Springer Proceedings in Mathematics & Statistics

Patrícia Gonçalves
Ana Jacinta Soares *Editors*

From Particle Systems to Partial Differential Equations

PSPDE IV, Braga, Portugal, December 2015

Springer

Springer Proceedings in Mathematics & Statistics

Volume 209

Springer Proceedings in Mathematics & Statistics

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at http://www.springer.com/series/10533

Patrícia Gonçalves · Ana Jacinta Soares
Editors

From Particle Systems to Partial Differential Equations

PSPDE IV, Braga, Portugal, December 2015

Springer

Editors
Patrícia Gonçalves
Center for Mathematics, Analysis,
Geometry and Dynamical Systems
Instituto Superior Técnico
Lisboa
Portugal

Ana Jacinta Soares
Centre of Mathematics
University of Minho
Braga
Portugal

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-319-66838-3 ISBN 978-3-319-66839-0 (eBook)
DOI 10.1007/978-3-319-66839-0

Library of Congress Control Number: 2017951991

Mathematics Subject Classification (2010): 35L67, 35R60, 35Q20, 35Q30, 35Q35, 60K35, 60G60, 60F17, 82C40

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume presents the proceedings of the fourth international conference on Particle Systems and Partial Differential Equations, "PS-PDEs IV", which was held at the Centre of Mathematics of the University of Minho in Braga, Portugal, from 16 to 18 December 2015.

The meeting was intended to bring together prominent active researchers working in the fields of probability and partial differential equations, so that they could present their latest scientific findings in both areas, and to promote discussion on some of their areas of expertise. Further, it was intended to introduce a vast and varied public, including young researchers, to the subject of interacting particle systems, its underlying motivation and its relation to partial differential equations.

This volume includes sixteen contributed papers written by conference participants on essential and intriguing topics in the fields of probability theory, partial differential equations and kinetic theory.

We believe that this volume will be of great interest to probabilists, analysts and also to those mathematicians with a general interest in mathematical physics, stochastic processes and differential equations, as well as those physicists whose work intersects with statistical mechanics, statistical physics and kinetic theory.

We would like to take this opportunity to extend our thanks to all the speakers, and to the participants, for contributing to the success of this meeting.

Lastly, we wish to gratefully acknowledge the financial support provided by Fundação para a Ciência e a Tecnologia through the FCT-FACC funds, to the Centre of Mathematics of the University of Minho, to the Centre of Mathematics, Fundamental Applications and Operations Research of the University of Lisbon, to the Center for Mathematical Analysis, Geometry and Dynamical Systems of the University of Lisbon and to the Co-Lab initiative UT Austin-Portugal.

We really hope that you enjoy reading this book!

Lisboa, Portugal
Braga, Portugal
May 2017

Patrícia Gonçalves
Ana Jacinta Soares

Contents

Construction of Φ^4_3 Diagrams for Pedestrians

Jean-Christophe Mourrat, Hendrik Weber and Weijun Xu

Abstract We aim to give a pedagogic and essentially self-contained presentation of the construction of various stochastic objects appearing in the dynamical Φ^4_3 model. The construction presented here is based on the use of paraproducts. The emphasis is on describing the stochastic objects themselves rather than introducing a solution theory for the equation.

Keywords Feynman diagrams · Singular stochastic PDEs
Stochastic quantisation · Paraproducts

1 Introduction

The purpose of this note is to give a pedagogic presentation of the construction of the various stochastic "basis" terms entering the construction of the dynamic Φ^4 theory in three space dimensions (Φ^4_3 for short). Formally, the dynamic Φ^4 model on the torus $\mathbb{T}^3 = [0, 1]^3$ is the solution $X(t, x)$ to the stochastic partial differential equation

$$\begin{cases} \partial_t X = \Delta X - X^3 + mX + \xi, & \text{on } \mathbb{R}_+ \times [0, 1]^3, \\ X(0, \cdot) = X_0, \end{cases} \tag{1}$$

where ξ denotes a white noise over $\mathbb{R} \times \mathbb{T}^3$, and m is a real parameter. Equation (1) describes the natural reversible dynamics for the "static" Φ^4_3 Euclidean field theory, which is formally given by the expression

J.-C. Mourrat (✉)
Ecole Normale Supérieure de Lyon, CNRS, Lyon, France
e-mail: jean-christophe.mourrat@ens-lyon.fr

H. Weber · W. Xu
University of Warwick, Coventry, United Kingdom
e-mail: hendrik.weber@warwick.ac.uk

W. Xu
e-mail: weijun.xu@warwick.ac.uk

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_1

$$\mu \propto \exp\left(-2\int_{\mathbb{T}^3}\left[\frac{1}{2}|\nabla X|^2 + \frac{1}{4}X^4 - \frac{m}{2}X^2\right]\right)\prod_{x\in\mathbb{T}^3} \mathrm{d}X(x). \tag{2}$$

The quartic potential X^4 in this energy gives the model its name (replacing X by Φ). Mathematically, neither (1) nor (2) make sense as they stand. While this problem is the main concern of this note, we postpone its discussion and first proceed heuristically. (Alternatively, we temporarily replace the continuous space $\mathbb{T}^3$ by a finite approximation, with a suitable interpretation of the gradient.)

The potential function $x \mapsto \frac{1}{4}x^4 - \frac{m}{2}x^2$ has a single minimum at $x = 0$ for $m \leq 0$. As m becomes positive, a pitchfork bifurcation occurs, with the appearance of two minima at $x = \pm\sqrt{m}$, while the point $x = 0$ becomes a local maximum. In the energy between square brackets appearing in (2), the part consisting of

$$\int_{\mathbb{T}^3}\left[\frac{1}{4}X^4 - \frac{m}{2}X^2\right]$$

favors fields X that take values close to those minima, while the part

$$\int_{\mathbb{T}^3}|\nabla X|^2$$

favors some agreement between nearby values of the field X. This description is highly reminiscent of that of the Ising model. Indeed, these two models are believed to have comparable phase transitions and large-scale properties. In one and two space dimensions (when $\mathbb{T}^3$ is replaced by $\mathbb{T}^d$, $d \in \{1, 2\}$), a precise link between the Glauber dynamics of an Ising model with long-range interactions and the dynamic Φ^4 model was obtained [2, 10, 20], and a similar result is conjectured to hold in our present three-dimensional setting.

Starting from the 60s, the Φ^4 model was the subject of an intense research effort from the perspective of quantum field theory. From this point of view, the construction of the so-called Euclidean Φ^4 measure (2) is a first step towards building the corresponding quantum field theory. This requires the verification of certain properties known as the Osterwalder–Schrader axioms [24, 25], among which the reflection positivity and the invariance under Euclidean isometry are the most important (we refer to [3] for a review on reflection positivity — in particular, the Ising measure is reflection positive, see [3, Corollary 5.4]). The whole endeavour was viewed as a test-bed for more complicated (and more physically relevant) quantum field theories. We stress that from this point of view, one of the directions of space becomes the time variable in the quantum field theory. The time variable appearing in (1) is then seen as an additional, physically fictitious variable, sometimes called the "stochastic time" in the literature. The construction of a quantum field theory based on the invariant measure of a random process is called "stochastic quantisation", and was proposed by Parisi and Wu [26]. We refer to [12, Sect. 20.1] for references and more precise explanations.

We now return to the problem that (1) and (2) do not actually make sense mathematically. In (1), the roughness of the noise requires X to be distribution-valued, and therefore the interpretation of the cubic power is unclear. In (2), one could interpret

$$\exp\left(-\int_{\mathbb{T}^3} |\nabla X|^2\right) \prod_{x\in\mathbb{T}^3} \mathrm{d}X(x)$$

as a formal notation to denote the law of a Gaussian free field. Again, the Gaussian free field is distribution-valued, and there is no canonical interpretation for X^4 or X^2.

From now on, we focus on making sense of (1). A naive attempt would consist of regularising the noise, defining the corresponding solution, and trying to pass to the limit. However, the progressive blow-up of the non-linearity forces the limit to be identically zero (see [17] for a rigorous justification in the case of two space dimensions). Thus, we need to take a step back and modify the original Eq. (1) in a way that is faithful to the intended "physics" of the model, as sketched above.

A formal scaling argument (see e.g. [6, Sect. 1.1]) shows that the non-linearity should become less and less relevant as we zoom in on the solution: the equation is said to be *subcritical*, or super-renormalisable. The basic idea for making sense of the equation is therefore to postulate a first-order expansion of X of the form $X = \varpi + Y$, where ϖ is the stationary solution to the linear equation

$$(\partial_t - \Delta + 1)\varpi = \xi. \tag{3}$$

In other words, letting $\{P_t = e^{t(\Delta-1)}\}_{t\geq 0}$ denote the heat semigroup on $\mathbb{T}^3$, we have

$$\varpi(t) = \int_{-\infty}^{t} P_{t-s}(\xi(s))\,\mathrm{d}s. \tag{4}$$

The "+1" in (3) serves to prevent the divergence of the low-frequency part of ϖ in the long-time limit (and to allow us to talk about a stationary solution over the whole time line $\mathbb{R}$).

Making the ansatz $X = \varpi + Y$ and formally rewriting (1) in terms of Y leads to the equation

$$\partial_t Y = \Delta Y - (Y + \varpi)^3 + m(Y + \varpi). \tag{5}$$

Solving this equation for Y requires us to make sense of quantities such as $(\varpi)^2$ or $(\varpi)^3$.

These are again ill-defined. We may regularise the noise, on scale $1/n$, and define the corresponding solution ϖ_n. While $(\varpi_n)^2$ and $(\varpi_n)^3$ still diverge as we remove the regularisation, the very explicit and simple structure of ϖ allows us now to identify a constant $\mathbf{c}_n$ such that $(\varpi_n)^2 - \mathbf{c}_n$ and $(\varpi_n)^3 - 3\mathbf{c}_n\varpi_n$ converge to non-trivial limits as n tends to infinity, which we denote by $\varv$ and $\varv\varpi$ respectively.

At this point, we can rewrite the Eq. (5) for $Y = X - \varpi$ using ϖ, $\varv$ and $\varv\varpi$. In two space dimensions, this equation has been solved in [8] with classical methods, without

Table 1 The list of relevant diagrams, together with their regularity exponent, where $\varepsilon > 0$ is arbitrary

τ						
$\lvert\tau\rvert$	$-\frac{1}{2} - \varepsilon$	$-1 - \varepsilon$	$\frac{1}{2} - \varepsilon$	$-\varepsilon$	$-\frac{1}{2} - \varepsilon$	$-\varepsilon$

further recourse to the probabilistic structure of the problem. Note that this approach shares the philosophy of rough path theory (see e.g. [11] for an introduction), in that one first constructs a few fundamental objects (here , and) by relying on the probabilistic structure of the problem, and then one builds the solution as a deterministic and continuous map of the enriched datum (, ,).

In three space dimensions, the equation one obtains for the remainder Y is still ill-defined, and we need to proceed further in the postulated "Taylor expansion" of the solution X. The procedure becomes more intricate, and was solved only recently by Hairer [15] within his ground-breaking theory of *regularity structures* (see also [14] for the treatment of the KPZ equation with rough paths). Catellier and Chouk [5] then showed how to recover the results of Hairer for the Φ^4_3 model, using the alternative theory of *para-controlled distributions* set up by Gubinelli, Imkeller and Perkowski [13]. We refer to [22, Sect. 1] for a presentation of the latter approach with notation consistent with the one we use here. Yet another approach based on Wilsonian *renormalisation group* analysis was given by Kupiainen in [19].

We work here in the para-controlled framework, as in [5, 13, 22]. As it turns out, six "basis" elements, that is, non-linear objects based on the solution to the linear equation (3) and built using the probabilistic structure, are required to define a solution to the Φ^4_3 equation. We call these processes "the diagrams". They are listed in Table 1; their precise meaning will be explained shortly. The purpose of this note is to review their construction.

Minor variants of these diagrams were built in [15] in the context of regularity structures. There, a very convenient graphical notation akin to Feynman diagrams was introduced to derive the bounds required for the construction of these diagrams (see [27] for an elementary introduction to Feynman diagrams. An earlier version of a graph-based method to bound stochastic terms using diagrams in the context of the KPZ equation was developed in [14]). In the context of para-controlled distributions, the exact same diagrams as those we consider here were also built in [5], albeit with a possibly less transparent notation. More recently, a remarkable machinery was developped in the context of regularity structures, which ensures the convergence of diagrams for a large class of models under extremely general assumptions; see [16, Theorem A.3] and [7].

This note is mostly expository: we aim to provide a gentle introduction to this graphical notation, and to make clear that it applies equally well in the para-controlled setting. We do not strive to capture the deep results in [7, 16], but rather to give a "pedestrian" exposition of the calculations involved.

We now introduce some notation. Let $\mathcal{S}'$ denote the space of Schwartz distributions over the torus $\mathbb{T}^3$, and define

$$I(f) : \begin{cases} \mathbb{R} \to \mathcal{S}' \\ t \mapsto \int_{-\infty}^{t} P_{t-s}(f(s))\, \mathrm{d}s, \end{cases} \tag{6}$$

for every $f \in C(\mathbb{R}, \mathcal{S}')$ for which the integral makes sense. In other words, $I(f)$ is the "ancient" solution to the heat equation with right-hand side f, that is, the one "started at time $-\infty$". We measure the regularity of distributions on $\mathbb{T}^3$ via a scale of function spaces which we denote by $\mathcal{C}^\alpha$, $\alpha \in \mathbb{R}$. The precise definition is recalled below, and is a natural extension of the notion of α-Hölder regular functions. We also recall below the definition of the resonant product $\circleddash$. Our goal is to identify suitable deterministic constants $\mathbf{c}_n, \mathbf{c}'_n \in \mathbb{R}$ and show the convergence as n tends to infinity of the following five processes:

$$\begin{aligned}
,\ \vee_n &:= (\uparrow_n)^2 - \mathbf{c}_n, \\
\Psi_n &:= I\left((\uparrow_n)^3 - 3\mathbf{c}_n \uparrow\right), \\
\tau_{\Psi\uparrow,n} &:= \Psi_n \circleddash \uparrow_n, \\
\tau_{\vee\vee,n} &:= I(\vee_n) \circleddash \vee_n - 2\mathbf{c}'_n, \\
\tau_{\Psi\vee,n} &:= \Psi_n \circleddash \vee_n - 6\mathbf{c}'_n \uparrow_n.
\end{aligned} \tag{7}$$

The interested reader is referred to the discussion in [22, Sect. 1.1] to see how these diagrams arise in the construction of solutions to (1). The stationarity in space and time of the white noise ξ as well as the fact that I defines ancient solutions to the inhomogeneous heat equation imply that these processes are stationary in space and time. Here is the main result on which we will focus.

Theorem 1 ([5, 15]) *Fix*

$$\mathbf{c}_n := \mathbb{E}\left[(\uparrow_n(t))^2\right] \quad \textit{and} \quad \mathbf{c}'_n := \mathbb{E}\left[I(\vee_n) \circleddash \vee_n(t)\right]. \tag{8}$$

For each pair $(\tau, |\tau|)$ as in Table 1, let τ_n be defined as in (7). There exists a stochastic process, denoted by τ and taking values in $C(\mathbb{R}, \mathcal{C}^{|\tau|})$, such that for every $p \in [1, +\infty)$, we have

$$\sup_{t \in \mathbb{R}} \mathbb{E}\left[\|\tau_n(t) - \tau(t)\|^p_{\mathcal{C}^{|\tau|}}\right] \xrightarrow[n \to +\infty]{} 0. \tag{9}$$

Moreover, for every $p \in [1, +\infty)$,

$$\sup_{t \in \mathbb{R}} \mathbb{E}\left[\|\tau(t)\|^p_{\mathcal{C}^{|\tau|}}\right] < +\infty, \tag{10}$$

and for every $p \in [1, \infty)$ and $\lambda \in [0, 1]$,

$$\sup_{s<t} \frac{\mathbb{E}\left[\|\tau(t)-\tau(s)\|^p_{\mathcal{C}^{|\tau|-2\lambda}}\right]}{|t-s|^{\lambda}} < +\infty. \tag{11}$$

Remark 1 Since the processes we consider are stationary in time, the constants in (8) do not depend on the time t. Moreover, the suprema in (9) and (10) are superfluous. We prefer to write them nonetheless, since the statements with the suprema are robust to perturbations of the stationarity property.

Remark 2 As will be seen below, the constants $\mathbf{c}_n$ and $\mathbf{c}'_n$ diverge at order n and $\log n$ respectively.

Remark 3 The bound (11) immediately implies the pathwise Hölder continuity of the symbols, by the Kolmogorov continuity test. In the construction of solutions to (1), this strong control on the Hölder regularity of τ is only needed for the symbol Ψ. For the remaining symbols, a weaker bound of the type

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|\tau(t)\|^p_{\mathcal{C}^{|\tau|}}\Big]$$

suffices. However, the proofs of (11) and (10) are relatively similar anyway, as we hope to convince the reader below.

This note is organised as follows. In Sect. 2, we introduce Besov spaces. The diagrams take values in these spaces. The definition of these function spaces is based on the Littlewood-Paley decomposition. This allows us to define paraproducts along the way, and to give relevant intuition for them. In Sect. 3, we introduce the white noise process and discuss the property of equivalence of moments. The latter is very convenient to reduce the bounds (10) and (11) to easy-to-check second moment computations. In Sect. 4, we construct the diagrams and prove the fixed time bound (10). In Sect. 5, we briefly discuss how the bound (11) for time differences follows easily from the fixed time one. Finally, in the appendix, we give an alternative proof of the equivalence of moments property exposed in Sect. 3.

2 Function Spaces and Paraproducts

In this section, we introduce the function spaces we will use, denoted by $\mathcal{C}^\alpha$, for $\alpha \in \mathbb{R}$. When $\alpha \in (0,1)$, they are (a separable version of) the usual Hölder spaces. For general α, they belong to the larger class of Besov spaces, and enjoy remarkable stability properties under multiplication. We choose to also give an informal presentation of these properties, although we will not refer to these in our actual construction of the diagrams. Our choice is motivated by the fact that the question of defining products of distributions is central to making sense of the Φ^4 model. It is therefore useful to survey first what can be achieved with purely deterministic methods (and what can be ultimately used to show well-posedness of the Φ^4 model).

Moreover, exploring this question naturally leads to the introduction of paraproducts. In this section, the space dimension d is arbitrary. For most results, we only provide a sketch of proof. A much more detailed treatment of the topics discussed in this section can be found in [1, Chap. 2].

We wish to extend the notion of α-Hölder regularity of a distribution f on $\mathbb{T}^d$ to exponents $\alpha < 0$. Roughly speaking, this should mean that for every $x \in \mathbb{T}^d$ and every test function $\varphi \in C_c^\infty(\mathbb{R}^d)$,

$$\left\langle f, \varepsilon^{-d}\varphi(\varepsilon^{-1}(\cdot - x))\right\rangle \lesssim \varepsilon^\alpha \quad \text{uniformly in } x \in \mathbb{T}^d \text{ and } \varepsilon \to 0, \tag{12}$$

where we interpret f as a periodic distribution on $\mathbb{R}^d$ in the duality pairing above.[1] A precise definition can be built along these lines (the interested reader can find it in [15, Definition 3.7]). However, we prefer to adopt here a point of view based on Fourier analysis, which allows for a more direct understanding of the stability of the spaces under multiplication.

Remark 4 For *positive* α, the condition (12) should be replaced by

$$\left\langle f - p_x(\cdot), \varepsilon^{-d}\varphi(\varepsilon^{-1}(\cdot - x))\right\rangle \lesssim \varepsilon^\alpha \quad \text{uniformly in } x \in \mathbb{T}^d \text{ and } \varepsilon \to 0, \tag{13}$$

where p_x is the Taylor approximation of order $\lfloor \alpha \rfloor$ of f. For negative α, there is no such polynomial, and this recentering is unnecessary.

For every $f \in L^1(\mathbb{T}^d)$ and $\omega \in \mathbb{Z}^d$, we write

$$\mathscr{F} f(\omega) = \widehat{f}(\omega) := \int_{\mathbb{T}^d} f(x) e^{-2i\pi\omega\cdot x}\, \mathrm{d}x, \tag{14}$$

for the Fourier coefficient of f with frequency ω, so that

$$f(x) = (\mathscr{F}^{-1}\widehat{f})(x) := \sum_{\omega\in\mathbb{Z}^d} \widehat{f}(\omega) e^{2i\pi\omega\cdot x},$$

where $\mathscr{F}^{-1}$ denotes the inverse Fourier transform.

The definition of Besov spaces rests on a decomposition of the Fourier series of a function along dyadic annuli, an idea due to Littlewood and Paley. More precisely, we think of splitting $\widehat{f}$ into

$$\widehat{f}\,\mathbf{1}_{B(0,1)} + \sum_{k=0}^{+\infty} \widehat{f}\,\mathbf{1}_{B(0,2^{k+1})\setminus B(0,2^k)}, \tag{15}$$

where $B(0, r) := \{\omega \in \mathbb{Z}^d \ : \ |\omega| < r\}$. In (15), the terms of the series associated with large k's measure the fast oscillations of the function; the general Besov norm

[1] Here and below we write $A \lesssim B$ to mean that there exists a constant C, which is independent of the quantities of interest, such that $A \leq CB$.

can be thought of as a weighted average of the L^p norm of these summands. This type of decomposition enjoys better analytical properties if we replace the indicator functions by smoothened versions thereof. More precisely, we can find functions $\widetilde{\chi}, \chi \in C_c^\infty(\mathbb{R}^d)$ both taking values in $[0,1]$, with supports

$$\mathrm{Supp}\tilde{\chi} \subseteq B\left(0, \frac{4}{3}\right), \qquad \mathrm{Supp}\chi \subseteq B\left(0, \frac{8}{3}\right) \setminus B\left(0, \frac{3}{4}\right),$$

and such that

$$\widetilde{\chi}(\zeta) + \sum_{k=0}^{+\infty} \chi(\zeta/2^k) = 1, \quad \forall \zeta \in \mathbb{R}^d. \tag{16}$$

We may furthermore choose these functions to be radially symmetric. We write

$$\chi_{-1} := \widetilde{\chi}, \qquad \chi_k(\cdot) := \chi(\cdot/2^k) \quad k \geq 0. \tag{17}$$

The supports of $\tilde{\chi}$ and χ ensure that χ_k and $\chi_{k'}$ overlap only if $|k - k'| \leq 1$ (see [1, Proposition 2.10] for more details). For every $f \in C^\infty(\mathbb{T})$ and $k \geq -1$, we let

$$\delta_k f := \mathscr{F}^{-1}\left(\chi_k \widehat{f}\right),$$

so that $\widehat{f} = \sum_{k\geq -1} \chi_k \widehat{f}$ (compare with (15)) and $f = \sum_{k\geq -1} \delta_k f$. Let

$$\eta_k = \mathscr{F}^{-1}(\chi_k), \qquad \eta = \eta_0. \tag{18}$$

For $k \geq 0$, we have

$$\eta_k \simeq 2^{dk} \eta(2^k \cdot), \tag{19}$$

up to a small error due to the fact that our phase space $\mathbb{Z}^d$ is discrete (since our state space $\mathbb{T}^d$ is compact).[2] For every $k \geq -1$, we have

$$\delta_k f = \eta_k \star f, \tag{20}$$

where $\star$ denotes the convolution on the torus $\mathbb{T}^d$. In agreement with (12), (19) and (20), we define the $\mathcal{C}^\alpha$ norm by

$$\|f\|_{\mathcal{C}^\alpha} := \sup_{k\geq -1} 2^{\alpha k} \|\delta_k f\|_{L^\infty}. \tag{21}$$

[2] One may estimate the error in (19) and show that it is negligible for our purpose, but the simplest way around this technical point is probably to interpret each periodic function on $\mathbb{T}^d$ as a periodic function on $\mathbb{R}^d$, and then use $L^p(\mathbb{R}^d)$ norms and the continuous Fourier transform throughout, so that (19) becomes exact. The continuous Fourier transform of any Schwartz distribution is well-defined, by duality. For a periodic $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, the Fourier transform is a sum of Dirac masses at every $\omega \in \mathbb{Z}^d$, each carrying a mass $\widehat{f}(\omega)$ as defined in (14).

One can check that this quantity is finite if $f \in C^\infty(\mathbb{T}^d)$. The space $\mathcal{C}^\alpha$ is the completion of $C^\infty(\mathbb{T}^d)$ with respect to this norm. This space can be realised as a subspace of the space of Schwartz distributions.

Remark 5 We depart slightly from the convention to define $\mathcal{C}^\alpha$ as the space of distributions with finite $\|\cdot\|_{\mathcal{C}^\alpha}$ norm. Our definition makes the space separable and allows us below to define products of distributions and functions via approximation. Moreover, one can check that if a distribution f satisfies $\|f\|_{\mathcal{C}^\alpha} < \infty$, then $f \in \mathcal{C}^\beta$ for every $\beta < \alpha$.

The most important property of Besov spaces for our purpose is the following multiplicative structure.

Proposition 1 *Let $\alpha < 0 < \beta$ be such that $\alpha + \beta > 0$. The multiplication $(f, g) \mapsto fg$ extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$.*

The proof of this proposition rests on the decomposition

$$fg = \sum_{k<l-1} \delta_k f\, \delta_l g + \sum_{|k-l|\le 1} \delta_k f\, \delta_l g + \sum_{k>l+1} \delta_k f\, \delta_l g,$$

which we will write suggestively in the form

$$fg = f \varolessthan g + f \varominus g + f \varogreaterthan g.$$

This is often called *Bony's decomposition* into the *paraproducts* $f \varolessthan g$, $f \varogreaterthan g = g \varolessthan f$, and the *resonant product* $f \varominus g$.

In order to prove Proposition 1, it suffices to show that each of these terms extends to a continuous bilinear map from $\mathcal{C}^\alpha \times \mathcal{C}^\beta$ to $\mathcal{C}^\alpha$. However, it is very important for our more general goal of making sense of the Φ^4 model to be precise about the behaviour of each term separately. We thus simply assume that $\alpha < 0 < \beta$ and $f \in \mathcal{C}^\alpha$, $g \in \mathcal{C}^\beta$ to begin with (but do not yet prescribe the sign of $\alpha + \beta$), and see whether and how we can estimate each of the terms in Bony's decomposition. We start with

$$f \varolessthan g = \sum_{k<l-1} \delta_k f\, \delta_l g. \tag{22}$$

In view of (21), in order to see which Hölder class $f \varolessthan g$ belongs to, we need to estimate

$$\big\|\delta_j(f \varolessthan g)\big\|_{L^\infty} = \left\| \sum_{l=-1}^{+\infty} \delta_j \left(\sum_{k=-1}^{l-2} \delta_k f \delta_l g \right) \right\|_{L^\infty}. \tag{23}$$

Recall that the Fourier transform of $\delta_k f$ (resp. $\delta_l g$) is supported on an annulus of both inner and outer radius of size about 2^k (resp. 2^l). The important observation is that as we sum over $k \le l-2$, the Fourier transform of $\delta_k f \delta_l g$ is still supported in an annulus of inner and outer radius both proportional to 2^l. Hence, only a finite

number of terms $l \simeq j$ contribute to the outer sum on the right hand side above. Now, since

$$\|\delta_k f\|_{L^\infty} \lesssim 2^{-\alpha k} \|f\|_{\mathcal{C}^\alpha}, \qquad \|\delta_l g\|_{L^\infty} \lesssim 2^{-\beta l} \|g\|_{\mathcal{C}^\alpha},$$

the right hand side of (23) can then be bounded by

$$\| \sum_{k=-1}^{j-2} \delta_k f \delta_j g \|_{L^\infty} \lesssim \sum_{k=-1}^{j-2} 2^{-\alpha k - \beta j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta} < C 2^{-(\alpha+\beta)j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta},$$

where we used that $\alpha < 0$, and C does not depend on f or g. This shows that

$$\|f \circledcirc g\|_{\mathcal{C}^{\alpha+\beta}} \lesssim C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

The same analysis applies for the term $g \circledcirc f$, except that since we have $\beta > 0$, we get

$$\sum_{l=-1}^{j-2} 2^{-\alpha j - \beta l} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta} \lesssim C 2^{-\alpha j} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta},$$

which implies that

$$\|g \circledcirc f\|_{\mathcal{C}^\alpha} \lesssim C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

Note that we have made no assumption on the sign of $\alpha + \beta$ so far. The term $f \circledcirc g$ inherits essential features of the small scale behaviour of g "modulated" by the low frequency modes of f. This is in agreement with the fact that $f \circledcirc g$ is more regular than $g \circledcirc f$ under our hypothesis $\alpha < \beta$.

We now turn to the resonant term $f \circledast g$, which for simplicity we think of as being

$$\sum_{k=-1}^{+\infty} \delta_k f \, \delta_k g.$$

The crucial difference in the analysis of this term, compared with the previous computations, is that the support of the Fourier transform of the summand indexed by k, which is the convolution of the annulus of radius about 2^k by itself, results in a *ball* of radius 2^k, as opposed to an *annulus*. Therefore, every summand indexed by $k \geq l$ contributes the the l-th Littlewood-Paley block $\delta_l(f \circledast g)$. The L^∞ norm of each summand is bounded by

$$2^{-(\alpha+\beta)k} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

If we want this to be summable over $k \geq l$, we need to assume $\alpha + \beta > 0$. In this case, the sum is of order $2^{-(\alpha+\beta)l} \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}$, which suggests that

$$\|f \circledast g\|_{\mathcal{C}^{\alpha+\beta}} \lesssim C \|f\|_{\mathcal{C}^\alpha} \|g\|_{\mathcal{C}^\beta}.$$

Table 2 Summary of the regularity properties of paraproducts for $\alpha < 0 < \beta$, $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$

	$f \varolessthan g$	$g \varolessthan f$	$f \varodot g$	fg
Regularity	$\alpha + \beta$	α	$\alpha + \beta$	α
Needs $\alpha + \beta > 0$	No	No	Yes	Yes

These computations can all be made rigorous (see e.g. [1]), and are summarised in Table 2. Note that these relations are relevant in the construction of Φ^4 since there, one needs to characterise the products between $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$ (though sometimes necessarily $\alpha + \beta < 0$, and renormalisations are then needed for the resonant term).

We point out that the regularising effect of the heat kernel can be conveniently measured using the spaces $\mathcal{C}^\alpha$. While we will not use this proposition in itself here, it is a useful guide to the intuition. In particular, the time singularity in (24) is integrable as long as the difference of regularity exponents is less than 2. In other words, the integration operator I in (6) brings a gain of 2 units of regularity.

Proposition 2 *If $\alpha \le \beta \in \mathbb{R}$, then there exists $C < \infty$ such that for every $t > 0$, we have*

$$\|e^{t\Delta} f\|_{\mathcal{C}^\beta} \le C\, t^{\frac{\alpha-\beta}{2}}\, \|f\|_{\mathcal{C}^\alpha}. \tag{24}$$

Sketch of proof The Laplacian Δ is a multiplication operator in Fourier space. As a consequence, we have $\delta_k(e^{t\Delta} f) = e^{t\Delta}(\delta_k f)$, and since $\widehat{\Delta}(\omega) = -|\omega|^2$, roughly speaking, we have $e^{t\Delta}(\delta_k f) \simeq e^{-t2^{2k}} \delta_k f$. This suggests that

$$\|\delta_k(e^{t\Delta} f)\|_{L^\infty} \le C \exp\left(-t2^{2k}\right) \|\delta_k f\|_{L^\infty} \le C \exp\left(-t2^{2k}\right) 2^{-\alpha k} \|f\|_{\mathcal{C}^\alpha},$$

and therefore

$$2^{\beta k} \|\delta_k(e^{t\Delta} f)\|_{L^\infty} \le C \left[(2^{2k} t)^{\frac{\beta-\alpha}{2}} \exp\left(-2^{2k} t\right) \right] t^{\frac{\alpha-\beta}{2}} \|f\|_{\mathcal{C}^\beta}.$$

Since the term between square brackets is bounded uniformly over k and t, the heuristic argument is complete. A rigorous proof can be derived using techniques similar to those exposed for Lemma 2 below (see also e.g. [21, Proposition 3.11]).

In view of (19), we expect that for every $p, p' \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{p'} = 1$,

$$\sup_{k \ge -1} 2^{-\frac{dk}{p}} \|\eta_k\|_{L^{p'}} < \infty. \tag{25}$$

Indeed, the relation (19), and therefore the inequality (25), are immediate by scaling if the torus $\mathbb{T}^d$ is replaced by the full space $\mathbb{R}^d$ (and therefore the discrete Fourier series is replaced by the continuous Fourier transform). A rigorous proof of (25) can be found e.g. in [20, Lemma B.1]. By (20) and Hölder's inequality, we deduce the following lemma.

Lemma 1 *Let $p \in [1, \infty]$. There exists $C < \infty$ such that*

$$\|\delta_k f\|_{L^\infty} \leq C 2^{\frac{dk}{p}} \|f\|_{L^p} \tag{26}$$

for every $f \in C^\infty(\mathbb{T}^d)$ and $k \geq -1$.

If we had chosen to define the decomposition $f = \sum_{k \geq -1} \delta_k f$ from the Fourier series decomposition displayed in (15) based on indicator functions, then we would have $\delta_k \delta_k f = \delta_k f$, and we could therefore replace f by $\delta_k f$ on the right side of (26). With our actual definition of δ_k, this replacement is also possible: instead of using that

$$\mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)} \mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)} = \mathbf{1}_{B(0,2^{k+1}) \setminus B(0,2^k)},$$

we choose a smooth function $\chi' \in C_c^\infty(\mathbb{R}^d)$ with support in an annulus and such that $\chi' \equiv 1$ on the support of χ, so that

$$\chi' \chi = \chi.$$

Setting $\delta'_k f := \mathscr{F}^{-1}(\chi'(\cdot/2^k) \widehat{f})$, the identity above translates into

$$\delta'_k \delta_k f = \delta_k f. \tag{27}$$

Next, we verify that the argument leading to Lemma 1 also applies if $\delta_k f$ is replaced by $\delta'_k f$ on the left side of (26). Using (27), we thus obtain the following lemma.

Lemma 2 (Bernstein inequality) *Let $p \in [1, \infty]$. There exists $C < \infty$ such that*

$$\|\delta_k f\|_{L^\infty} \leq C 2^{\frac{dk}{p}} \|\delta_k f\|_{L^p} \tag{28}$$

for every $f \in C^\infty(\mathbb{T}^d)$ and $k \geq -1$.

The following proposition uses the previous lemma to bound p-th moments of the Hölder norm of a random distribution in terms of the p-th moments of its decomposition in Fourier space.

Proposition 3 (Boundedness criterion) *Let $\beta < \alpha - \frac{d}{p}$. There exists $C < \infty$ such that for every random distribution f on $\mathbb{T}^d$, we have*

$$\mathbb{E}\left[\|f\|_{\mathcal{C}^\beta}^p\right] \leq C \sup_{k \geq -1} 2^{\alpha k p} \, \mathbb{E}\left[\|\delta_k f\|_{L^p}^p\right]. \tag{29}$$

Proof By definition of the $\mathcal{C}^\beta$ norm and then Lemma 2, we have

$$\|f\|_{\mathcal{C}^\beta}^p = \sup_{k \geq -1} 2^{\beta k p} \|\delta_k f\|_{L^\infty}^p \leq C \sup_{k \geq -1} 2^{k(\beta p + d)} \|\delta_k f\|_{L^p}^p.$$

In order to take the expectation of $\|\delta_k f\|_{L^p}^p$ directly, we enlarge the supremum on the right side above to a sum, and get

$$\mathbb{E}\|f\|_{\mathcal{C}^\beta}^p \le C \sum_{k\ge -1} 2^{k(\beta p+d)} \mathbb{E}\|\delta_k f\|_{L^p}^p = C \sum_{k\ge -1} 2^{kp(\beta+\frac{d}{p}-\alpha)} 2^{\alpha k p} \mathbb{E}\|\delta_k f\|_{L^p}^p .$$

The announced estimate then follows, since $\alpha > \beta + \frac{d}{p}$.

Remark 6 With our definition of the space $\mathcal{C}^\beta$ as a completion, the fact that a distribution f satisfies $\|f\|_{\mathcal{C}^\beta} < \infty$ does not imply that $f \in \mathcal{C}^\beta$. However, in the context of Proposition 3, when the right side of (29) is finite, we do have $f \in \mathcal{C}^\beta$ with probability one. Indeed, we can always pick $\beta' \in (\beta, \alpha - \frac{d}{p})$, deduce that $\|f\|_{\mathcal{C}^{\beta'}}$ is finite with probability one, and conclude by Remark 5.

3 White Noise and Nelson's Estimate

Proposition 3 gives a criterion for determining whether a random distribution belongs to $\mathcal{C}^\beta$ by looking at the p-th moment of its Paley-Littlewood blocks. However, it is often difficult to get sharp bounds of high moments of a random distribution. On the other hand, fortunately, the objects we encounter in the construction of Φ^4_3 (and many other Gaussian models) are all built from multiplications of finitely many Gaussian random variables. These objects belong to a Wiener chaos of finite order, and we can therefore leverage on the equivalence of moments property, also often called Nelson's estimate, to deduce high-moment estimates from second-moment calculations. The purpose of this section is to present these arguments, and state the implied simpler criterion for belonging to $\mathcal{C}^\beta$. In this section, we continue to work in arbitrary space dimension d. We only sketch some well-known arguments concerning iterated integrals and Wiener chaos, and refer the interested reader to [18, Chap. 9] for a more detailed exposition of these topics.

We start by introducing the space-time white noise. Formally, the space-time white noise ξ is a centred Gaussian distribution on $\mathbb{R} \times \mathbb{T}^d$ with covariance

$$\mathbb{E}\xi(t,x)\,\xi(t',x') = \delta(t-t')\,\delta^d(x-x') \,, \tag{30}$$

where $\delta(\cdot)$ and $\delta^d(\cdot)$ denote Dirac delta functions over $\mathbb{R}$ and $\mathbb{T}^d$ respectively. Testing (30) against a function $\varphi : \mathbb{R} \times \mathbb{T}^d \to \mathbb{R}$ leads us to postulate that

$$\mathbb{E}\left[\xi(\varphi)^2\right] = \|\varphi\|^2_{L^2(\mathbb{R}\times\mathbb{T}^d)} \,. \tag{31}$$

Definition 1 A space-time white noise over $\mathbb{R} \times \mathbb{T}^d$ is a family of centred Gaussian random variables $\{\xi(\varphi),\ \varphi \in L^2(\mathbb{R} \times \mathbb{T}^d)\}$ such that (31) holds.

The existence of a space-time white noise follows from Kolmogorov's extension theorem. We prefer here to take a more constructive approach, based on Fourier analysis.

Let $(W(\cdot,\omega))_{\omega\in\mathbb{Z}^d}$ be a family of complex-valued Brownian motions over $\mathbb{R}$. These Brownian motions are independent for different values of ω except for the constraint $W(t,\omega)=\overline{W(t,-\omega)}$ (so that the white noise we are building is real-valued). The magnitude of the variance is fixed by the condition

$$\mathbb{E}\left[W(t,\omega)W(t,-\omega')\right]=\begin{cases}|t| & \text{if } \omega=\omega',\\ 0 & \text{otherwise.}\end{cases}$$

We then set ξ to be the time derivative of the cylindrical Wiener process

$$(t,x)\mapsto\sum_{\omega\in\mathbb{Z}^d}W(t,\omega)e^{2i\pi\omega\cdot x}.$$

More precisely, for every $\varphi\in L^2(\mathbb{R}\times\mathbb{T}^d)$, we set

$$\xi(\varphi):=\sum_{\omega\in\mathbb{Z}^d}\int_{t=-\infty}^{+\infty}\widehat{\varphi}(t,\omega)\,\mathrm{d}W(t,\omega),$$

where the integral is interpreted in Itô's sense, and the notation $\widehat{\varphi}(t,\omega)$ stands for $\widehat{\varphi(t,\cdot)}(\omega)$. By Itô's and Fourier's isometries, this expression is well-defined and the relation (31) is satisfied. Since $\xi(\varphi)$ is also a centered Gaussian, this provides us with a construction of white noise. We use the somewhat informal notation

$$\xi(\varphi):=\int_{\mathbb{R}\times\mathbb{T}^d}\varphi(z)\,\xi(\mathrm{d}z)\ ,$$

although ξ is almost surely not a measure. In particular, for a given $\varphi\in L^2(\mathbb{R}\times\mathbb{T}^d)$, the random variable $\xi(\varphi)$ is only defined outside of a set of measure zero, and a priori this set depends on the choice of φ.

