions, mainly the theorem on difference quotients (Lemma 2.5) and Albertis rank-one theorem (Theorem 2.6). Then we proved some convergence lemmata, especially on convergence in measure and defined the convolution of distributions. After this we had a distributional form of Gronwalls inequality (Lemma 2.17), Youngs inequality and a Lemma of Bouchut.
In Section 3 we started considering the actual topic of this thesis, the fokker-planck-equation with BV-drift. The main theorem is Theorem 3.2. In the definition of weak solutions we had the term $\sigma^* \nabla p$, which is a priori not well defined because p is only assumed to be in L^p. Hence there is a weak definition of this term (Remark 3.3). Section 3.1 deals with all problems arising in this context.
Then we proven existence of solutions with an approximation-ansatz using some a-priori-estimates (Theorem 3.7 and Lemma 3.8).
Then the main part of the thesis started, the proof of uniqueness of solutions using the theory of renormalized solutions and commutator estimates.
At first we assumed the renormalization assumption and proved uniqueness under this assumption (Theorem 3.11). Then we defined commutators (Definition 3.12) and proved the commuator estimates from the DiPerna and Lions (Lemma 3.14) and from Ambrosio (Theorem 3.18)
Both were used to prove then the renormalization assumption (Theorem 3.20)

References

[1] ALBERTI, Giovanni: Rank one property for derivatives of functions with bounded variation. In: *Proceedings of the Royal Society of Edinburgh Section A: Mathematics* 123 (1993), Nr. 2, S. 239–274. http://dx.doi.org/10.1017/S030821050002566X. – DOI 10.1017/S030821050002566X

[2] AMBROSIO, Luigi: Transport equation and Cauchy problem for BV vector fields. In: *Inventiones mathematicae* 158 (2004), Nr. 2, 227–260. http://dx.doi.org/10.1007/s00222-004-0367-2. – DOI 10.1007/s00222–004–0367–2. – ISSN 1432–1297

[3] AMBROSIO, Luigi ; FUSCO, Nicola ; PALLARA, Diego: *Functions of Bounded Variation and Free Discontinuity Problems (Oxford Mathematical Monographs)*. Oxford University Press, USA, 2000. – 452 S. – ISBN 9780198502456

[4] ARNOLD, Vladimir I.: *Ordinary differential equations*. 1992

[5] BRIS, C. L. ; LIONS, P.-L.: Existence and Uniqueness of Solutions to Fokker–Planck Type Equations with Irregular Coefficients. In: *Communications in Partial Differential Equations* 33 (2008), Nr. 7, 1272-1317. http://dx.doi.org/10.1080/03605300801970952. – DOI 10.1080/03605300801970952

[6] DE LELLIS, Camillo ; COLLECTIF: Ordinary differential equations with rough coefficients and the renormalization theorem of Ambrosio [after Ambrosio, DiPerna, Lions]. Version: 2008. http://www.numdam.org/item/AST_2008__317__175_0/. In: *Séminaire Bourbaki - Volume 2006/2007 - Exposés 967-981*. Société mathématique de France, 2008 (Astérisque 317). – talk:972

[7] DIPERNA, R. J. ; LIONS, P. L.: Ordinary differential equations, transport theory and Sobolev spaces. In: *Inventiones Mathematicae* 98 (1989), oct, Nr. 3, S. 511–547. http://dx.doi.org/10.1007/bf01393835. – DOI 10.1007/bf01393835

[8] ELSTRODT, Jürgen: *Maß- und Integrationstheorie*. 2011

[9] EVANS, Lawrence C.: *Partial differential equations*. 2010

[10] EVANS, Lawrence C. ; GARIEPY, Ronald F.: *Measure Theory and Fine Properties of Functions, Revised Edition*. Taylor and Francis Group, 2015. – 313 S. – ISBN 9781482242393

[11] IGARI, Katsuju: Degenerate Parabolic Differential Equations. In: *Publications of the research institute for mathematical sciences* (1973)

[12] J. J. DUISTERMAAT, J.A. C. K.: *Distributions*. Birkhäuser, 2010

[13] KRYLOV, N. V.: *Lectures on elliptic and parabolic equations in Sobolev spaces*. American Mathematical Society, 2008. – ISBN 9780821846841

[14] LUO, Dejun: Fokker-Planck type equations with Sobolev diffusion coefficients and BV drift coefficients. In: *Acta Mathematica Sinica, English Series* 29 (2011), 05. http://dx.doi.org/10.1007/s10114-012-1302-x. – DOI 10.1007/s10114–012–1302–x

[15] SOGGE, Christopher D.: *Fourier integrals in classical analysis*. Cambridge University Press, 1993. – 236 S. – ISBN 0521434645

Declaration

I hereby certify that I have written this thesis independently and that I have not used any sources or aids other than those indicated, that all passages of the work which have been taken over verbatim or in spirit from other sources from other sources have been marked as such and that the work has not yet been has not yet been submitted to any examination authority in the same or a similar form.

Erlangen, September 13, 2023 my signature

CV

Christian Lange

Personal data

Name	Christian Lange
Born	22.08.2000
Adress	Fürstenweg 29
	91058 Erlangen

Education

09/2006 - 08/2010	Grundschule Zedernstraße Fürth-Vach
09/2010 - 06/2018	Emmy-Noether-Gymnasium Erlangen

Study

10/2018 - 09/2023	FAU Erlangen-Nürnberg