As explained for example in [18, Chap. 9] or [23, Sect. 1.1.2], we can define iterated Wiener-Itô integrals based on ξ. For each $k\geq 1$ and $\varphi\in L^2((\mathbb{R}\times\mathbb{T}^d)^k)$, we denote the iterated integral of φ by

$$\xi^{\otimes k}(\varphi)=\int_{(\mathbb{R}\times\mathbb{T}^d)^k}\varphi(z_1,\ldots,z_k)\,\xi(\mathrm{d}z_1)\cdots\xi(\mathrm{d}z_k).$$

Denoting by $\widetilde{\varphi}$ the symmetrized function obtained from φ:

$$\widetilde{\varphi}(z_1,\ldots,z_k):=\frac{1}{k!}\sum_{\sigma\in S_k}\varphi(z_{\sigma(1)},\ldots,z_{\sigma(k)}),\tag{32}$$

where S_k denotes the permutation group over $\{1, \dots, k\}$, we have

$$\xi^{\otimes k}(\varphi) = \xi^{\otimes k}(\widetilde{\varphi}) = k! \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}(z_1, \dots, z_k)\, \mathbf{1}_{\{t_1<\cdots<t_k\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_k), \quad (33)$$

where t_i is the time component of z_i. Moreover, we have the isometry property

$$\mathbb{E}\left[\xi^{\otimes k}(\varphi)^2\right] = \mathbb{E}\left[\xi^{\otimes k}(\widetilde{\varphi})^2\right] = \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}^2(z_1, \dots, z_k)\, \mathrm{d}z_1 \cdots \mathrm{d}z_k, \quad (34)$$

and by Jensen's inequality,

$$\int_{(\mathbb{R}\times\mathbb{T}^d)^k} \widetilde{\varphi}^2(z_1, \dots, z_k)\, \mathrm{d}z_1 \cdots \mathrm{d}z_k \le \int_{(\mathbb{R}\times\mathbb{T}^d)^k} \varphi^2(z_1, \dots, z_k)\, \mathrm{d}z_1 \cdots \mathrm{d}z_k. \quad (35)$$

Assuming now for notational convenience that φ is a symmetric function, that is, $\varphi = \widetilde{\varphi}$, we may rewrite the expression (33) for the iterated integral $\xi^{\otimes k}(\varphi)$ as a series of finite-dimensional iterated Wiener-Itô integrals:

$$\xi^{\otimes k}(\varphi) = k! \sum_{\omega_1,\dots,\omega_k\in\mathbb{Z}^d} \int_{t_1<\cdots<t_k} \widehat{\varphi}(t_1, \omega_1, \dots, t_k, \omega_k)\, \mathrm{d}W(t_1, \omega_1) \cdots \mathrm{d}W(t_k, \omega_k), \quad (36)$$

where

$$\widehat{\varphi}(t_1, \omega_1, \dots, t_k, \omega_k) := \int_{(\mathbb{T}^d)^k} \varphi(t_1, x_1, \dots, t_k, x_k) e^{-2i\pi(\omega_1\cdot x_1+\cdots+\omega_k\cdot x_k)}\, \mathrm{d}x_1 \cdots \mathrm{d}x_k.$$

See also [18, Sect. 9.6] for a definition of iterated integrals. We let

$$\mathcal{H}_k := \{\xi^{\otimes k}(\varphi),\ \varphi \in L^2((\mathbb{R}\times\mathbb{T}^d)^k)\}$$

denote the k-th *Wiener chaos*, with $\mathcal{H}_0 = \mathbb{R}$. Denote by $(\Omega, \mathcal{F}, \mathbb{P})$ the probability space on which ξ is defined. The spaces $\mathcal{H}_k$ are orthogonal in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, although we will not use it, we recall that if $\mathcal{F}$ is the σ-algebra generated by $\mathcal{H}_1$, then

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \bigoplus_{k=0}^{+\infty} \mathcal{H}_k$$

(the interested reader may find this result in [18, Sect. 9.5]). The more important property for our purpose is the following.

Lemma 3 *For each $n \in \mathbb{N}$, the closure in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ of the linear span of the set*

$$\left\{\xi(\varphi_1) \cdots \xi(\varphi_k),\ k \le n,\ \varphi_1, \dots, \varphi_k \in L^2(\mathbb{R}\times\mathbb{T}^d)\right\}. \quad (37)$$

coincides with

$$\mathcal{H}_{\leq n} := \bigoplus_{k=0}^{n} \mathcal{H}_k. \tag{38}$$

Sketch of proof Let $H_n = H_n(X, T)$ be the Hermite polynomials, defined recursively by

$$\begin{cases} H_0 = 1, \\ H_n = X H_{n-1} - T\, \partial_X H_{n-1} \qquad (n \in \mathbb{N}), \end{cases} \tag{39}$$

so that $H_1 = X$, $H_2 = X^2 - T$, $H_3 = X^3 - 3XT$, etc. By a recursive application of Itô's formula, see [18, Theorem 9.6.9], we have, for every $\varphi \in L^2(\mathbb{R} \times \mathbb{T}^d)$ and $n \in \mathbb{N}$,

$$\xi^{\otimes n}(\varphi^{\otimes n}) = H_n(\xi(\varphi), \|\varphi\|^2_{L^2(\mathbb{R}\times\mathbb{T}^d)}). \tag{40}$$

This relation already shows that every n-fold iterated integral is a linear combination of elements of (37). Conversely, it also shows that every Hermite polynomial in $\xi(\varphi)$ of degree at most n — and therefore every polynomial in $\xi(\varphi)$ of degree at most n — belongs to $\mathcal{H}_{\leq n}$. The full proof of Lemma 3 can be derived from [18, Theorem 9.5.4].

By extension, we say that a stochastic process $\tau : \mathbb{R} \to \mathcal{S}'(\mathbb{T}^d)$ belongs to $\mathcal{H}_n$ (resp. $\mathcal{H}_{\leq n}$) if for every t and smooth test function ϕ, we have

$$\langle \tau(t), \phi \rangle \in \mathcal{H}_n \text{ (resp. } \mathcal{H}_{\leq n}).$$

If $\tau(t)$ is in fact a continuous function of the space variable, this boils down to asking that $\tau(t, x) \in \mathcal{H}_n$ (resp. $\mathcal{H}_{\leq n}$) for each $x \in \mathbb{T}^d$. By Lemma 3, the approximations to the diagrams we want to construct, see (7), all belong to $\mathcal{H}_{\leq 5}$. Since Wiener chaoses are closed, this remains true of their candidate limits (see also Sect. 4 for explicit representations).

Since $\delta_k f$ is linear in f for every k, the fact that f belongs to some Wiener chaos implies that $\delta_k f$ belongs to the Wiener chaos of the same order. Thus, in view of Proposition 3, the possibility to estimate arbitrarily high moments of elements of a fixed Wiener chaos from their L^2 moments will be very convenient.

Proposition 4 (Nelson's estimate) *For every $n \geq 1$ and $p \in [2, \infty)$, there exists a constant $C < \infty$ such that for every $X \in \mathcal{H}_{\leq n}$,*

$$\mathbb{E}\left[|X|^p\right]^{\frac{1}{p}} \leq C \mathbb{E}\left[X^2\right]^{\frac{1}{2}}.$$

If $X \in \mathcal{H}_n$, then we can take $C = (p-1)^{\frac{n}{2}}$.

We now give a proof of Proposition 4 based on the Burkholder-Davis-Gundy inequality ([28, Sect. 4.4]). The appendix contains an alternative argument based on the logarithmic Sobolev inequality.

Lemma 4 (BDG inequality) *Let $p \in (0, \infty)$. There exists $C < \infty$ such that if $(M_t)_{t \geq 0}$ is a continuous local martingale starting from 0, then*

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} |M_s|^p\right] \leq C\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right],$$

where $\langle M \rangle$ denotes the quadratic variation of M.

Proof (*First proof of Proposition* 4) We show the result by induction on n. For $n = 0$, the space $\mathcal{H}_0$ only contains constants, so the result is obvious. We now fix $n \geq 1$. We need to verify the property for random variables of the form $\xi^{\otimes n}(\varphi)$, $\varphi \in L^2((\mathbb{R} \times \mathbb{T}^d)^n)$. By (33), we may assume that φ is symmetric in its variables. Let $\mathcal{F}_t$ be the σ-algebra generated by

$$\left\{\xi(\varphi),\ \operatorname{Supp} \varphi \subseteq (-\infty, t] \times \mathbb{T}^d\right\}.$$

The process

$$M_t := \int_{(\mathbb{R}\times\mathbb{T}^d)^n} \varphi(z_1, \ldots, z_n)\, \mathbf{1}_{\{t_1 < \cdots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_n) \qquad (t \in \mathbb{R})$$

is an $(\mathcal{F}_t)$-martingale. This can be justified either by approximation of φ by elementary functions which vanish on the diagonal (see [23, (1.10)], and take the A_i there of product form), or by appealing to the representation (36). Moreover,

$$\langle M \rangle_t = \int_{\mathbb{R}\times\mathbb{T}^d} \left(\int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi(z_1, \ldots, z_n)\, \mathbf{1}_{\{t_1 < \cdots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_{n-1}) \right)^2 \mathrm{d}z_n.$$

By Minkowski's triangle inequality for the exponent $\frac{p}{2} \geq 1$,

$$\begin{aligned} &\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right]^{\frac{2}{p}} \\ &\leq \int_{\mathbb{R}\times\mathbb{T}^d} \mathbb{E}\left[\left(\int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi(z_1, \ldots, z_n)\, \mathbf{1}_{\{t_1 < \cdots < t_n < t\}}\, \xi(\mathrm{d}z_1) \cdots \xi(\mathrm{d}z_{n-1}) \right)^p \right]^{\frac{2}{p}} \mathrm{d}z_n. \end{aligned}$$

By symmetrizing φ as in (32), the induction hypothesis and (34), we infer that

$$\mathbb{E}\left[\langle M \rangle_t^{\frac{p}{2}}\right]^{\frac{2}{p}} \leq C \int_{\mathbb{R}\times\mathbb{T}^d} \int_{(\mathbb{R}\times\mathbb{T}^d)^{n-1}} \varphi^2(z_1, \ldots, z_n)\, \mathbf{1}_{\{t_1 < \cdots < t_n < t\}}\, \mathrm{d}z_1 \cdots \mathrm{d}z_{n-1}\, \mathrm{d}z_n.$$

The conclusion then follows from the Lemma 4, used with $t = +\infty$.

Remark 7 As hinted at in the introduction, certain Ising-type Markov processes converge (or are expected to converge) to the Φ^4 model (see [20]). The proof of Nelson's estimate based on the BDG inequality is sufficiently robust to allow for

a generalization to these Markov processes. Indeed, loosely speaking, a Markov process can be thought of as an evolution equation with a random forcing that is white in time. In more precise words, a Markov process comes with a martingale structure indexed by time, and the possibly surprisingly special role played by the time variable in the proof we presented above becomes very natural in this context. Using versions of Itô's formula and the BDG inequality for processes with jumps ([20, Appendix C]), one can show that (40) still holds approximately ([20, Proposition 5.3]) and prove a version of Nelson's estimate ([20, Lemma 4.1]) by following essentially the same reasoning as above.

With Nelson's estimate, we are now ready to provide a simple criterion to check the main convergence result in Theorem 1. Since for most of the processes defined in (7), their limits can be characterised explicitly without referring to the limiting procedure as $n \to +\infty$ (and in the cases when the limiting procedure is necessary, it is also obvious what the limit should be), we only give detailed characterisations of the limiting processes τ's themselves. Once all the properties of the limits are well understood, the convergence does not pose any further problem.

Proposition 5 *Let $n \in \mathbb{N}$, and let $\tau : \mathbb{R} \to \mathcal{S}'(\mathbb{T}^d)$ be a random process in $\mathcal{H}_{\leq n}$ which is stationary in space, in the sense that for every $x \in \mathbb{T}^d$,*

$$\text{the processes } (\tau(t,\cdot))_{t\in\mathbb{R}} \text{ and } (\tau(t,x+\cdot))_{t\in\mathbb{R}} \text{ have the same law.} \tag{41}$$

Let $(\widehat{\tau}(t,\omega))_{\omega\in\mathbb{Z}^d}$ denote the Fourier coefficients of $\tau(t)$. If for some $t \in \mathbb{R}$, there exists $C < \infty$ and $\alpha \in \mathbb{R}$ such that for every $\omega \in \mathbb{Z}^d$,

$$\mathbb{E}\left[|\widehat{\tau}(t,\omega)|^2\right] \leq C(1+|\omega|)^{-d-2\alpha}, \tag{42}$$

then for every $\beta < \alpha$, we have $\tau(t) \in \mathcal{C}^\beta(\mathbb{T}^3)$, and moreover,

$$\mathbb{E}\left[\|\tau(t)\|_{\mathcal{C}^\beta}^p\right] < +\infty. \tag{43}$$

If, in addition to (42), there exists $\lambda \in (0,1)$ such that

$$\mathbb{E}\left[|\widehat{\tau}(t,\omega)-\widehat{\tau}(s,\omega)|^2\right] \leq C|t-s|^\lambda(1+|\omega|)^{-d-2\alpha+2\lambda} \tag{44}$$

uniformly in $0 < |t-s| < 1$ and $\omega \in \mathbb{Z}^d$, then for every $\beta < \alpha - \lambda$, we have $\tau \in C(\mathbb{R}, \mathcal{C}^\beta(\mathbb{T}^3))$, and moreover,

$$\sup_{0<|t-s|<1} \frac{\mathbb{E}\left[\|\tau(t)-\tau(s)\|_{\mathcal{C}^\beta}^p\right]}{|t-s|^{\frac{\lambda p}{2}}} < +\infty. \tag{45}$$

Proof Step 1. We first show that by the stationarity assumption (41), for every $\omega, \omega' \in \mathbb{Z}^d$,

$$\omega + \omega' \neq 0 \quad \Longrightarrow \quad \mathbb{E}\left[\widehat{\tau}(s,\omega)\widehat{\tau}(t,\omega')\right] = 0. \tag{46}$$

Indeed, we have, using a slightly informal integral-sign notation,

$$\begin{aligned}\mathbb{E}\left[\widehat{\tau}(s,\omega)\widehat{\tau}(t,\omega')\right] &= \iint_{(\mathbb{T}^d)^2} \mathbb{E}[\tau(s,x)\tau(t,y)]e^{-2\pi i(\omega\cdot x+\omega'\cdot y)}\,dx\,dy \\ &= \iint_{(\mathbb{T}^d)^2} \mathbb{E}[\tau(s,x)\tau(t,y)]e^{-2\pi i\left[(\omega+\omega')\cdot x+\omega'\cdot(y-x)\right]}\,dx\,dy.\end{aligned}$$

By the stationarity assumption, the expectation above is a function of $(y-x)$ only. Integrating in y first and then in x, we therefore obtain (46).

Step 2. We now focus on the proof of (45); the proof of (43) is only simpler. Let $\tau_{s,t} := \tau(t) - \tau(s)$. We have

$$(\delta_k\tau_{s,t})(x) = \sum_{\omega} \chi_k(\omega)\widehat{\tau}_{s,t}(\omega)e^{2\pi i\omega\cdot x}.$$

We implicitly assume here that the processes under consideration are real-valued, and therefore that for every $\omega\in\mathbb{Z}^d$,

$$\widehat{\tau}(t,-\omega) = \overline{\widehat{\tau}(t,\omega)}.$$

Since χ_k is an even function, we deduce that

$$\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right] = \sum_{\omega,\omega'\in\mathbb{Z}^d} \chi_k(\omega)\chi_k(\omega')\mathbb{E}\left[\widehat{\tau}_{s,t}(\omega)\widehat{\tau}_{s,t}(\omega')\right]e^{2\pi i(\omega+\omega')\cdot x}.$$

Using (46) and the bound (44), we get

$$\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right] = \sum_{\omega\in\mathbb{Z}^d} |\chi_k(\omega)|^2\mathbb{E}\left[|\widehat{\tau}_{s,t}(\omega)|^2\right] \lesssim |t-s|^{\lambda}2^{-2k(\alpha-\lambda)}.$$

The above bound holds uniformly in $k\geq -1$, $0<|t-s|<1$ and $x\in\mathbb{T}^d$. Also, since τ belongs to $\mathcal{H}_{<n}$, Nelson's estimate implies

$$\begin{aligned}\mathbb{E}\left[\|\delta_k\tau_{s,t}\|^p_{L^p}\right] &\leq \sup_{x\in\mathbb{T}^d}\mathbb{E}\left[\left|(\delta_k\tau_{s,t})(x)\right|^p\right] \\ &\lesssim \sup_{x\in\mathbb{T}^d}\left(\mathbb{E}\left[|(\delta_k\tau_{s,t})(x)|^2\right]\right)^{\frac{p}{2}} \\ &\lesssim |t-s|^{\frac{\lambda p}{2}}2^{-kp(\alpha-\lambda)}.\end{aligned}$$

By Proposition 3, we deduce that for each $\beta<\alpha-\lambda-\frac{d}{p}$,

$$\mathbb{E}\left[\|\tau_{s,t}\|^p_{\mathcal{C}^\beta}\right] \lesssim |t-s|^{\frac{\lambda p}{2}},$$

uniformly over all $0<|t-s|<1$.

4 Construction of the Diagrams

We are now ready to construct the diagrams listed in Table 1, and prove the relevant bounds appearing in Theorem 1. We will focus on the bound (10) for fixed times, and will only briefly discuss in the next section how the continuity in time follows from there. We also omit the detailed proof of the convergence (9), since once the bounds (10) for the limit diagrams are established, the convergence of approximations follows in essentially the same way. Since all our processes τ belong to $\mathcal{H}_{\leq 5}$ and are stationary both in space and time, we can invoke Proposition 5 and reduce the proof of (10) to showing the second moment bound (42) for each τ.

The derivation of these bounds involves the estimation of several nested integrals and sums. We use a graphical notation to represent these operations. This has the advantage of making the manipulations with potentially very long expressions shorter and more transparent. Naturally, the price to pay is to get used to the notation. One of the aims of these notes is to convince the reader that this investment is worth their while.

The graphical notation is heavily inspired by the treatment of the stochastic terms in [15, Sect. 10]. One difference is that there the calculations are performed in "real space", while we prefer to work with the spatial Fourier transform, and to keep the time variable fixed. It turns out that despite this change, the graphs we encounter in our approach are very similar to the ones in [15, Sect. 10] — only the interpretation changes slightly. Another difference is that we work with resonant parts of products, while Hairer considers increments of processes. We comment on this difference below (see also [14, Sect. 5] for an earlier graphical approach to bounding stochastic quantities which are represented using the spatial Fourier transform).

The presentation is separated into two parts. We first show how to represent the various processes as iterated stochastic integrals, and then derive the bounds on these.

4.1 Iterated Integral Representation

We start by showing how all of the stochastic terms τ in Table 1 can be represented as sums of iterated stochastic integrals. As stated above, each of the terms τ is an element of the inhomogeneous Wiener chaos $\mathcal{H}_{\leq 5}$, and this sum yields the explicit decomposition into its components in the homogeneous Wiener chaoses $\mathcal{H}_0, \dots, \mathcal{H}_5$. This representation as iterated integrals reduces the proof of the required moment bounds to an application of the isometry property (34). As will be shown below, this representation also makes the choice of infinite renormalisation constants transparent.

Case $\tau = \bullet$. We take the simplest process $\tau = \bullet$, the solution to the stochastic heat equation (3), as the starting point of our discussion. For each $t \in \mathbb{R}$ and $\omega \in \mathbb{Z}^3$, we can write

$$\widehat{\bullet}(t,\omega) = \int_{u=-\infty}^{t} \widehat{P}_{t-u}(\omega)\,\mathrm{d}W(u,\omega), \tag{47}$$

where $(W(\cdot,\omega))_\omega$ is the family of complex valued Brownian motions introduced in Sect. 3, and, for $t \geq 0$, $\widehat{P}_t$ is the Fourier transform of the heat kernel for $P_t = e^{t(\Delta-1)}$, that is,

$$\widehat{P}_t(\omega) = e^{-t(1+4\pi^2|\omega|^2)} = e^{-t\langle\omega\rangle^2}, \tag{48}$$

where we set

$$\langle\omega\rangle := \sqrt{1+4\pi^2|\omega|^2}$$

for concision. It is convenient to extend P_t and $\widehat{P}_t$ to every time $t \in \mathbb{R}$, by setting

$$\text{for every } t < 0, \qquad P_t \equiv 0 \quad \text{and} \quad \widehat{P}_t \equiv 0, \tag{49}$$

so that (47) can be rewritten as

$$\widehat{\bullet}(t,\omega) = \int_{u\in\mathbb{R}} \widehat{P}_{t-u}(\omega)\,\mathrm{d}W(u,\omega).$$

In future expressions of integrals against $\mathrm{d}W(u,\omega)$, we always understand that the variable ω is fixed, and that the variable u is the variable of integration. For instance, we simply write

$$\widehat{\bullet}(t,\omega) = \int_{\mathbb{R}} \widehat{P}_{t-u}(\omega)\,\mathrm{d}W(u,\omega). \tag{50}$$

The graphical version of (47) or (50) is

$\widehat{\bullet}(t,\omega) =$  . (51)

Here the root ● represents the pair (t,ω), i.e., the time and frequency at which we seek to evaluate $\widehat{\bullet}$. The leaf ○ represents an instance of the noise $dW(u,\omega)$, and the line connecting them is the kernel $\widehat{P}_{t-u}$. The time variable u associated to the node ○ is integrated out.

Case $\tau = \bullet\!\!\bullet$. We now proceed to represent the process $\tau = \bullet\!\!\bullet$, the limit of $\bullet\!\!\bullet_n := (\bullet_n)^2 - \mathbf{c}_n$. We start with the product $\bullet_n^2$, writing

$$\begin{aligned}\widehat{\bullet_n^2}(t,\omega) &= \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\leq n}} \widehat{\bullet}(t,\omega_1)\widehat{\bullet}(t,\omega_2)\\ &= \sum_{\substack{\omega_1+\omega_2=\omega\\ |\omega_i|\leq n}} \Big(\int_{-\infty}^{t} \widehat{P}_{t-u_1}(\omega_1)\,\mathrm{d}W(u_1,\omega_1)\Big)\Big(\int_{-\infty}^{t} \widehat{P}_{t-u_2}(\omega_2)\,\mathrm{d}W(u_2,\omega_2)\Big)\end{aligned}$$

$$= \sum_{\substack{\omega_1+\omega_2=\omega \\ |\omega_i|\le n}} \Big(2 \int_{-\infty}^{t} \Big[\int_{-\infty}^{u_1} \widehat{P}_{t-u_1}(\omega_1) \widehat{P}_{t-u_2}(\omega_2) \, \mathrm{d}W(u_2, \omega_2) \Big] \mathrm{d}W(u_1, \omega_1) + \mathbf{1}_{\{\omega_1=-\omega_2\}} \int_{-\infty}^{t} \widehat{P}_{t-u}(\omega_1) \widehat{P}_{t-u}(\omega_2) \, \mathrm{d}u \Big), \tag{52}$$

where the last equality follows from Itô's formula. The last term on the right side vanishes for $\omega \neq 0$, because in this case the conditions $\omega_1 + \omega_2 = \omega$ and $\omega_1 = -\omega_2$ are incompatible. For $\omega = 0$, the sum of these terms can be rewritten as

$$\sum_{|\omega_1|\le n} \int_{-\infty}^{t} |\widehat{P}_{t-u}(\omega_1)|^2 \, \mathrm{d}u = \sum_{|\omega_1|\le n} \frac{1}{2\langle \omega_1 \rangle^2}.$$

This is precisely the term $\mathbf{c}_n := \mathbb{E}\big[(\text{🙂}_n(t))^2\big]$, which is of the order of n as n goes to infinity, and is removed in the renormalisation procedure. Below we will show that we can pass to the limit $n \to \infty$ in the first term on the right side of (52). We denote the limit by $\widehat{\text{🙂}}(t, \omega)$. It is instructive to translate the expressions back into "real space", and to check that in the notation introduced in Sect. 3, we have

$$\begin{aligned} \widehat{\text{🙂}}(t, \omega) &= \sum_{\omega_1+\omega_2=\omega} 2 \Big(\int_{-\infty}^{t} \Big[\int_{-\infty}^{u_1} \widehat{P}_{t-u_1}(\omega_1) \widehat{P}_{t-u_2}(\omega_2) \, \mathrm{d}W(u_2, \omega_2) \Big] \mathrm{d}W(u_1, \omega_1) \Big) \\ &= \int_{(\mathbb{R}\times\mathbb{T}^3)^2} \int_{\mathbb{T}^3} P_{t-u_1}(y - x_1) P_{t-u_2}(y - x_2) e^{-i2\pi\omega\cdot y} \, \mathrm{d}y \, \xi(\mathrm{d}u_1, \mathrm{d}x_1) \xi(\mathrm{d}u_2, \mathrm{d}x_2), \end{aligned} \tag{53}$$

where we identify the operator P_t with its kernel, which we interpret as being null for $t \le 0$. This expression shows in particular that $\widehat{\text{🙂}}(t, \omega)$ is an element of the homogeneous Wiener chaos $\mathcal{H}_2$ as defined in Sect. 3.

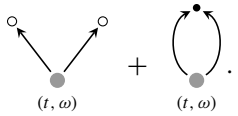
By the definition of iterated stochastic integrals, we may rewrite the identity in (52) as

$$\widehat{\text{🙂}}^2_n(t, \omega) = \sum_{\substack{\omega_1+\omega_2=\omega \\ |\omega_i|\le n}} \Big(\int_{-\infty}^{t} \int_{-\infty}^{t} \widehat{P}_{t-u_1}(\omega_1) \widehat{P}_{t-u_2}(\omega_2) \, \mathrm{d}W(u_2, \omega_2) \, \mathrm{d}W(u_1, \omega_1) + \mathbf{1}_{\{\omega_1=-\omega_2\}} \int_{-\infty}^{t} \widehat{P}_{t-u}(\omega_1) \widehat{P}_{t-u}(\omega_2) \, \mathrm{d}u \Big),$$

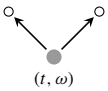
or, with our convention (49),

$$\widehat{\text{🙂}}^2_n(t, \omega) = \sum_{\substack{\omega_1+\omega_2=\omega \\ |\omega_i|\le n}} \Big(\int_{\mathbb{R}^2} \widehat{P}_{t-u_1}(\omega_1) \widehat{P}_{t-u_2}(\omega_2) \, \mathrm{d}W(u_2, \omega_2) \, \mathrm{d}W(u_1, \omega_1) + \mathbf{1}_{\{\omega_1=-\omega_2\}} \int_{-\infty}^{t} \widehat{P}_{t-u}(\omega_1) \widehat{P}_{t-u}(\omega_2) \, \mathrm{d}u \Big). \tag{54}$$

As the reader can see, this expression is already quite bulky — and much worse is to come. This motivates to introduce a graphical notation which encodes these expressions in a much more transparent way, and which we will be able to manipulate directly. At this stage, we disregard the truncation, and perform all calculations in the formal limit $n = \infty$. The expression (54) then becomes

$\widehat{\bullet^2}(t,\omega) =$  .

As before the root ● of the graph represents the pair (t, ω), and each of the leaves ○ represents one occurrence of the white noise and carries a pair (u_i, ω_i) itself. The kernel $\widehat{P}_t$ is represented by the arrow connecting the nodes evaluated at time $(t - u_i)$, and the arrow points towards the node whose time variable is "earlier". Then all time variables u_i except for the one t at the node are integrated out and the ω_i are summed over, subject to $\omega_1 + \omega_2 = \omega$. We will see below that this last rule corresponds to Kirchhoff's law that the "ingoing" variable ω must coincide with the sum of all "outgoing" ω_i's. The second graph on the right side is obtained by "contracting" the two nodes of the first graph. In this second graph, we may associate a frequency ω_1 to the left arrow, and a frequency ω_2 to the right arrow. Kirchhoff's law for the bottom node then imposes $\omega = \omega_1 + \omega_2$, and for the top node, $\omega_1 + \omega_2 = 0$, and we recover that this term is zero unless $\omega = 0$. This contracted graph is removed in the renormalisation procedure, so that

$\widehat{\bullet\bullet}(t,\omega) =$  . (55)

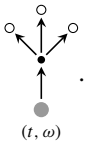
Case $\tau = \bullet\bullet$. We now discuss the next term $\bullet\bullet\bullet$, which as announced arises as the limit of $\bullet\bullet\bullet_n := I\left((\bullet_n)^3 - 3\mathbf{c}_n \bullet_n\right)$. As above in (52) (see also (40) with $n = 3$) we can use Itô's formula to obtain an iterated integral representation for $\widehat{\bullet^3_n}$, which takes the form

$$\widehat{\bullet^3_n}(t,\omega) = \sum_{\substack{\omega_1+\omega_2+\omega_3=\omega \\ |\omega_i|\le n}} 6 \int_{-\infty}^{t}\int_{-\infty}^{u_1}\int_{-\infty}^{u_2} \widehat{P}_{t-u_1}(\omega_1)\,\widehat{P}_{t-u_2}(\omega_2)\,\widehat{P}_{t-u_3}(\omega_3)\, \mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_2,\omega_2)\,\mathrm{d}W(u_1,\omega_1) + 3\mathbf{c}_n \widehat{\bullet}_n(t,\omega). \tag{56}$$

We had already seen above that $\mathbf{c}_n$ diverges as n goes to ∞, which motivates to remove the second term in the renormalisation procedure. The reason why we choose to work with the integrated object $\bullet\bullet\bullet$ rather than $\bullet\bullet$ is that (as we will see below) although the latter can be defined as a space time distribution, it cannot be evaluated at any fixed t.

(This is similar to temporal white noise which can also only be interpreted when tested against a function of time and space, but never pointwise in t). In fact, this is an instance of the well-known fact that Wick powers of order ≥ 3 over the three dimensional Gaussian free field do not exist (since the covariance function would not be integrable; see e.g. [9, Sect. 2.7]), and for readers familiar with this fact, it may be surprising that Wick powers up to order 4 can be constructed as space-time objects. Using a similar graphical notation to the one used for Ξ, we arrive at

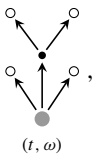
$$\widehat{\Xi}(t,\omega) =$$



.

Case $\tau = \Xi\Xi$. At this point we want to start to think more systematically about the derivation of the diagrammatic expressions and their interpretation, and we illustrate this with the diagram $\Xi\Xi$. This particular expansion follows from an iterated application of Itô's formula, see [23, Propositions 1.1.2 and 1.1.3] and [15, Sect. 10].

For the moment we ignore the additional complexity introduced by the resonant products $\ominus$ in (7), and give a graphical representation for $\Xi\Xi$ defined as in (7) with $\ominus$ replaced by the usual product. To begin with, we give the graphical representation of this symbol without taking into account the renormalisation procedure, i.e. we work with the random function $I(\uparrow_n^2)\uparrow_n^2$. This random function takes values in $\mathcal{H}_{\leq 4}$, with components in $\mathcal{H}_4$, $\mathcal{H}_2$ and $\mathcal{H}_0$. The component in the highest Wiener chaos $\mathcal{H}_4$ is represented by the graph



, (57)

i.e. precisely the graph we use as a symbol to represent this object. This graph can be interpreted as random variable either in "real space" coordinates or in "Fourier coordinates". Both interpretations are equivalent and the former is closer in spirit to [15, Sect. 10] and also (53) above, while the latter is closer to the spirit of the present notes. Here we present both interpretations, starting with the "real space" interpretation because it is slightly easier to explain. The Fourier interpretation then follows by turning multiplication into convolution in the space coordinates: For the "real space" interpretation we assign a space-time point to each of the vertices of this graph (e.g. $(u_1, x_1), \ldots, (u_4, x_4)$ to the four leaves, (u_5, x_5) to the internal vertex, and (t, x_6) to the root), an instance of the heat kernel P evaluated at the difference of the variables associated to the adjacent vertices and the arrow pointing towards the "earlier" time variable to each of the arrows (e.g. $P_{u_5-u_1}(x_5 - x_1)$ to the upper right arrow), and multiply all of these kernels. Finally, the variable corresponding to

the internal variable (u_5, x_5) is integrated out over space-time, the variables (u_i, x_i) for the leaves are integrated against the white noise ξ and we take the spatial Fourier transform with respect to the variable x_6 at the root, yielding the expression[3]

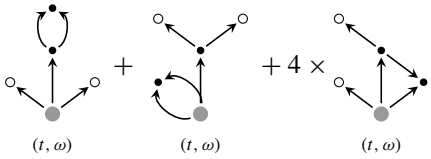
$$\int_{(\mathbb{R}\times\mathbb{T}^3)^4} \Big(\int_{\mathbb{T}^3} \mathrm{d}x_6 \int_{\mathbb{R}\times\mathbb{T}^3} \mathrm{d}u_5\, \mathrm{d}x_5\, P_{u_5-u_1}(x_5 - x_1) P_{u_5-u_2}(x_5 - x_2) \\ \times P_{t-u_5}(x_6 - x_5) P_{t-u_3}(x_6 - x_3) P_{t-u_4}(x_6 - x_4) e^{-i2\pi\omega\cdot x_6} \Big) \\ \xi(\mathrm{d}u_1, \mathrm{d}x_1)\xi(\mathrm{d}u_2, \mathrm{d}x_2)\xi(\mathrm{d}u_3, \mathrm{d}x_3)\xi(\mathrm{d}u_4, \mathrm{d}x_4). \tag{58}$$

Translating the previous expression and interpretation into Fourier variables can be done as follows: each of the vertices is equipped with a time-frequency variable in $\mathbb{R} \times \mathbb{Z}^3$, say $(u_1, \omega_1), \dots, (u_4, \omega_4)$ for the leaves, (u_5, ω_5) for the internal vertex, and (t, ω) for the root. The formula (58) then becomes

$$\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3 \\ \omega_1+\omega_2=\omega_5 \\ \omega_3+\omega_4+\omega_5=\omega}} \int_{\mathbb{R}^4} \Big(\int_{\mathbb{R}} \mathrm{d}u_5\, \widehat{P}_{u_5-u_1}(\omega_1) \widehat{P}_{u_5-u_2}(\omega_2) \widehat{P}_{t-u_5}(\omega_5) \widehat{P}_{t-u_3}(\omega_3) \widehat{P}_{t-u_4}(\omega_4) \Big) \\ dW(u_1, \omega_1)\, \mathrm{d}W(u_2, \omega_2)\, \mathrm{d}W(u_3, \omega_3)\, \mathrm{d}W(u_4, \omega_4), \tag{59}$$

that is, each arrow now corresponds to an instance of $\widehat{P}$, where the time variables stay the same as before, but the difference of the space variables is replaced by the frequency variable corresponding to the top of the arrow. The fact that the product turns into convolution under the Fourier transform is reflected in the "Kirchhoff rule" that at each internal vertex, the sum of "incoming" frequency variables equals the sum of "outgoing" frequency variables. The same rule applies at the root, with the understanding that ω is an "ingoing" frequency.

The terms in the lower order Wiener chaoses $\mathcal{H}_2$ and $\mathcal{H}_0$ arise from Itô's formula, in the same spirit to our discussion of $\Vvee$ above. For $\mathcal{H}_2$ we get



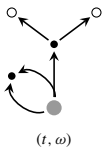
These are precisely the graphs that can be obtained by picking a pair of leaves in (57) and "gluing" them together. The pre-factor "4" is combinatorial and corresponds to the fact that there are four different ways of picking one vertex from the "top level" and one from the "bottom level" of the graph, each of which giving rise to the same

[3] Actually, for finite n the heat kernels connecting to the leaves, i.e. $P_{u_5-u_1}$, $P_{u_5-u_2}$, P_{t-u_3} and P_{t-u_4} (but not P_{t-u_5}) should be replaced by the regularised heat kernel $(t, x) \mapsto \sum_{|\omega|\le n} \widehat{P}_t(\omega) e^{i2\pi\omega\cdot x}$. Similarly, in (59) and after we will leave implicit the constraint $|\omega_1|, \dots, |\omega_4| \le n$. Here and below we drop the regularisations for convenience.

iterated stochastic integral. The interpretation for these graphs is the same as before. For instance, for the first of these graphs corresponds to the expression

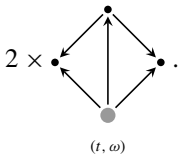
$$\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5\\ \omega_3+\omega_4+\omega_5=\omega\\ \omega_1+\omega_2=0}} \int_{\mathbb{R}^2} \Big(\int_{\mathbb{R}^2} \mathrm{d}u_5\,\mathrm{d}u_1\,\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_1}(\omega_2)\widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_3}(\omega_3)\widehat{P}_{t-u_4}(\omega_4)\Big)$$
$$\mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_4,\omega_4)$$
$$= \mathbf{c}_n \sum_{\substack{\omega_3,\omega_4\in\mathbb{Z}^3\\ \omega_3+\omega_4=\omega}} \int_{\mathbb{R}^2} \widehat{P}_{t-u_3}(\omega_3)\widehat{P}_{t-u_4}(\omega_4)\,\mathrm{d}W(u_3,\omega_3)\,\mathrm{d}W(u_4,\omega_4)$$
$$= \mathbf{c}_n \widehat{\Vee}(t,\omega), \tag{60}$$

which arises from (59) by replacing the white noises $\mathrm{d}W(du_1,\omega_1)\,\mathrm{d}W(du_2,\omega_2)$ from the vertices which are "glued together" by $\delta_0(u_1-u_2)\mathbf{1}_{\omega_1=-\omega_2}$. In the same way we get the identity



$$= \mathbf{c}_n \times \widehat{I(\Vee)}(t,\omega). \tag{61}$$

Finally, the term in the zero-th Wiener chaos (that is, the constant) is given by the only graph that can be obtained from two contractions, namely



(62)

We now move on to discussing the renormalisation of these terms. As the reader can verify, working with $I(\Vee)\Vee$ instead of $I(\bullet^2)\bullet^2$ (i.e. performing the Wick renormalisation of the product $\bullet^2$ as above) corresponds exactly to removing the graphs in (60) and (61). The logarithmic sub-divergence corresponding to $\mathbf{c}'_n$ arises in (62). Indeed, evaluating this expression yields

$$\sum_{\substack{\omega_1,\dots,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5\\ \omega_3+\omega_4+\omega_5=\omega\\ \omega_1+\omega_3=0\\ \omega_2+\omega_4=0}} \int_{\mathbb{R}^3} \mathrm{d}u_5\,\mathrm{d}u_1\,\mathrm{d}u_2\,\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_2}(\omega_2)\widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_1}(\omega_3)\widehat{P}_{t-u_2}(\omega_4)$$

$$= \mathbf{1}_{\{\omega=0\}} \sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \int_{\mathbb{R}^3} \mathrm{d}u_5\,\mathrm{d}u_1\,\mathrm{d}u_2\,\widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{u_5-u_2}(\omega_2) \times \widehat{P}_{t-u_5}(\omega_5)\widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{t-u_2}(-\omega_2)$$

$$= \mathbf{1}_{\{\omega=0\}} \sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \int_{\mathbb{R}} \mathrm{d}u_5\,\widehat{P}_{t-u_5}(\omega_5)\frac{e^{-|t-u_5|\langle\omega_1\rangle^2}}{2\langle\omega_1\rangle^2}\frac{e^{-|t-u_5|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2}$$

$$= \mathbf{1}_{\{\omega=0\}}\frac{1}{4}\sum_{\substack{\omega_1,\omega_2,\omega_5\in\mathbb{Z}^3\\ \omega_1+\omega_2=\omega_5}} \frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^2}\frac{1}{\langle\omega_1\rangle^2+\langle\omega_2\rangle^2+\langle\omega_5\rangle^2}, \tag{63}$$

In the first identity above, we have used the fact that the four restrictions on the ω_i are incompatible unless $\omega = 0$. The fact that the expression vanishes for non-zero ω could also be deduced from the fact that (62) represents the expectation of $I(\varv_n^2)\varv_n^2$, which is constant in space by stationarity. In the second identity, we have made use of the symmetry of $\widehat{P}$ in ω and of the fact that $\int_{\mathbb{R}} \widehat{P}_{u_5-u_1}(\omega_1)\widehat{P}_{t-u_1}(\omega_1)\,\mathrm{d}u_1 = \frac{e^{-|t-u_5|\langle\omega_1\rangle^2}}{2\langle\omega_1\rangle^2}$, as well as the corresponding identity for u_2.

The sum in (63), with cutoffs $|\omega_i| \le n$, diverges logarithmically as n tends to infinity. It is therefore necessary to remove the diagram in (62) in the renormalisation procedure. We arrive at the expression $I(\varv_n)\varv_n - 2\mathbf{c}'_n$,[4] which is already very close to the definition of $\varv\!\!\varv_n$ in (7).

It only remains to re-introduce the resonant product $\circledcirc$ in our construction, in place of the full product. We start by briefly explaining why this is actually necessary, going back to the discussion of product estimates in Sect. 2. In the solution theory of (1), the term $\varv\!\!\varv$ plays the role of a product of $\varv = I(\varv)$ and $\varv$. Now, as we have already discussed at length, $\varv$ is a random function of class $\mathcal{C}^{1-}$, $\varv$ is a random distribution of class $\mathcal{C}^{-1-}$, and products (more specifically the resonant part of products) are not well defined in this regularity class. (Of course, the purpose of the present article is to explain how to define these products as probabilistic limits of renormalised approximations). As we have just seen, a renormalisation procedure can be used to define the products. Yet, it will not improve the regularity of the resulting object (predicted in Table 2), i.e. the product $\varv\varv\varv$ will inherit the bad regularity $\mathcal{C}^{-1-}$ from $\varv$. It is however crucial, both in Hairer's theory of regularity structures and in the theory of paracontrolled distributions, to obtain a bound which reflects the "good"

[4] In fact, the sum represented by the diagram (62) does not coincide exactly with the constant $\mathbf{c}'_n$ as defined in (8) because the latter is defined as the expectation of the resonant product $\mathbb{E}\left[I\left(\varv_n\right)\circledcirc\varv_n(t)\right]$ while the former coincides with $\mathbb{E}\left[I\left(\varv_n\right)\varv_n(t)\right]$. However, as the reader can check, the difference between these constants remains bounded as n tends to infinity.

regularity of Υ, i.e. we need to get a bound of regularity $(-1-)+(1-)=0-$. In Hairer's theory, this is accomplished by working with "increments": there the fundamental object is

$$(\Upsilon(y)-\Upsilon(x))\mathcal{V}(y),$$

which indeed behaves like $(y-x)^{0-}$ as $y\to x$. One key observation in [13] was that the same effect can be obtained by working with the resonant product $\Upsilon \circledcirc \mathcal{V}$, which is of class $\mathcal{C}^{0-}$.

After this short detour, it remains to incorporate the resonant product $\circledcirc$ into the graphical notation. For this we recall from Sect. 2 that for arbitrary f, g (e.g. $\in \mathcal{C}^\infty$) we have

$$f \circledcirc g(x) = \sum_{|k-l|\le 1} \delta_k f(x)\delta_l g(x) = \sum_{\omega_1,\omega_2\in\mathbb{Z}^3} e^{2i\pi(\omega_1+\omega_2)\cdot x}\,\widehat{f}(\omega_1)\widehat{g}(\omega_2) \sum_{|k-l|\le 1}\chi_k(\omega_1)\chi_l(\omega_2).$$

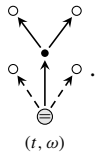
For the last sum appearing in this expression, we have

$$\sum_{|k-l|\le 1}\chi_k(\omega_1)\chi_l(\omega_2) \begin{cases} \in [0,1] & \text{for all } \omega_1,\omega_2 \\ = 0 & \text{if } (|\omega_1|>\tfrac{8}{3} \text{ or } |\omega_2|>\tfrac{8}{3}) \text{ and } \frac{|\omega_1|}{|\omega_2|}\notin [c,c^{-1}], \end{cases} \tag{64}$$

where $c=\frac{9}{64}$. Roughly speaking, and in agreement with the intuition in (15), this term acts as a smooth indicator function, which only selects pairs (ω_1,ω_2) for which $|\omega_1|$ and $|\omega_2|$ are close to one another on the logarithmic scale. This intuition justifies the slightly abusive notation

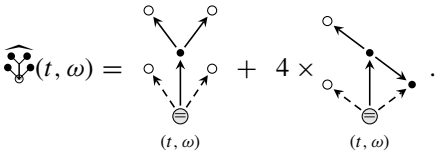
$$\widehat{f\circledcirc g}(\omega) = \sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1\sim\omega_2}} \widehat{f}(\omega_1)\widehat{g}(\omega_2) := \sum_{\omega_1+\omega_2=\omega} \widehat{f}(\omega_1)\widehat{g}(\omega_2) \sum_{|k-l|\le 1}\chi_k(\omega_1)\chi_l(\omega_2). \tag{65}$$

We represent these "restricted convolutions" in the graphical notation by dotted lines. For example, when the product is replaced by $\circledcirc$, the first graph (57) becomes



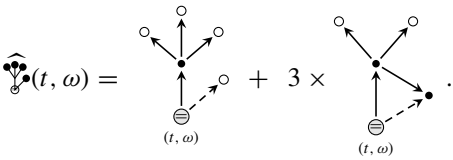
The interpretation of this diagram is the same as in (59), with the only exception that the additional restriction $\{\omega_3+\omega_4\sim\omega_5\}$ is enforced. The convention is that next to a node $\circledcirc$, the sum of the frequency variables corresponding to the dotted arrows is similar to the sum of frequency variables from the regular arrows. Summarising all

of this discussion, we finally arrive at the graphical expression



(66)

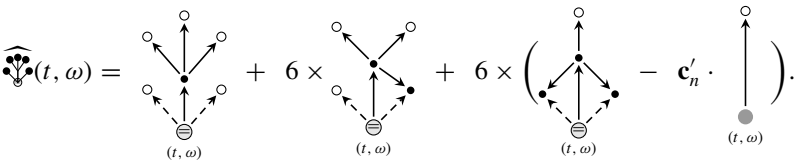
Remaining two diagrams. The graphical representation/decomposition of its components in different Wiener chaoses for the remaining terms follows the same line of reasoning, and we omit the details. For $\tau =$  , we get



(67)

We only mention that as above in the discussion for  , the Wick renormalisation (i.e. the fact that we work with $I($$) \ominus$ rather than $I($  $) \ominus$) corresponds exactly to removing the graphs in the second Wiener chaos which arise by contracting two of the three leaves at the top of the graph. The logarithmic sub-divergence plays no role for this term.

Finally, for $\tau =$  , we have



It is worth pointing out here that for this diagram, the renormalisation with the logarithmically diverging constant $\mathbf{c}'_n$ does not result in the complete removal of a diagram, but only in its modification. In fact, when performing the renormalisation procedure for more complicated equations, it is common that the renormalisation of a graph results in the subtraction of divergent substructures, rather than the removal of the whole graph. It is rather a peculiarity that in the graphs (60) and (61) the removal of the divergent substructure amounts to removing the whole graph. See [7] for a discussion of this point in a much more general framework.

4.2 Bounds on Iterated Integrals

We now proceed to explain how to derive the bound (42) for the various stochastic integrals introduced in the previous subsection. The core ingredient is the isometry identity (34), which permits to bound the second moment of an iterated stochastic integral by the L^2 norm of the corresponding kernel.

The symbols $\tau = \text{𝟙}, \text{𝕍}, \text{𝕍}^{\bullet}$. As before, we first treat the symbol 𝟙. For this symbol, Eq. (47) together with the standard Itô isometry yields

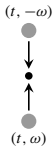
$$\mathbb{E}\big[\widehat{\text{𝟙}}(t,\omega)\,\widehat{\text{𝟙}}(t',-\omega)\big] = \int_{\mathbb{R}} \widehat{P}_{t-u}(\omega)\,\widehat{P}_{t'-u}(-\omega)\,\mathrm{d}u = \frac{e^{-|t-t'|\langle\omega\rangle^2}}{2\langle\omega\rangle^2}, \tag{68}$$

and in particular the bound (42) (for $\alpha = -\frac{1}{2}$) follows from the trivial bound $e^{-|t-t'|\langle\omega\rangle^2} \le 1$. For later use, we record the following immediate corollary of the previous bound: for any $\gamma \ge 0$,

$$\big|\mathbb{E}\big[\widehat{\text{𝟙}}(t,\omega)\,\widehat{\text{𝟙}}(t',-\omega)\big]\big| \lesssim \frac{1}{\langle\omega\rangle^2}\Big(\frac{1}{|t-t'|\langle\omega\rangle^2}\Big)^{\gamma}, \tag{69}$$

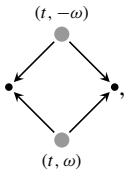
where the implicit constant depends only the choice of γ. In graphical notation, the previous calculation with $t = t'$ becomes

$$\mathbb{E}\big[|\widehat{\text{𝟙}}(t,\omega)|^2\big] \;=\; \lesssim \langle\omega\rangle^{-2}, \tag{70}$$



which can be obtained from (51) by "doubling" the graph and by contracting the leaves $\circ$ to a vertex $\bullet$. The interpretation of the resulting graph then remains the same as in the previous section, i.e. the time variable is integrated out. This algorithm for producing the diagram is very natural: first, taking the square corresponds to the doubling of the graph; and then taking the expectation results in collapsing each pair of leaves represented by a vertex $\circ$ to a single vertex $\bullet$, for each possible pairing; this being due to the trivial covariance structure of the instances of white noise represented by the leaves $\circ$ (in other words, this being due to Itô's isometry). In the same "graphical" way, we obtain the formula

$$\mathbb{E}\big[|\widehat{\text{𝕍}}(t,\omega)|^2\big] \;=\; 2$$



(71)

which should be read as

$$\mathbb{E}\Big[|\widehat{\Psi}(t,\omega)|^2\Big] = 2\sum_{\omega_1+\omega_2=\omega}\int_{\mathbb{R}^2}\Big(\widehat{P}_{t-u_1}(\omega_1)\widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{t-u_2}(\omega_2)\widehat{P}_{t-u_2}(-\omega_2)\Big)\,\mathrm{d}u_1\,\mathrm{d}u_2$$
$$= 2\sum_{\omega_1+\omega_2=\omega}\frac{1}{2\langle\omega_1\rangle^2}\frac{1}{2\langle\omega_2\rangle^2}. \tag{72}$$

Of course, this expression could now be bounded directly, but we prefer a slightly more general approach which will allow us to systematise the calculations to come. The following lemma, which gives a bound on discrete convolutions, is essentially contained in [15, Lemma 10.14]. We formulate it in arbitrary space dimension d, although we are only interested in the case $d = 3$ here.

Lemma 5 *Let $d \geq 1$ and $\alpha, \beta \in \mathbb{R}$ satisfy*

$$\alpha + \beta > d \quad \textit{and} \quad \alpha, \beta < d. \tag{73}$$

We have, uniformly over $\omega \in \mathbb{Z}^d$,

$$\sum_{\substack{\omega_1,\omega_2\in\mathbb{Z}^d\\ \omega_1+\omega_2=\omega}} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta}. \tag{74}$$

Proof We subdivide the index set $\mathcal{A} = \{(\omega_1, \omega_2) \in (\mathbb{Z}^d)^2 \colon \omega_1 + \omega_2 = \omega\}$ of the summation into

$$\mathcal{A}_1 := \Big\{(\omega_1,\omega_2)\in\mathcal{A}\colon \quad |\omega_1| \geq 2|\omega|\Big\}$$
$$\mathcal{A}_2 := \Big\{(\omega_1,\omega_2)\in\mathcal{A}\colon \quad |\omega_1| \leq \frac{1}{2}|\omega|\Big\}$$
$$\mathcal{A}_3 := \Big\{(\omega_1,\omega_2)\in\mathcal{A}\colon \quad |\omega_2| \leq \frac{1}{2}|\omega|\Big\}$$
$$\mathcal{A}_4 := \mathcal{A}\setminus\Big(\bigcup_{j=1}^{3}\mathcal{A}_j\Big),$$

and bound the sums over the individual $\mathcal{A}_j$ separately. For $(\omega_1, \omega_2) \in \mathcal{A}_1$, we make use of the triangle inequality in the form $|\omega_2| = |\omega - \omega_1| \geq |\omega_1| - |\omega| \geq \frac{1}{2}|\omega_1|$ to get

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_1}\langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \sum_{(\omega_1,\omega_2)\in\mathcal{A}_1}\langle\omega_1\rangle^{-\alpha}\left(\frac{1}{2}\langle\omega_1\rangle\right)^{-\beta}$$
$$\lesssim \sum_{|\omega_1|\geq 2|\omega|}\langle\omega_1\rangle^{-\alpha-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

where we have used the first condition in (73). For $\mathcal{A}_2$, we use the triangle inequality in the form $|\omega_2| = |\omega - \omega_1| \geq |\omega| - |\omega_1| \geq \frac{1}{2}|\omega|$ to get

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_2} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \left(\frac{1}{2}\langle\omega\rangle\right)^{-\beta} \sum_{|\omega_1|\leq\frac{1}{2}|\omega|} \langle\omega_1\rangle^{-\alpha} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

where this time we have used the second assumption (on α) in (73). Exchanging the role of ω_1 and ω_2, the same bound follows for the sum over $\mathcal{A}_3$. Finally, on $\mathcal{A}_4$ we have $|\omega_1|, |\omega_2| \geq \frac{1}{2}|\omega|$, so that

$$\sum_{(\omega_1,\omega_2)\in\mathcal{A}_4} \langle\omega_1\rangle^{-\alpha}\langle\omega_2\rangle^{-\beta} \leq \left(\frac{1}{2}\langle\omega\rangle\right)^{-\alpha} \sum_{|\omega_2|\geq\frac{1}{2}|\omega|} \left(\frac{1}{2}\langle\omega_2\rangle\right)^{-\beta} \lesssim \langle\omega\rangle^{d-\alpha-\beta},$$

and the statement follows.

We briefly discuss the conditions (73) on the exponents α, β. The first condition $\alpha + \beta > d$ is necessary to obtain a bound of the type (74), because without it even the convergence of the sum cannot be guaranteed. The second condition $\alpha, \beta < d$ may seem more surprising. It states that a decay beyond summability for $\langle\omega_1\rangle^{-\alpha}$ or $\langle\omega_2\rangle^{-\beta}$ does not improve the behaviour of the convolution. We will see below that this restriction corresponds exactly to the fact that a (renormalised) product of a random function f and a random distribution g cannot have better regularity than g itself. We will show below how the use of the resonant product $\circleddash$ instead of the usual product translates into a Lemma 6, which can be understood as a variant of Lemma 5 for which this restriction is removed.

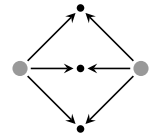
Applying this Lemma to the right side of (72) yields

$$\mathbb{E}\left[|\widehat{\Psi}(t,\omega)|^2\right] \lesssim \frac{1}{\langle\omega\rangle},$$

which is the desired bound (42) with $\alpha = -1$ for this symbol. This implies that this process belongs to C^β for every $\beta < -1$.

We move on to the symbol Ψ. We first discuss why the need for the extra integration against the heat kernel arises. If we tried to work with Ψ, that is to say, to define the limit of $\Psi_n^3 - 3\mathbf{c}_n\Psi_n$ (see (56)), then the same calculation as (71) and (72) would become

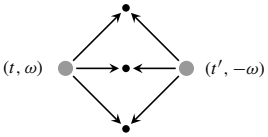
$$\mathbb{E}\left[|\widehat{\Psi}(t,\omega)|^2\right] = 6 \times$$



$$= 6 \sum_{\omega_1+\omega_2+\omega_3=\omega} \frac{1}{2\langle\omega_1\rangle^2}\frac{1}{2\langle\omega_2\rangle^2}\frac{1}{2\langle\omega_3\rangle^2},$$

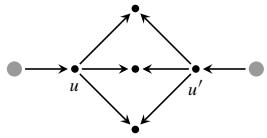
but then Lemma 5 does not apply to this situation (applying the Lemma once for the variables ω_1 and ω_2 would yields the bound $\lesssim \sum_{\tilde{\omega}+\omega_3=\omega}\langle\tilde{\omega}\rangle^{-1}\langle\omega_3\rangle^{-2}$, but then the

resulting exponents $\alpha = 1$ and $\beta = 2$ just fail the summability condition $\alpha + \beta > 3$.) We leave it for the reader to check that indeed, this sum diverges logarithmically for every ω. However, this problem can be fixed by considering different times $t \neq t'$. Then recalling (69) we get for any $\gamma \geq 0$

$$\mathbb{E}\left[\widehat{\Psi}(t,\omega)\widehat{\Psi}(t',-\omega)\right] = 6 \times$$ 

$$\lesssim \frac{1}{|t-t'|^\gamma} \sum_{\omega_1+\omega_2+\omega_3=\omega} \frac{1}{\langle\omega_1\rangle^{2+2\gamma}} \frac{1}{\langle\omega_2\rangle^2} \frac{1}{\langle\omega_3\rangle^2}$$
$$\lesssim \frac{1}{|t-t'|^\gamma \langle\omega\rangle^{2\gamma}},$$

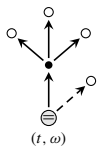
which (for $\gamma < 1$) can be taken as a basis for defining Ψ as a space-time distribution. We prefer, to integrate it once more against a heat kernel (because this is the way it enters the solution theory for (1)), yielding an object which can be evaluated at fixed time. More precisely, for every $\gamma \in (0, 1)$,

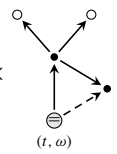
$$\mathbb{E}\left[|\widehat{\Psi}(t,\omega)|^2\right] = 6$$ 

$$\lesssim \int_{\mathbb{R}^2} \widehat{P}_{t-u}(\omega)\widehat{P}_{t-u'}(\omega) \frac{1}{|u-u'|^\gamma \langle\omega\rangle^{2\gamma}} \,\mathrm{d}u\,\mathrm{d}u' \lesssim \frac{1}{\langle\omega\rangle^4}, \tag{75}$$

which proves that this symbol satisfies (42) for $\alpha = \frac{1}{2}$.

Case $\tau = \Psi\,\mathcal{I}(\Psi)$. We now turn to the case $\tau = \Psi\,\mathcal{I}(\Psi^3)$. Recall that its decomposition into components in homogeneous Wiener chaoses was given in (67). We write

$$\widehat{\Psi\,\mathcal{I}(\Psi^3)}(t,\omega) =$$ 

$$+ \; 3 \times$$ 

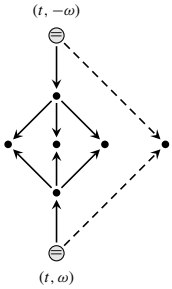
$$=: \widehat{\Psi\,\mathcal{I}(\Psi^3)}^{(4)}(t,\omega) + \widehat{\Psi\,\mathcal{I}(\Psi^3)}^{(2)}(t,\omega). \tag{76}$$

It is then clear that $\widehat{\Psi\,\mathcal{I}(\Psi^3)} \in \mathcal{H}_n$, and as a consequence of the orthogonality of the $\mathcal{H}_n$'s in L^2, we have

$$\mathbb{E}\left[|\widehat{\Psi\,\mathcal{I}(\Psi^3)}(t,\omega)|^2\right] = \mathbb{E}\left[|\widehat{\Psi\,\mathcal{I}(\Psi^3)}^{(4)}(t,\omega)|^2\right] + \mathbb{E}\left[|\widehat{\Psi\,\mathcal{I}(\Psi^3)}^{(2)}(t,\omega)|^2\right]. \tag{77}$$

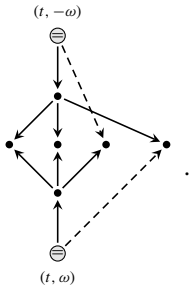
In our graphical notation, the first term above can be bounded by

$$\mathbb{E}\left[\left|\widehat{\Psi}^{(4)}(t,\omega)\right|^2\right] \lesssim$$



. (78)

There is a slightly subtle point worth underlying here: standard Gaussian calculus (the Wick formula) yields an explicit identity for the quantity on the left hand side of this expression, in terms of contractions of all of the different leaves on the original diagram representing $\widehat{\Psi}^{(4)}$. This formula includes additional graphs such as



.

However, using (35) to bound the term on the left side of (78) greatly simplifies the ensuing argument, as opposed to relying on the exact formula involving the asymmetric trees. (This idea was first used in this context in [15, Sect. 10].)

Going back to bounding (78), we use (75) on the "left part" and (70) on the "right part" of the graph to obtain the bound

$$\mathbb{E}\left[\left|\widehat{\Psi}^{(4)}(t,\omega)\right|^2\right] \lesssim \sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1\sim\omega_2}} \frac{1}{\langle\omega_1\rangle^4}\frac{1}{\langle\omega_2\rangle^2}. \tag{79}$$

Lemma 5 on the decay of convolutions is not enough to bound the remaining sum. Indeed, this is precisely a case as discussed below Lemma 5, where the second condition in (73) fails (here, $4 \geq 3$). Of course, the same estimate could be used by "forgetting" some of the good decay of $\langle\omega_1\rangle^{-4}$, and replacing it by $\langle\omega_1\rangle^{-(3-)}$, but this would only yield a bound of order $\langle\omega\rangle^{-(2-)}$ corresponding to a regularity of index $-\frac{1}{2}-$ instead of $0-$. The following lemma shows that the additional condition $\omega_1 \sim \omega_2$, which arises from our use of the resonant product $\ominus$ in the definition of the diagram, resolves this problem.

Lemma 6 *Let $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha + \beta > d$. We have, uniformly over $\omega \in \mathbb{Z}^d$,*

$$\sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1 \sim \omega_2}} \langle \omega_1 \rangle^{-\alpha} \langle \omega_2 \rangle^{-\beta} \lesssim \langle \omega \rangle^{d-\alpha-\beta}.$$

Proof Recall (65) and the definition of the "smooth cut-off" (64). For small $|\omega|$ (say, $|\omega| \leq \frac{16}{3}$) there is only a bounded number of admissible ω_1, ω_2 with $\omega_1 + \omega_2 = \omega$ for which $\sum_{|k-l|\leq 1} \chi_k(\omega_1)\chi_l(\omega_2) \neq 0$. Hence, for such ω, we have

$$\sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1 \sim \omega_2}} \langle \omega_1 \rangle^{-\alpha} \langle \omega_2 \rangle^{-\beta} \lesssim 1.$$

We can thus now assume that $|\omega| > \frac{16}{3}$. For such ω, the conditions $\omega_1 + \omega_2 = \omega$ and $\omega_1 \sim \omega_2$ enforce that $\frac{|\omega_1|}{|\omega_2|} \in [c, c^{-1}]$ for $c = \frac{9}{64}$, by (64). Hence, on the one hand, we have

$$|\omega| \leq |\omega_1| + |\omega_2| \leq |\omega_1| + c^{-1}|\omega_1|,$$

that is, $|\omega_1| \geq \frac{1}{1+c^{-1}}|\omega|$, and on the other hand, $|\omega_2| \geq c|\omega_1|$. These considerations allow us to write

$$\sum_{\substack{\omega_1+\omega_2=\omega \\ \omega_1 \sim \omega_2}} \langle \omega_1 \rangle^{-\alpha} \langle \omega_2 \rangle^{-\beta} \lesssim \sum_{|\omega_1| \geq \frac{1}{1+c^{-1}}|\omega|} \langle \omega_1 \rangle^{-\alpha-\beta} \lesssim \langle \omega \rangle^{d-\alpha-\beta},$$

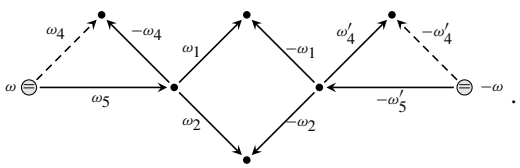
as desired.

Applying this Lemma to (79) immediately yields

$$\mathbb{E}\left[|\widehat{\Psi}^{(4)}(t,\omega)|^2 \right] \lesssim \langle \omega \rangle^{-3}.$$

We now turn to the variance of $\widehat{\Psi}^{(2)}$, which has the expression

$$\mathbb{E}\left[|\widehat{\Psi}^{(2)}(t,\omega)|^2 \right] \lesssim$$



For later reference, we have labelled all the edges with their frequency variables (but dropped the time variables) in the diagram above. In addition to the identities already implicit in this diagram, Kirchhoff's law enforces that we have to sum over the indices $\omega_1, \omega_2, \omega_4, \omega_5, \omega_4', \omega_5'$ satisfying

$$\omega_4 + \omega_5 = \omega, \qquad \omega_4' + \omega_5' = \omega, \qquad \omega_1 + \omega_2 = \omega,$$

so that we ultimately have to sum over three free variables. The fact that we have a resonant product $\ominus$ at the roots yields the additional constraints $\omega_4 \sim \omega_5$ and $\omega_4' \sim \omega_5'$, but we will not rely on these additional constraints to bound this diagram. With this notation in place, we proceed to analyse each part of this diagram separately. The inner square corresponds to the integral

$$\sum_{\omega_1+\omega_2=\omega} \Big(\int_{\mathbb{R}} P_{u-u_1}(\omega_1) P_{u'-u_1}(-\omega_1)\, \mathrm{d}u_1 \Big) \Big(\int_{\mathbb{R}} P_{u-u_2}(\omega_2) P_{u'-u_1}(-\omega_2)\, \mathrm{d}u_2 \Big)$$
$$\lesssim \sum_{\omega_1+\omega_2=\omega} \frac{1}{\langle \omega_1 \rangle^2} \frac{1}{\langle \omega_2 \rangle^2} \lesssim \frac{1}{\langle \omega \rangle},$$

where we have used Lemma 5 in the last inequality. Note that this bound is slightly sub-optimal, because we have not used the fact that the time variables u and u' corresponding to the nodes at the left and right corners of the square are different. In principle, this would yield extra factors $\exp(-|u-u'|\langle \omega_i \rangle)$ for each term, but we simply bound these factors by 1. Similarly, we get for the left-most triangle

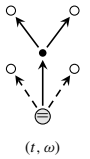
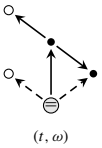
$$\sum_{\omega_4+\omega_5=\omega} \int P_{t-u}(\omega_5) \Big(\int_{\mathbb{R}} P_{t-u_4}(\omega_4) P_{u-u_4}(-\omega_4)\, \mathrm{d}u_4 \Big)\, \mathrm{d}u$$
$$\lesssim \sum_{\omega_4+\omega_5=\omega} \int P_{t-u}(\omega_5) \frac{1}{\langle \omega_4 \rangle^2}\, \mathrm{d}u \lesssim \sum_{\omega_4+\omega_5=\omega} \frac{1}{\langle \omega_4 \rangle^2} \frac{1}{\langle \omega_5 \rangle^2} \lesssim \frac{1}{\langle \omega \rangle},$$

where again we have used Lemma 5. The right-most triangle in the diagram is bounded in the same way, resulting in the final bound

$$\mathbb{E}\Big[|\hat{\bullet}^{(2)}(t,\omega)|^2 \Big] \lesssim \langle \omega \rangle^{-3},$$

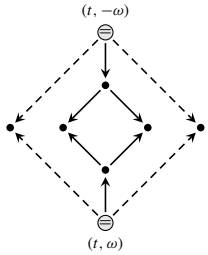
as desired.

Case $\tau = \hat{\bullet}$. For this symbol, we recall from (66) the Wiener chaos decomposition

$\hat{\bullet}(t,\omega) =$  $+ \; 4 \times$  $=: \hat{\bullet}^{(4)}(t,\omega) + \hat{\bullet}^{(2)}(t,\omega)\,.$

We start with $\Xi^{(4)}$. Similarly to the case for Ξ, we have the bound

$$\mathbb{E}\left[|\Xi^{(4)}(t,\omega)|^2\right] \lesssim$$



. (80)

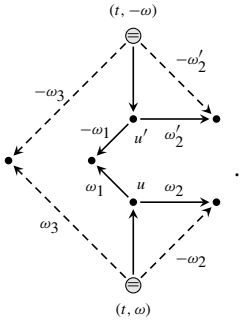
After all our preparation, this diagram poses no additional difficulty. First, the same calculation as in (75) allows to bound the "inner part" of the diagram by $\langle\omega_2\rangle^{-5}$, and the integrals corresponding to the "outer parts" can be bounded by $\langle\omega_1\rangle^{-2}$ and $\langle\omega_3\rangle^{-2}$ immediately, yielding the bound

$$\mathbb{E}\left[|\Xi^{(4)}(t,\omega)|^2\right] \lesssim \sum_{\substack{\omega_1+\omega_2+\omega_3=\omega \\ \omega_1+\omega_3\sim\omega_2}} \frac{1}{\langle\omega_1\rangle^2}\frac{1}{\langle\omega_2\rangle^5}\frac{1}{\langle\omega_3\rangle^2} \lesssim \sum_{\substack{\tilde{\omega}+\omega_3=\omega \\ \tilde{\omega}\sim\omega_2}} \frac{1}{\langle\tilde{\omega}\rangle}\frac{1}{\langle\omega_2\rangle^5} \lesssim \frac{1}{\langle\omega\rangle^3}, \tag{81}$$

where we have used Lemma 5 in the first and Lemma 6 in the second inequality.

We now turn to the bound for $\Xi^{(2)}$, for which we have

$$\mathbb{E}\left[|\Xi^{(2)}(t,\omega)|^2\right] \lesssim$$



. (82)

This graph is more complicated, and requires more careful treatment. We integrate first the innermost time variable, with two arrows pointing to it labelled ω_1 and $-\omega_1$ respectively. We bound this contribution by $\langle\omega_1\rangle^{-2}$, uniformly over u and u', as in (70). We are more careful with the integration of the time variable with incoming arrows labelled ω_2 and $-\omega_2$, and evaluate its contribution to be

$$\frac{e^{-|t-u|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2},$$

as in (68). We next compute the contribution of the triangle in the lower part of the diagram by integrating over u:

$$\int_{\mathbb{R}} \widehat{P}_{t-u}(\omega_1+\omega_2)\,\frac{e^{-|t-u|\langle\omega_2\rangle^2}}{2\langle\omega_2\rangle^2}\,\mathrm{d}u = \frac{1}{2\langle\omega_2\rangle^2}\,\frac{1}{\langle\omega_2\rangle^2+\langle\omega_1+\omega_2\rangle^2}.$$

A similar calculation applies to the upper part of the diagram, and we therefore obtain the bound

$$\mathbb{E}\left[\left|\widehat{\Xi}^{(2)}(t,\omega)\right|^2\right] \lesssim \sum \left(\langle\omega_1\rangle\langle\omega_2\rangle\langle\omega_2'\rangle\langle\omega_3\rangle\big(\langle\omega_2\rangle+\langle\omega_1+\omega_2\rangle\big)\big(\langle\omega_2'\rangle+\langle\omega_2'-\omega_1\rangle\big)\right)^{-2}, \tag{83}$$

where the sum is over all $(\omega_1,\omega_2,\omega_2',\omega_3)$ satisfying

$$\omega_1+\omega_3=\omega, \quad \omega_1+\omega_2\sim\omega_3-\omega_2, \quad -\omega_1+\omega_2'\sim-\omega_3-\omega_2'. \tag{84}$$

The first requirement comes from the "Kirchhoff" law in the bottom node, while the other two constraints come from the paraproduct in the bottom and upper-most nodes.

We proceed to estimate this sum. Note first that the first two conditions in (84) imply that

$$\omega_1+\omega_2\sim\omega-(\omega_1+\omega_2)\,.$$

By (64), if $|\omega_1+\omega_2|>\frac{8}{3}$, we deduce that

$$|\omega_1+\omega_2|\geq c\,(|\omega|-|\omega_1+\omega_2|)\,,$$

where $c=\frac{9}{64}$, and therefore that

$$|\omega_1+\omega_2|\geq\frac{c}{1+c}|\omega|.$$

After reducing the constant $c>0$ as necessary, if follows that in every case, the first two conditions in (84) imply that

$$\langle\omega_1+\omega_2\rangle\geq 2c\langle\omega\rangle. \tag{85}$$

(The factor of 2 is of course a matter of convenience only.) The same argument also shows that under the conditions in (84), we have

$$\langle\omega_1-\omega_2'\rangle\geq 2c\langle\omega\rangle. \tag{86}$$

Define the sets of indices $E_1(\omega)$ and $E_2(\omega)$ by

$$E_1(\omega) := \big\{(\omega_1,\omega_2,\omega_2',\omega_3) \ : \ (84) \text{ holds and } (|\omega_1| < c|\omega| \ \text{ or } \ |\omega_3| < c|\omega|)\big\},$$
$$E_2(\omega) := \big\{(\omega_1,\omega_2,\omega_2',\omega_3) \ : \ (84) \text{ holds and } |\omega_1| \geq c|\omega| \text{ and } |\omega_3| \geq c|\omega|\big\}.$$

We first estimate the sum on the right side of (83) over the set $E_1(\omega)$. By (85), (86) and the constraint $\omega_3 - \omega_2 \sim \omega_1 + \omega_2$, for variables in the set $E_1(\omega)$, we must have

$$\langle \omega_2\rangle \geq \widetilde{c}\,\langle\omega\rangle \quad \text{and} \quad \langle \omega_2'\rangle \geq \widetilde{c}\,\langle\omega\rangle,$$

for some $\widetilde{c} > 0$. Thus, using the simple bounds

$$\langle \omega_2\rangle + \langle \omega_1+\omega_2\rangle \geq \langle\omega_2\rangle, \qquad \langle \omega_2'\rangle + \langle \omega_2'-\omega_1\rangle \geq \langle\omega_2'\rangle,$$

and summing over ω_1 and $\omega_3 = \omega - \omega_1$ using Lemma 5, we arrive at

$$\sum_{E_1(\omega)} \cdots \lesssim \langle\omega\rangle^{-1} \sum_{|\omega_2|,|\omega_2'|\gtrsim|\omega|} \Big(\langle\omega_2\rangle^{-4}\cdot\langle\omega_2'\rangle^{-4}\Big) \lesssim \langle\omega\rangle^{-3}, \tag{87}$$

where the left side above stands for the sum on the right side of (83) restricted to the index set $E_1(\omega)$. As for the the index set $E_2(\omega)$, we start by summing over ω_2 to get

$$\sum_{\omega_2}\langle\omega_2\rangle^{-2}\big(\langle\omega_2\rangle+\langle\omega_1+\omega_2\rangle\big)^{-2} \lesssim \langle\omega_1\rangle^{-2}\sum_{|\omega_2|\leq\frac{|\omega_1|}{2}}\langle\omega_2\rangle^{-2} + \sum_{|\omega_2|\geq\frac{|\omega_1|}{2}}\langle\omega_2\rangle^{-4} \lesssim \langle\omega_1\rangle^{-1},$$

and similarly we have

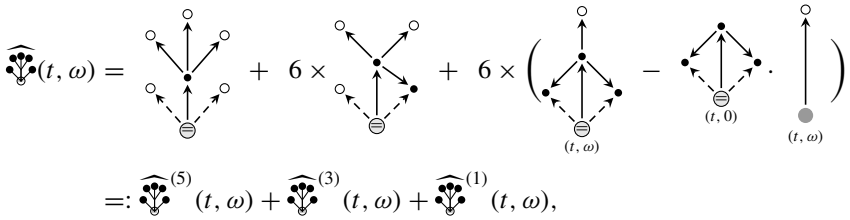
$$\sum_{\omega_2'}\langle\omega_2'\rangle^{-2}\big(\langle\omega_2'\rangle+\langle\omega_2'-\omega_1\rangle\big)^{-2} \lesssim \langle\omega_1\rangle^{-1}.$$

We deduce that

$$\sum_{E_2(\omega)} \cdots \lesssim \sum_{\substack{\omega_1+\omega_3=\omega\\ |\omega_1|,|\omega_3|\geq\frac{|\omega|}{16}}} \langle\omega_1\rangle^{-4}\langle\omega_3\rangle^{-2} \lesssim \langle\omega\rangle^{-2}\sum_{|\omega_1|\geq\frac{|\omega|}{16}}\langle\omega_1\rangle^{-4} \lesssim \langle\omega\rangle^{-3},$$

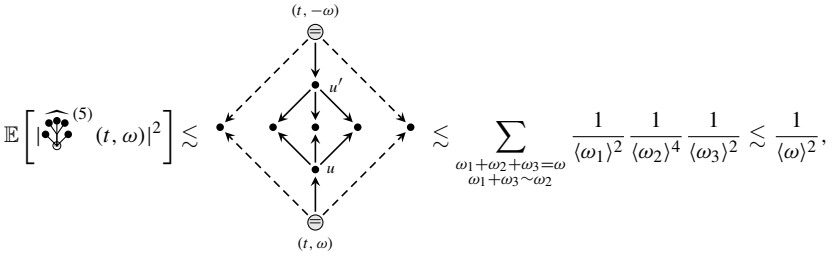
with the same notational convention as in (87). Combining this with (83) and (87) gives the desired bound for  .

Case $\tau =$  . This is the last diagram. We have the Wiener chaos decomposition



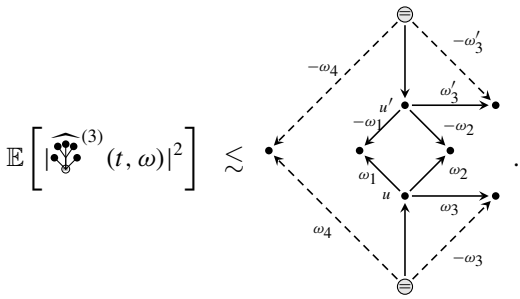
where the last term subtracted in the parenthesis above corresponds to the renormalisation $\mathfrak{c}'_n$ in (7) (in the limit as $n \to +\infty$). As we will see, neither term in the parenthesis above makes sense separately: they both represent some divergent object, but their difference converges to a well-defined limit as the Fourier mode cut-off n goes to infinity. In addition, this limit can be characterised explicitly without referring to a limiting procedure, so this justifies the notation $\Xi^{(1)}(t,\omega)$.

We start with $\Xi^{(5)}$. Proceeding as in (80) and (81), using Lemmas 5 and 6, we immediately have



where the term $\frac{1}{\langle\omega_2\rangle^4}$ comes from the previous bound for the tree Ξ.

We now turn to the component in the third Wiener chaos, whose second moment is bounded by the graph



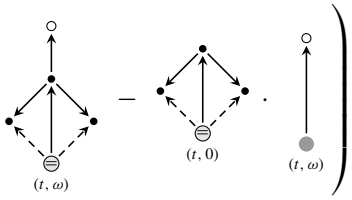
The bound for this graph is similar to the one in (82), and one can proceed essentially in the same way to get

$$\mathbb{E}\left[\left|\widehat{\Xi}^{(3)}(t,\omega)\right|^2\right] \lesssim \langle\omega\rangle^{-2},$$

which is the desired bound.

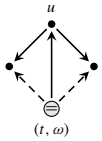
We now turn to the last term

$$\widehat{\Xi}^{(1)}(t,\omega) = 6 \times \left(\quad - \quad \cdot \quad \right).$$



As mentioned before, the two terms in the parenthesis are both ill-defined, but their difference is well-defined as the limit when the regularisation parameter n tends to infinity. To see this, we introduce a notation for the "lower square" which both of the expressions have in common, namely

$$K_{t-u}(\omega) := \quad = \sum_{\omega_1\sim\omega_2} \widehat{P}_{t-u}(\omega+\omega_1+\omega_2)\Big(\int_{-\infty}^{u} \widehat{P}_{t-u_1}(-\omega_1)\widehat{P}_{u-u_1}(\omega_1)\mathrm{d}u_1\Big)$$



$$\times \Big(\int_{-\infty}^{u} \widehat{P}_{t-u_2}(-\omega_2)\widehat{P}_{u-u_1}(\omega_2)\mathrm{d}u_2\Big)$$

$$= \sum_{\omega_1\sim\omega_2} \frac{e^{-(t-u)(\langle\omega_1\rangle^2+\langle\omega_2\rangle^2+\langle\omega+\omega_1+\omega_2\rangle^2)}}{4\langle\omega_1\rangle^2\langle\omega_2\rangle^2}.$$

Note that the divergent constant $\mathbf{c}'$ coincides with $\int_{-\infty}^{t} K_{t-u}(0)\,\mathrm{d}u$.[5] The kernel K_{t-u} is clearly well defined and controlled uniformly over the regularisation for any fixed $t-u>0$. We use this notation to represent $\widehat{\Xi}^{(1)}(t,\omega)$ as

$$\widehat{\Xi}^{(1)}(t,\omega) = \int_{-\infty}^{t} K_{t-u}(\omega)\big(\widehat{\Psi}(u,\omega) - \widehat{\Psi}(t,\omega)\big)\,\mathrm{d}u, \tag{88}$$

so that

$$\mathbb{E}\left[\left|\widehat{\Xi}^{(1)}(t,\omega)\right|^2\right] = \int_{-\infty}^{t}\int_{-\infty}^{t} K_{t-u}(\omega)K_{t-u'}(\omega) \times \mathbb{E}\left[\big(\widehat{\Psi}(u,\omega) - \widehat{\Psi}(t,\omega)\big)\big(\widehat{\Psi}(u',-\omega) - \widehat{\Psi}(t,-\omega)\big)\right]\mathrm{d}u\,\mathrm{d}u'.$$

[5] This is slightly formal - here $\mathbf{c}'$ should denote the limit of $\mathbf{c}'_n$ as $n\to\infty$, which is infinite as discussed before. We are implicitly assuming that a regularisation is present, although we do not capture it in the notation.

Now, since by (68) we have for any $\lambda > 0$

$$\mathbb{E}\left[|\widehat{\bullet}(u,\omega) - \widehat{\bullet}(t,\omega)|^2\right]^{\frac{1}{2}} \lesssim (t-u)^{\lambda}\langle\omega\rangle^{-1+2\lambda},$$

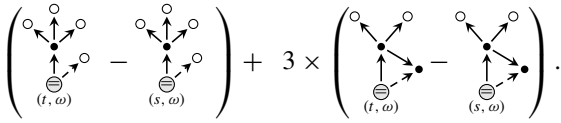
an application of the Cauchy-Schwarz inequality yields

$$\mathbb{E}\left[|\widehat{\bullet\!\!\bullet}^{(1)}(t,\omega)|^2\right] \lesssim \langle\omega\rangle^{-2+4\lambda}\left(\int_{-\infty}^{t}(t-u)^{\lambda}K_{t-u}(\omega)\mathrm{d}u\right)^2 \lesssim \langle\omega\rangle^{-2},$$

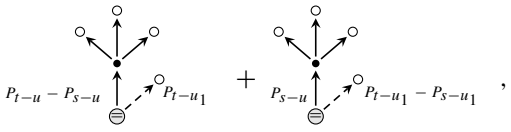
where we have used the fact that for any (small) strictly positive λ, the integral can be bounded uniformly over the regularisation by $\langle\omega\rangle^{-2\lambda}$. This latter point can be checked as follows: for $\lambda = 0$ and without the condition $\omega_1 \sim \omega_2$, we had already calculated the integral in (63); the factor $(t-u)^{\lambda}$ makes an extra power $(\langle\omega_1\rangle^2 + \langle\omega_2\rangle^2 + \langle\omega+\omega_1+\omega_2\rangle^2)^{-\lambda}$ appear in the sum, which permits to invoke Lemma 5 and conclude. This completes the bound for $\tau = \bullet\!\!\bullet$.

5 Bounds for Time Differences

We finally discuss briefly how the reasoning in Sect. 4 should be modified to establish the bound (44) on the time differences $\mathbb{E}|\widehat{\tau}(t,\omega) - \widehat{\tau}(s,\omega)|^2$. We illustrate the (simple) modification necessary for the graph $\tau = \bullet\!\!\bullet$. First, recall the Wiener chaos decomposition (76) for this graph, which yields the following expression for its time-differences:

$\widehat{\bullet\!\!\bullet}(t,\omega) - \widehat{\bullet\!\!\bullet}(s,\omega) =$ 

The differences of graphs can be bounded separately. We only discuss the first difference here. We can rewrite this difference as



,

where we have made explicit which kernels are associated to the lower edges in the graph. (A more systematic treatment would suggest the use of new graphical notation for these!) The variance of each of these terms can then be bounded, by "glueing" two copies of each graph together, as in (78). We then use the elementary bounds

$$\int_{-\infty}^{t} |\widehat{P}_{t-u}(\omega) - \widehat{P}_{s-u}(\omega)|\,\mathrm{d}u \lesssim \frac{1}{\langle\omega\rangle^2}\big(1 \wedge |t-s|\langle\omega\rangle^2\big),$$
$$\int_{-\infty}^{t} (\widehat{P}_{t-u_1}(\omega) - \widehat{P}_{s-u_1}(\omega))^2\,\mathrm{d}u_1 \lesssim \frac{1}{\langle\omega\rangle^2}\big(1 \wedge |t-s|\langle\omega\rangle^2\big).$$

By interpolation, the right side can be replaced by $\lesssim \langle\omega\rangle^{-2+2\lambda}|t-s|^\lambda$, for any $\lambda \in [0,1]$. In other words, an extra factor $|t-s|^\lambda$ can be obtained by sacrificing a bit of the decay of the integral in ω. Then all of the arguments based on convolutions can be performed exactly as before, only with a slightly worse factor of ω at one place. We do not go through the details here, but leave it to the interested reader to check that this does not change the arguments in Sect. 4 in any significant way.

Acknowledgements We are grateful to the very careful referees for their detailed and constructive criticisms, which led to many improvements over the whole paper. We were particularly impressed that one of the referees' reports turned out to be longer than our paper itself! JCM is partially supported by the ANR Grant LSD (ANR-15-CE40-0020-03). HW is supported by the Royal Society through the University Research Fellowship UF140187. WX is supported by the EPSRC through the fellowship EP/N021568/1.

Appendix

We now give a second proof of Proposition 4, based on the following logarithmic Sobolev inequality (see [4, Sect. 1.6] and [23, Sects. 1.1 and 1.5]).

Lemma 7 (log-Sobolev inequality) *Let μ be a Gaussian measure and $X \in W^{1,2}(\mu)$. We have*

$$\mathbb{E}(|X|^2 \log|X|) \le \mathbb{E}|\mathbf{D}X|^2 + \frac{1}{2}\mathbb{E}|X|^2 \log(\mathbb{E}|X|^2), \tag{89}$$

where $\mathbf{D}$ *is the Malliavin derivative, and* $\mathbb{E}$ *is the expectation taken with respect to* μ.

Now, let T_t be the Ornstein-Uhlenbeck semigroup defined by

$$T_t X = \sum_{n=0}^{+\infty} e^{-nt} X_n, \tag{90}$$

where X_n is the component of X in $\mathcal{H}_n$. The Ornstein-Uhlenbeck semigroup is closely related to the Malliavin derivative, because it determines the quadratic form associated with the infinitesimal generator L of T_t. More precisely, for sufficiently nice random variables X, Y, we have

$$\partial_t \mathbb{E}[(T_t X)Y] = \mathbb{E}[(LX)Y] = -\mathbb{E}[\langle \mathbf{D}X, \mathbf{D}Y \rangle]. \tag{91}$$

See [23, Sect. 1.4] for a more detailed discussion of these objects. The main use of the logarithmic Sobolev inequality will be to show the following hypercontractivity estimate.

Proposition 6 (Hypercontractivity) *Let T_t be the Ornstein-Uhlenbeck semigroup. We have*

$$\left(\mathbb{E}|T_t X|^q\right)^{\frac{1}{q}} \leq \left(\mathbb{E}|X|^p\right)^{\frac{1}{p}}, \tag{92}$$

for all $p \geq 2$ and $q = 1 + (p-1)e^{2t}$.

Proof (*Second proof of Proposition 4*) If $X \in \mathcal{H}_n$, then $T_t X = e^{-nt} X$, and we can see that Proposition 4 is an immediate consequence of Proposition 6. It then remains to prove Proposition 6. We can assume $X \geq 0$ without loss of generality.

Fix $p \geq 2$. Let $q(t) = 1 + (p-1)e^{2t}$, and let

$$F(t) = \mathbb{E}|T_t X|^{q(t)}, \qquad G(t) = F(t)^{\frac{1}{q(t)}}.$$

Our aim is to show that $G'(t) \leq 0$ for all $t > 0$, and (92) will follow. In fact, we have

$$G'(t) = G(t)\left[-\frac{q'(t)}{q^2(t)} \log F(t) + \frac{F'(t)}{q(t)F(t)} \right].$$

Since $q'(t) \geq 0$, it suffices to show that

$$-\frac{1}{q(t)} F(t) \log F(t) + \frac{F'(t)}{q'(t)} \leq 0. \tag{93}$$

Noting that

$$F'(t) = \mathbb{E}\left[(T_t X)^{q(t)} \left(q'(t) \log(T_t X) + q(t) \frac{L T_t X}{T_t X} \right) \right],$$

we see that (93) is equivalent to

$$-\frac{1}{q(t)} F(t) \log F(t) + \mathbb{E}\left[(T_t X)^{q(t)} \log(T_t X) \right] + \frac{q(t)}{q'(t)} \mathbb{E}\left[(T_t X)^{q(t)-1} (L T_t X) \right]. \tag{94}$$

Applying the log-Sobolev inequality to the random variable $(T_t X)^{\frac{q(t)}{2}}$ and using the integration by parts formula (91) which in the current context becomes

$$\mathbb{E}\langle \mathbf{D}Y, \mathbf{D}Z\rangle = -\mathbb{E}\big(Y(LZ)\big),$$

we see that (94) follows immediately.

References

1. Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren der Mathematischen Wissenschaften, vol. 343. Springer, Heidelberg (2011)
2. Bertini, L., Presutti, E., Rüdiger, B., Saada, E.: Dynamical fluctuations at the critical point: convergence to a nonlinear stochastic PDE. Teor. Veroyatnost. i Primenen. **38**(4), 689–741 (1993)
3. Biskup, M.: Reflection positivity and phase transitions in lattice spin models. In: Methods of Contemporary Mathematical Statistical Physics. Lecture Notes in Mathematics, vol. 1970, pp. 1–86. Springer, Berlin (2009)
4. Bogachev, V.: Gaussian Measures. Mathematical Surveys and Monographs, vol. 62, 1st edn. American Mathematical Society, Providence (1998)
5. Catellier, R., Chouk, K.: Paracontrolled distributions and the 3-dimensional stochastic quantization equation (2013). arXiv:1310.6869
6. Chandra, A., Weber, H.: Stochastic PDEs, Regularity Structures, and Interacting Particle Systems (2015). arXiv:1508.03616
7. Chandra, A., Hairer, M.: An analytic BPHZ Theorem for Regularity Structures (2016). arXiv:1612.08138
8. Da Prato, G., Debussche, A.: Strong solutions to the stochastic quantization equations. Ann. Probab. **31**(4), 1900–1916 (2003)
9. Jentzen, W.E.A., Shen, H.: Renormalized powers of Ornstein–Uhlenbeck processes and well-posedness of stochastic Ginzburg–Landau equations. Nonlinear Anal. **142**, 152–193 (2016)
10. Fritz, J., Rüdiger, B.: Time dependent critical fluctuations of a one-dimensional local mean field model. Probab. Theory Relat. Fields **103**(3), 381–407 (1995)
11. Friz, P.K., Hairer, M.: A Course on Rough Paths, with An Introduction to Regularity Structures. Universitext. Springer, Berlin (2014)
12. Glimm, J., Jaffe, A.: Quantum Physics, 2nd edn. Springer, New York (1987)
13. Gubinelli, M., Imkeller, P., Perkowski, N.: Paraproducts, Rough Paths and Controlled Distributions (2012). arXiv:1210.2684
14. Hairer, M.: Solving the KPZ equation. Ann. Math. **2**(178), 559–664 (2013)
15. Hairer, M.: A theory of regularity structures. Invent. Math. **198**(2), 269–504 (2014)
16. Hairer, M., Quastel, J.: A Class of Growth Models Rescaling to KPZ (2015). arXiv:1512.07845
17. Hairer, M., Ryser, M., Weber, H.: Triviality of the $2d$ stochastic Allen–Cahn equation. Electron. J. Probab. **17**(39), 1–14 (2012)
18. Kuo, H.-H.: Introduction to Stochastic Integration. Universitext. Springer, New York (2006)
19. Kupiainen, A.: Renormalization Group and Stochastic PDE's (2014). arXiv:1410.3094
20. Mourrat, J.-C., Weber, H.: Convergence of the two-dimensional dynamic Ising–Kac model to Φ^4_2. Commun. Pure Appl. Math. (in press). arXiv:1410.1179
21. Mourrat, J.-C., Weber, H.: Global Well-Posedness of the Dynamic Φ^4 Model in the Plane (2015). arXiv:1501.06191
22. Mourrat, J.-C., Weber, H.: Global Well-Posedness of the Dynamic Φ^4_3 Model on the Torus (2016). arXiv:1601.01234
23. Nualart, D.: The Malliavin Calculus and Related Topics. Probability and its Applications (New York), 2nd edn. Springer, Berlin (2006)
24. Osterwalder, K., Schrader, R.: Axioms for Euclidean Green's functions. Commun. Math. Phys. **31**, 83–112 (1973)

25. Osterwalder, K., Schrader, R.: Axioms for Euclidean Green's functions. II. Commun. Math. Phys. **42**, 281–305 (1975). With an appendix by Stephen Summers
26. Parisi, G., Wu, Y.S.: Perturbation theory without gauge fixing. Sci. Sinica **24**(4), 483–496 (1981)
27. Polyak, M.: Feynman diagrams for pedestrians and mathematicians. In: Graphs and Patterns in Mathematics and Theoretical Physics. Proceedings of Symposia in Pure Mathematics, vol. 73, pp. 15–42. American Mathematical Society, Providence (2005)
28. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Grundlehren der Mathematischen Wissenschaften, vol. 293, 2nd edn. Springer, Berlin (1994)

Two Classes of Nonlocal Evolution Equations Related by a Shared Traveling Wave Problem

Franz Achleitner

Abstract We consider nonlocal reaction-diffusion equations and nonlocal Korteweg-de Vries-Burgers (KdVB) equations, i.e. scalar conservation laws with diffusive-dispersive regularization. We review the existence of traveling wave solutions for these two classes of evolution equations. For classical equations the traveling wave problem (TWP) for a local KdVB equation can be identified with the TWP for a reaction-diffusion equation. In this article we study this relationship for these two classes of evolution equations with nonlocal diffusion/dispersion. This connection is especially useful, if the TW equation is not studied directly, but the existence of a TWS is proven using one of the evolution equations instead. Finally, we present three models from fluid dynamics and discuss the TWP via its link to associated reaction-diffusion equations.

Keywords Nonlocal evolution equations · Traveling wave solutions
Reaction-diffusion equations · Korteweg-de Vries-Burgers equation

1 Introduction

We will consider two classes of (nonlocal) evolution equations and study the associated traveling wave problems in parallel: On the one hand, we consider scalar conservation laws with (nonlocal) diffusive-dispersive regularization

$$\partial_t u + \partial_x f(u) = \varepsilon \mathscr{L}_1[u] + \delta \partial_x \mathscr{L}_2[u] , \quad t > 0 , \quad x \in \mathbb{R} , \tag{1}$$

for some nonlinear function $f : \mathbb{R} \to \mathbb{R}$, Lévy operators $\mathscr{L}_1$ and $\mathscr{L}_2$, as well as constants $\varepsilon, \delta \in \mathbb{R}$. The Fourier multiplier operators $\mathscr{L}_1$ and $\partial_x \mathscr{L}_2$ model diffusion and dispersion, respectively. On the other hand, we consider scalar reaction-diffusion equations

$$\partial_t u = \sigma \mathscr{L}_3[u] + r(u) , \quad t > 0 , \quad x \in \mathbb{R} , \tag{2}$$

F. Achleitner (✉)
Faculty of Mathematics, University of Vienna,
Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria
e-mail: franz.achleitner@univie.ac.at

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_2

for some positive constant σ, as well as a nonlinear function $r : \mathbb{R} \to \mathbb{R}$ and a Lévy operator $\mathscr{L}_3$ modeling reaction and diffusion, respectively.

Definition 1 A traveling wave solution (TWS) of an evolution equation–such as (1) and (2)–is a solution $u(x,t) = \bar{u}(\xi)$ whose *profile* $\bar{u}$ depends on $\xi := x - ct$ for some *wave speed* c. Moreover, the profile $\bar{u} \in C^2(\mathbb{R})$ is assumed to approach distinct endstates $u_\pm$ such that

$$\lim_{\xi \to \pm\infty} \bar{u}(\xi) = u_\pm \,, \qquad \lim_{\xi \to \pm\infty} \bar{u}^{(n)}(\xi) = 0 \quad \text{with } n = 1, 2. \tag{3}$$

Such a TWS is also known as a *front* in the literature. A TWS $(\bar{u}, c)$ is called monotone, if its profile $\bar{u}$ is a monotone function.

Definition 2 The traveling wave problem (TWP) associated to an evolution equation is to study for some distinct endstates $u_\pm$ the existence of a TWS $(\bar{u}, c)$ in the sense of Definition 1.

We want to identify classes of evolutions equations of type (1) and (2), which lead to the same TWP. This connection is especially useful, if the TWP is not studied directly, but the existence of a TWS is proven using one of the evolution equations instead. A classical example of (1) is a scalar conservation law with local diffusive-dispersive regularization

$$\partial_t u + \partial_x f(u) = \varepsilon \partial_x^2 u + \delta \partial_x^3 u \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{4}$$

for some nonlinear function $f : \mathbb{R} \to \mathbb{R}$ and some constants $\varepsilon > 0$ and $\delta \in \mathbb{R}$. Equation (4) with Burgers flux $f(u) = u^2$ is known as Korteweg-de Vries-Burgers (KdVB) equation; hence we refer to Eq. (4) with general f as *generalized* KdVB equation and Eq. (1) as *nonlocal generalized* KdVB equation. A TWS $(\bar{u}, c)$ satisfies the traveling wave equation (TWE)

$$-c\bar{u}' + f'(\bar{u})\,\bar{u}' = \varepsilon \bar{u}'' + \delta \bar{u}''' \,, \quad \xi \in \mathbb{R} \,, \tag{5}$$

or integrating on $(-\infty, \xi]$ and using (3),

$$h(\bar{u}) := f(\bar{u}) - c\bar{u} - (f(u_-) - c\,u_-) = \varepsilon \bar{u}' + \delta \bar{u}'' \,, \quad \xi \in \mathbb{R} \,. \tag{6}$$

However, the TW ansatz $v(x,t) = \bar{u}(x - \varepsilon t)$ for the scalar reaction-diffusion equation

$$\partial_t v = -h(v) + \delta \partial_x^2 v \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{7}$$

leads to the same TWE (6) except for a different interpretation of the parameters. The traveling wave speeds in the TWP of (4) and (7) are c and ε, respectively. For fixed parameters c, ε, and δ, the existence of a traveling wave profile $\bar{u}$ satisfying (3) and (6) reduces to the existence of a heteroclinic orbit for this ODE. This is an example, where the existence of TWS is studied directly via the TWE.

An example, where the TWE is not studied directly, is the TWP for a nonlocal KdVB equation (1) with $\mathscr{L}_1[u] = \partial_x^2 u$ and $\mathscr{L}_2[u] = \phi_\varepsilon * u - u$ for some even non-

negative function $\phi \in L^1(\mathbb{R})$ with compact support and unit mass, where $\phi_\varepsilon(\cdot) := \phi(\cdot/\varepsilon)/\varepsilon$ with $\varepsilon > 0$. It has been derived as a model for phase transitions with long range interactions close to the surface, which supports planar TWS associated to undercompressive shocks of (B.1), see [52]. In particular, the TWP for a cubic flux function $f(u) = u^3$ is related to the TWP for a reaction-diffusion equation (2) with $\mathscr{L}_3[u] = \mathscr{L}_2[u]$. The existence of TWS for this reaction-diffusion equation has been proven via a homotopy of (2) to a classical reaction-diffusion model (7), see [14].

Outline. In Sect. 2 we collect background material on Lévy operators $\mathscr{L}$, which will model diffusion in our nonlocal evolution equations. Special emphasize is given to convolution operators and Riesz–Feller operators. In Sect. 3 we review the classical results on the TWP for reaction-diffusion equations (7) and generalized Korteweg-de Vries-Burgers equation (4). We study their relationship in detail, especially the classification of function $h(u)$, which will be used again in Sect. 4. In Sect. 4, first we review the results on TWP for nonlocal reaction-diffusion equations (2) with operators $\mathscr{L}_3$ of convolution type and Riesz–Feller type, respectively. Finally, we study the example of nonlocal generalized Korteweg-de Vries-Burgers equation (1) with $\mathscr{L}_1[u] = \mathscr{D}_+^{1/3} u$ and $\mathscr{L}_2[u] = \partial_x^2 u$ modeling a shallow water flow [44], and Fowler's equation

$$\partial_t u + \partial_x u^2 = \partial_x^2 u - \partial_x \mathscr{D}_+^{1/3} u \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{8}$$

modeling dune formation [36], where $\mathscr{D}_+^\alpha$ is a Caputo derivative. In the Appendix, we collect background material on Caputo derivatives $\mathscr{D}_+^\alpha$ and the shock wave theory for scalar conservation laws, which will explain the importance of the TWP for KdVB equations.

Scalar conservation laws with fractional Laplacian are another example of equation (1) with $\mathscr{L}_1[u] = -(-\partial_x^2)^{\alpha/2} u$, $0 < \alpha < 2$, and $\mathscr{L}_2[u] \equiv 0$. However, its traveling wave problem can not be related to a nonlocal reaction-diffusion problem like our examples. Therefore, instead of discussing its traveling wave problem, we refer the interested reader to the literature [7, 8, 10, 15, 23, 26, 30–33, 43] and references therein.

Notations. We use the conventions in probability theory, and define the Fourier transform $\mathscr{F}$ and its inverse $\mathscr{F}^{-1}$ for $g \in L^1(\mathbb{R})$ and $x, k \in \mathbb{R}$ as

$$\mathscr{F}[g](k) := \int_{\mathbb{R}} \mathrm{e}^{+\mathrm{i}kx}\, g(x) \;\; \mathrm{d}x \;; \qquad \mathscr{F}^{-1}[g](x) := \tfrac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}kx}\, g(k) \;\; \mathrm{d}k \;.$$

In the following, $\mathscr{F}$ and $\mathscr{F}^{-1}$ will denote also their respective extensions to $L^2(\mathbb{R})$.

2 Lévy Operators

A Lévy process is a stochastic process with independent and stationary increments which is continuous in probability [12, 40, 53]. Therefore a Lévy process is charac-

terized by its transition probabilities $p(t,x)$, which evolve according to an evolution equation

$$\partial_t p = \mathscr{L} p \tag{9}$$

for some operator $\mathscr{L}$, also called a *Lévy operator*. First, we define Lévy operators on the function spaces $C_0(\mathbb{R}) := \{f \in C(\mathbb{R}) \mid \lim_{|x|\to\infty} f(x) = 0\}$ and $C_0^2(\mathbb{R}) := \{f, f', f'' \in C_0(\mathbb{R})\}$.

Definition 3 The family of Lévy operators in one spatial dimension consists of operators $\mathscr{L}$ defined for $f \in C_0^2(\mathbb{R})$ as

$$\mathscr{L} f(x) = \tfrac{1}{2} A f''(x) + \gamma\ f'(x) + \int_{\mathbb{R}} \left(f(x+y) - f(x) - y\ f'(x) 1_{(-1,1)}(y) \right) \nu(\mathrm{d}y) \tag{10}$$

for some constants $A \geq 0$ and $\gamma \in \mathbb{R}$, and a measure ν on $\mathbb{R}$ satisfying

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}} \min(1, |y|^2)\ \nu(\mathrm{d}y)\ < \infty\,.$$

Remark 1 The function $f(x+y) - f(x) - y\ f'(x) 1_{(-1,1)}(y)$ is integrable with respect to ν, because it is bounded outside of any neighborhood of 0 and

$$f(x+y) - f(x) - y\ f'(x) 1_{(-1,1)}(y) = O(|y|^2) \quad \text{as} \quad |y| \to 0$$

for fixed x. The indicator function $c(y) = 1_{(-1,1)}(y)$ is only one possible choice to obtain an integrable integrand. More generally, let $c(y)$ be a bounded measurable function from $\mathbb{R}$ to $\mathbb{R}$ satisfying $c(y) = 1 + o(|y|)$ as $|y| \to 0$, and $c(y) = O(1/|y|)$ as $|y| \to \infty$. Then (10) is rewritten as

$$\mathscr{L} f(x) = \tfrac{1}{2} A f''(x) + \gamma_c\ f'(x) + \int_{\mathbb{R}} \left(f(x+y) - f(x) - y\ f'(x) c(y) \right) \nu(\mathrm{d}y)\ , \tag{11}$$

with $\gamma_c = \gamma + \int_{\mathbb{R}} y\ (c(y) - 1_{(-1,1)}(y))\ \nu(\mathrm{d}y)$.

Alternative choices for c:

(c 0) If a Lévy measure ν satisfies $\int_{|y|<1} |y|\ \nu(\mathrm{d}y)\ < \infty$ then $c \equiv 0$ is admissible.
(c 1) If a Lévy measure ν satisfies $\int_{|y|>1} |y|\ \nu(\mathrm{d}y)\ < \infty$ then $c \equiv 1$ is admissible.

We note that A and ν are invariant no matter what function c we choose.

Examples

(a) The Lévy operators

$$\mathscr{L} f = \int_{\mathbb{R}} \left(f(x+y) - f(x) \right) \nu(\mathrm{d}y) \tag{12}$$

are infinitesimal generators associated to a compound Poisson process with finite Lévy measure ν satisfying (c 0). The special case of $\nu(\mathrm{d}y) = \phi(-y)\ \mathrm{d}y$ for some function $\phi \in L^1(\mathbb{R})$ yields

$$\mathscr{L} f(x) = \int_{\mathbb{R}} \big(f(x+y) - f(x)\big)\, \phi(-y)\ \mathrm{d}y = \left(\phi * f - \int_{\mathbb{R}} \phi\ \ \mathrm{d}y\ \ f\right)(x)\,. \tag{13}$$

(b) *Riesz–Feller operators*. The Riesz–Feller operators of order a and asymmetry θ are defined as Fourier multiplier operators

$$\mathscr{F}[D_\theta^a f](k) = \psi_\theta^a(k)\ \mathscr{F}[f](k)\,, \qquad k \in \mathbb{R}\,, \tag{14}$$

with symbol $\psi_\theta^a(k) = -|k|^a \exp\big[\mathrm{i}\ \mathrm{sgn}(k)\ \theta\pi/2\big]$ such that $(a,\theta) \in \mathfrak{D}_{a,\theta}$ and

$$\mathfrak{D}_{a,\theta} := \{\,(a,\theta) \in \mathbb{R}^2 \,|\, 0 < a \le 2\,, \quad |\theta| \le \min\{a, 2-a\}\,\}\,,$$

see also (Fig. 1)
Special cases of Riesz–Feller operators are

- Fractional Laplacians $-(-\Delta)^{a/2}$ on $\mathbb{R}$ with Fourier symbol $-|k|^a$ for $0 < a \le 2$. In particular, fractional Laplacians are the only symmetric Riesz–Feller operators with $-(-\Delta)^{a/2} = D_0^a$ and $\theta \equiv 0$.
- Caputo derivatives $-\mathscr{D}_+^\alpha$ with $0 < \alpha < 1$ are Riesz–Feller operators with $a = \alpha$ and $\theta = -\alpha$, such that $-\mathscr{D}_+^\alpha = D_{-\alpha}^\alpha$, see also Sect. A.
- Derivatives of Caputo derivatives $\partial_x \mathscr{D}_+^\alpha$ with $0 < \alpha < 1$ are Riesz–Feller operators with $a = 1+\alpha$ and $\theta = 1-\alpha$, such that $\partial_x \mathscr{D}_+^\alpha = D_{1-\alpha}^{1+\alpha}$.

Next we consider the Cauchy problem

$$\partial_t u(x,t) = D_\theta^a[u(\cdot,t)](x)\,, \quad u(x,0) = u_0(x)\,, \tag{15}$$

for $(x,t) \in \mathbb{R} \times (0,\infty)$ and initial datum u_0.

Proposition 1 *For $(a,\theta) \in \mathfrak{D}_{a,\theta}$ with $\theta \neq \pm 1$ and $1 \le p < \infty$, the Riesz–Feller operator D_θ^a generates a strongly continuous L^p-semigroup*

$$S_t : L^p(\mathbb{R}) \to L^p(\mathbb{R})\,, \quad u_0 \mapsto S_t u_0 = G_\theta^a(\cdot,t) * u_0\,,$$

with heat kernel $G_\theta^a(x,t) = \mathscr{F}^{-1}[\exp(t\ \psi_\theta^a(\cdot))](x)$. In particular, $G_\theta^a(x,t)$ is the probability measure of a Lévy strictly a-stable distribution.

The proof of this proposition for a subclass $1 < \alpha \le 2$ in [6, Proposition 2.2] can be extended to cover all cases $(a,\theta) \in \mathfrak{D}_{a,\theta}$ with $\theta \neq \pm 1$. For $(a,\theta) \in \{(1,1),(1,-1)\}$, the probability measure G_θ^a is a delta distribution, e.g. $G_1^1(x,t) = \delta_{x+t}$ and $G_{-1}^1(x,t) = \delta_{x-t}$, and is called trivial [53, Definition 13.6]. However, we are interested in non-trivial probability measures G_θ^a for

$$(a,\theta) \in \mathfrak{D}_{a,\theta}^\diamond := \{\,(a,\theta) \in \mathfrak{D}_{a,\theta} \,|\, |\theta| < 1\,\}\,,$$

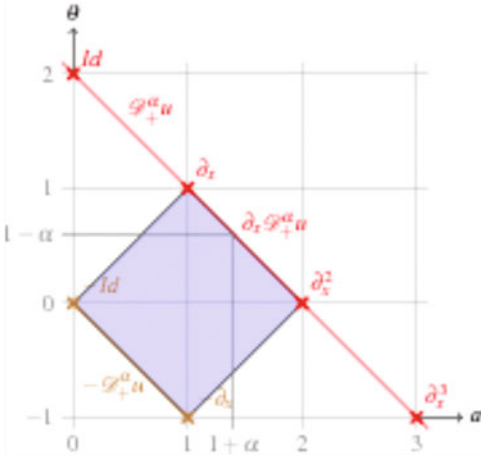


Fig. 1 The family of Fourier multipliers $\psi_\theta^a(k) = -|k|^a \exp\big[\mathrm{i}\,\mathrm{sgn}(k)\theta\pi/2)\big]$ has two parameters a and θ. Some Fourier multiplier operators $\mathscr{F}[Tf](k) = \psi_\theta^a(k)\,\mathscr{F}[f](k)$ are inserted in the parameter space (a, θ): partial derivatives and Caputo derivatives $\mathscr{D}_+^\alpha$ with $0 < \alpha < 1$. The Riesz–Feller operators D_θ^a are those operators with parameters $(a, \theta) \in \mathfrak{D}_{a,\theta}$. The set $\mathfrak{D}_{a,\theta}$ is also called *Feller-Takayasu diamond* and depicted as a *shaded region*, see also [47]

such that $\mathfrak{D}_{a,\theta} = \mathfrak{D}_{a,\theta}^\diamond \cup \{(1,1),\ (1,-1)\}$. Note, nonlocal Riesz–Feller D_θ^a operators are those with parameters

$$(a,\theta) \in \mathfrak{D}_{a,\theta}^\bullet := \{\,(a,\theta) \in \mathfrak{D}_{a,\theta}\,|\,0 < a < 2\,,\quad |\theta| < 1\,\},$$

such that $\mathfrak{D}_{a,\theta}^\diamond = \mathfrak{D}_{a,\theta}^\bullet \cup \{(2,0)\}$.

Proposition 2 ([6, Lemma 2.1]) *For $(a,\theta) \in \mathfrak{D}_{a,\theta}^\diamond$ the probability measure G_θ^a is absolutely continuous with respect to the Lebesgue measure and possesses a probability density which will be denoted again by G_θ^a. For all $(x,t) \in \mathbb{R} \times (0,\infty)$ the following properties hold;*

(a) $G_\theta^a(x,t) \geq 0$. *If $\theta \neq \pm a$ then* $G_\theta^a(x,t) > 0$;
(b) $\|G_\theta^a(\cdot,t)\|_{L^1(\mathbb{R})} = 1$;
(c) $G_\theta^a(x,t) = t^{-1/a} G_\theta^a(xt^{-1/a}, 1)$;
(d) $G_\theta^a(\cdot,s) * G_\theta^a(\cdot,t) = G_\theta^a(\cdot,s+t)$ *for all* $s,t \in (0,\infty)$;
(e) $G_\theta^a \in C_0^\infty(\mathbb{R} \times (0,\infty))$.

The Lévy measure ν of a Riesz–Feller operator D_θ^a with $(a,\theta) \in \mathfrak{D}_{a,\theta}^\bullet$ is absolutely continuous with respect to Lebesgue measure and satisfies

$$\nu(\mathrm{d}y) = \begin{cases} c_-(\theta) y^{-1-a}\ \mathrm{d}y & \text{on } (0,\infty)\,, \\ c_+(\theta)|y|^{-1-a}\ \mathrm{d}y & \text{on } (-\infty,0)\,, \end{cases} \tag{16}$$

with $c_\pm(\theta) = \Gamma(1+a)\sin((a \pm \theta)\pi/2)/\pi$, see [47, 54].

To study a TWP for evolution equations involving Riesz–Feller operators, it is necessary to extend the Riesz–Feller operators to $C_b^2(\mathbb{R})$. Their singular integral representations (10) may be used to accomplish this task.

Theorem 1 ([6]) *If $(a,\theta) \in \mathfrak{D}_{a,\theta}^{\bullet}$ with $a \neq 1$, then for all $f \in \mathscr{S}(\mathbb{R})$ and $x \in \mathbb{R}$*

$$\begin{aligned} D_\theta^a f(x) = & \frac{c_+(\theta) - c_-(\theta)}{1-a} f'(x) \\ & + c_+(\theta) \int_0^\infty \frac{f(x+y) - f(x) - f'(x)\, y 1_{(-1,1)}(y)}{y^{1+a}} \, \mathrm{d}y \\ & + c_-(\theta) \int_0^\infty \frac{f(x-y) - f(x) + f'(x)\, y 1_{(-1,1)}(y)}{y^{1+a}} \, \mathrm{d}y \end{aligned} \tag{17}$$

with $c_\pm(\theta) = \Gamma(1+a)\sin((a \pm \theta)\pi/2)/\pi$. Alternative representations are

- *If $0 < a < 1$, then*

$$D_\theta^a f(x) = c_+(\theta) \int_0^\infty \frac{f(x+y) - f(x)}{y^{1+a}} \, \mathrm{d}y + c_-(\theta) \int_0^\infty \frac{f(x-y) - f(x)}{y^{1+a}} \, \mathrm{d}y \, .$$

- *If $1 < a < 2$, then*

$$\begin{aligned} D_\theta^a f(x) = & c_+(\theta) \int_0^\infty \frac{f(x+y) - f(x) - f'(x)\, y}{y^{1+a}} \, \mathrm{d}y \\ & + c_-(\theta) \int_0^\infty \frac{f(x-y) - f(x) + f'(x)\, y}{y^{1+a}} \, \mathrm{d}y \, . \end{aligned} \tag{18}$$

These representations allow to extend Riesz–Feller operators D_θ^a to $C_b^2(\mathbb{R})$ such that $D_\theta^a C_b^2(\mathbb{R}) \subset C_b(\mathbb{R})$. For example, one can show

Proposition 3 ([6, Proposition 2.4]) *For $(a,\theta) \in \mathfrak{D}_{a,\theta}$ with $1 < a < 2$, the integral representation (18) of D_θ^a is well-defined for functions $f \in C_b^2(\mathbb{R})$ with*

$$\sup_{x \in \mathbb{R}} |D_\theta^a f(x)| \leq \mathscr{K} \|f''\|_{C_b(\mathbb{R})} \frac{M^{2-a}}{2-a} + 4\mathscr{K} \|f'\|_{C_b(\mathbb{R})} \frac{M^{1-a}}{a-1} < \infty \tag{19}$$

for some positive constants M and $\mathscr{K} = \frac{\Gamma(1+a)}{\pi} |\sin((a+\theta)\frac{\pi}{2}) + \sin((a-\theta)\frac{\pi}{2})|$.

Estimate (19) is a key estimate, which is used to adapt Chen's approach [24] to the TWP for nonlocal reaction-diffusion equations with Riesz–Feller operators [6].

3 TWP for Classical Evolution Equations

In this section we review the importance of the TWP for reaction-diffusion equations and scalar conservation laws with higher-order regularizations, respectively.

3.1 Reaction-Diffusion Equations

A scalar reaction-diffusion equation is a partial differential equation

$$\partial_t u = \sigma \partial_x^2 u + r(u) \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{20}$$

for some positive constant $\sigma > 0$, as well as a nonlinear function $r : \mathbb{R} \to \mathbb{R}$ and second-order derivative $\partial_x^2 u$ modeling reaction and diffusion, respectively. The TWP for given endstates $u_\pm$ is to study the existence of a TWS $(\bar{u}, c)$ for (20) in the sense of Definition 1. If the profile $\bar{u} \in C^2(\mathbb{R})$ is bounded, then it satisfies $\lim_{\xi \to \pm\infty} \bar{u}^{(n)}(\xi) = 0$ for $n = 1, 2$. A TWS $(\bar{u}, c)$ satisfies the TWE

$$-c\bar{u}' = r(\bar{u}) + \sigma \bar{u}'' \,, \quad \xi \in \mathbb{R} \,. \tag{21}$$

Phase plane analysis. A traveling wave profile $\bar{u}$ is a heteroclinic orbit of the TWE (21) connecting the endstates $u_\pm$. To identify necessary conditions on the existence of TWS, TWE (21) is written as a system of first-order ODEs for $u, v := u'$:

$$\frac{\mathrm{d}}{\mathrm{d}\xi} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v \\ (-r(u) - cv)/\sigma \end{pmatrix} =: F(u, v) \,, \quad \xi \in \mathbb{R} \,. \tag{22}$$

First, an endstate (u_s, v_s) of a heteroclinic orbit has to be a stationary state of F, i.e. $F(u_s, v_s) = 0$, which implies $v_s \equiv 0$ and $r(u_s) = 0$. Second, $(u_-, 0)$ has to be an unstable stationary state of (22) and $(u_+, 0)$ either a saddle or a stable node of (22). As long as a stationary state (u_s, v_s) is hyperbolic, i.e. the linearization of F at (u_s, v_s) has only eigenvalues λ with non-zero real part, the stability of (u_s, v_s) is determined by these eigenvalues. The linearization of F at (u_s, v_s) is

$$\mathrm{D}F(u_s, v_s) = \begin{pmatrix} 0 & 1 \\ -r'(u_s)/\sigma & -c/\sigma \end{pmatrix} \,. \tag{23}$$

Eigenvalues $\lambda_\pm$ of the Jacobian $\mathrm{D}F(u_s, v_s)$ satisfy the characteristic equation $\lambda^2 + \lambda c/\sigma + r'(u_s)/\sigma = 0$. Moreover, $\lambda_- + \lambda_+ = -c/\sigma$ and $\lambda_-\lambda_+ = r'(u_s)/\sigma$. The eigenvalues $\lambda_\pm$ of the Jacobian $\mathrm{D}F(u_s, v_s)$ are

$$\lambda_\pm = -\frac{c}{2\sigma} \pm \sqrt{\frac{c^2}{4\sigma^2} - \frac{r'(u_s)}{\sigma}} = \frac{-c \pm \sqrt{c^2 - 4\sigma r'(u_s)}}{2\sigma} \,. \tag{24}$$

Thus $r'(u_s) < 0$ ensures that $(u_s, 0)$ is a saddle point, i.e. with one positive and one negative eigenvalue.

Balance of potential. The potential R (of the reaction term r) is defined as $R(u) := \int_0^u r(\upsilon) \, \mathrm{d}\upsilon$. The potentials of the endstates $u_\pm$ are called *balanced* if $R(u_+) = R(u_-)$ and *unbalanced* otherwise. A formal computation reveals a connection between the sign of c and the balance of the potential $R(u)$: Multiplying TWE (21) with $\bar{u}'$, integrating on $\mathbb{R}$ and using (3), yields

$$-c\|\bar{u}'\|_{L^2}^2 = \int_{u_-}^{u_+} r(v)\ \mathrm{d}v = R(u_+) - R(u_-)\,, \tag{25}$$

since $\int_{\mathbb{R}} \bar{u}''\bar{u}'\ \mathrm{d}\xi = 0$ due to (3). Thus $-\operatorname{sgn} c = \operatorname{sgn}(R(u_+) - R(u_-))$. In case of a balanced potential the wave speed c is zero, hence the TWS is stationary.

Definition 4 Assume $u_- > u_+$. A function $r \in C^1(\mathbb{R})$ with $r(u_\pm) = 0$ is

- *monostable* if $r'(u_-) < 0$, $r'(u_+) > 0$ and $r(u) > 0$ for $u \in (u_+, u_-)$.
- *bistable* if $r'(u_\pm) < 0$ and

$$\exists u_* \in (u_+, u_-)\ :\ r(u) \begin{cases} < 0 & \text{for } u \in (u_+, u_*)\,, \\ > 0 & \text{for } u \in (u_*, u_-)\,. \end{cases}$$

- *unstable* if $r'(u_\pm) > 0$.

We chose a very narrow definition compared to [56]. Moreover, in most applications of reaction-diffusion equations a quantity u models a density of a substance/population. In these situations only nonnegative states $u_\pm$ and functions u are of interest.
If a TWS $(\bar{u}, c)$ exists, then a closer inspection of the eigenvalues (24) at $(u_+, 0)$ indicates the geometry of the profile $\bar{u}$ for large ξ:

$$c^2 - 4\sigma r'(u_+) \begin{cases} \geq 0 & \text{TWS with monotone decreasing profile } \bar{u} \text{ for large } \xi; \\ < 0 & \text{TWS with oscillating profile } \bar{u} \text{ for large } \xi. \end{cases}$$

Proposition 4 ([56, Sect. 2.2]) *Assume $\sigma > 0$ and $u_- > u_+$.*

- *If r is monostable, then there exists a positive constant c_* such that for all $c \geq c_*$ there exists a monotone TWS $(\bar{u}, c)$ of (20) in the sense of Definition 1. For $c < c_*$ no such monotone TWS exists (however oscillatory TWS may exist).*
- *If r is bistable, then there exists an (up to translations) unique monotone TWS $(\bar{u}, c)$ of (20) in the sense of Definition 1.*
- *If r is unstable, then there does not exist a monotone TWS $(\bar{u}, c)$ of (20).*

3.2 Korteweg-de Vries-Burgers Equation (KdVB)

A generalized KdVB equation is a scalar partial differential equation

$$\partial_t u + \partial_x f(u) = \varepsilon \partial_x^2 u + \delta \partial_x^3 u, \quad x \in \mathbb{R}, \quad t > 0, \tag{26}$$

for some flux function $f : \mathbb{R} \to \mathbb{R}$ as well as constants $\varepsilon > 0$ and $\delta \in \mathbb{R}$. The TWP for given endstates $u_\pm$ is to study the existence of a TWS $(\bar{u}, c)$ for (26) in the sense of Definition 1. The importance of the TWP for KdVB equations in the shock wave

theory of (scalar) hyperbolic conservation laws is discussed in Sect. B. A TWS $(\bar{u}, c)$ satisfies the TWE

$$-c\bar{u}' + f'(\bar{u})\,\bar{u}' = \varepsilon\bar{u}'' + \delta\bar{u}''' \,, \quad \xi \in \mathbb{R} \,, \tag{27}$$

or integrating on $(-\infty, \xi]$ and using (3),

$$h(\bar{u}) := f(\bar{u}) - c\bar{u} - (f(u_-) - c\,u_-) = \varepsilon\bar{u}' + \delta\bar{u}'' \,, \quad \xi \in \mathbb{R} \,. \tag{28}$$

Connection with reaction-diffusion equation. A TWS $u(x,t) = \bar{u}(x - ct)$ of a generalized Korteweg-de Vries-Burgers equation (26) satisfies TWE (28). Thus $v(x,t) = \bar{u}(x - \varepsilon t)$ is a TWS $(\bar{u}, \varepsilon)$ of the reaction-diffusion equation

$$\partial_t v = -h(v) + \delta\partial_x^2 v \,, \quad x \in \mathbb{R} \,, \quad t > 0 \,. \tag{29}$$

Phase plane analysis. Following the analysis of TWE (21) for a reaction-diffusion equation (20) with $r(u) = -h(u)$ and $\sigma = \delta$, necessary conditions on the parameters can be identified. First, a TWE is rewritten as a system of first-order ODEs with vector field F. Then the condition on stationary states implies that endstates $u_\pm$ and wave speed c have to satisfy

$$f(u_+) - f(u_-) = c(u_+ - u_-) \,. \tag{30}$$

This condition is known in shock wave theory as Rankine–Hugoniot condition (B.4) on the shock triple $(u_-, u_+; c)$. The (nonlinear) stability of hyperbolic stationary states (u_s, v_s) of F is determined by the eigenvalues

$$\lambda_\pm = -\frac{1}{2}\frac{\varepsilon}{\delta} \pm \frac{\sqrt{\varepsilon^2 + 4\delta h'(u_s)}}{2|\delta|} \tag{31}$$

of the Jacobian $\mathrm{D}F(u_s, v_s)$. If $\varepsilon, \delta > 0$, then $(u_+, 0)$ is always either a saddle or stable node, and $h'(u_-) = f'(u_-) - c > 0$ ensures that $(u_-, 0)$ is unstable. For example, Lax' entropy condition (B.5), i.e. $f'(u_+) < c < f'(u_-)$, implies the latter condition.

Convex Flux Functions

Theorem 2 *Suppose $f \in C^2(\mathbb{R})$ is a strictly convex function. Let ε, δ be positive and let $(u_-, u_+; c)$ satisfy the Rankine–Hugoniot condition* (B.4) *and the entropy condition* (B.5)*, i.e. $u_- > u_+$. Then, there exists an (up to translations) unique TWS $(\bar{u}, c)$ of* (26) *in the sense of Definition 1.*

Proof We consider the associated reaction-diffusion equation (29), i.e. $\partial_t u = r(u) + \delta\partial_x^2 u$ with $r(u) = -h(u)$. Due to (B.4) and (B.5), $r(u)$ is monostable in the sense of Definition 4. Moreover, function r is strictly concave, since $r''(u) = -f''(u)$ and $f \in C^2(\mathbb{R})$ is strictly convex. In fact, $(u_\pm, 0)$ are the only stationary points of system (22), where $(u_-, 0)$ is a saddle point and $(u_+, 0)$ is a stable node. Thus, for all wave speeds ε there exists a TWS $(\bar{u}, \varepsilon)$ – with possibly oscillatory pro-

file $\bar{u}$ – of reaction-diffusion equation (29). Moreover, $(\bar{u}, c)$ is a TWS of (26), due to (27)–(29). □

The TWP for KdVB equations (26) with Burgers' flux $f(u) = u^2$ has been investigated in [16]. The sign of δ in (26) is irrelevant, since it can be changed by a transformation $\tilde{x} = -x$ and $\tilde{u}(\tilde{x}, t) = -u(x, t)$, see also [41]. First, the results in Theorem 2 on the existence of TWS and geometry of its profiles are proven. More importantly, the authors investigate the convergence of profiles $\bar{u}(\xi; \varepsilon, \delta)$ in the limits $\varepsilon \to 0, \delta \to 0$, as well as ε and δ tending to zero simultaneously. Assuming that the ratio δ/ε^2 remains bounded, they show that the TWS converge to the classical Lax shocks for this vanishing diffusive-dispersive regularization [16].

Concave-Convex Flux Functions

Definition 5 ([45]) A function $f \in C^3(\mathbb{R})$ is called *concave-convex* if

$$u f''(u) > 0 \quad \forall u \neq 0\,, \quad f'''(0) \neq 0\,, \quad \lim_{u \to \pm\infty} f'(u) = +\infty\,. \tag{32}$$

Here the single inflection point is shifted without loss of generality to the origin. We consider a cubic flux function $f(u) = u^3$ as the prototypical concave-convex flux function with a single inflection point, see [39, 45].

Proposition 5 ([38, 41]) *Suppose* $f(u) = u^3$ *and* $\varepsilon > 0$.

(a) If $\delta \leq 0$ *then a TWS* $(\bar{u}, c)$ *of* (26) *exists if and only if* $(u_-, u_+; c)$ *satisfy the Rankine–Hugoniot condition* (B.4) *and the entropy condition* (B.5).

(b) If $\delta > 0$ *then a TWS* $(\bar{u}, c)$ *of* (26) *exists for* $u_- > 0$ *if and only if* $u_+ \in S(u_-)$ *with*

$$S(u_-) = \begin{cases} [-\frac{u_-}{2}, u_-) & \text{if } u_- \leq 2\beta\,, \\ \{-u_- + \beta\} \cup [-\beta, u_-) & \text{if } u_- > 2\beta\,, \end{cases} \tag{33}$$

where the coefficient β *is given by* $\beta = \frac{\sqrt{2}}{3}\frac{\varepsilon}{\sqrt{\delta}}$.

Proof Following the discussion from (26)–(29), we consider the associated reaction-diffusion equation (29), i.e. $\partial_t u = r(u) + \delta \partial_x^2 u$ with $r(u) = -h(u)$. From this point of view, we need to classify the reaction term $r(u) = -h(u)$: Whereas $r(u_-) = 0$ by definition, $r(u_+) = 0$ if and only if $(u_-, u_+; c)$ satisfies the Rankine–Hugoniot condition (B.4). The Rankine–Hugoniot condition implies $c = u_+^2 + u_+ u_- + u_-^2$. Hence, the reaction term $r(u)$ has a factorization

$$r(u) = -(u^3 - u_-^3 - c(u - u_-)) = -(u - u_-)\,(u - u_+)\,(u + u_+ + u_-) \tag{34}$$

Thus, $r(u)$ is a cubic polynomial with three roots $u_1 \leq u_2 \leq u_3$, such that $r(u) = -(u - u_1)(u - u_2)(u - u_3)$. In case of distinct roots $u_1 < u_2 < u_3$ we deduce $r'(u_1) < 0$, $r'(u_2) > 0$ and $r'(u_3) < 0$. The ordering of the roots $u_\pm$ and $u_* = -u_- - u_+$ depending on $u_\pm$ is visualized in Fig. 2. Next, we will discuss the results

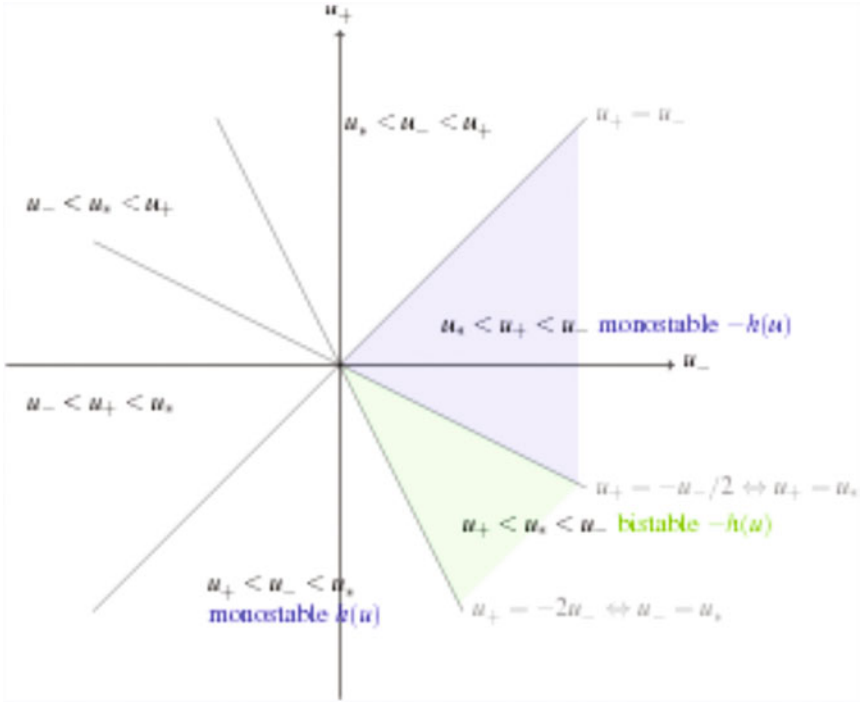


Fig. 2 Classification of the cubic reaction function $r(u) = -h(u)$ in (34) depending on its roots u_-, u_+ and $u_* = -u_- - u_+$ according to Definition 4

in Proposition 5(b) (for $u_- > 0$ and $\delta > 0$) via results on the existence of TWS for a reaction-diffusion equation (29).

1. For $u_+ < u_* < u_-$, function $r(u)$ is bistable, see also Fig. 2. Due to Proposition 4, there exists an (up to translations) unique TWS $(\bar{u}, \varepsilon)$ with possibly negative wave speed. Under our assumption that the wave speed ε is positive, relation (25) yields the restriction $-u_+ > u_-$. In fact, for $u_- > 2\beta$ and $u_+ = -u_- + \beta$ there exists a TWS $(\bar{u}, \varepsilon)$ for reaction-diffusion equation (29), see [41, Theorem 3.4]. The function r is bistable with $u_* = -u_- - u_+ = -\beta$, hence $f'(u_\pm) > c$. This violates Lax' entropy condition (B.5) and is known in the shock wave theory as a slow undercompressive shock [45].
2. For $u_* < u_+ < u_-$, function $r(u)$ is monostable, see Fig. 2. Due to Proposition 4, there exists a critical wave speed c_*, such that monotone TWS $(\bar{u}, \varepsilon)$ for (29) exist for all $\varepsilon \geq c_*$. However, not all endstates (u_-, u_+) in the subset defined by $u_* < u_+ < u_-$ admit a TWS $(\bar{u}, c)$, see (33) and Fig. 3b. The TWS $(\bar{u}, c)$ associated to non-classical shocks appear again, with reversed roles for the roots u_+ and u_*: For $u_- > 2\beta$ and $u_+ = -\beta$, there exists a TWS $(\bar{u}, \varepsilon)$ for reaction-diffusion equation (29), see [41, Theorem 3.4]. These TWS form a horizontal halfline in Fig. 3b and divides the set defined by $u_* < u_+ < u_-$ into two subsets.

In particular, TWS exist only for endstates (u_-, u_+) in the subset above this halfline.

3. For $u_+ < u_- < u_*$, function $r(u) = -h(u)$ satisfies $r(u) < 0$ for all $u \in (u_+, u_-)$, see also Fig. 2. Thus the necessary condition (25) can not be fulfilled for positive $c = \varepsilon$, hence there exists no TWS $(\bar{u}, \varepsilon)$ for the reaction-diffusion equation.
4. For $u_* < u_- < u_+$, function $r(u)$ is monostable with reversed roles of the endstates $u_\pm$, see Fig. 2. Due to Proposition 4, there exists a TWS $(\bar{u}, \varepsilon)$ however satisfying $\lim_{\xi \to \mp\infty} \bar{u}(\xi) = u_\pm$.

If $\delta = 0$, then equation (26) is a viscous conservation law, and its TWE (28) is a simple ODE $-\varepsilon \bar{u}' = r(\bar{u})$ with $r(u) = -h(u)$. Thus a heteroclinic orbit exists only for monostable $r(u)$, i.e. if the unstable node u_- and the stable node u_+ are not separated by any other root of r.

If $\delta < 0$, then we rewrite TWE (28) as $\varepsilon \bar{u}' = h(u) + |\delta| \bar{u}''$. It is associated to a reaction-diffusion equation $\partial_t u = h(u) + |\delta| \partial_x^2 u$ via a TWS ansatz $u(x,t) = \bar{u}(x - (-\varepsilon)t)$; note the change of sign for the wave speed. If $u_+ < u_* < u_-$ then $h(u)$ is an unstable reaction function. Thus there exists no TWS $(\bar{u}, -\varepsilon)$ according to Proposition 4. If $u_* < u_+ < u_-$ then function $h(u) = -r(u)$ satisfies $h(u) < 0$ for all $u \in (u_+, u_-)$, see also Fig. 2. The necessary condition (25) is still fine, since also the sign of the wave speed changed. In contrast to the case $\delta > 0$, there exists no TWS connecting u_- with u_*, which would indicate a bifurcation. Thus, the existence of TWS for all pairs (u_-, u_+) in the subset defined by $u_* < u_+ < u_-$ can be proven. The TWP for other pairs (u_-, u_+) is discussed similarly. □

4 TWP for Nonlocal Evolution Equations

4.1 Reaction-Diffusion Equations

The first example of a reaction-diffusion equation with nonlocal diffusion is the integro-differential equation

$$\partial_t u = J * u - u + r(u)\,, \quad t > 0\,, \quad x \in \mathbb{R}\,, \tag{35}$$

for some even, non-negative function J with mass one, i.e. for all $x \in \mathbb{R}$

$$J \in C(\mathbb{R})\,, \quad J \geq 0\,, \quad J(x) = J(-x)\,, \quad \int_{\mathbb{R}} J(y)\ \mathrm{d}y = 1\,, \tag{36}$$

and some function r. The operator $\mathscr{L}[u] = J * u - u$ is a Lévy operator, see (13), which models nonlocal diffusion. It is the infinitesimal generator of a compound Poisson stochastic process, which is a pure jump process.

The TWP for given endstates $u_\pm$ is to study the existence of a TWS $(\bar{u}, c)$ for (35) in the sense of Definition 1. Such a TWS $(\bar{u}, c)$ satisfies the TWE $-c\bar{u}' = J * \bar{u} - \bar{u} + r(\bar{u})$ for $\xi \in \mathbb{R}$. Next, we recall some results on the TWP for (35), which will

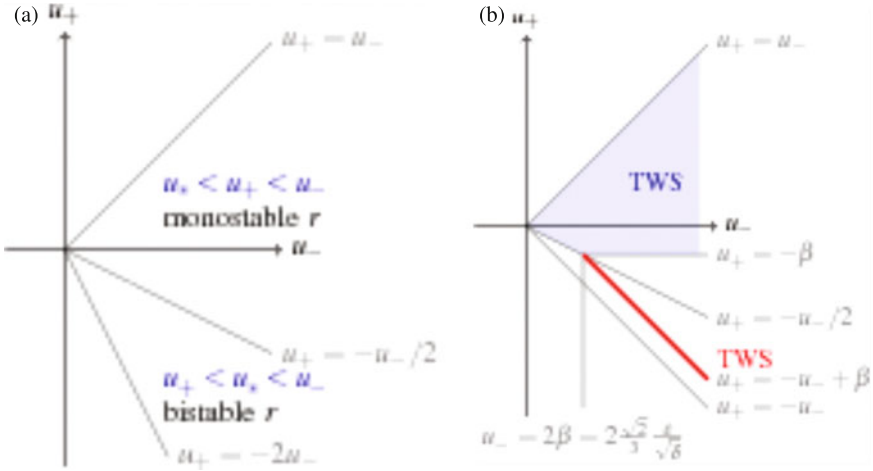


Fig. 3 **a** Classification of reaction function r in (34) depending on its roots u_-, u_+ and $u_* = -u_- - u_+$ according to Definition 4; **b** Endstates $u_\pm$ in the *shaded region* and on the *thick line* can be connected by TWS of the cubic KdVB equation; TWS in the *shaded region* and on the *thick line* are associated to classical and non-classical shocks of $\partial_t u + \partial_x u^3 = 0$, respectively. For a classical shock the shock triple satisfies Lax' entropy condition $f'(u_-) > c > f'(u_+)$; i.e. characteristics in the Riemann problem meet at the shock. In contrast, the non-classical shocks are of slow undercompressive type, i.e. characteristics in the Riemann problem cross the shock

depend crucially on the type of reaction function r and the tail behavior of a kernel function J. We will present the existence of TWS with monotone decreasing profiles $\bar{u}$, which will follow from the cited literature after a suitable transformation.

Proposition 6 ((monostable [27]), (bistable [14, 24])) *Suppose $u_- > u_+$ and consider reaction functions r in the sense of Definition 4. Suppose $J \in W^{1,1}(\mathbb{R})$ and its continuous representative satisfies* (36).

- *If r is monostable and there exists $\lambda > 0$ such that $\int_{\mathbb{R}} J(y) \exp(\lambda y) \, \mathrm{d}y < \infty$ then there exists a positive constant c_* such that for all $c \geq c_*$ there exists a monotone TWS $(\bar{u}, c)$ of* (35). *For $c < c_*$ no such monotone TWS exists.*
- *If r is bistable and $\int_{\mathbb{R}} |y| J(y) \, \mathrm{d}y < \infty$, then there exists an (up to translations) unique monotone TWS $(\bar{u}, c)$ of* (35).

For monostable reaction functions, the tail behavior of kernel function J is very important. There exist kernel functions J, such that TWS exist only for bistable – but not for monostable – reaction functions r, see [58]. The prime example are kernel functions J which decay more slowly than any exponentially decaying function as $|x| \to \infty$ in the sense that $J(x) \exp(\eta |x|) \to \infty$ as $|x| \to \infty$ for all $\eta > 0$.

For reaction-diffusion equations of bistable type, Chen established a unified approach [24] to prove the existence, uniqueness and asymptotic stability with exponential decay of traveling wave solutions. The results are established for a subclass of nonlinear nonlocal evolution equations

$$\partial_t u(x,t) = \mathscr{A}[u(\cdot,t)](x) \quad \text{for } (x,t) \in \mathbb{R} \times (0,T]\,,$$

where the nonlinear operator $\mathscr{A}$ is assumed to

(a) be independent of t;
(b) generate a L^∞ semigroup;
(c) be translational invariant, i.e. $\mathscr{A}$ satisfies for all $u \in \operatorname{dom} \mathscr{A}$ the identity

$$\mathscr{A}[u(\cdot + h)](x) = \mathscr{A}[u(\cdot)](x + h) \quad \forall x\,, h \in \mathbb{R}\,.$$

Consequently, there exists a function $r : \mathbb{R} \to \mathbb{R}$ which is defined by $\mathscr{A}[\upsilon \mathbf{1}] = r(\upsilon)\mathbf{1}$ for $\upsilon \in \mathbb{R}$ and the constant function $\mathbf{1} : \mathbb{R} \to \mathbb{R}$, $x \mapsto 1$. This function r is assumed to be bistable in the sense of Definition 4;
(d) satisfy a comparison principle: If $\partial_t u \geq \mathscr{A}[u]$, $\partial_t v \leq \mathscr{A}[v]$ and $u(\cdot,0) \gneq v(\cdot,0)$, then $u(\cdot,t) > v(\cdot,t)$ for all $t > 0$.

Chen's approach relies on the comparison principle and the construction of sub- and supersolutions for any given traveling wave solution. Importantly, the method does not depend on the balance of the potential. More quantitative versions of the assumptions on $\mathscr{A}$ are needed in the proofs. Finally integro-differential evolution equations

$$\partial_t u = \varepsilon \partial_x^2 u + G(u, J_1 * S^1(u), \ldots, J_n * S^n(u)) \tag{37}$$

are considered for some diffusion constant $\varepsilon \geq 0$, smooth functions G and S^k, and kernel functions $J_k \in C^1(\mathbb{R}) \cap W^{1,1}(\mathbb{R})$ satisfying (36) where $k = 1, \ldots, n$. Additional assumptions on the model parameters guarantee that (37) can be interpreted as a reaction-diffusion equation with bistable reaction function including Eqs. (20) and (35) as special cases.

Another example of reaction-diffusion equations with nonlocal diffusion are the integro-differential equations

$$\partial_t u = D_\theta^a u + r(u)\,, \quad t > 0\,, \quad x \in \mathbb{R}\,, \tag{38}$$

for a (particle) density $u = u(x,t)$, some function $r = r(u)$, and a Riesz–Feller operator D_θ^a with $(a,\theta) \in \mathfrak{D}_{a,\theta}$. The nonlocal Riesz–Feller operators are models for superdiffusion, where from a probabilistic view point a cloud of particle is assumed to spread faster than by following Brownian motion. Integro-differential equation (38) can be derived as a macroscopic equation for a particle density in the limit of modified Continuous Time Random Walk (CTRW), see [48]. In the applied sciences, Eq. (38) has found many applications, see [54, 57] for extensive reviews on modeling, formal analysis and numerical simulations.

The TWP for given endstates $u_\pm$ is to study the existence of a TWS $(\bar{u}, c)$ for (38) in the sense of Definition 1. Such a TWS $(\bar{u}, c)$ satisfies the TWE

$$-c\bar{u}' = D_\theta^a \bar{u} + r(\bar{u})\,, \quad \xi \in \mathbb{R}\,. \tag{39}$$

First we collect mathematical rigorous results about the TWP associated to (38) in case of the fractional Laplacian $D_0^a = -(-\Delta)^{a/2}$ for $a \in (0, 2)$, i.e. a Riesz–Feller operator D_θ^a with $\theta = 0$.

Proposition 7 ((monostable [17, 18, 34]), (bistable [19–21, 25, 37, 50])) *Suppose $u_- > u_+$. Consider the TWP for reaction-diffusion equation (38) with functions r in the sense of Definition 4 and fractional Laplacian D_0^a, i.e. symmetric Riesz–Feller operators D_θ^a with $0 < a < 2$ and $\theta = 0$.*

- *If r is monostable then there does not exist any TWS $(\bar{u}, c)$ of (38).*
- *If r is bistable then there exists an (up to translations) unique monotone TWS $(\bar{u}, c)$ of (38).*

For monostable reaction functions, Cabré and Roquejoffre prove that a front moves exponentially in time [17, 18]. They note that the genuine algebraic decay of the heat kernels G_0^a associated to fractional Laplacians is essential to prove the result, which implies that no TWS with constant wave speed can exist. Engler [34] considered the TWP for (38) for a different class of monostable reaction functions r and non-extremal Riesz–Feller operators D_θ^a with $(a, \theta) \in \mathfrak{D}_{a,\theta}^+$ and $\mathfrak{D}_{a,\theta}^+ := \{ (a, \theta) \in \mathfrak{D}_{a,\theta} \,|\, |\theta| < \min\{a, 2 - a\} \}$. Again the associated heat kernels $G_\theta^a(x, t)$ with $(a, \theta) \in \mathfrak{D}_{a,\theta}^+$ decay algebraically in the limits $x \to \pm\infty$, see [47].

To our knowledge, we established the first result [6] on existence, uniqueness (up to translations) and stability of traveling wave solutions of (38) with Riesz–Feller operators D_θ^a for $(a, \theta) \in \mathfrak{D}_{a,\theta}$ with $1 < a < 2$ and bistable functions r. We present our results for monotone decreasing profiles, which can be inferred from our original result after a suitable transformation.

Theorem 3 ([6]) *Suppose $u_- > u_+$, $(a, \theta) \in \mathfrak{D}_{a,\theta}$ with $1 < a < 2$, and $r \in C^\infty(\mathbb{R})$ is a bistable reaction function. Then there exists an (up to translations) unique monotone decreasing TWS $(\bar{u}, c)$ of (38) in the sense of Definition 1.*

The technical details of the proof are contained in [6], whereas in [5] we give a concise overview of the proof strategy and visualize the results also numerically. In a forthcoming article [4], we extend the results to all non-trivial Riesz–Feller operators D_θ^a with $(a, \theta) \in \mathfrak{D}_{a,\theta}^\diamond$. The smoothness assumption on r is convenient, but not essential. To prove Theorem 3, we follow – up to some modifications – the approach of Chen [24]. It relies on a strict comparison principle and the construction of sub- and supersolutions for any given TWS. His quantitative assumptions on operator $\mathscr{A}$ are too strict, such that his results are not directly applicable. A modification allows to cover the TWP for (38) for all Riesz–Feller operators D_θ^a with $1 < a < 2$ also for non-zero θ, and all bistable functions r regardless of the balance of the potential.

Next, we quickly review different methods to study the TWP of reaction-diffusion equation (38) with bistable function r and fractional Laplacian. In case of a classical reaction-diffusion equation (20), the existence of a TWS can be studied via phase-plane analysis [13, 35]. This method has no obvious generalization to our TWP for (38), since its traveling wave equation (39) is an integro-differential equation. The variational approach has been focused – so far – on symmetric diffusion operators such as fractional Laplacians and on balanced potentials, hence covering only

stationary traveling waves [50]. Independently, the same results are achieved in [19–21] by relating the stationary TWE $(39)_{\theta=0,c=0}$ via [22] to a boundary value problem for a nonlinear partial differential equation. The homotopy to a simpler TWP has been used to prove the existence of TWS in case of (35), and $(38)_{\theta=0}$ with unbalanced potential [37].

Chmaj [25] also considers the TWP for $(38)_{\theta=0}$ with general bistable functions r. He approximates a given fractional Laplacian by a family of operators $J_\varepsilon * u - (\int J_\varepsilon)u$ such that $\lim_{\varepsilon\to 0} J_\varepsilon * u - (\int J_\varepsilon)u = D_0^a u$ in an appropriate sense. This allows him to obtain a TWS of $(38)_{\theta=0}$ with general bistable function r as the limit of the TWS u_ε of (35) associated to $(J_\varepsilon)_{\varepsilon\geq 0}$. It might be possible to modify Chmaj's approach to study reaction-diffusion equation (38) with asymmetric Riesz–Feller operators. This would give an alternative existence proof for TWS in Theorem 3. However, Chen's approach allows to establish uniqueness (up to translations) and stability of TWS as well.

4.2 Nonlocal Korteweg-de Vries-Burgers Equation

First we consider the integro-differential equation in multi-dimensions $d \geq 1$

$$\partial_t u + \partial_x f(u) = \varepsilon \Delta_x u + \gamma \varepsilon^2 \sum_{j=1}^{d} \left(\phi_\varepsilon * \partial_{x_j} u - \partial_{x_j} u\right), \quad x \in \mathbb{R}^d, \quad t > 0, \quad (40)$$

for parameters $\varepsilon > 0$, $\gamma \in \mathbb{R}$, a smooth even non-negative function ϕ with compact support and unit mass, i.e. $\int_{\mathbb{R}^d} \phi(x)\ \mathrm{d}x = 1$, and the rescaled kernel function $\phi_\varepsilon(x) = \phi(x/\varepsilon)/\varepsilon^d$. It has been derived as a model for phase transitions with long range interactions close to the surface, which supports planar TWS associated to undercompressive shocks of (B.1), see [52]. A planar TWS $(\bar{u}, c)$ is a solution $u(x,t) = \bar{u}(x - cte)$ for some fixed vector $e \in \mathbb{R}^d$, such that the profile is transported in direction e. The existence of planar TWS is proven by reducing the problem to a one-dimensional TWP for $(40)_{d=1}$, identifying the associated reaction-diffusion equation (35) and using results in Proposition 6. For cubic flux function u^3, the existence of planar TWS associated to undercompressive shocks of (B.1) is established. Moreover, the well-posedness of its Cauchy problem and the convergence of solutions u^ε as $\varepsilon \searrow 0$ have been studied [52].

Another example is the **fractal Korteweg-de Vries-Burgers equation**

$$\partial_t u + \partial_x f(u) = \varepsilon \partial_x \mathscr{D}_+^\alpha u + \delta \partial_x^3 u, \quad x \in \mathbb{R}, \quad t > 0, \quad (41)$$

for some $\varepsilon > 0$ and $\delta \in \mathbb{R}$.

Equation (41) with $\alpha = 1/3$ has been derived as a model for shallow water flows, by performing formal asymptotic expansions associated to the triple-deck (boundary layer) theory in fluid mechanics, e.g. see [44, 55]. In particular, the situations of

one-layer and two-layer shallow water flows have been considered, which yield a quadratic (one layer) and cubic flux function (two layer), respectively. In the monograph [49], similar models are considered and the well-posedness of the initial value problem and possible wave-breaking are studied.

The TWP for given endstates $u_\pm$ is to study the existence of a TWS $(\bar{u}, c)$ for (41) in the sense of Definition 1. Such a TWS $(\bar{u}, c)$ satisfies the TWE

$$h(\bar{u}) := f(\bar{u}) - f(u_-) - c(\bar{u} - u_-) = \varepsilon \mathscr{D}_+^\alpha \bar{u} + \delta \bar{u}'' \,. \tag{42}$$

We obtain a necessary condition for the existence of TWS – see also (25) – by multiplying the TWE with $\bar{u}'$ and integrating on $\mathbb{R}$,

$$\int_{u_-}^{u_+} h(u) \;\mathrm{d}u \;= \varepsilon \int_{-\infty}^{\infty} \bar{u}' \; \mathscr{D}_+^\alpha \bar{u}(\xi) \;\mathrm{d}\xi \;\geq 0 \,, \tag{43}$$

where the last inequality follows from (A.1).

Connection with reaction-diffusion equation. If a TWS $(\bar{u}, c)$ for (41) exists, then $u(x,t) = \bar{u}(x)$ is a stationary TWS $(\bar{u}, 0)$ of the evolution equation

$$\partial_t u = -\varepsilon \mathscr{D}_+^\alpha u - \delta \partial_x^2 u + h(u), \quad x \in \mathbb{R}, \quad t > 0. \tag{44}$$

To interpret Eq. (44) as a reaction-diffusion equation, we need to verify that $-\varepsilon \mathscr{D}_+^\alpha u - \delta \partial_x^2 u$ is a diffusion operator, e.g. that $-\varepsilon \mathscr{D}_+^\alpha u - \delta \partial_x^2 u$ generates a positivity preserving semigroup.

Lemma 1 *Suppose $0 < \alpha < 1$ and $\gamma_1, \gamma_2 \in \mathbb{R}$. The operator $\gamma_1 \mathscr{D}_+^\alpha u + \gamma_2 \partial_x^2 u$ is a Lévy operator if and only if $\gamma_1 \leq 0$ and $\gamma_2 \geq 0$. Moreover, the associated heat kernel is strictly positive if and only if $\gamma_2 > 0$.*

Proof For $\alpha \in (0,1)$, the operator $-\mathscr{D}_+^\alpha$ is a Riesz–Feller operator $D_{-\alpha}^\alpha$ and generates a positivity preserving convolution semigroup with a Lévy stable probability distribution $G_{-\alpha}^\alpha$ as its kernel. The probability distribution is absolutely continuous with respect to Lebesgue measure and its density has support on a half-line [47]. For example the kernel associated to $-\mathscr{D}^{1/2}$ is the Lévy–Smirnov distribution. Thus, for $\gamma_1 \leq 0$ and $\gamma_2 \geq 0$, the operator $\gamma_1 \mathscr{D}_+^\alpha u + \gamma_2 \partial_x^2 u$ is a Lévy operator, because it is a linear combination of Lévy operators. Using the notation for Fourier symbols of Riesz–Feller operators, the partial Fourier transform of equation

$$\partial_t u = -|\gamma_1| \mathscr{D}^\alpha[u] + \gamma_2 \partial_x^2 u$$

is given by $\partial_t \mathscr{F}[u](k) = (|\gamma_1| \psi_{-\alpha}^\alpha(k) - \gamma_2 k^2) \mathscr{F}[u](k)$. Therefore, the operator generates a convolution semigroup with heat kernel

$$\mathscr{F}^{-1}[\exp\{(|\gamma_1| \psi_{-\alpha}^\alpha(k) - \gamma_2 k^2)\, t\}](x) = G_{-\alpha}^\alpha(\cdot, |\gamma_1| t) * G_0^2(\cdot, \gamma_2 t)\,(x) \,,$$

which is the convolution of two probability densities. The kernel is positive on $\mathbb{R}$ since probability densities are non-negative on $\mathbb{R}$ and the normal distribution G_0^2 is positive on $\mathbb{R}$ for positive $\gamma_2 t$.

The operator $\mathscr{D}_+^\alpha$ for $\alpha \in (0, 1)$ is not a Riesz–Feller operator, see Fig. 1, and it generates a semigroup which is not positivity preserving. Therefore, this operator and any linear combination with it for $\gamma_1 > 0$ is not a Lévy operator. □

Convex Flux Functions

Proposition 8 *Consider* (41) *with* $0 < \alpha < 1$, $\delta \in \mathbb{R}$ *and strictly convex flux function* $f \in C^3(\mathbb{R})$. *For a shock triple* $(u_-, u_+; c)$ *satisfying the Rankine–Hugoniot condition (B.4), a non-constant TWS* $(\bar{u}, c)$ *can exist if and only if Lax' entropy condition (B.5) is fulfilled, i.e.* $u_- > u_+$.

Proof The Rankine–Hugoniot condition (B.4) ensures that $h(u)$ in (42) has exactly two roots $u_\pm$. If Lax' entropy condition (B.5) is fulfilled, then $u_- > u_+$ and $-h(u)$ is monostable in the sense of Definition 4. Thus, the necessary condition (43) is satisfied. If $u_- = u_+$ then (43) implies that $\bar{u}$ is a constant function satisfying $\bar{u} \equiv u_\pm$. If $u_- < u_+$ then $-h(u)$ is monostable in the sense of Definition 4 with reversed roles of $u_\pm$. Thus, the necessary condition (43) is not satisfied. □

Next, we recall some existence result which have been obtained by directly studying the TWE. In an Addendum [28], we removed an initial assumption on the solvability of the linearized TWE.

Theorem 4 ([3]) *Consider* (41) *with* $\delta = 0$ *and convex flux function* $f(u)$. *For a shock triple* $(u_-, u_+; c)$ *satisfying* (B.4) *and* (B.5), *there exists a monotone TWS of* (41) *in the sense of Definition 1, whose profile* $\bar{u} \in C_b^1(\mathbb{R})$ *is unique (up to translations) among all functions* $u \in u_- + H^2(-\infty, 0) \cap C_b^1(\mathbb{R})$.

This positive existence result is consistent with the negative existence result in Proposition 7 and Engler [34] for (38) with non-extremal Riesz–Feller operators D_θ^a for $(a, \theta) \in \mathfrak{D}_{a,\theta}^+$. The reason is that $-\mathscr{D}_+^\alpha$ for $0 < \alpha < 1$ is the generator of a convolution semigroup with a one-sided strictly stable probability density function as its heat kernel; in contrast to heat kernels with genuine algebraic decay [17, 18, 34].

Theorem 5 ([2]) *Consider* (41) *with flux function* $f(u) = u^2/2$. *For a shock triple* $(u_-, u_+; c)$ *satisfying (B.4) and (B.5), there exists a TWS of* (41) *in the sense of Definition 1, whose profile* $\bar{u}$ *is unique (up to translations) among all functions* $u \in u_- + H^4(-\infty, 0) \cap C_b^3(\mathbb{R})$.

If dispersion dominates diffusion then the profile of a TWS $(\bar{u}, c)$ will be oscillatory in the limit $\xi \to \infty$. For a classical KdVB equation this geometry of profiles depends on the ratio ε^2/δ and the threshold can be determined explicitly.

Concave-convex flux functions. We consider a cubic flux function $f(u) = u^3$ as the prototypical concave-convex flux function. Again the necessary condition (43) and the classification of function $h(u) = -r(u)$ in Fig. 2 can be used to identify non-admissible shock triples $(u_-, u_+; c)$ for the TWP of (41).

We conjecture that a statement analogous to Proposition 5 holds true. Of special interest is again the occurrence of TWS $(\bar{u}, c)$ associated to non-classical shocks, which are only expected in case of (41) with $\varepsilon > 0$ and $\delta > 0$.

Proposition 9 *Suppose* $f(u) = u^3$ *and* $\varepsilon > 0$.

1. *If* $\delta \leq 0$ *then a TWS* $(\bar{u}, c)$ *of* (41) *exists if and only if* $(u_-, u_+; c)$ *satisfy the Rankine–Hugoniot condition (B.4) and the entropy condition (B.5).*
2. *Conjecture: If* $\delta > 0$ *then a TWS* $(\bar{u}, c)$ *of* (41) *exists for* $u_- > 0$ *if and only if* $u_+ \in S(u_-)$ *for some set* $S(u_-)$ *similar to* (33).

Proof (*Sketch of proof*) If $\delta = 0$, then Eq. (41) is a viscous conservation law, and its TWE (42) is a fractional differential equation $\varepsilon \mathscr{D}_+^\alpha \bar{u} = h(\bar{u})$. Thus a heteroclinic orbit exists only for monostable $-h(u)$, i.e. if the unstable node u_- and the stable node u_+ are not separated by any other root of h. This follows from Theorem 4 and its proof in [3, 28].

If $\delta < 0$, then the TWE (42) is associated to a reaction-diffusion equation (44) via a stationary TWS ansatz $u(x, t) = \bar{u}(x)$. First we note that a stronger version of the necessary condition (43) is available

$$\int_{-\infty}^{\xi} h(\bar{u})\bar{u}'(y) \ \mathrm{d}y = \varepsilon \int_{-\infty}^{\xi} \bar{u}' \, \mathscr{D}_+^\alpha \bar{u}(y) \ \mathrm{d}y \geq 0 \,, \quad \forall \xi \in \mathbb{R} \,, \tag{45}$$

see [2]. If $u_+ < u_* < u_-$ then $h(u)$ is an unstable reaction function, see Fig. 2. Thus there exists no TWS in the sense of Definition 1 satisfying the necessary condition (45). If $u_* < u_+ < u_-$ then function $-h(u)$ is monostable in the sense of Definition 4 and the necessary condition (43) can be satisfied. The existence of a TWS $(\bar{u}, c)$ can be proven by following the analysis in [2, 28]. The TWP for other pairs (u_-, u_+) is discussed similarly.

If $\delta > 0$ then the occurrence of TWS $(\bar{u}, c)$ associated to non-classical shocks is possible. Unlike in our previous examples, the associated evolution Eq. (44) is not a reaction-diffusion equation, since $-\varepsilon \mathscr{D}_+^\alpha \bar{u} - \delta \bar{u}''$ is not a Lévy operator. Especially, the results on existence of TWS for reaction-diffusion equations with bistable reaction function can not be used to prove the existence of TWS $(\bar{u}, c)$ associated to undercompressive shocks. Instead, we investigate the TWP directly [1], extending the analysis in [2, 28] for Burgers' flux to the cubic flux function $f(u) = u^3$. □

4.3 Fowler's Equation

Fowler's equation (8) for dune formation is a special case of the evolution equation

$$\partial_t u + \partial_x f(u) = \delta \partial_x^2 u - \varepsilon \partial_x \mathscr{D}_+^\alpha u \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{46}$$

with $0 < \alpha < 1$, positive constant $\varepsilon, \delta > 0$ and flux function f. Here the fractional derivative appears with the negative sign, but this instability is regularized by the second order derivative. The initial value problem for (8) is well-posed in L^2 [9]. However, it does not support a maximum principle, which is intuitive in the context of the application due to underlying erosions [9]. The existence of TWS of (8) – without assumptions (3) on the far-field behavior – has been proven [11].

For given endstates $u_\pm$, the TWP for (46) is to study the existence of a TWS $(\bar{u}, c)$ for (46) in the sense of Definition 1. Such a TWS $(\bar{u}, c)$ satisfies the TWE

$$h(\bar{u}) := f(\bar{u}) - f(u_-) - c(\bar{u} - u_-) = \delta \bar{u}' - \varepsilon \mathscr{D}_+^\alpha \bar{u} \,, \quad \xi \in \mathbb{R} \,. \tag{47}$$

For $\delta = 0$, the TWE reduces to a fractional differential equation $\varepsilon \mathscr{D}_+^\alpha \bar{u} = -h(\bar{u})$, which has been analyzed in [3, 28] for monostable functions $-h(u)$.

Equation (47) is also the TWE for a TWS $(\bar{u}, \delta)$ of an evolution equation

$$\partial_t u = -\varepsilon \mathscr{D}_+^\alpha u - h(u), \quad x \in \mathbb{R}, \quad t > 0. \tag{48}$$

For $\varepsilon > 0$, the operator is $-\varepsilon \mathscr{D}_+^\alpha \bar{u}$ is a Riesz–Feller operator $\varepsilon D_{-\alpha}^\alpha$ whose heat kernel $G_{-\alpha}^\alpha$ has only support on a halfline. For a shock triple $(u_-, u_+; c)$ satisfying the Rankine–Hugoniot condition (B.4), at least $h(u_\pm) = 0$ holds. Under these assumptions, Eq. (48) is a reaction-diffusion equation with a Riesz–Feller operator modeling diffusion.

The abstract method in [11] does not provide any information on the far-field behavior. Thus, assume the existence of a TWS $(\bar{u}, c)$ in the sense of Definition 1, for some shock triple $(u_-, u_+; c)$ satisfying the Rankine–Hugoniot condition (B.4). Again, a necessary condition is obtained by multiplying TWE (47) with $\bar{u}'$ and integrating on $\mathbb{R}$; hence,

$$\int_{u_-}^{u_+} h(u) \;\mathrm{d}u \;=\; \int_{\mathbb{R}} (\bar{u}')^2 \;\mathrm{d}\xi \;-\; \int_{\mathbb{R}} \bar{u}' \mathscr{D}_+^\alpha \bar{u} \;\mathrm{d}\xi \;. \tag{49}$$

The sign of the right hand side is not pre-determined since each integral is non-negative, see also (A.1).

For a cubic flux function $f(u) = u^3$ and a shock triple $(u_-, u_+; c)$ satisfying the Rankine–Hugoniot condition (B.4), we deduce a bistable reaction function $r(u) = -h(u)$ as long as $u_+ < -u_+ - u_- < u_-$ see Fig. 2. However, since the heat kernel has only support on a halfline, we can not obtain a strict comparison principle as needed in Chen's approach [4, 6, 24].

Acknowledgements The author was partially supported by Austrian Science Fund (FWF) under grant P28661 and the FWF-funded SFB #F65.

Appendix A: Caputo Fractional Derivative on $\mathbb{R}$

For $\alpha > 0$, the (Gerasimov–)Caputo derivatives are defined as, see [42, 54],

$$(\mathscr{D}_+^\alpha f)(x) = \begin{cases} f^{(n)}(x) & \text{if } \alpha = n \in \mathbb{N}_0 \,, \\ \frac{1}{\Gamma(n-\alpha)} \int_{-\infty}^x \frac{f^{(n)}(y)}{(x-y)^{\alpha-n+1}} \, \mathrm{d}y & \text{if } n-1 < \alpha < n \text{ for some } n \in \mathbb{N}_0 \,. \end{cases}$$

$$(\mathscr{D}_-^\alpha f)(x) = \begin{cases} f^{(n)}(x) & \text{if } \alpha = n \in \mathbb{N}_0 \,, \\ \frac{(-1)^n}{\Gamma(n-\alpha)} \int_x^\infty \frac{f^{(n)}(y)}{(y-x)^{\alpha-n+1}} \, \mathrm{d}y & \text{if } n-1 < \alpha < n \text{ for some } n \in \mathbb{N}_0 \,. \end{cases}$$

Properties:

- For $\alpha > 0$ and $\lambda > 0$

$$(\mathscr{D}_+^\alpha \exp(\lambda \cdot))(x) = \lambda^\alpha \exp(\lambda x) \,, \quad (\mathscr{D}_-^\alpha \exp(-\lambda \cdot))(x) = \lambda^\alpha \exp(-\lambda x)$$

- For $\alpha > 0$ and $f \in \mathscr{S}(\mathbb{R})$, a Caputo derivative is a Fourier multiplier operator with $(\mathscr{F} \mathscr{D}_+^\alpha f)(k) = (\mathrm{i}\, k)^\alpha (\mathscr{F} f)(k)$ where $(\mathrm{i}\, k)^\alpha = \exp(\alpha \pi \, \mathrm{i} \, \mathrm{sgn}(k)/2)$.
- If $\bar{u}$ is the profile of a TWS $(\bar{u}, c)$ in the sense of Definition 1, then

$$\int_{-\infty}^\infty \bar{u}'(y) \, \mathscr{D}_+^\alpha \bar{u}(y) \, \mathrm{d}y = \tfrac{1}{2} \int_{\mathbb{R}} \bar{u}'(x) \int_{\mathbb{R}} \frac{\bar{u}'(y)}{|x-y|^\alpha} \, \mathrm{d}y \, \mathrm{d}x \geq 0 \,, \tag{A.1}$$

  where the last inequality follows from [46, Theorem 9.8].

Appendix B: Shock Wave Theory for Scalar Conservation Laws

A standard reference on the theory of conservation laws is [29], whereas [45] covers the special topic of non-classical shock solutions. A scalar conservation law is a partial differential equation

$$\partial_t u + \partial_x f(u) = 0 \,, \quad t > 0 \,, \quad x \in \mathbb{R} \,, \tag{B.1}$$

for some flux function $f : \mathbb{R} \to \mathbb{R}$. For nonlinear functions f, it is well known that the initial value problem (IVP) for (B.1) with smooth initial data may not have a classical solution for all time $t > 0$ (due to shock formation). However, weak solutions may not be unique. The *Riemann problems* are a subclass of IVPs for (B.1), and especially important in some numerical algorithms: For given $u_-, u_+ \in \mathbb{R}$, find a weak solution $u(x,t)$ for the initial value problem of (B.1) with initial condition

$$u(x,0) = \begin{cases} u_- , & x < 0 , \\ u_+ , & x > 0 . \end{cases} \tag{B.2}$$

Weak solutions of a Riemann problem that are discontinuous for $t > 0$ may not be unique.

Example 1 A *shock wave* is a discontinuous solution of the Riemann problem,

$$u(x,t) = \begin{cases} u_- , & x < ct , \\ u_+ , & x > ct , \end{cases} \tag{B.3}$$

if the *shock triple* $(u_-, u_+; c)$ satisfies the Rankine–Hugoniot condition

$$f(u_+) - f(u_-) = c(u_+ - u_-) . \tag{B.4}$$

The Rankine–Hugoniot condition (B.4) is a necessary condition that $u_\pm$ are stationary states of an associated TWE (28), see (30).

Shock Admissibility

Classical approaches to select a unique weak solution of the Riemann problem are

(a) *Lax' entropy condition:*

$$f'(u_+) < c < f'(u_-) . \tag{B.5}$$

It ensures that in the method of characteristics all characteristics enter the shock/discontinuity of a shock solution (B.3). For convex flux function f, condition (B.5) reduces to $u_- > u_+$. Shocks satisfying (B.5) are also called Lax or classical shocks. For non-convex flux functions f, also non-classical shocks can arise in experiments, called slow undercompressive shocks if $f'(u_\pm) > c$, and fast undercompressive shocks if $f'(u_\pm) < c$.

(b) *Oleinik's entropy condition.*

$$\frac{f(w) - f(u_-)}{w - u_-} \geq \frac{f(u_+) - f(u_-)}{u_+ - u_-} \text{ for all } w \text{ between } u_- \text{ and } u_+ . \tag{B.6}$$

(c) *Entropy solutions* satisfying integral inequalities based on entropy-entropy flux pairs, such as Kruzkov's family of entropy-entropy flux pairs.

(d) *Vanishing viscosity.* In the classical vanishing viscosity approach, instead of (B.1) one considers for $\varepsilon > 0$ equation

$$\partial_t u + \partial_x f(u) = \varepsilon \partial_x^2 u , \quad t > 0 , \quad x \in \mathbb{R} , \tag{B.7}$$

where $\varepsilon\partial_x^2 u$ models diffusive effects such as friction. Equation (B.7) is a parabolic equation, hence the Cauchy problem has global smooth solutions u^ε for positive times, especially for Riemann data (B.2). An admissible weak solution of the Riemann problem is identified by studying the limit of u^ε as $\varepsilon \searrow 0$.
In other applications, different higher order effects may be important. For example, a nonlocal generalized KdVB equation (1) can be interpreted as a scalar conservation law (B.1) with higher-order effects $\mathscr{R}[u] := \varepsilon\mathscr{L}_1[u] + \delta\partial_x\mathscr{L}_2[u]$.
Already for convex functions f, the convergence of solutions of the regularized equations (e.g. (1)) to solutions of (B.1) reveals a diverse solution structure. The solutions of viscous conservation laws (B.7) converge for $\varepsilon \searrow 0$ to Kruzkov entropy solutions of (B.1). In contrast, in case of KdVB equation (4) with $f(u) = u^2$ the limit $\varepsilon, \delta \to 0$ depends on the relative strength of diffusion and dispersion:

- **Weak dispersion**: $\delta = O(\varepsilon^2)$ for $\varepsilon \to 0$ e.g. $\delta = \beta\varepsilon^2$ for some $\beta > 0$.
  TWS converge strongly to entropy solution of Burgers equation.
- **Moderate dispersion**: $\delta = o(\varepsilon)$ for $\varepsilon \to 0$ includes weak dispersion.
  TWS converge strongly to entropy solution of Burgers equation, see [51].
- **Strong dispersion**: weak limit of TWS for $\varepsilon, \delta \to 0$ may not be a weak solution of Burgers equation.

For non-convex flux functions f, a TWS may converge to a weak solution of (B.1) which is not an Kruzkov entropy solution, but a non-classical shock.

A simplistic shock admissibility criterion based on the vanishing viscosity approach is the existence of TWS for a given shock triple:

Definition 6 (*compare with* [41]) A solution u of the Riemann problem is called *admissible* (with respect to a fixed regularization $\mathscr{R}$), if there exists a TWS $(\bar{u}, c)$ in the sense of Definition 1 of the regularized equation (e.g. (1)) for every shock wave with shock triple $(u_-, u_+; c)$ in the solution u.

References

1. Achleitner, F., Cuesta, C.M.: Non-classical shocks in a non-local generalised Korteweg-de Vries-Burgers equation. Work in progress
2. Achleitner, F., Cuesta, C.M., Hittmeir, S.: Travelling waves for a non-local Korteweg-de Vries-Burgers equation. J. Differ. Equ. **257**(3), 720–758 (2014)
3. Achleitner, F., Hittmeir, S., Schmeiser, C.: On nonlinear conservation laws with a nonlocal diffusion term. J. Differ. Equ. **250**(4), 2177–2196 (2011)
4. Achleitner, F., Kuehn, C.: Traveling waves for a bistable reaction-diffusion equation with nonlocal Riesz-Feller operator. Work in progress
5. Achleitner, F., Kuehn, C.: Analysis and numerics of traveling waves for asymmetric fractional reaction-diffusion equations. Commun. Appl. Ind. Math. **6**(2), e–532, 25 (2015)
6. Achleitner, F., Kuehn, C.: Traveling waves for a bistable equation with nonlocal diffusion. Adv. Differ. Equ. **20**(9–10), 887–936 (2015)

7. Alibaud, N.: Entropy formulation for fractal conservation laws. J. Evol. Equ. **7**(1), 145–175 (2007)
8. Alibaud, N., Andreianov, B.: Non-uniqueness of weak solutions for the fractal Burgers equation. Ann. Inst. H. Poincaré Anal. Non Linéaire **27**(4), 997–1016 (2010)
9. Alibaud, N., Azerad, P., Isèbe, D.: A non-monotone nonlocal conservation law for dune morphodynamics. Differ. Integral Equ. **23**(1/2), 155–188 (2010)
10. Alibaud, N., Droniou, J., Vovelle, J.: Occurrence and non-appearance of shocks in fractal Burgers equations. J. Hyperbolic Differ. Equ. **4**(3), 479–499 (2007)
11. Alvarez-Samaniego, B., Azerad, P.: Existence of travelling-wave solutions and local well-posedness of the Fowler equation. Discret. Contin. Dyn. Syst. Ser. B **12**(4), 671–692 (2009)
12. Applebaum, D.: Lévy Processes and Stochastic Calculus. Cambridge Studies in Advanced Mathematics, vol. 116, 2nd edn. Cambridge University Press, Cambridge (2009)
13. Aronson, D.G., Weinberger, H.F.: Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. Partial Differential Equations and Related Topics. Lecture Notes in Mathematics, vol. 446, pp. 5–49. Springer, Berlin (1975)
14. Bates, P.W., Fife, P.C., Ren, X., Wang, X.: Traveling waves in a convolution model for phase transitions. Arch. Ration. Mech. Anal. **138**, 105–136 (1997)
15. Biler, P., Funaki, T., Woyczynski, W.A.: Fractal Burgers equations. J. Differ. Equ. **148**(1), 9–46 (1998)
16. Bona, J.L., Schonbek, M.E.: Travelling-wave solutions to the Korteweg-de Vries-Burgers equation. Proc. R. Soc. Edinb. Sect. A **101**(3–4), 207–226 (1985)
17. Cabré, X., Roquejoffre, J.-M.: Propagation de fronts dans les équations de Fisher-KPP avec diffusion fractionnaire. Comptes Rendus Mathématique **347**(23), 1361–1366 (2009)
18. Cabré, X., Roquejoffre, J.-M.: The influence of fractional diffusion in Fisher-KPP equations. Commun. Math. Phys. **320**(3), 679–722 (2013)
19. Cabré, X., Sire, Y.: Nonlinear equations for fractional Laplacians, I: Regularity, maximum principles, and Hamiltonian estimates. Ann. Inst. H. Poincaré Anal. Non Linéaire **31**(1), 23–53 (2014)
20. Cabré, X., Sire, Y.: Nonlinear equations for fractional Laplacians II: Existence, uniqueness, and qualitative properties of solutions. Trans. Am. Math. Soc. **367**(2), 911–941 (2015)
21. Cabré, X., Solà-Morales, J.: Layer solutions in a half-space for boundary reactions. Commun. Pure Appl. Math. **58**, 1678–1732 (2005)
22. Caffarelli, L., Silvestre, L.: An extension problem related to the fractional Laplacian. Commun. Partial Differ. Equ. **32**(7–9), 1245–1260 (2007)
23. Chan, C.H., Czubak, M., Silvestre, L.: Eventual regularization of the slightly supercritical fractional Burgers equation. Discret. Contin. Dyn. Syst. **27**(2), 847–861 (2010)
24. Chen, X.: Existence, uniqueness, and asymptotic stability of travelling waves in nonlocal evolution equations. Adv. Differ. Equ. **2**, 125–160 (1997)
25. Chmaj, A.: Existence of traveling waves in the fractional bistable equation. Arch. der Math. **100**(5), 473–480 (2013)
26. Chmaj, A.: Existence of traveling waves in the fractional Burgers equation. Preprint, 6 pp (2014). http://vixra.org/abs/1701.0324
27. Coville, J., Dupaigne, L.: On a non-local equation arising in population dynamics. Proc. R. Soc. Edinb. Sect. A **137**(4), 727–755 (2007)
28. Cuesta, C.M., Achleitner, F.: Addendum to "Travelling waves for a non-local Korteweg–de Vries–Burgers equation" [J. Differ. Equ. **257**(3), 720–758 (2014)]. J. Differ. Equ. **262**(2), 1155–1160 (2017)
29. Dafermos, C.M.: Hyperbolic conservation laws in continuum physics. Grundlehren der Mathematischen Wissenschaften, vol. 325. Springer, Berlin (2010)
30. Dong, H., Du, D., Li, D.: Finite time singularities and global well-posedness for fractal Burgers equations. Indiana Univ. Math. J. **58**(2), 807–821 (2009)
31. Droniou, J., Gallouet, T., Vovelle, J.: Global solution and smoothing effect for a non-local regularization of a hyperbolic equation. J. Evol. Equ. **3**(3), 499–521 (2003)

32. Droniou, J., Imbert, C.: Fractal first-order partial differential equations. Arch. Ration. Mech. Anal. **182**(2), 299–331 (2006)
33. Endal, J., Jakobsen, E.R.: L^1 contraction for bounded (nonintegrable) solutions of degenerate parabolic equations. SIAM J. Math. Anal. **46**(6), 3957–3982 (2014)
34. Engler, H.: On the speed of spread for fractional reaction-diffusion equations. Int. J. Differ. Equ. pages Art. ID 315421, 16 (2010)
35. Fife, P., McLeod, J.: The approach of solutions of nonlinear diffusion equations to travelling front solutions. Arch. Ration. Mech. Anal. **65**, 335–361 (1977)
36. Fowler, A.C.: Evolution equations for dunes and drumlins. RACSAM. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. **96**(3), 377–387 (2002)
37. Gui, C., Zhao, M.: Traveling wave solutions of Allen-Cahn equation with a fractional Laplacian. Ann. Inst. H. Poincaré Anal. Non Linéaire **32**(4), 785–812 (2015)
38. Hayes, B.T., LeFloch, P.G.: Non-classical shocks and kinetic relations: scalar conservation laws. Arch. Ration. Mech. Anal. **139**(1), 1–56 (1997)
39. Hayes, B.T., Shearer, M.: Undercompressive shocks and Riemann problems for scalar conservation laws with non-convex fluxes. Proc. R. Soc. Edinb. Sect. A **129**, 733–754 (1999)
40. Jacob, N.: Pseudo Differential Operators and Markov Processes, vol. I. Imperial College Press, London (2001). Fourier analysis and semigroups
41. Jacobs, D., McKinney, B., Shearer, M.: Travelling wave solutions of the modified Korteweg-de Vries-Burgers equation. J. Differ. Equ. **166**, 448–467 (1995)
42. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J.: Theory and Applications of Fractional Differential Equations. Elsevier Science B.V., Amsterdam (2006)
43. Kiselev, A., Nazarov, F., Shterenberg, R.: Blow up and regularity for fractal Burgers equation. Dyn. Partial Differ. Equ. **5**(3), 211–240 (2008)
44. Kluwick, A., Cox, E.A., Exner, A., Grinschgl, C.: On the internal structure of weakly nonlinear bores in laminar high Reynolds number flow. Acta Mech. **210**(1), 135–157 (2010)
45. LeFloch, P.G.: Hyperbolic Systems of Conservation Laws. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2002). The theory of classical and nonclassical shock waves
46. Lieb, E.H., Loss, M.: Analysis. Graduate Studies in Mathematics, vol. 14. American Mathematical Society, Providence (1997)
47. Mainardi, F., Luchko, Y., Pagnini, G.: The fundamental solution of the space-time fractional diffusion equation. Fract. Calc. Appl. Anal. **4**(2), 153–192 (2001)
48. Méndez, V., Fedotov, S., Horsthemke, W.: Reaction-Transport Systems. Springer Series in Synergetics. Springer, Heidelberg (2010)
49. Naumkin, P.I., Shishmarëv, I.A.: Nonlinear Nonlocal Equations in the Theory of Waves. American Mathematical Society, Providence (1994)
50. Palatucci, G., Savin, O., Valdinoci, E.: Local and global minimizers for a variational energy involving a fractional norm. Ann. Mat. Pura Appl. (4) **192**(4), 673–718 (2013)
51. Perthame, B., Ryzhik, L.: Moderate dispersion in scalar conservation laws. Commun. Math. Sci. **5**, 473–484 (2007)
52. Rohde, C.: Scalar conservation laws with mixed local and nonlocal diffusion-dispersion terms. SIAM J. Math. Anal. **37**(1), 103–129 (2005)
53. Sato, K.: Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics, vol. 68. Cambridge University Press, Cambridge (1999)
54. Uchaikin, V.V.: Fractional derivatives for physicists and engineers. Nonlinear Physical Science, vols. I & II. Higher Education Press, Beijing; Springer, Heidelberg (2013)
55. Viertl, N.: Viscous regularisation of weak laminar hydraulic jumps and bore in two layer shallow water flow. Ph.d. thesis, Technische Universität Wien (2005)
56. Volpert, A.I., Volpert, V.A., Volpert, V.A.: Traveling Wave Solutions of Parabolic Systems. American Mathematical Society, Providence (1994)
57. Volpert, V.A., Nec, Y., Nepomnyashchy, A.A.: Fronts in anomalous diffusion-reaction systems. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **371**(1982), 20120179, 18 (2013)
58. Yagisita, H.: Existence and nonexistence of traveling waves for a nonlocal monostable equation. Publ. Res. Inst. Math. Sci. **45**(4), 925–953 (2009)

On a Vlasov–Poisson Plasma with Infinite Charge and Velocities

Silvia Caprino

Abstract In this article I will overview some recent results on the Vlasov and the Vlasov–Poisson equations, for a given initial particle distribution which is not L^1 in space and has infinite support in the velocities.

Keywords Vlasov–Poisson equation · Magnetically confined plasma

1 Introduction

We consider a one-species, positively charged plasma whose time evolution is ruled by the following Vlasov equations

$$\begin{cases} \partial_t f(x,v,t) + v \cdot \nabla_x f(x,v,t) + (E(x,t) + F(x,v)) \cdot \nabla_v f(x,v,t) = 0 \\ E(x,t) = \displaystyle\int \frac{x-y}{|x-y|^3}\, \rho(y,t)\, dy \\ \rho(x,t) = \displaystyle\int f(x,v,t)\, dv \\ f(x,v,0) = f_0(x,v) \end{cases} \tag{1}$$

where $f(x,v,t)$ represents the plasma density, that is the distribution of charged particles at the point of the phase space $(x,v) \in \mathbb{R}^3 \times \mathbb{R}^3$ at time t, $\rho(x,t)$ is the charge in x at time t, E is the self-induced electric field and F is an external force acting on the system. Notice that for $F = 0$ we get the usual Vlasov–Poisson equation.

S. Caprino (✉)
Dipartimento di Matematica, Università Tor Vergata, Via della Ricerca Scientifica, Roma, Italy
e-mail: caprino@mat.uniroma2.it

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_3

Equation (1) is a conservation equation for the density f along the characteristics of the system, that is the solutions to the following problem:

$$\begin{cases} \dot{X}(t) = V(t) \\ \dot{V}(t) = E(X(t), t) + F((X(t), V(t)) \\ (X(0), V(0)) = (x, v), \end{cases} \tag{2}$$

where $(X(t), V(t)) = (X(x, v, t), V(x, v, t))$ denote position and velocity of a particle starting at time $t = 0$ from (x, v) and the particles are initially distributed by $f_0(x, v)$.

Since f is time-invariant along this motion, it is:

$$\|f(t)\|_{L^\infty} = \|f_0\|_{L^\infty} = C. \tag{3}$$

As it is well known, solutions of (2) produce weak solutions of Eq. (1), which become classical solutions whenever $f(x, v, 0)$ is assumed smooth.

The above equations have been widely studied in the case $F = 0$ and $f_0 \in L^1$, and the problem of the existence and uniqueness of its solutions is completely solved, as it can be seen in many papers (we quote [7–9] and a nice review of many such results in [6] for all). The subsequent problem of a spatial density not belonging to L^1 has been investigated since many years by many authors, but here we only quote papers [1–3], which will be discussed in our exposition. All the results in these papers have been obtained in the hypothesis that the density has compact support in the velocities, which actually seems somewhat unphysical. For this reason, in very recent times the same authors have faced also the problem of a plasma whose particles may have infinite velocities, even if "with low probability", that is having the density a sufficiently strong decay in the velocities. This has been done in papers [4, 5], and our aim here is to briefly describe the problems and the results in both situations, of infinite charge and of infinite velocities. This will be done in subsequent Sects. 2 and 3 in which we will illustrate the results in papers [2–5] respectively. In the spirit of a review, we will only sketch the proofs of the results, which can be found in extensive form in the papers.

2 The Case with Infinite Charge

In [2] and in [3] we have considered an infinitely extended plasma, confined in a cylinder by an external magnetic field ($F \neq 0$) and in the whole space respectively ($F = 0$). In both cases the density is assumed to have compact support in the velocities, while the spatial density is not supposed to be in L^1. In this case it is of primary relevance to achieve a good position of the equations, since the electric field could be infinite even at time $t = 0$. To avoid this problem, we have assumed that the spatial density, even if not integrable, is slightly decaying at infinity. Another problem coming from the infinite charge of the plasma is that the total energy of the system,

usually employed to get estimates on the spatial density, is infinite. We bypass this difficulty by introducing the *local energy*, which is defined as follows.

For $\mu \in \mathbb{R}$ and $R > 0$ we introduce the function,

$$\varphi^{\mu,R}(x) = \varphi\left(\frac{|x-\mu|}{R}\right) \tag{4}$$

with φ a smooth function such that:

$$\varphi(r) = 1 \text{ if } r \in [0,1] \tag{5}$$

$$\varphi(r) = 0 \text{ if } r \in [2,+\infty) \tag{6}$$

$$-2 \le \varphi'(r) \le 0. \tag{7}$$

Then the local energy is defined as

$$\begin{aligned} W(\mu,R,t) = &\frac{1}{2}\int dx\ \varphi^{\mu,R}(x)\int dv\ |v|^2 f(x,v,t) \\ &+\frac{1}{2}\int dx\ \varphi^{\mu,R}(x)\rho(x,t)\int dy\frac{\rho(y,t)}{|x-y|}. \end{aligned} \tag{8}$$

Notice that the function W can be seen as a sort of energy of a bounded region, which however takes into account the whole interaction with the rest of the particles. We will make assumptions on f_0 such that it will be finite at time $t = 0$, and this will enable us to prove, in both papers, the existence and uniqueness of the solution globally in time, for initial data which are not L^1 in space.

2.1 A Magnetically Confined Plasma

In [2] the equation we consider is (2), in which we have chosen the external force F to be a magnetic Lorentz force, confining the plasma in an infinite cylinder, that is, if $x = (x_1, x_2, x_3)$ then

$$F(x,v) = v \wedge B(x) \tag{9}$$

with

$$B = \left(h(x_2^2 + x_3^2), 0, 0\right). \tag{10}$$

The function h is defined in the cylinder

$$D = \left\{x \in \mathbb{R}^3 : -\infty < x_1 < +\infty,\ 0 \le \sqrt{x_2^2 + x_3^2} < L\right\} \tag{11}$$

and it is assumed to be non negative, smooth in D and diverging together with its primitive as $\sqrt{x_2^2 + x_3^2} \to L$.

We consider a cylinder $D_0 \subset D$

$$D_0 = \left\{ x \in \mathbb{R}^3 : -\infty < x_1 < +\infty,\ 0 \le \sqrt{x_2^2 + x_3^2} \le L_0 \right\} \tag{12}$$

for some $L_0 < L$ and define the set

$$S_0 = \left\{ (x, v) : x \in D_0,\ |v| \le \mathscr{V}_0 \right\} \tag{13}$$

with $\mathscr{V}_0$ a positive constant. Moreover, for any $\alpha > 0$ we define the following family of states:

$$\mathscr{F}^\alpha = \left\{ \rho : D_0 \to \mathbb{R}^+ \text{ such that } M_i \le C|i|^{-\alpha} \right\} \tag{14}$$

with C a positive constant, $i \in \mathbb{Z}/\{0\}$ and

$$M_i = \int_{i \le x_1 \le i+1} \rho(x)\, dx. \tag{15}$$

In the following we set

$$\rho_0(x) = \int f_0(x, v)\, dv.$$

Notice that, if we restrict the domain of $\rho_0(x)$ in D_0, then the magnetic field at time $t = 0$ is bounded.

Now we have all the ingredients to state the main result in [2].

Theorem 1 *Let us fix an arbitrary positive time T. Let $f_0 \in L^\infty$ be supported on the set S_0 defined in (13) and $\rho_0 \in \mathscr{F}^\alpha$ with $\alpha > 0$ arbitrary. Then there exists a solution to system (2) in $[0, T]$ and continuous functions $\mathscr{V}(t)$ and $L(t)$ from $[0, T] \to \mathbb{R}^+$, satisfying $\mathscr{V}(0) = \mathscr{V}_0$, $L(0) = L_0$ and $0 \le \sup_{t \in [0,T]} L(t) < L$ such that, for all times $t \in [0, T]$, $f(t)$ is supported on the set:*

$$S_t = \{(x, v) : x \in D_t,\ |v| \le \mathscr{V}(t)\} \tag{16}$$

with

$$D_t = \left\{ x \in \mathbb{R}^3 : -\infty < x_1 < +\infty,\ 0 \le \sqrt{x_2^2 + x_3^2} \le L(t) \right\} \tag{17}$$

and moreover $\rho(t) \in \mathscr{F}_t^\alpha$ where

$$\mathscr{F}_t^\alpha = \left\{ \rho : D_t \to \mathbb{R}^+ \text{ such that } M_i \le C|i|^{-\alpha} \right\}. \tag{18}$$

This solution is unique in the class of the characteristics distributed with $f(t) \in L^\infty$, *supported on* S_t *and belonging to* $\mathscr{F}_t^\alpha$ $\forall t \in [0, T]$.

Remark 1 The assumption for ρ_0 to belong to the class (14) makes the electric field E finite at time $t = 0$. It can be satisfied in case that the spatial density, even being not integrable, has a suitable decay at infinity, but also whenever $\rho_0(x)$ is constant, or has an oscillatory character, provided it has suitable support properties. Hence this hypothesis allows for spatial densities which possibly do not belong to any L^p space.

The theorem is proved by considering a cut-off system, obtained by putting in Eq. (1) the initial condition

$$f_0^N(x, v) = f_0(x, v)\chi_{D_0^N}(x), \tag{19}$$

being $\chi_{\mathscr{A}}$ the characteristic function of the set $\mathscr{A}$ and

$$D_0^N = D_0 \cap \{x \in \mathbb{R}^3 : |x_1| \leq N\}.$$

Let us set $\left(X^N(x, v, t), V^N(x, v, t)\right)$ to indicate the time evolved characteristics of system (2) with initial distribution f_0^N. Then existence and uniqueness of solutions to (2), together with property (18) are ensured, being the spatial support of f_0^N compact. Moreover one can see that, provided we have a bound on the velocities, then the confinement in the cylinder comes automatically. The following argument proves this fact. For simplicity we omit the index N, which would indicate that we are considering the time evolution of the cut-off system.

We put $l(t) = \sqrt{X_2^2(t) + X_3^2}$. Writing by components Eq. (2) with initial datum (44), after elementary calculation we get:

$$(V_2X_2 + V_3X_3)h(l^2) = \dot{V}_2X_3 - \dot{V}_3X_2 + X_2E_3 - X_3E_2.$$

Denoting by H a primitive of h and integrating in time we obtain,

$$\begin{aligned}
&\frac{1}{2}\int_0^t \frac{d}{ds}H(l^2(s))ds \\
&\quad = \frac{1}{2}\left[H(l^2(t)) - H(l^2(0))\right] \\
&\quad = \int_0^t ds\Big[\dot{V}_2(s)X_3(s) - \dot{V}_3(s)X_2(s) + X_2(s)E_3(s) - X_3(s)E_2(s)\Big].
\end{aligned} \tag{20}$$

$H(l^2(0))$ is bounded by the hypotheses on f_0^N, while the assumptions on h imply that $H(l^2(t)) \to \infty$ as $l \to L$. Hence, the left hand side in Eq. (20) goes to infinity as $l \to L$. On the other hand, by integrating by parts the right hand side we have,

$$\int_0^t ds \left[\dot{V}_2(s)X_3(s) - \dot{V}_3(s)X_2(s) + X_2(s)E_3(s) - X_3(s)E_2(s)\right]$$
$$= [V_2(s)X_3(s) - V_3(s)X_2(s)]_0^t$$
$$+ \int_0^t [X_2(s)E_3(X(s), s) - X_3(s)E_2(X(s), s)]\, ds. \tag{21}$$

If one has a bound on the velocities over the time interval $[0, T]$, it can be proven that also the field E is bounded, so that by (21) the latter term in (20) is bounded. Hence, not to get an absurd, the plasma does remain confined in a cylinder properly contained in D, as stated before.

The previous argument shows that the main effort has to be done to find a bound on the particle velocities, which is independent of the cut-off N. To this aim an important role is played by the local energy. Putting W^N for the function defined in (8) relative to the evolution of the cut-off system, and setting

$$Q^N(R, t) = \max\left\{1, \sup_{\mu\in\mathbb{R}} W^N(\mu, R, t)\right\} \tag{22}$$

$$\mathscr{V}^N(t) = \sup_{s\in[0,t]} \sup_{(x,v)\in D\times\mathbb{R}^3} |V^N(x, v, s)| \tag{23}$$

and

$$R^N(t) = 1 + \int_0^t \mathscr{V}^N(s)ds \tag{24}$$

we can state the following result:

Proposition 1 *There exists a constant C independent of N such that*

$$Q^N(R^N(t), t) \le C\, Q^N(R^N(t), 0).$$

Since the assumptions on f_0^N ensure that the function Q^N is finite at time $t = 0$, the above proposition shows that it remains finite at later times, and enables us to get estimates on the spatial density $\rho(x, t)$ and consequently on the electric field $E(x, t)$. This is sufficient to prove the theorem.

2.2 The Vlasov–Poisson Plasma in $\mathbb{R}^3$

In this case we consider a plasma in the whole of $\mathbb{R}^3$ and its time evolution given by the Vlasov–Poisson equation, that is Eq. (1) in which $F = 0$. In analogy with the preceding case, we define W as in (8) and we set

$$Q(R,t) = \max\left\{1, \sup_{\mu\in\mathbb{R}} W(\mu, R, t)\right\}. \tag{25}$$

As before, for any $\alpha > 0$ we define the following family of states:

$$\mathscr{F}^{\alpha}(C) = \left\{\rho : \mathbb{R}^3 \to \mathbb{R}^+ \text{ such that } \ M_i \le C|i|^{-\alpha}\right\} \tag{26}$$

with C a positive constant, $i \in \mathbb{Z}^3/\{0\}$ and

$$M_i = \int_{|i-x|\le 1} \rho(x)\, dx. \tag{27}$$

Theorem 2 *Let us fix an arbitrary positive time T. Let $f_0 \in L^\infty$ be supported on the set*

$$S_0 = \left\{(x,v) : x \in \mathbb{R}^3, \quad |v| \le \mathscr{V}_0\right\} \tag{28}$$

with $\mathscr{V}_0$ a positive constant. Moreover assume that $\rho_0 \in \mathscr{F}^{1+\varepsilon}(C_0)$, with ε arbitrary and positive, and that for any $R > 1$

$$Q(R,0) \le C R^{\beta}, \qquad \beta < \frac{7}{5} \tag{29}$$

with C a positive constant. Then there exists a solution to system (2) in $[0,T]$. Moreover this solution is unique among those distributed according to $f(t) \in L^\infty$, which enjoys the following properties: it is supported at any time $t \in [0,T]$ on the set

$$S_t = \left\{(x,v) : x \in \mathbb{R}^3, \quad |v| \le \mathscr{V}(t)\right\},$$

$\rho(t) \in \mathscr{F}^{1+\varepsilon}(C(t))$ with ε arbitrary and positive, being $\mathscr{V}(t)$ and $C(t)$ positive continuous functions on $[0,T]$ such that $\mathscr{V}(0) = \mathscr{V}_0$ and $C(0) = C_0$ and finally, for any $R > 1$

$$Q(R,t) \le C R^{\beta} \qquad \beta < \frac{7}{5} \tag{30}$$

with C a positive constant.

3 The Case with Infinite Charge and Velocities

In this section we illustrate some results obtained removing, in the two preceding systems, the assumption for the density to be compactly supported in the velocities. More precisely, we consider a plasma having infinite charge and velocities, and we assume that its density is slowly decaying in space (not integrably) and Gaussian

in the velocities. As in the previous cases, these hypotheses allow to well pose the problem, and make the local energy bounded at time $t = 0$. Because of this new setup, we will look for bounds on the velocity of any single particle, for a fixed initial (x, v), over an arbitrary time interval $[0, T]$.

3.1 A Magnetically Confined Plasma

In [4] we have considered the same model as in [2]. In order to have the confinement of the plasma we have to assume that the magnetic field is sufficiently strongly diverging on the walls of the cylinder. More precisely, we define as before

$$F(x, v) = v \wedge B(x)$$

$$B = \left(h(x_2^2 + x_3^2), 0, 0\right). \tag{31}$$

but we ask for the function h to satisfy

$$h(r^2) = \frac{1}{(L^2 - r^2)^\theta} \qquad \text{with} \qquad \theta > 2 \tag{32}$$

where $r = \sqrt{x_2^2 + x_3^2}$. As before we consider a plasma initially confined in a cylinder $D_0 \subset D$, in order to have a finite magnetic field at time $t = 0$, and we introduce the family of states (14). We prove the following result:

Theorem 3 *Let us fix an arbitrary positive time T. Assume that the field B satisfies (31). Let $f_0(x, v)$ be supported on $D_0 \times \mathbb{R}^3$ and satisfy the two following hypotheses:*

$$0 \le f_0(x, v) \le C_0 e^{-\lambda |v|^2} \tag{33}$$

$$\rho_0(x) \in \mathscr{F}^\alpha \qquad \alpha > \frac{5}{9} \tag{34}$$

for some positive constants C_0 and λ.
Then there exists a solution to system (2) in $[0, T]$ such that, for any $(x, v) \in D_0 \times \mathbb{R}^3$, $\sup_{0 \le t \le T} \sqrt{X_2(t)^2 + X_3(t)^2} < L$. Moreover

$$0 \le f(x, v, t) \le C e^{-\bar{\lambda} |v|^2} \tag{35}$$

for some positive constants C and $\bar{\lambda}$ and, for any $i \in \mathbb{Z}/\{0\}$,

$$\int_{i \le x_1 \le i+1} \rho(x, t)\, dx \le C \frac{\log(1 + |i|)}{|i|^\alpha}. \tag{36}$$

This solution is unique in the class of those satisfying (35) and (36).

3.2 The Vlasov–Poisson Plasma in $\mathbb{R}^3$

In [5] we prove the following two theorems.

Theorem 4 *Let us fix an arbitrary positive time T. Let f_0 satisfy the following hypotheses:*

$$0 \leq f_0(x,v) \leq C_0 e^{-\lambda|v|^2} g(|x|) \tag{37}$$

where is g a bounded, continuous function satisfying, for any $i \in \mathbb{Z}^3/\{0\}$,

$$\int_{|i-x|\leq 1} g(|x|)\, dx \leq C \frac{1}{|i|^{2+\varepsilon}} \qquad \forall\, \varepsilon > 1/15 \tag{38}$$

being λ, C_0 and C positive constants. Then there exists a solution to Eq. (2) on $[0,T]$ and positive constants C and $\bar{\lambda}$ such that

$$0 \leq f(x,v,t) \leq C e^{-\bar{\lambda}|v|^2} \tag{39}$$

and for any $i \in \mathbb{Z}^3/\{0\}$

$$\int_{|i-x|\leq 1} \rho(x,t)\, dx \leq C \frac{\log^{\frac{3}{2}}(1+|i|)}{|i|^{2+\varepsilon}}. \tag{40}$$

This solution is unique in the class of those satisfying (39) and (40).

Remark 2 Making the more strict assumption on the spatial density to be point-wise decreasing for large $|x|$ in place of hypothesis (38), allows us to improve the thesis of Theorem 4. Indeed, we are able to show that at any time $t \in [0,T]$ $\rho(t)$ keeps the same decreasing property, which is the object of the next theorem.

Theorem 5 *Let us fix an arbitrary positive time T. Let f_0 satisfy the following hypotheses:*

$$0 \leq f_0(x,v) \leq C_0 e^{-\lambda|v|^2} g(|x|) \tag{41}$$

where is g a bounded, continuous, not increasing function such that, for $|x| \geq 1$

$$g(|x|) \leq C \frac{1}{|x|^{2+\varepsilon}} \qquad \forall\, \varepsilon > 1/15 \tag{42}$$

being λ, C_0 and C positive constants. Then there exists a solution to Eq. (2) on $[0,T]$ and positive constants C and $\bar{\lambda}$ such that

$$0 \leq f(x,v,t) \leq C e^{-\bar{\lambda}|v|^2} g(|x|). \tag{43}$$

This solution is unique in the class of those satisfying (43).

3.3 *Proof of Theorem 5*

We want to briefly sketch the proof of this theorem, as a paradigm of the kind of proofs of Theorems 3 and 4.

We introduce a cut-off system, obtained by putting in Eq. (1) the initial condition

$$f_0^N(x,v) = f_0(x,v)\chi\{|v| \le N\}. \tag{44}$$

We consider Eq. (2) for the characteristics, with initial distribution f_0^N, and set $(X^N(t), V^N(t))$ for position and velocity in this dynamics of a point particle which is in (x,v) at time $t=0$. The solutions for such system are known to exist by the preceding results in Sect. 2, being the velocity support of the density compact. From now on all the functions relative to this dynamics will be indexed by N.

The main job is to get a fine bound on the electric field. Recalling the definition

$$\mathscr{V}^N(t) = \sup_{s\in[0,t]} \sup_{(x,v)\in\mathbb{R}^3\times\mathbb{R}^3} |V^N(x,v,s)|$$

by means of the local energy we can prove the following estimate:

Proposition 2 *There exists a positive constant C such that, for any $t \in [0,T]$:*

$$\int_0^t |E^N(x,s)|ds \le C\left[\mathscr{V}^N(t)\right]^\alpha \qquad \frac{5-\varepsilon}{9} < \alpha < \frac{2}{3}$$

being ε the one in (38).

The proof of this proposition is quite lengthy and involved, and is devoted to find the exponent α, which is suited for our purposes. A consequence of this result is the following

Corollary 1

$$\mathscr{V}^N(T) \le CN \tag{45}$$

$$\rho^N(x,t) \le CN^{3\alpha} \tag{46}$$

$$Q^N(R^N(t),t) \le CN^{1-\varepsilon}. \tag{47}$$

At this point we consider the time evolved of the point (x,v) in the phase space, in the N-th and in the $(N+1)$-th dynamics, that is we consider $\left(X^N(t), V^N(t)\right)$ and $\left(X^{N+1}(t), V^{N+1}(t)\right)$, the solutions to Eq. (2) with initial condition f_0^N and f_0^{N+1} respectively given by (44). We set

$$\delta^N(t) = \sup_{(x,v)\in\mathbb{R}^3\times B(0,N)} |X^N(t) - X^{N+1}(t)|$$

where $B(0, N) = \{v : |v| \le N\}$, and we prove the following result:

Proposition 3 *For any $t \in [0, T]$ it holds*

$$\delta^N(x, v, t) \le C \int_0^t dt_1 \int_o^{t_1} dt_2 \left[N^{3\alpha} \delta^N(t_2) \left(\left| \log \delta^N(t_2) \right| + 1 \right) + e^{-\frac{\lambda}{2} N^2} \right]. \quad (48)$$

We do not give its proof, but the above result enables us to prove what follows. Put

$$\eta^N(t) = \sup_{(x,v)\in\mathbb{R}^3\times B(0,N)} |V^N(t) - V^{N+1}(t)|$$

and

$$\sigma^N(t) = \max\{\delta^N(t), \eta^N(t)\}.$$

Hence we have

Proposition 4 *There exists a positive constant C such that*

$$\sup_{t\in[0,T]} \sigma^N(t) \le C^{-CN}. \quad (49)$$

This result is obviously conclusive for the proof of Theorem 5.

We prove the proposition for the quantity $\delta^N(t)$, being the proof for $\eta^N(t)$ similar and easier.

Notice that estimate (48) is not linear in $\delta^N(t)$, while we would like to iterate it in time. However, it is easily seen that if $r \in (0, 1)$, for any $a \in (0, 1)$ the following inequality holds by convexity:

$$r(|\log r| + 1) \le r|\log a| + a.$$

Hence, for $\delta^N(t) < 1$, Proposition 3 gives us, for any $a < 1$,

$$\delta^N(x, v, t) \le C \int_0^t dt_1 \int_0^{t_1} dt_2 \left[N^{3\alpha} \left(\delta^N(t_2) |\log a| + a \right) + e^{-\frac{\lambda}{2} N^2} \right]. \quad (50)$$

On the other side, in case $\delta^N(t) \ge 1$, it easily seen that estimate (48) would be linear, that is

$$\delta^N(t) \le C \int_0^t dt_1 \int_0^{t_1} dt_2 \left[N^{3\alpha} \delta^N(t_2) + e^{-\frac{\lambda}{2} N^2} \right]. \quad (51)$$

Hence, in case $\delta^N(t) < 1$, choosing $a = e^{-\frac{\lambda}{2}N^2}$ in (50), we get

$$\delta^N(t) \leq C \int_0^t dt_1 \int_0^{t_1} dt_2 \left[N^{3\alpha+2} \delta^N(t_2) + e^{-\frac{\lambda}{4}N^2} \right]. \tag{52}$$

Now we are in condition to iterate inequality (48), by inserting in the integral the same estimate for $\delta^N(t_2)$. We make k iterations up to time t_k and at the last step we use the estimate $\sup_{t\in[0,T]} \delta^N(t) \leq CN$, which comes from (45). Notice that at any step we get a double factorial at the denominator, due to the double time integration, obtaining

$$\delta^N(t) \leq CNe^{-\frac{\lambda}{4}N^2} \sum_{i=1}^{k-1} \frac{C^i N^{(3\alpha+2)i} T^{2i}}{(2i)!} + \frac{C^k N^{(3\alpha+2)k} T^{2k}}{(2k)!} \tag{53}$$

where C stands as ever for a positive constant. The latter term is exponentially vanishing as $N \to \infty$, provided $k > N^{(3\alpha+2)}$. For the former term we have

$$\sum_{i=1}^{k-1} \frac{C^i N^{(3\alpha+2)i} T^{2i}}{(2i)!} \leq e^{\sqrt{CT} N^{\left(\frac{3}{2}\alpha+1\right)}}. \tag{54}$$

Hence

$$\delta^N(t) \leq CNe^{-\frac{\lambda}{4}N^2} e^{\sqrt{CT} N^{\left(\frac{3}{2}\alpha+1\right)}} + C^{-N}. \tag{55}$$

Since in Proposition 2 we have found the "right" exponent, that is $\frac{3}{2}\alpha + 1 < 2$, we have

$$\delta^N(t) \leq C^{-CN}. \tag{56}$$

References

1. Caprino, S., Cavallaro, G., Marchioro, C.: On a magnetically confined plasma with infinite charge. SIAM J. Math. Anal. **46**, 133–164 (2014)
2. Caprino, S., Cavallaro, G., Marchioro, C.: Remark on a magnetically confined plasma with infinite charge. Rend. Mat. Appl. **35**, 69–98 (2014)
3. Caprino, S., Cavallaro, G., Marchioro, C.: Existence and uniqueness of the time evolution for a plasma with infinite charge in $\mathbb{R}^3$. Commun. Partial Differ. Equ. **40**, 1–29 (2015)
4. Caprino S., Cavallaro G., Marchioro C.: A Vlasov–Poisson plasma with unbounded mass and velocities confined in a cylinder by a magnetic mirror (To appear on Kin. Rel. Mod)
5. Caprino S., Cavallaro G., Marchioro C.: The Vlasov–Poisson equation in $\mathbb{R}^3$ with infinite charge and velocities (Submitted to Commun. Partial Differ. Equ.)
6. Glassey, R.: The Cauchy problem in kinetic theory. SIAM, Philadelphia (1996)
7. Lions, P.L., Perthame, B.: Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. Invent. Math. **105**, 415–430 (1991)
8. Pfaffelmoser, K.: Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. J. Differ. Equ. **95**, 281–303 (1992)
9. Schaeffer, J.: Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. Commun. Partial Differ. Equ. **16**(8–9), 1313–1335 (1991)

A Note on Local Well-Posedness of Generalized KdV Type Equations with Dissipative Perturbations

Xavier Carvajal and Mahendra Panthee

Abstract In this note we report some local well-posedness results for the Cauchy problems associated to generalized KdV (gKdV) type equations with dissipative perturbation for given data in the low regularity L^2-based Sobolev spaces. The method of proof is based on the *contraction mapping principle* employed in some appropriate time weighted spaces.

Keywords Initial value problem · Well-posedness
Dispersive-dissipative models · Kdv equation

1 Introduction

In this brief note we are interested in reporting well-posedness results for the Cauchy problems associated to generalized Korteweg-de Vries (gKdV) type equations with dissipative perturbations, viz.,

$$\begin{aligned} &v_t + v_{xxx} + \eta Lv + (v^{k+1})_x = 0, \quad x \in \mathbb{R},\ t \geq 0, \quad k \in \mathbb{N}, \\ &v(x,0) = v_0(x), \end{aligned} \tag{1}$$

X. Carvajal
Instituto de Matemática - UFRJ Av. Horácio Macedo,
Centro de Tecnologia Cidade Universitária, Ilha Do Fundão,
Rio de Janeiro, RJ 21941-972, Brazil
e-mail: carvajal@im.ufrj.br

M. Panthee (✉)
Departamento de Matemática, UNICAMP, Rua Sergio Buarque de Holanda 651,
Campinas, Sp 13083-859, Brazil
e-mail: mpanthee@ime.unicamp.br

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_4

and

$$\begin{aligned} &u_t + u_{xxx} + \eta Lu + (u_x)^{k+1} = 0, \quad x \in \mathbb{R},\ t \geq 0, \quad k \in \mathbb{N}, \\ &u(x,0) = u_0(x), \end{aligned} \tag{2}$$

where $\eta > 0$ is a constant; $u = u(x,t)$, $v = v(x,t)$ are real valued functions and the linear operator L is defined via the Fourier transform by $\widehat{Lf}(\xi) = -\Phi(\xi)\widehat{f}(\xi)$. For $p \in \mathbb{R}^+$, the Fourier symbol $\Phi(\xi) \in \mathbb{R}$ is measurable and has the form

$$\Phi(\xi) = -|\xi|^p + \Phi_1(\xi), \tag{3}$$

where $|\Phi_1(\xi)| \leq C(1 + |\xi|^q)$ with $0 \leq q < p$.

The models under consideration are posed for $t \geq 0$. In order to deal with some estimates involving semi-group, we extend these models for all $t \in \mathbb{R}$. For this, we define

$$\eta(t) \equiv \eta\, \mathrm{sgn}(t) = \begin{cases} \eta & \text{if } t \geq 0, \\ -\eta & \text{if } t < 0. \end{cases} \tag{4}$$

and write (1) and (2) respectively in the form

$$\begin{aligned} &v_t + v_{xxx} + \eta(t) Lv + (v^{k+1})_x = 0, \quad x,\ t \in \mathbb{R}, \\ &v(0) = v_0, \end{aligned} \tag{5}$$

and

$$\begin{aligned} &u_t + u_{xxx} + \eta(t) Lu + (u_x)^{k+1} = 0, \quad x,\ t \in \mathbb{R}, \\ &u(x,0) = u_0(x). \end{aligned} \tag{6}$$

This sort of extension to consider the models posed for all $t \in \mathbb{R}$ has already been appeared in earlier works [2, 3]. In what follows we will study the well-posedness issues for the Caucy problems (5) and (6) for given data in the low regularity Sobolev spaces $H^s(\mathbb{R})$. We start by introducing two spaces that are suitable to handle the models in question.

Let $k \in \mathbb{N}$, $|t| \leq T \leq 1$, $p > \frac{3}{2}k + 1$ and $\gamma_k := \frac{3k+2}{2(k+1)}$. To deal with the Cauchy problem (5), we define, for $s \geq 0$,

$$\begin{aligned} \|f\|_{\mathfrak{X}_T^s} := \sup_{t\in[-T,T]} \Big\{ &\|f(t)\|_{H^s(\mathbb{R})} \\ &+ |t|^{\frac{\gamma_k}{p}} \Big(\|f(t)\|_{L^{2(k+1)}} + \|\partial_x f(t)\|_{L^{2(k+1)}} + \|D_x^s \partial_x f(t)\|_{L^{2(k+1)}} \Big) \Big\}, \end{aligned} \tag{7}$$

for $-1 < s < 0$,

$$\|f\|_{\mathfrak{X}_T^s} := \sup_{t\in[-T,T]} \Big\{ \|f(t)\|_{H^s(\mathbb{R})} + |t|^{\frac{\gamma_k}{p}} \Big(\|f(t)\|_{L^{2(k+1)}} + \|D_x^s \partial_x f(t)\|_{L^{2(k+1)}} \Big) \Big\}, \tag{8}$$

and introduce the Banach space

$$\mathcal{X}_T^s := \big\{ f \in C([-T,T]; H^s(\mathbb{R})) : \|f\|_{\mathcal{X}_T^s} < \infty \big\}, \tag{9}$$

where $H^s(\mathbb{R})$ is the usual L^2-based Sobolev space of order s.

Similarly, to work on the Cauchy problem (6), we define, for $s \in \mathbb{R}$,

$$\|f\|_{\mathcal{Y}_T^s} := \sup_{t\in[-T,T]} \left\{ \|f(t)\|_{H^s} + |t|^{\frac{\gamma_k}{p}} \left(\|\partial_x f(t)\|_{L^{2(k+1)}} + \|D_x^s \partial_x f(t)\|_{L^{2(k+1)}} \right) \right\}, \tag{10}$$

and introduce, accordingly, the Banach space

$$\mathcal{Y}_T^s := \big\{ f \in C([-T,T]; H^s(\mathbb{R})) : \|f\|_{\mathcal{Y}_T^s} < \infty \big\}. \tag{11}$$

Motivation behind the definition of these sort of spaces is the work of Dix [5] where the author uses space with time weight to prove sharp well-posedness result for the Cauchy problem associated to the Burgers' equation in $H^s(\mathbb{R})$, $s > -\frac{1}{2}$ by showing that uniqueness fails for data Sobolev regularity less than $-\frac{1}{2}$. In our recent work [4], similar spaces are used to get sharp local well-posedness results for KdV type equations. We notice that the space used in [4] needs restriction of the Sobolev regularity of the initial data not only from below but also from above. To overcome the restriction of Sobolev regularity from above, we used extra norms in the definitions of $\mathcal{X}_T^s$ and $\mathcal{Y}_T^s$ above. Although we could remove the regularity restriction from above with the introduction of extra norms, we could not lower the regularity requirement of the initial data as much as we desired. This question will be addressed elsewhere.

Now, we state the main results of this work.

Theorem 1 *Let $\eta > 0$ be fixed, $k > 0$ and $\Phi(\xi)$ be as given by (3) with $p > \frac{3}{2}k + 1$ as the order of the leading term. Then the Cauchy problem (5) is locally well-posed for any data $v_0 \in H^s(\mathbb{R})$, whenever $s \geq -1$. Moreover, the map $v_0 \mapsto v$ is smooth from $H^s(\mathbb{R})$ to $C([-T,T]; H^s(\mathbb{R})) \cap \mathcal{X}_T^s$ and $v \in C([-T,T]\backslash\{0\}; H^\infty(\mathbb{R}))$.*

Theorem 2 *Let $\eta > 0$ be fixed, $k > 0$ and $\Phi(\xi)$ be as given by (3) with $p > \frac{3}{2}k + 1$ as the order of the leading term. Then the Cauchy problem (6) is locally well-posed for any data $u_0 \in H^s(\mathbb{R})$, whenever $s \geq 0$. Moreover, the map $u_0 \mapsto u$ is smooth from $H^s(\mathbb{R})$ to $C([-T,T]; H^s(\mathbb{R})) \cap \mathcal{Y}_T^s$ and $u \in C([-T,T]\backslash\{0\}; H^\infty(\mathbb{R}))$.*

Remark 1 The results in Theorems 1 and 2 improve the known well-posedness results for the Cauchy problems (5) and (6) proved in [2, 3]. However, in view of our recent work [4], we believe that there is still room to get better results than the ones presented here. Due to the presence of higher order nonlinearity, the time weighted spaces introduced in [4] are not good enough in this case. So, one expects to introduce some new sort of spaces to cater higher order nonlinearity. Currently, we are working in this direction.

In earlier works [2] (see also [3]) we proved that the IVPs (5) and (6) (with $k = 1$) for given data in $H^s(\mathbb{R})$ are locally well-posed whenever $s > -\frac{3}{4}$ and $s > \frac{1}{4}$

respectively. The IVPs (5) and (6) with $k = 2, 3, 4$ were also considered in [3] to get local well-posedness results respectively for $s > \frac{1}{4}, -\frac{1}{6}, 0$ and $s > \frac{5}{4}, \frac{5}{6}, 1$. To obtain these results we followed the techniques developed by Bourgain [1] and Kenig, Ponce and Vega [7] (see also [8]). The main ingredients in the proof are estimates of the integral equation associated to an extended IVP that is defined for all $t \in \mathbb{R}$. The principal tool in [2, 3] is the usual Bourgain space associated to the KdV equation instead of that associated to the linear part of the IVPs (5) and (6).

A detailed explanation about the particular examples that belong to the class considered in (5) and (6) and the known well-posedness results about them are presented in our earlier works [2, 4].

This paper is organized as follows: In Sect. 2, we prove some linear and nonlinear estimates. Section 3 is dedicated to prove the main results.

2 Basic Estimates

In this section we derive some linear and nonlinear estimates that will be used to prove the main results of this work. Consider the Cauchy problem associated to the linear parts of (5) and (6),

$$\begin{aligned} &w_t + w_{xxx} + \eta(t) Lw = 0, \quad x,\ t \in \mathbb{R}, \\ &w(0) = w_0. \end{aligned} \tag{12}$$

The solution to (12) is given by $w(x,t) = V(t)w_0(x)$ where the semi-group $V(t)$ is defined as

$$\widehat{V(t)w_0}(\xi) = e^{it\xi^3 + \eta|t|\Phi(\xi)}\widehat{w_0}(\xi). \tag{13}$$

In what follows, we prove some estimates satisfied by the semi-group defined in (13). Without loss of generality, we suppose $\eta = 1$ from here onwards. We start with the following lemmas.

Lemma 1 (Hausdorff-Young) *Let $p_1 \geq 2$ and $\frac{1}{p_1} + \frac{1}{q_1} = 1$. Then*

$$\|f\|_{L^{p_1}} \leq C\|\widehat{f}\|_{L^{q_1}}. \tag{14}$$

Lemma 2 *There exists a sufficiently large number $M > 0$ such that the following estimates*

$$\Phi(\xi) = -|\xi|^p + \Phi_1(\xi) < -1, \tag{15}$$

$$\frac{|\Phi_1(\xi)|}{|\xi|^p} \leq \frac{1}{2}, \tag{16}$$

and

$$|\Phi(\xi)| \geq \frac{|\xi|^p}{2}, \tag{17}$$

hold true for all $|\xi| \geq M$.

Proof Note that,

$$\lim_{|\xi|\to\infty} \frac{\Phi_1(\xi)+1}{|\xi|^p} = 0 \quad \text{and} \quad \lim_{|\xi|\to\infty} \frac{|\Phi_1(\xi)|}{|\xi|^p} = 0. \tag{18}$$

The estimates (15) and (16) follow respectively from the first and the second relation in (18).

For $|\xi| \geq M$, using (15) and (16), one has that

$$|\Phi(\xi)| = |\xi|^p - \Phi_1(\xi) \geq \frac{|\xi|^p}{2}, \tag{19}$$

which proves (17). □

Lemma 3 *Let $|t| \leq T$ and $\Phi(\xi)$ be as defined in (3). Then $\Phi(\xi)$ is bounded from above and*

$$\|e^{|t|\Phi(\xi)}\|_{L^\infty} \leq e^{TC_M}. \tag{20}$$

Proof Using (15) in Lemma 2, we see that there is $M > 1$ large enough such that $\Phi(\xi) < -1$. Therefore,

$$e^{|t|\Phi(\xi)} \leq e^{-|t|} \leq 1, \qquad \forall \ |\xi| \geq M. \tag{21}$$

For $|\xi| < M$, it is easy to show that $\Phi(\xi) < C_M$, and consequently

$$e^{|t|\Phi(\xi)} \leq e^{TC_M} \quad \forall \ |\xi| < M. \tag{22}$$

Combining (21) and (22) we conclude the proof. □

Lemma 4 *Let $v_0 \in H^s(\mathbb{R})$ and $V(t)$ be defined as in (13), then*

$$V(\cdot)v_0 \in C(\mathbb{R}; H^s(\mathbb{R})) \cap C(\mathbb{R}\backslash\{0\}; H^\infty(\mathbb{R})).$$

Proof It is enough to show that $V(t)v_0 \in H^{s'}(\mathbb{R})$ whenever $s' > s$. We have,

$$\begin{aligned} \|V(t)v_0\|_{H^{s'}} &= \|\langle\xi\rangle^{s'} e^{it\xi^3+|t|\Phi(\xi)}\widehat{v_0}(\xi)\|_{L^2} \\ &= \|\langle\xi\rangle^{s}\widehat{v_0}(\xi)\langle\xi\rangle^{s'-s} e^{|t|\Phi(\xi)}\|_{L^2} \\ &\leq \|\langle\xi\rangle^{s'-s} e^{|t|\Phi(\xi)}\|_{L^\infty}\|v_0\|_{H^s}, \end{aligned} \tag{23}$$

where we have used the notation $\langle x\rangle := 1 + |x|$.

Let $M \gg 1$ be as in Lemma 2, then one gets

$$\begin{aligned}\|\langle\xi\rangle^{s'-s}e^{|t|\Phi(\xi)}\|_{L^\infty} &\le \|\langle\xi\rangle^{s'-s}e^{|t|\Phi(\xi)}\|_{L^\infty(|\xi|\le M)} + \|\langle\xi\rangle^{s'-s}e^{|t|\Phi(\xi)}\|_{L^\infty(|\xi|>M)}\\ &\le C_M + \|e^{-\frac{|\xi|^p}{2}|t|}\langle\xi\rangle^{s'-s}\|_{L^\infty} < \infty, \qquad t\in\mathbb{R}\backslash\{0\}. \end{aligned} \tag{24}$$

One can use the dominated convergence theorem to prove continuity. □

Lemma 5 *Let $s > -1 - \frac{k}{2(k+1)}$, $|t| \le T \le 1$ and $p > 0$, then we have*

$$|t|^{\frac{\gamma_k}{p}}\|\partial_x D_x^s V(t)w_0\|_{L^{2(k+1)}} \le C\|w_0\|_{H^s} \tag{25}$$

and

$$|t|^{\frac{\gamma_k}{p}}\|V(t)w_0\|_{L^{2(k+1)}} \le C\|w_0\|_{H^{-1}}. \tag{26}$$

Proof By Hausdorff-Young and Hölder inequalities, we obtain

$$\begin{aligned}\|\partial_x D_x^s V(t)w_0\|_{L^{2(k+1)}} &\le C\|e^{|t|\Phi(\xi)}\xi|\xi|^s\widehat{w_0}\|_{L^{\frac{2(k+1)}{2k+1}}}\\ &\le C\left\|e^{|t|\Phi(\xi)}\frac{\xi|\xi|^s}{\langle\xi\rangle^s}\right\|_{L^{\frac{2(k+1)}{k}}}\|\langle\xi\rangle^s\widehat{w_0}\|_{L^2}.\end{aligned} \tag{27}$$

Let M as in Lemma 2, then we have

$$\begin{aligned}\left\|\frac{\xi|\xi|^s}{\langle\xi\rangle^s}e^{|t|\Phi(\xi)}\right\|_{L^{\frac{2(k+1)}{k}}} &\le \left\|\frac{\xi|\xi|^s}{\langle\xi\rangle^s}e^{|t|\Phi(\xi)}\chi_{\{|\xi|\le M\}}\right\|_{L^{\frac{2(k+1)}{k}}}\\ &\quad+\left\|\frac{\xi|\xi|^s}{\langle\xi\rangle^s}e^{|t|\Phi(\xi)}\chi_{\{|\xi|>M\}}\right\|_{L^{\frac{2(k+1)}{k}}}\\ &\le C_M + \left(\int_{\mathbb{R}}|\xi|^{\frac{2(k+1)}{k}}e^{-|t|(\frac{k+1}{k})|\xi|^p}d\xi\right)^{\frac{k}{2(k+1)}}.\end{aligned} \tag{28}$$

Using a change of variable $|t|^{-1/p}\xi := x$, we get

$$\left\|\frac{\xi|\xi|^s}{\langle\xi\rangle^s}e^{|t|\Phi(\xi)}\right\|_{L^{\frac{2(k+1)}{k}}} \le C_M + \frac{1}{|t|^{\frac{\gamma_k}{p}}}\left(\int_{\mathbb{R}}|x|^{\frac{2(k+1)}{k}}e^{-(\frac{k+1}{k})|x|^p}dx\right)^{\frac{k}{2(k+1)}}. \tag{29}$$

Since the integral in (29) is finite, inserting (29) in (27) and multiplying by $|t|^{\frac{\gamma_k}{p}}$ we get the required estimate (25).

With a similar argument as above (considering $s = 0$ in (25), one can obtain

$$|t|^{\frac{\gamma_k}{p}}\|D_x V(t)w_0\|_{L^{2(k+1)}} \le C\|w_0\|_{L^2}. \tag{30}$$

Analogously, we can prove

$$|t|^{\frac{\gamma_k}{p}}\|V(t)w_0\|_{L^{2(k+1)}} \le C\|w_0\|_{L^2}. \tag{31}$$

Now, using (30) and (31), we obtain

$$|t|^{\frac{\gamma_k}{p}} \|(1+D_x)V(t)w_0\| \leq C\|w_0\|_{L^2}. \tag{32}$$

This final estimate (32) is equivalent to (26) and this completes the proof. □

Corollary 1 *Let $s > -1 - \frac{k}{2(k+1)}$, $|\tau| \leq T$, $|t| \leq T \leq 1$ and $p > 0$, then we have*

$$|t-\tau|^{\frac{\gamma_k}{p}} \|\partial_x D_x^s V(t-\tau) f(t,\cdot)\|_{L^{2(k+1)}} \leq C\|f(t,\cdot)\|_{H^s} \tag{33}$$

and

$$|t-\tau|^{\frac{\gamma_k}{p}} \|V(t-\tau) f(t,\cdot)\|_{L^{2(k+1)}} \leq C\|f(t,\cdot)\|_{H^{-1}}. \tag{34}$$

Lemma 6 *Let $V(t)$ be as defined in (13), $|t| \leq T \leq 1$ and $p > \frac{3}{2}k+1$ then for all $s \geq -1$, we have*

$$\|V(t)w_0\|_{\mathfrak{X}_T^s} \leq C\|w_0\|_{H^s} \tag{35}$$

and for all $s \geq 0$, we have

$$\|V(t)w_0\|_{\mathcal{Y}_T^s} \leq C\|w_0\|_{H^s}. \tag{36}$$

Proof We provide a detailed proof for (35), the proof of (36) follows similarly. We start with the first component of $\mathfrak{X}_T^s$-norm. First, note that

$$\|V(t)w_0\|_{H^s} = \|\langle\xi\rangle^s e^{|t|\Phi(\xi)}\widehat{w_0}(\xi)\|_{L^2} \leq \|e^{|t|\Phi(\xi)}\|_{L^\infty}\|w_0\|_{H^s}. \tag{37}$$

From (20) and (37), we can conclude

$$\|V(t)w_0\|_{H^s} \leq e^{TC_M}\|w_0\|_{H^s}. \tag{38}$$

For the next components of the $\mathfrak{X}_T^s$-norm, we use (25) and (26) from Lemma 5 to get, for $s \geq -1$

$$|t|^{\frac{\gamma_k}{p}} \|D_x^s \partial_x V(t)w_0\|_{L^{2(k+1)}} + |t|^{\frac{\gamma_k}{p}} \|V(t)w_0\|_{L^{2(k+1)}} \leq C\|w_0\|_{H^s}. \tag{39}$$

Combining (38) and (39) we get the required estimate (35). □

Lemma 7 *Let $V(t)$ be as defined in (13), $|t| \leq T \leq 1$ and $p > \frac{3}{2}k+1$, then the following estimates hold true.*

$$\left\| \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau \right\|_{\mathfrak{X}_T^s} \leq T^{\omega_k}\|v\|_{\mathfrak{X}_T^s}^{k+1}, \quad \forall\, s \geq -1, \tag{40}$$

and

$$\left\| \int_0^t V(t-\tau)(\partial_x u)^{k+1}(\tau)d\tau \right\|_{\mathcal{Y}_T^s} \leq T^{\omega_k} \|u\|_{\mathcal{Y}_T^s}^{k+1} \quad \forall\, s \geq 0, \tag{41}$$

where $\omega_k = \frac{2p-3k-2}{2p} > 0$.

Proof We start by proving (40). First consider the case when $s \geq 0$. Using the definition of $\mathcal{X}_T^s$–norm and Lemma 5 for $s \geq 0$ (see (7)), we get, for the first component

$$\sup_{t\in[-T,T]} \left\| \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau \right\|_{H_T^s} \leq \sup_{t\in[-T,T]} \int_0^{|t|} \|\partial_x(v^{k+1})(\tau)\|_{H^s} d\tau. \tag{42}$$

To estimate the other components of $\mathcal{X}_T^s$–norm, we use Minkowski's integral inequality and Corollary 1. For the sake of brevity and clarity, we give details for the fourth component only, because others follow similarly.

$$\begin{aligned} |t|^{\frac{\gamma_k}{p}} \left\| D_x^s \partial_x \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau \right\|_{L^{2(k+1)}} & \\ = |t|^{\frac{\gamma_k}{p}} \left\| \int_0^t D_x^s \partial_x V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau \right\|_{L^{2(k+1)}} & \\ \leq C\,|t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \left\| D_x^s \partial_x V(t-\tau)\partial_x(v^{k+1})(\tau) \right\|_{L^{2(k+1)}} d\tau & \\ \leq C\,|t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|\partial_x(v^{k+1})(\tau)\|_{H^s} d\tau. & \end{aligned} \tag{43}$$

Combining (42), (43) and similar estimates for the other components of the $\mathcal{X}_T^s$–norm, we obtain

$$\begin{aligned} &\left\| \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau \right\|_{\mathcal{X}_T^s} \\ &\quad \leq \sup_{t\in[-T,T]} \int_0^{|t|} \left(\|\partial_x(v^{k+1})\|_{L^2} + \|D_x^s \partial_x(v^{k+1})\|_{L^2} \right) d\tau \\ &\qquad + C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \left(\|\partial_x(v^{k+1})\|_{L^2} + \|D_x^s \partial_x(v^{k+1})\|_{L^2} \right) d\tau \\ &\quad \leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \left(\|\partial_x(v^{k+1})\|_{L^2} + \|D_x^{\tilde{s}}(v^{k+1})\|_{L^2} d\tau \right) \\ &\quad =: I_1 + I_2, \end{aligned} \tag{44}$$

where we used $\tilde{s} = 1 + s$ and the estimate

$$1 = \frac{|t-\tau|}{|t-\tau|} \leq \frac{|t|+|\tau|}{|t-\tau|} \leq \frac{2|t|}{|t-\tau|}, \quad 0 \leq \tau < |t|.$$

In what follows, we will obtain an estimate for I_2. Now, using fractional chain rule (see Tao [9] (A.15) p. 338), for $\tilde{s} \geq 0$, we have

$$\begin{aligned} I_2 &\leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|v^k\|_{L^{\frac{2(k+1)}{k}}} \|D_x^{\tilde{s}} v\|_{L^{2(k+1)}} d\tau \\ &\leq C \|v\|_{\mathcal{X}_T^s}^{k+1} \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \frac{1}{\tau^{\frac{k\gamma_k}{p}}} \frac{1}{\tau^{\frac{\gamma_k}{p}}} d\tau \\ &\leq C \|v\|_{\mathcal{X}_T^s}^{k+1} \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \frac{1}{\tau^{\frac{3k+2}{2p}}} d\tau \\ &\leq C_{p,k} \|v\|_{\mathcal{X}_T^s}^{k+1} T^{\omega_k}. \end{aligned} \tag{45}$$

The estimate for I_1 will follow from the one of I_2 by considering $s = 0$. In fact,

$$\begin{aligned} I_1 &\leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{t} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|v^k\|_{L^{\frac{2(k+1)}{k}}} \|\partial_x v\|_{L^{2(k+1)}} d\tau \\ &\leq C \|v\|_{\mathcal{X}_T^s}^{k+1} \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{t} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \frac{1}{\tau^{\frac{3k+2}{2p}}} d\tau \\ &\leq C_{p,k} \|v\|_{\mathcal{X}_T^s}^{k+1} T^{\omega_k}. \end{aligned} \tag{46}$$

Inserting estimates (45) and (46) in (44) we obtain the required estimate (40) in the case $s \geq 0$.

Next we consider the case $-1 \leq s < 0$. Using a similar argument as above, one gets

$$\begin{aligned} &\left\| \int_0^t V(t-\tau)\ \partial_x(v^{k+1})(\tau) d\tau \right\|_{\mathcal{X}_T^s} \\ &\qquad \leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|\partial_x(v^{k+1})\|_{H^s} d\tau \\ &\qquad \leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|D_x^s \partial_x(v^{k+1})\|_{L^2} d\tau \\ &\qquad \leq C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \|D_x^{\tilde{s}}(v^{k+1})\|_{L^2} d\tau \\ &\qquad \leq C_{p,k} \|v\|_{\mathcal{X}_T^s}^{k+1} T^{\omega_k}, \end{aligned} \tag{47}$$

where $\tilde{s} = 1 + s \geq 0$, for $s \geq -1$.

Now, we move to prove the estimate (41). By using (36) from Lemma 6 and fractional chain rule as in (45), for $s \geq 0$, we get

$$
\begin{aligned}
&\Big\| \int_0^t V(t-\tau)(\partial_x u)^{k+1}(\tau)d\tau \Big\|_{\mathcal{Y}_T^s} \\
&\quad \le C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \left(\|(u_x)^{k+1}\|_{L^2} + \|D_x^s (u_x)^{k+1}\|_{L^2} \right) d\tau \qquad (48)\\
&\quad \le C \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \left(\|u_x\|_{L^{2(k+1)}}^{k+1} + \|(u_x)^k\|_{L^{\frac{2(k+1)}{k}}} \|D_x^s (u_x)\|_{L^{2(k+1)}} \right) d\tau \\
&\quad \le C\|u\|_{\mathcal{Y}_T^s}^{k+1} \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \left[\frac{1}{\tau^{\frac{(k+1)\gamma_k}{p}}} + \frac{1}{\tau^{\frac{k\gamma_k}{p}}} \frac{1}{\tau^{\frac{\gamma_k}{p}}} \right] d\tau \\
&\quad \le C\|u\|_{\mathcal{Y}_T^s}^{k+1} \sup_{t\in[-T,T]} |t|^{\frac{\gamma_k}{p}} \int_0^{|t|} \frac{1}{|t-\tau|^{\frac{\gamma_k}{p}}} \frac{1}{\tau^{\frac{3k+2}{2p}}} d\tau \\
&\quad \le C_{p,k}\, \|u\|_{\mathcal{Y}_T^s}^{k+1} T^{\omega_k},
\end{aligned}
$$

as required. In the last inequality of (48) the condition $p > \frac{3}{2}k + 1$ has been used. □

Now, we move to prove some more estimates that are useful in our analysis.

Lemma 8 *Let $\theta \ge 0$, $p > 0$ and $\tau \in [-1, 1]\backslash\{0\}$. Then we have*

$$
\|\langle \xi \rangle^{\theta} e^{|\tau|\Phi(\xi)}\|_{L^\infty_\xi} \lesssim \frac{1}{|\tau|^{\frac{\theta}{p}}}. \qquad (49)
$$

Proof Considering M large as in Lemma 2, one can obtain

$$
\|\langle \xi \rangle^{\theta} e^{|\tau|\Phi(\xi)}\|_{L^\infty} \le \|\langle \xi \rangle^{\theta} e^{|\tau|\Phi(\xi)}\|_{L^\infty(|\xi|\le M)} + \|\langle \xi \rangle^{\theta} e^{|\tau|\Phi(\xi)}\|_{L^\infty(|\xi|> M)} =: A_1 + A_2. \qquad (50)
$$

For $\tau \in [-1, 1]\backslash\{0\}$ and $\frac{\theta}{p} \ge 0$, we have

$$
A_1 \le C_M e^{|\tau| C_M} \lesssim \frac{1}{|\tau|^{\frac{\theta}{p}}}. \qquad (51)
$$

In what follows, we obtain and estimate for the high frequency part A_2. From (17) we have

$$
A_2 \lesssim \frac{\| \, |\, |\tau|^{\frac{1}{p}} \xi|^{\theta} e^{-||\tau|^{\frac{1}{p}} \xi|^p} \|_{L^\infty(|\xi|>M)}}{|\tau|^{\frac{\theta}{p}}} \lesssim \frac{1}{|\tau|^{\frac{\theta}{p}}}, \qquad (52)
$$

where in the last inequality $x^{\frac{\theta}{p}} e^{-x} \le C_{\theta,p}$, if $x \ge 0$ has been used. The proof of the lemma follows inserting (51) and (52) in (50). □

Proposition 1 *Let $s > -1$, $p > \frac{3k}{2} + 1$, $k \in \mathbb{N}$. There exists $\mu := \mu(s, p, k) > 0$ such that if*

$$
\|f\|_{\mathcal{Z}_T^s} := \sup_{t\in[-T,T]\backslash\{0\}} \left\{ \|f(t)\|_{H^s} + |t|^{\frac{\gamma_k}{p}} \|f(t)\|_{L^{2(k+1)}} \right\} < \infty, \qquad (53)
$$

then the application

$$t \mapsto \mathcal{L}(f)(t) := \int_0^t V(t-t')\partial_x(f^{k+1})(t')dt', \qquad 0 < |t| \leq T \leq 1, \tag{54}$$

is continuous from $[-T, T]\backslash\{0\}$ *to* $H^{s+\mu}$.

Proof We prove this proposition considering $0 < t \leq T \leq 1$. The case $-1 \leq -T \leq t < 0$ follows similarly. First we show that $\mathcal{L}(f)(t) \in H^{s+\mu}(\mathbb{R})\ \forall\ f \in Z_T^s$. We do this considering two different cases.

Case I, $-1 \leq s < p-1$**:** Let $0 < t \leq T \leq 1$ and $f \in Z_T^s$. In this case, we have

$$\begin{aligned}
\|\mathcal{L}(f)(t)\|_{H^{s+\mu}} &= \left\| \langle\xi\rangle^{s+\mu} \int_0^t \left(e^{i(t-t')\xi^3+|t-t'|\Phi(\xi)}\right) i\xi\widehat{f^{k+1}}(\xi,t')dt' \right\|_{L^2} \\
&\leq \int_0^t \|\langle\xi\rangle^{s+\mu}\, \xi\, e^{|t-t'|\Phi(\xi)}\|_{L^\infty} \|\widehat{f^{k+1}}(\cdot,t')\|_{L^2} dt' \\
&\leq \int_0^t \|\langle\xi\rangle^{s+\mu+1} e^{|t-t'|\Phi(\xi)}\|_{L^\infty} \|f(\cdot,t')\|_{L^{2(k+1)}}^{k+1} dt' \\
&\lesssim \|f\|_{\mathcal{Z}_T^s}^{k+1} \int_0^t \frac{1}{|t-t'|^{\frac{1+s+\mu}{p}}} \frac{1}{|t'|^{\frac{3k+2}{2p}}} dt' < \infty,
\end{aligned} \tag{55}$$

where $0 < \mu < p-1-s$, the definition of $\mathcal{Z}_T^s$-norm, Minkowski's inequality and inequality (49) from Lemma 8 are used. Note that, the condition $s \geq -1$ guarantees $s+\mu+1 \geq 0$ to use (49).

Case II, $s \geq p-1$**:** Similarly as above, using the fact that, in this case H^s is an algebra, we obtain

$$\begin{aligned}
\|\mathcal{L}(f)(t)\|_{H^{s+\mu}} &= \left\| \langle\xi\rangle^{s+\mu} \int_0^t \left(e^{i(t-t')\xi^3+|t-t'|\Phi(\xi)}\right) i\xi\widehat{f^{k+1}}(\xi,t')dt' \right\|_{L^2} \\
&\leq \int_0^t \|\langle\xi\rangle^{\mu}\, \xi\, e^{|t-t'|\Phi(\xi)} \left(J^s(f^{k+1})(\cdot,t')\right)^\wedge(\xi)\|_{L^2} dt' \\
&\leq \int_0^t \|\langle\xi\rangle^{\mu} \xi \left(e^{|t-t'|\Phi(\xi)}\right)\|_{L^\infty} \|J^s(f^{k+1})(\cdot,t')\|_{L^2} dt' \\
&\leq \int_0^t \|\langle\xi\rangle^{\mu+1} \left(e^{|t-t'|\Phi(\xi)}\right)\|_{L^\infty} \|f(\cdot,t')\|_{H^s}^{k+1} dt' \\
&\lesssim \|f\|_{\mathcal{Z}_T^s}^{k+1} \int_0^t \frac{1}{|t-t'|^{\frac{1+\mu}{p}}} dt' < \infty.
\end{aligned} \tag{56}$$

The last estimate in (56) follows by choosing $0 < \mu < p-1$.

To prove the continuity part, let $t_0 \in (0, T]$ fixed, and let $f \in Z_T^s$. We will prove that

$$\lim_{t\to t_0}\|\mathcal{L}(f)(t)-\mathcal{L}(f)(t_0)\|_{H^{s+\mu}}=0. \tag{57}$$

We divide the proof in two different cases.
Case a, $0<t\leq t_0$**:** In this case, using (54) and the additive property of the integral, we obtain

$$\begin{aligned}
\|\mathcal{L}(f)(t) &- \mathcal{L}(f)(t_0)\|_{H^{s+\mu}} = \\
&= \left\| \int_0^{t_0} V(t_0-t')\partial_x(f^{k+1})(t')d\tau - \int_0^{t} V(t-t')\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\
&\leq \left\| \int_0^{t} \left(V(t_0-t')-V(t-t')\right)\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\
&\quad + \left\| \int_t^{t_0} V(t_0-t')\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\
&=: I_1(t,t_0)+I_2(t,t_0).
\end{aligned} \tag{58}$$

We write

$$I_1(t,t_0)=\left\|\langle\xi\rangle^{s+\mu}\int_0^t\left(e^{i(t_0-t')\xi^3+(t_0-t')\Phi(\xi)}-e^{i(t-t')\xi^3+(t-t')\Phi(\xi)}\right)i\xi\widehat{f^{k+1}}(\xi,t')dt'\right\|_{L^2}, \tag{59}$$

and note that

$$(e^{i(t_0-t')\xi^3+(t_0-t')\Phi(\xi)}-e^{i(t-t')\xi^3+(t-t')\Phi(\xi)})=\int_t^{t_0}(i\xi^3+\Phi(\xi))e^{i(\tau-t')\xi^3+(\tau-t')\Phi(\xi)}d\tau.$$

We analyse I_1, considering two different cases.
Case a.1, $-1\leq s<\min\{2(p-2),\,p-1\}$**:** In this case, we have

$$\begin{aligned}
I_1(t,t_0) &= \left\| \int_0^t\int_t^{t_0}\langle\xi\rangle^{s+\mu}(i\xi^3+\Phi(\xi))e^{i(\tau-t')\xi^3+(\tau-t')\Phi(\xi)}\xi\widehat{f^{k+1}}(\xi,t')d\tau dt' \right\|_{L^2} \\
&\leq \left\| \int_0^t\int_t^{t_0}\langle\xi\rangle^{s+\mu+1}|i\xi^3+\Phi(\xi)|e^{(\tau-t')\Phi(\xi)}|\widehat{f^{k+1}}(\xi,t')|d\tau dt' \right\|_{L^2} \\
&\leq \int_0^t\int_t^{t_0}\|\langle\xi\rangle^{s+\mu+1}|i\xi^3+\Phi(\xi)|e^{(\tau-t')\Phi(\xi)}\widehat{f^{k+1}}(\xi,t')\|_{L^2}d\tau dt' \\
&\lesssim \int_0^t\int_t^{t_0}\|\langle\xi\rangle^{s+\mu+r}e^{(\tau-t')\Phi(\xi)}\widehat{f^{k+1}}(\xi,t')\|_{L^2}d\tau dt' \\
&\lesssim \|f\|_{\mathcal{Z}_T^s}^{k+1}\int_0^t\int_t^{t_0}\frac{1}{|\tau-t'|^{\frac{\mu+s+r}{p}}}\frac{1}{|t'|^{\frac{3k+2}{2p}}}d\tau dt' \\
&=: \|f\|_{\mathcal{Z}_T^s}^{k+1}J,
\end{aligned} \tag{60}$$

where $r=\max\{3,\,p\}+1$. In the domain of integration one has $0\leq t'\leq t\leq\tau\leq t_0$, and consequently $|\tau-t'|=(\tau-t')\geq t-t'=|t-t'|$. Therefore, we can estimate

J as follows

$$J = \int_0^t \frac{1}{|t'|^{\frac{3k+2}{2p}}} \int_t^{t_0} \frac{1}{|\tau - t'|^{\frac{\mu+s+r}{p}}} d\tau dt' \sim \int_0^t \frac{1}{|t'|^{\frac{3k+2}{2p}}} \left(\frac{1}{(t_0 - t')^{\alpha_\mu}} - \frac{1}{(t - t')^{\alpha_\mu}} \right) dt', \tag{61}$$

where

$$\alpha_\mu = \alpha_\mu(s, p) = \begin{cases} \frac{\mu+s+1}{p}, & \text{if} \quad p \geq 3 \quad \text{and} \quad 0 < \mu < p - 1 - s, \\ \frac{\mu+s+4-p}{p}, & \text{if} \quad p \leq 3 \quad \text{and} \quad 0 < \mu < 2(p-2) - s. \end{cases} \tag{62}$$

Note that, for the choices of p, s and μ, we have $\alpha_\mu < 1$.

Now, making change of variables $t' = t_0 x$ and $t' = tx$ respectively in the first and second integrals in (61) and taking in account that $\alpha_\mu < 1$, we obtain

$$J \sim t_0^{1-\alpha_\mu - \frac{3k+2}{2p}} \int_0^{t/t_0} \frac{1}{|x|^{\frac{3k+2}{2p}}} \frac{1}{(1-x)^{\alpha_\mu}} dx - t^{1-\alpha_\mu - \frac{3k+2}{2p}} \int_0^1 \frac{1}{|x|^{\frac{3k+2}{2p}}} \frac{1}{(1-x)^{\alpha_\mu}} dx \to 0, \tag{63}$$

whenever $t \to t_0$.

Case a.2, $s \geq \min\{2(p-2), p-1\}$**:** In this case, with a similar argument as above, and the fact that H^s is an algebra, one can get

$$I_1(t, t_0) \lesssim \|f\|_{\mathcal{Z}_T^s}^{k+1} \int_0^t \int_t^{t_0} \frac{1}{|\tau - t'|^{\frac{\mu+r}{p}}} d\tau dt' := \|f\|_{\mathcal{Z}_T^s}^{k+1} \tilde{J}. \tag{64}$$

Similarly to J in (61), we get

$$\tilde{J} \sim \int_0^t \left(\frac{1}{(t_0 - t')^{\alpha_\mu}} - \frac{1}{(t - t')^{\alpha_\mu}} \right) dt', \tag{65}$$

where

$$\alpha_\mu = \alpha_\mu(s, p) = \begin{cases} \frac{\mu+1}{p}, & \text{if} \quad p \geq 3 \quad \text{and} \quad 0 < \mu < p - 1, \\ \frac{\mu+4-p}{p}, & \text{if} \quad p \leq 3 \quad \text{and} \quad 0 < \mu < 2(p-2). \end{cases} \tag{66}$$

In this case too, for the choices of p, s and μ, we have that $\alpha_\mu < 1$.

As in (63), making change of variables and the fact that $\alpha_\mu < 1$, we obtain

$$\tilde{J} \sim t_0^{1-\alpha_\mu} \int_0^{t/t_0} \frac{1}{(1-x)^{\alpha_\mu}} dx - t^{1-\alpha_\mu} \int_0^1 \frac{1}{(1-x)^{\alpha_\mu}} dx \to 0, \tag{67}$$

whenever $t \to t_0$.

Therefore, in the light of (63) and (67), we get

$$I_1(t,t_0)\to 0, \qquad \text{if} \quad t\to t_0. \tag{68}$$

Analogously, since

$$\int_0^{t_0}\|V(t-t')\partial_x(f^{k+1})(t')dt'\|_{H^{s+\mu}}<\infty$$

we get

$$I_2(t,t_0)\to 0, \qquad \text{if} \quad t\to t_0. \tag{69}$$

Hence, using (68) and (69) in (58), we conclude the proof of the proposition in this case.

Case b, $0<t_0<t$**:** In this case too, using (54) and the additive property of the integral, we obtain

$$\begin{aligned}\|\mathcal{L}(f)(t) &- \mathcal{L}(f)(t_0)\|_{H^{s+\mu}} = \\ &= \left\| \int_0^{t_0} V(t_0-t')\partial_x(f^{k+1})(t')d\tau - \int_0^{t} V(t-t')\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\ &\le \left\| \int_0^{t_0} \big(V(t_0-t')-V(t-t')\big)\,\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\ &\quad + \left\| \int_{t_0}^{t} V(t-t')\partial_x(f^{k+1})(t')dt' \right\|_{H^{s+\mu}} \\ &=: I_1(t,t_0)+I_2(t,t_0).\end{aligned} \tag{70}$$

The rest follows analogously as in **Case a**, and this completes the proof of the proposition. □

Proposition 2 *Let* $s\ge 0$, $p>\frac{3k}{2}+1$. *There exists* $\mu:=\mu(s,p,k)>0$ *such that if*

$$\|f\|_{\tilde{\mathcal{Z}}_T^s}:=\sup_{t\in[-T,T]\setminus\{0\}}\left\{\|f(t)\|_{H^s}+|t|^{\frac{\gamma_k}{p}}\|\partial_x f(t)\|_{L^{2(k+1)}}\right\}<\infty, \tag{71}$$

then the application

$$t\mapsto \tilde{\mathcal{L}}(f)(t):=\int_0^t V(t-t')(f_x)^{k+1}(t')dt', \qquad 0<|t|\le T\le 1, \tag{72}$$

is continuous from $[-T,T]\setminus\{0\}$ *to* $H^{s+\mu}$.

Proof The proof follows by using (49) and a similar procedure applied in the proof of Proposition 2.9. □

3 Proof of the Main Results

Having the linear and nonlinear estimates at hand from the previous section, now we provide proof of the main results of this work.

Proof (*Proof of Theorem* 1) Consider the Cauchy problem (5) in its equivalent integral form

$$v(t) = V(t)v_0 - \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau, \tag{73}$$

where $V(t)$ is the semi-group associated with the linear part given by (13).

Let us define an application

$$\Psi(v)(t) = V(t)u_0 - \int_0^t V(t-\tau)\partial_x(v^{k+1})(\tau)d\tau. \tag{74}$$

For $s \geq -1$, $r > 0$ and $0 < T \leq 1$, let

$$B_r^T = \{f \in \mathcal{X}_T^s;\ \|f\|_{\mathcal{X}_T^s} \leq r\},$$

be a ball in $\mathcal{X}_T^s$ with center at origin and radius r. We will show that there exists $r > 0$ and $0 < T \leq 1$ such that the application Ψ maps B_r^T into B_r^T and is a contraction. For this, let $v \in B_r^T$. Using the estimates (35) and (40), one can obtain

$$\|\Psi(v)\|_{\mathcal{X}_T^s} \leq C\|v_0\|_{H^s} + C_{p,k}T^{\omega_k}\|v\|_{\mathcal{X}_T^s}^{k+1}, \tag{75}$$

where $\omega_k = \frac{2p-3k-2}{2p} > 0$.

For $v \in B_r^T$ let us choose $r = 4c\|v_0\|_{H^s}$ in such a way that $cT^{\omega_k}r^k = 1/4$. For this choice, one can easily obtain

$$\|\Psi(v)\|_{X_T^s} \leq \frac{r}{4} + cT^{\omega_k}r^{k+1} \leq \frac{r}{2}. \tag{76}$$

From (76) we conclude that Ψ maps B_r^T into itself. A similar argument proves that Ψ is a contraction. Hence Ψ has a fixed point v which is a solution of the Cauchy problem (5) such that $v \in C([0,T], H^s(\mathbb{R}))$. The rest of the proof follows from standard argument, see for example [6].

The regularity part follows using Lemma 4 and Proposition 1 as in [3]. □

Proof (*Proof of Theorem* 2) Proof of this theorem is very similar to that of Theorem 1. In this case, we use the estimate (36) from Lemma 6 and estimate (41) from Lemma 7 to perform the contraction mapping argument. □

Acknowledgements The first author thanks supports from CNPq under grants 304036/2014-5 and 481715/2012-6. The second author acknowledges supports from FAPESP under grants 2012/

20966-4, and CNPq under grants 479558/2013-2 and 305483/2014-5. The authors would like to thank anonymous referees whose comments helped to improve the original manuscript.

References

1. Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation. Geom. Funct. Anal. **3**, 209–262 (1993)
2. Carvajal, X., Panthee, M.: Well-posedness for some perturbations of the KdV equation with low regularity data. Electron. J. Differ. Equations **2**, 1–18 (2008)
3. Carvajal, X., Panthee, M.: Well-posedness of KdV type equations. Electron. J. Differ. Equations **40**, 1–15 (2012)
4. Carvajal, X., Panthee, M.: Sharp local well-posedness of KdV type equations with dissipative perturbations. Q. Appl. Math. **74**, 571–594 (2016)
5. Dix, D.B.: Nonuniqueness and uniqueness in the initial value problem for Burgers' equation. SIAM J. Math. Anal. **27**, 708–724 (1996)
6. Kenig, C.E., Ponce, G., Vega, L.: Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle. Commun. Pure Appl. Math. **46**(4), 527–620 (1993)
7. Kenig, C.E., Ponce, G., Vega, L.: A bilinear estimate with applications to the KdV equation. J. Am. Math. Soc. **9**(2), 573–603 (1996)
8. Tao, T.: Multilinear weighted convolution of L^2 functions, and applications to nonlinear dispersive equations. Am. J. Math. **123**, 839–908 (2001)
9. Tao, T.: Nonlinear Dispersive Equations, Local and Global Analysis. CBMS Regional Conference Series in Mathematics, vol. 106. AMS, Providence (2006)

A Kinetic Approach to Steady Detonation Waves and Their Linear Stability

Filipe Carvalho

Abstract Detonation waves have a relevant role in many engineering processes, namely those related to propulsion. Most studies, regarding the stability of detonation waves, have been carried out by computer simulations, notwithstanding their multi-scale nature and unstable behaviour makes it difficult to achieve accurate results. In this paper we propose a kinetic approach to this problem, explain the constraints and the difficulties that this choice entails as well as its advantages. Numerical methods are needed to obtain the solutions of the stability. The ones in use imply that a regular computer takes a long period of time to provide the solutions. In this paper, taking into account the developments proposed by others, we present a numerical procedure that helps to overcome the difficulties of the current methods and allows us to answer some questions related to the stability problem.

Keywords Kinetic theory · Boltzmann equation · Chemical reactions
Steady detonation waves · Hydrodynamic stability

1 Introduction

In kinetic theory of gases the description of a gas system evolution is obtained using distribution functions in the mesoscopic level. These distributions are defined in the phase space composed by the position x, and the velocity c for a given time t in such a way that $f(x, c, t)$ represents the number of particles that, at time t, are in the volume element $dxdc$ around position x and velocity c.

The basic assumption of kinetic theory modelling is that the number of particles of a gas is so large that it can be treated as a continuum. If we consider that particles obey

F. Carvalho (✉)
Centro de Matemática, Instituto Politécnico de Viana do Castelo,
Universidade do Minho, Braga, Portugal
e-mail: filipecarvalho@esce.ipvc.pt

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_5

classical laws of mechanics and we neglect macroscopic forces acting on the particles and interactions between them, according to Newton's Principle, each particle travels at constant velocity along a straight line, therefore, $f(x, c, t) = f(x - ct, c, 0)$ for any time t. In these conditions, f is a weak solution of the free transport equation

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} c_i \frac{\partial}{\partial x_i} f = 0. \tag{1}$$

In 1872, see Ref. [1] Ludwig Boltzmann derived an equation, the so called Boltzmann Equation, that is used to describe the evolution of a gas, considering that particles interact. This derivation was based on some physical assumptions (for more details see for instance [1] or [2]):

(a) the collision time is much smaller in comparison with the free travelling time of a particle;
(b) collisions between more than two particles may be neglected;
(c) collisions are micro-reversible;
(d) the velocities of two particles that are about to collide are uncorrelated;
(e) there are some physical quantities that do not change during a collision, such as mass, linear momentum and kinetic energy.

The Boltzmann Equation is an integro-differential equation that, considering a single gas specie without macroscopic forces acting on the particles, has the following expression [3]:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} c_i \frac{\partial}{\partial x_i} f = \mathcal{Q}(f, f), \tag{2}$$

where $\mathcal{Q}(f, f)$ is the integral collisional operator. An extension of the Boltzmann equation to chemical reactive gases, will be presented with more detail in the next section.

The Boltzmann Equation was then extended to many different situations, such as gases with macroscopic forces acting on the particles, gases with more than one constituent, gases with more than one constituent where particles may interact not only in elastic collisions but also in reactive collisions, polyatomic gases, relativistic systems, among others.

The equilibrium solution is the only exact solution of the Boltzmann Equation, that is currently known. The wide range of applications of this equation and the difficulties of finding its solutions led many researchers to develop simplified variants of the Boltzmann equation which allowed them to find solutions, while still retaining its main features. The BGK model [4], the Kac model [5] and the discrete-velocity models [6] are some of the examples. Others, such as Grad, Chapman and Enskog, developed different methods to obtain approximate solutions of the full Boltzmann Equation.

In this work we will address the problem of the detonation wave propagation, that will be presented in Sect. 3 and its stability analysis that will be presented in Sect. 4. The kinetic approach to this problem, using the Boltzmann equation to describe the evolution of the gas system, allows a discussion around the microscopic features of the gas, such as the molecular potential, the reaction heat of the considered chemical reaction and the activation energy needed for the reaction to occur. It also allows the study on their influence in different macroscopic regimes, such as different detonation velocities and different initial concentrations of the gas constituents.

In what follows, we will present some of our work on this subject and make reference to a few other works that were important in the development of the current state of art. The paper is organized as follows. The kinetic description of the gas system is presented in Sect. 2. In Sects. 3 and 4 we present the detonation wave problem and the stability analysis modelling. Finally, in Sect. 5 we discuss some numerical techniques that were used to obtain solutions for the stability analysis and present recent developments on the study of stability.

2 Kinetic Framework and Macroscopic Equations

In this section we present the main microscopic features of the gas system and synthetically explain how to go from the microscopic description to the macroscopic equations that describe the gas evolution.

2.1 Microscopic Modelling

We consider a binary gaseous mixture whose constituents, A and B, have equal molecular mass m and binding energies E_A and E_B. Vibrational and rotational molecular degrees of freedom are not taken into account. The gas particles can undergo binary elastic collisions as well as collisions with chemical reaction according to the single reversible symmetric law

$$A + A \rightleftharpoons B + B. \tag{3}$$

At the mesoscopic scale, the thermodynamical behaviour of the mixture is modelled by the following system of Boltzmann equations for the constituent distribution functions $f_\alpha(x, c_\alpha, t)$, with $x, c_\alpha \in \mathbb{R}^3$ and $t \in \mathbb{R}^+$,

$$\frac{\partial f_\alpha}{\partial t} + c_\alpha \nabla_x f_\alpha = \mathcal{Q}(f_\alpha, f_\alpha), \qquad \alpha = A, B, \tag{4}$$

where $\mathcal{Q}(f_\alpha, f_\alpha) = \mathcal{Q}_\alpha^E + \mathcal{Q}_\alpha^R$, and

$$\mathcal{Q}_\alpha^E = \sum_{\beta=A}^{B} \int \left(f'_\alpha f'_\beta - f_\alpha f_\beta\right) \mathrm{d}^2\left(g_{\beta\alpha} \cdot \mathbf{k}_{\beta\alpha}\right) d\mathbf{k}_{\beta\alpha} d\mathbf{c}_\beta, \tag{5}$$

$$\mathscr{Q}_\alpha^R = \int \left[f_\beta f'_\beta - f_\alpha f'_\alpha \right] \sigma_\alpha^{\star 2} \left(g_\alpha \cdot \mathbf{k}_\alpha \right) d\mathbf{k}_\alpha d\mathbf{c}'_\alpha. \tag{6}$$

Above, in Eq. (5), the primes denote post collisional distribution functions, d the elastic particle diameter, $g_{\beta\alpha}$ the relative velocity between the α and β particles, $\mathbf{k}_{\beta\alpha}$ the unit collision vector and $d\mathbf{k}_{\beta\alpha}$ the element of solid angle for elastic collisions and in Eq. (6), the primes are used to distinguish two identical particles that participate in the reactive collision, $\mathbf{k}_\alpha$ is the unit collision vector, g_α the relative velocity between two identical particles of constituent α, with $(\alpha, \beta) \in \{(A, B), (B, A)\}$, $d\mathbf{k}_\alpha$ the element of solid angle for reactive collisions, and the term $\sigma_\alpha^{\star 2}$ is the differential reactive cross section.

The collisional operator $\mathscr{Q}(f_\alpha, f_\alpha)$, was used by Boltzmann to introduce the influence of the encounters between particles in the description of the evolution of the gas system. In the considered kinetic framework, this operator must consider elastic and reactive collisions, therefore, it is split into two operators: the elastic operator $\mathscr{Q}_\alpha^E$ and the reactive operator $\mathscr{Q}_\alpha^R$. As we can see in Eq. (5), in the elastic operator all collisions are considered, even those between two constituents A or two constituents B that end up being reactive and thus considered in the reactive operator. This double counting shouldn't create any major problem in situations where the number of elastic encounters is much larger than the number of reactive encounters. Although this is what happens more frequently, some recent works introduced a correction term to avoid this problem, see for instance [7].

In what follows, the reactive cross section $\sigma_\alpha^{\star 2}$ is defined as

$$\sigma_\alpha^{\star 2} = \begin{cases} 0 & \text{for } \gamma_\alpha < \varepsilon_\alpha^\star \\ d^2 & \text{for } \gamma_\alpha > \varepsilon_\alpha^\star \end{cases} \quad \alpha = A, B, \tag{7}$$

where the relative translational energy $\gamma_\alpha = \frac{m g_\alpha^2}{4kT}$ and the activation energies $\varepsilon_\alpha^\star$ are written in units of kT, with k being the Boltzmann constant and T the temperature of the mixture. There are many works on reacting gases that choose different collision potential. For more details on the implications of these choices see for instance [3].

2.2 Hydrodynamical Limit

Macroscopic state variables may be defined as mean values of microscopic quantities. For instance, the number densities n_α of the constituents, the mean velocity components v_i and temperature T of the mixture, may be defined by

$$n_\alpha = \int f_\alpha dc_\alpha, \qquad v_i = \frac{1}{n} \sum_{\alpha=A}^{B} \int c_i^\alpha f_\alpha dc_\alpha, \qquad T = \frac{m}{3kn} \sum_{\alpha=A}^{B} \int (c_\alpha - \mathbf{v})^2 f_\alpha dc_\alpha, \tag{8}$$

respectively.

As was explained before, there are some physical microscopic quantities that must be preserved during a collision and a proper microscopic model must respect those physical characteristics. It is possible to prove that the model that is used in this paper has the desired properties. Therefore, if f_α is a solution of the Boltzmann equation, and $\Phi_\alpha \in \left\{1, m_\alpha, m_\alpha c_1^\alpha, m_\alpha c_2^\alpha, m_\alpha c_3^\alpha, \frac{1}{2} m_\alpha c_\alpha^2 + E_\alpha\right\}$ then, multiplying both sides of the Boltzmann equation (2.1) by the proper Φ_α and integrating over c_α leads to the following macroscopic equations

$$\frac{\partial n_\alpha}{\partial t} + \sum_{i=1}^{3} \frac{\partial n_\alpha}{\partial x_i} \left(n_\alpha v_i + n_\alpha u_i^\alpha\right) = \tau_\alpha, \tag{9}$$

$$\frac{\partial}{\partial t}(\rho v_i) + \sum_{j=1}^{3} \frac{\partial}{\partial x_i} \left(p_{ij} + \rho v_i v_j\right) = 0, \tag{10}$$

$$\frac{\partial}{\partial t}\left(\frac{3}{2} nkT + \sum_{\alpha=A}^{B} n_\alpha E_\alpha + \frac{1}{2}\rho v^2\right) + \sum_{i=1}^{3} \frac{\partial}{\partial x_i}\left\{q_i + \sum_{j=1}^{3} p_{ij} v_j + \left(\frac{3}{2} nkT + \sum_{\alpha=A}^{B} n_\alpha E_\alpha + \frac{1}{2}\rho v^2\right) v_i\right\} = 0. \tag{11}$$

Above, u_i^α and τ_α are the diffusion velocity components and the reaction rate of the constituent α, and ρ, p_{ij}, q_i are the mass density, pressure tensor and heat flux of the gas system.

These six Eqs. (9)–(11) do not form a closed system and we must pass to the hydrodynamic limit to close it. Here we adopt the solution obtained in [8] using the asymptotic method of Chapman-Enskog and a second order Sonine expansion of the distribution functions. This approximate solution was used to describe the complete reactive process, starting from its early stage and going towards the equilibrium final state. Depending on the chemical regimes that we want to model, we may consider other regimes or even other approximation methods. In paper [9], for example, the author derived a different hydrodynamical limit where elastic encounters are the dominant interactions between particles. The subject of deriving hydrodynamical limits from microscopic descriptions is a research area for itself, see for instance [10, 11]. Although interesting, this subject will not be addressed here.

With the distribution function obtained in [8] we may evaluate some macroscopic quantities of Eqs. (9)–(11) such as the reaction rate τ_α, constituent diffusion velocities u_i^α, mixture pressure tensor p_{ij} and heat flux q_i. The resulting equations, in the one dimensional form, may be rearranged in the following way:

$$\frac{\partial n_A}{\partial t} + \frac{\partial}{\partial x}(n_A v) = \tau_A\,, \tag{12}$$

$$\frac{\partial}{\partial t}(n_A + n_B) + \frac{\partial}{\partial x}\Big[(n_A + n_B)v\Big] = 0, \tag{13}$$

$$\frac{\partial v}{\partial t} + \frac{1}{\rho}\frac{\partial p}{\partial x} + v\frac{\partial v}{\partial x} = 0, \tag{14}$$

$$\frac{\partial p}{\partial t} + v\frac{\partial p}{\partial x} + \frac{5}{3}p\frac{\partial v}{\partial x} + \frac{2}{3}\sum_{\alpha=A}^{B} E_\alpha \tau_\alpha = 0, \tag{15}$$

where v represents now the x-component of the gas system velocity. This choice was motivated by the fact that these equations' aim is to study the one dimensional detonation wave. This subject is the main topic of the next section.

3 One Dimensional Steady Detonation

The one dimensional detonation wave model is based on the assumption that ahead of a shock there is a gas system in a meta stable equilibrium. This means that in this location, chemical reactions between particles may be neglected but, if a trigger appears, such as a shock wave with a large velocity, then a chemical reaction occurs until the gas system reaches total, elastic and chemical, equilibrium. This phenomenon may be described by the simple ZND (Zel'dovich, John Von Neumann and Werner Döring) model in the case of a detonation wave with velocity greater then the CJ (Chapman-Jouguet) velocity. For more detail see for instance [12, 13].

Many researchers developed studies concerning the detonation wave problem. In paper [14] the authors did so by exploring the reaction zone thickness for different wave velocities and different activation energies. In [15] the influence of the chemical reaction velocity in the detonation wave profiles was studied.

In paper [16], in collaboration with Ana Jacinta Soares, we studied the influence of the reaction heat on the detonation wave profiles and its relation with the activation energy of the chemical reaction. We considered the gas system described before with the chemical reaction defined in Eq. (3) and with the governing Eqs. (12)–(15). Then, transforming these equations' frame to the shock front we obtained the steady state equations:

$$\frac{d}{dx}\Big[(v-D)\,n_A\Big] = D\tau_A, \tag{16}$$

$$\frac{d}{dx}\Big[(v-D)\,(n_A+n_B)\Big] = 0, \tag{17}$$

$$\frac{d}{dx}\Big[(v-D)\,\rho v + nkT\Big] = 0, \tag{18}$$

$$\frac{d}{dx}\left[(v-D)\left(\frac{3}{2}nkT + \frac{\rho v^2}{2} + E_A n_A + E_B n_B\right) + nkTv\right] = 0. \tag{19}$$

These equations with the Rankine-Hugoniot conditions

$$(v - D)\, n_A = -D n_{A0}, \tag{20}$$

$$(v - D)\, n_B = -D n_{B0}, \tag{21}$$

$$(v - D)\, \rho v + nkT = n_0 k T_0, \tag{22}$$

$$(v - D)\left(\frac{3}{2} nkT + \frac{\rho v^2}{2} + E_A n_A + E_B n_B\right) + nkTv = \\ = -D\left(\frac{3}{2} n_0 k T_0 + E_A n_{A0} + E_B n_{B0}\right), \tag{23}$$

constitute a closed equation system that is used to obtain the steady state detonation wave solution for a given set of parameters such as the initial values for the macroscopic variables noted with the subscript 0 and the detonation wave velocity D.

With this system it was possible to observe, for the input parameters, that for an exothermic reaction the steady detonation solution is a rarefaction wave, which means that the pressure decreases along the reaction zone, and for an endothermic reaction the steady detonation solution is a compression wave, which means that the pressure increases along the reaction zone. In addition, it was also possible to observe, among other expected and essential physical features, that the temperature increases for an exothermic chemical reaction and decreases for an endothermic reaction and that the extent of the reaction zone decreases when the reaction heat increases.

These are examples of works where the authors started from a kinetic description of the gas to analyse, using an appropriate hydrodynamic limit, the behaviour of the detonation wave for different values of hider microscopic, such as the reaction heat and activation energy, or macroscopic variables, such as initial concentrations and wave velocity. Provided we have experimental data, we may determine approximations for unknown parameters, such as transport coefficients of diffusion, viscosity and thermal conductivity and use the resulting Euler equations or Navier-Stokes equations to describe the spatio-temporal evolution of the gas system. However, for a number of reasons, we do not have experimental data for all cases. Hence, a kinetic theoretical approach constitutes an important step in the understanding of a phenomenon and may be used to predict the result of an experiment, regarding specific values of the variables in consideration [17].

4 Linear Stability Analysis

In this section we will discuss the linear stability of the detonation wave solution obtained with the ZND model, which is an idealised model of the detonation problem. Moreover, linear stability analysis only allows us to evaluate the response of the detonation wave solution to small disturbances. Although this is a simplified problem, it reflects some of the important features of more realistic models and simultaneously allows an accurate mathematical approach.

As far as we know, the study of more complex and complete situations, such as multidimensional flows, is nearly non-existent. Furthermore, the difficulties of this approach already constitute an interesting mathematical challenge and, although some recent contributions and approaches to this problem have been made, there is still much to be done, namely in the efficiency of the numerical methods that are used to search for solutions.

The modern theory of detonation stability started with Erpenbeck, in 1962 [18, 19] and it is still the basis for the current stability work. Erpenbeck described the stability problem as an initial-value problem considering a small perturbation from the steady state. Along with other mathematical results, Erpenbeck used the Nyquist winding theorem to develop a numerical procedure that, for a given set of parameters, would determine the number of zeros of an analytic function in the complex plan. The zeros of this function corresponded to unstable solutions of the stability problem.

In 1990 Lee and Stewart [20] introduced a normal-mode approach and a numerical shooting method to develop a simpler and more efficient search for unstable solutions. After Erpenbeck's work, others, besides Lee and Stewart, contributed to the development of the stability analysis, namely [12, 13, 18, 21–23]. For more details on the state of detonation stability see [24].

In the next section, we present in more detail the procedure used to develop a linear stability analysis, using the normal-mode approach proposed in [20].

4.1 Stability Macroscopic Equations

It is well known from theoretical studies, as well as from experimental investigations, that steady solutions may degenerate into an oscillatory solution in the long-time limit. The first required test of the steady solutions should be the evaluation of its response to small rear boundary perturbations. To do that, we introduce a perturbation that induces a deviation on the shock wave position, giving rise to small perturbations on the state variables that propagate into the reaction zone. The evolution of the state variables perturbations over time determines the stability of the steady detonation solution. In fact, when some perturbation grows over time, the steady solution is said to be hydrodynamically unstable and if all perturbations decay in time, the steady solution is stable.

The linear stability problem is formulated as an initial-boundary value problem in terms of the stability differential equations, with initial conditions at the Von Neumann state and an additional closure boundary condition at the final state.

The equations for the study of the linear stability are obtained from the steady state equations attached to the shock front (16)–(19) and the initial Rankine–Hugoniot conditions (20)–(23). The shock front of a steady detonation wave with constant velocity D is placed in $x - Dt$ and if we introduce a perturbation on the shock front that depends on time we may write

$$x = x^{\ell} - \psi(t), \qquad \text{with} \qquad \psi(t) = Dt + \widetilde{\psi}(t), \tag{24}$$

where x^ℓ is the laboratory frame coordinate, $\widetilde{\psi}(t)$ the displacement of the shock wave from the unperturbed position due to a small perturbation, and $\psi(t)$ the location of the perturbed wave. In the new shock attached coordinate system, the instantaneous position of the perturbed shock wave is $x = 0$ and its velocity is $D(t) = D + \widetilde{\psi}'(t)$. Furthermore, a normal mode expansion with exponential time dependent perturbations is assumed for the steady state variables,

$$z(x,t) = z^*(x) + e^{at}\,\overline{z}(x), \qquad \psi(t) = \overline{\psi}\,e^{at}, \qquad a,\ \overline{\psi} \in \mathbb{C}, \tag{25}$$

where $z = [n_A\ n_B\ v\ p]^T$ is the state vector, $z^*(x)$ indicates the one-dimensional steady solution, $\overline{z}(x)$ is the vector of complex eigenfunctions representing the unknown spatially disturbances, $\overline{\psi}$ is a complex disturbance amplitude parameter and a is the complex eigenvalue, with Re a and Im a being the disturbance growth rate and frequency, respectively. The transformed governing equations in the perturbed shock frame are then linearized about the steady solution $z^*(x)$, by means of the expansions (25). Performing a further normalization of the state variables, with respect to the complex amplitude parameter $\overline{\psi}$, namely $\overline{w} = \overline{z}/\overline{\psi}$, one obtains the evolution equations in the wave coordinate x, for the complex disturbances. Rewriting $\overline{z}$ instead of $\overline{w}$, the resulting equations, for $x \in]x_F, 0[$, are

$$Da\overline{n}_\alpha + \left(v^* - D\right)\frac{d\overline{n}_\alpha}{dx} + \frac{n_\alpha^*}{dx}\left(\overline{v} - Da\right) + \frac{dv^*}{dx}\overline{n}_\alpha + n_\alpha^*\frac{d\overline{v}}{dx} = \overline{\tau}_\alpha, \qquad \alpha = A, B, \tag{26}$$

$$\rho^* Da\overline{v} + \frac{d\overline{p}}{dx} + \rho^*\frac{dv^*}{dx}\left(\overline{v} - Da\right) + \left(v^* - D\right)\frac{dv^*}{dx}\overline{\rho} + \rho^*\left(v^* - D\right)\frac{d\overline{v}}{dx} = 0, \tag{27}$$

$$Da\overline{p} + \frac{5}{3}\left(p^*\frac{d\overline{v}}{dx} + \overline{p}\frac{dv^*}{dx}\right) + \left(v^* - D\right)\frac{d\overline{p}}{dx} + \left(\overline{v} - Da\right)\frac{dp^*}{dx} = \frac{Q_R^* D\overline{\tau}_A}{3}, \tag{28}$$

with $\overline{\tau}_\alpha$ being the linearized reaction rates, explicitly presented in [16].

The initial conditions to be added to the stability Eqs. (26)–(28) are obtained by introducing the same expansion (25) and linearization about the steady solution as before, followed by a normalization with respect to $\overline{\psi}$. Hence, the resulting equations are:

$$\overline{n}_\alpha(0) = \frac{\left(n_\alpha^* - n_{\alpha 0}\right)a - n_\alpha^*\overline{v}(0)}{v^* - D}, \qquad \alpha = A, B, \tag{29}$$

$$\overline{v}(0) = \frac{3\rho_0 v^{*2} + \frac{3}{2}\left(p^* - p_0\right) - \frac{3}{2}D\rho_0 v^* + 2E_A n_0 + Q_R^* n_{B0}}{-\rho^*\left(v^* - D\right)^2 + \frac{5}{2}p^*}\,a\,, \tag{30}$$

$$\overline{p}(0) = -\rho_0 a v^* - \left(v^* - D\right)\rho^*\overline{v}(0)\,. \tag{31}$$

Equations (26)–(28) constitute the stability equations for the present modelling, giving the spatial evolution of the complex perturbations $\overline{z}(x)$ in the reaction zone, from the perturbed shock position $x = 0$, with initial conditions given by Eqs. (29)–(31), to the equilibrium final state $x = x_F$.

They form a system of eight first-order homogeneous linear ordinary differential equations with spatially varying coefficients, and eight initial conditions, for the real and imaginary parts of the complex perturbations and of the eigenvalue a. This system, which henceforward will be called stability system, is not closed and more information is needed to close it.

4.2 Radiation Condition

The radiation condition is called, in literature, the closure condition since it is used to close the stability equations system obtained in the previous subsection. This condition was adopted in many previous works on detonation stability as, for example, in papers [20, 24–28] and it is needed for physical purposes. This is a system of eight equations and ten variables therefore it is possible to consider two of those variables, for instance the real and the imaginary part of the eigenvalue a, as parameters and solve the system for the remaining eight variables. By doing so, we would obtain a solution for the stability system. However, this may not be a solution for the stability problem. In fact, the probability that the solution obtained in this manner could represent a solution for the stability problem is very low.

There are, in the above cited papers, some interesting explanations for the need to include this additional condition in order to obtain a physical solution to the problem. One argues that the detonation wave solution results exclusively from the interplay between the leading shock and the reaction zone and can not be affected by further disturbances traveling towards the shock from a great distance from the reaction zone. Regarding the adopted model, it has the expression:

$$\overline{v}(x_F) + a = \frac{-1}{\gamma \rho^*_{eq} c^*_{eq}} \, \overline{p}(x_F), \tag{32}$$

where γ is the ratio of specific heats, c^*_{eq} and ρ^*_{eq} the isentropic sound speed and gas density at the equilibrium final state, for $x = x_F$.

A solution of the linear stability problem of the steady detonation in terms of the complex disturbances $\overline{z}(x)$ and eigenvalue a must be obtained using the ordinary differential equations (26)–(28) for $x \in]x_F, 0[$, with initial conditions (29)–(31) at $x = 0$ and closure condition (32) at $x = x_F$. This problem was addressed in [16] as described in the next subsections.

4.3 Numerical Method

The stability problem is addressed numerically, with an iterative shooting technique based on the numerical method proposed by Lee and Stewart in paper [20], with the aim of obtaining the stability spectrum for the eigenfunction perturbations $\overline{z}$ and

eigenvalue perturbation parameter a, in terms of the parameters characterizing the steady solution. We choose a trial value of a in a fixed bounded domain $\mathscr{R}$ of the complex plane and then integrate Eqs. (26)–(28) in the reaction zone $]x_F, 0[$ with initial conditions (29)–(31) at $x = 0$, using a fourth order Runge–Kutta routine. The solution $\overline{z}(x)$, $x \in [x_F, 0]$, obtained for the considered trial value of a is then evaluated for $x = x_F$ to check if the boundary condition (32) is verified.

As was mentioned before, for a given steady detonation solution, an arbitrary value of a does not satisfy the closure condition (32) and thus the outcome is not a solution of the stability problem. To overcome this mishap, we may use the residual function $\mathscr{H}(a)$, defined from the closure condition (32) by the expression

$$\mathscr{H}(a) = \overline{v}(x_F) + a + \frac{1}{\gamma \rho^*_{eq} c^*_{eq}} \overline{p}(x_F), \quad a \in \mathbb{C}. \tag{33}$$

Notice that the radiation condition is verified if and only if $\mathscr{H}(a) = 0$, therefore to search for solutions considering the eigenvalue a to be in a limited region $\mathscr{R}$ of the complex plan, we may choose a large number of trial values a in that region and search for the zeros of $\mathscr{H}$. Although it looks like a simple task, the search for the zeros of $\mathscr{H}$ is a very time consumer task. There are some straightforward procedures introduced by Lee and Stewart in paper [20] but they require a huge number of trial values in a region of the complex plan, which implies that a regular computer needs a long period of time to do the calculations.

In paper [16] we introduced a numerical procedure to reduce the number of trial values that are needed to search for the zeros of $\mathscr{H}$. This procedure recovered the Erpenbeck's idea of counting the number of zeros of a function in a fixed domain of the complex plan and combined it with the shooting method proposed by Lee and Stewart. We believe that it is important to develop new and more efficient methods to search for zeros of the residual function, in order to help researchers in the development of a complete stability analysis on the detonation wave propagation. For more realistic cases, the calculations are even longer and those resources are all the more necessary.

In what follows, we present the numerical method that we used to count the number of zeros in a limited region of the complex plan.

Numerical Procedure

First we have to decide the region $\mathscr{R}$, of the complex plan, where we want to search for eigenvalues. With the expansion about the steady state defined in Eq. (25), we know that a stability solution is unstable if and only if $\mathrm{Re}\, a > 0$. We also know that it takes only one unstable solution to classify a steady detonation solution as unstable and that the existence of stable solutions bring no information about the stability of the steady detonation solution. Therefore, we must search for solutions in the right half of the complex plan. On the other hand, since these modes occur in conjugate pairs, it is enough to choose a domain $\mathscr{R}$ in the upper-right quarter of the complex plan.

The argument principle states that the difference between the number Z of zeros and P of poles of the function $\mathscr{H}$ within the region $\mathscr{R}$, provided that there are no

zeros in its contour, is given by

$$Z - P = \frac{1}{2\pi i} \int_{\zeta} \frac{\mathscr{H}'(u)}{\mathscr{H}(u)}\, du, \tag{34}$$

or equivalently by

$$Z - P = \frac{1}{2\pi i} \int_{k}^{\ell} \frac{\mathscr{H}'(\zeta(t))}{\mathscr{H}(\zeta(t))} \parallel \zeta'(t) \parallel dt, \tag{35}$$

where $\zeta : [k, \ell] \to \mathbb{C}$ is a path smooth by parts, describing the contour of $\mathscr{R}$ in the positive sense. We may consider that $\mathscr{H}$ does not have any poles, since we expect that the disturbancies do not blow in a finite time consistent with a linear analysis of the stability. Therefore we have $P = 0$ and the expression (35) gives the number of zeros of $\mathscr{H}$ inside the region $\mathscr{R}$,

$$Z = \frac{1}{2\pi i} \int_{k}^{\ell} \frac{\mathscr{H}'(\zeta(t))}{\mathscr{H}(\zeta(t))} \parallel \zeta'(t) \parallel dt. \tag{36}$$

The mean value theorem states that the integral in expression (36) is equal to the integral

$$\int_{k}^{\ell} \frac{\mathscr{H}'(\zeta(t))}{\mathscr{H}(\zeta(t))} \parallel \zeta'(t) \parallel dt = \mu(\ell - k), \tag{37}$$

where μ represents the mean value in the interval $[k, \ell]$ of the function h defined by

$$h(t) = \frac{\mathscr{H}'(\zeta(t))}{\mathscr{H}(\zeta(t))} \parallel \zeta'(t) \parallel, \quad t \in [k, \ell]\ . \tag{38}$$

Hence, to count the number of zeros in the region $\mathscr{R}$, all we need to compute is the value of μ.

It is well known that if n is large enough, then the mean value of the sample S, μ_S, can be treated as a statistical variable following a normal distribution with mean value μ and standard deviation $\sigma_S/\sqrt{n}$, where σ_S is the standard deviation of S. Therefore, the mean value μ of the function h can be inferred in a confidence interval by the mean value μ_S of the sample S. In this situation we considered

$$S = \left\{h(t_j) :\ i = 1, 2, \ldots, n\right\}, \tag{39}$$

with $t_j \in [k, \ell]$. To obtain the value of each $h(t_j)$, since we know path ζ, we directly obtain the values for $\zeta'(t_j)$ and $\zeta(t_j)$ and this last one is a point in the complex plan that might be used as a trial value to the eigenvalue a_j and thus solving the stability equation system, we obtain the value of $\mathscr{H}(\zeta(t_j))$. The value of $\mathscr{H}'(\zeta(t_j))$ is obtained choosing a suitable point close enough to $\zeta(t_j)$, say b_j, with $\mathrm{Re}\, b_j = \mathrm{Re}\, a_j + 10^{-6}$ and $\mathrm{Im}\, a_j = \mathrm{Im}\, b_j$, as follows

$$\mathscr{H}'(a_j) \approx \frac{\mathscr{H}(b_j) - \mathscr{H}(a_j)}{b_j - a_j}, \quad i = 1, 2, \ldots, n. \tag{40}$$

Consequently, the number of zeros of the residual function $\mathscr{H}$ inside the domain $\mathscr{R}$ is estimated as follows

$$\frac{\ell - k}{2\pi i}\left(\mu_S - 2.58\,\frac{\sigma_S}{\sqrt{n}}\right) < Z < \frac{\ell - k}{2\pi i}\left(\mu_S + 2.58\,\frac{\sigma_S}{\sqrt{n}}\right), \tag{41}$$

where the number 2.58 is used to assure the 99% confidence of the interval.

5 Discussion

The numerical procedure presented in the previous section requires a much smaller number of trial values than those needed to compute the methods presented by Lee and Stewart and is sufficient to determine if, given a specific set of parameters, a steady detonation is stable or not. This is precisely the information we have to have if we want to determine the stability boundary in some parameter plan such as the plan defined by the reaction heat and forward activation energy. On the other hand, if we want to determine the value of the eigenvalues, we need to proceed with the computation, considering subregions of $\mathscr{R}$ and using the numerical procedure considered above or applying the methods proposed by Lee and Stewart or even a combination of both.

We consider that it is important to improve this method or to create new ones to obtain solutions of the stability problem in order to increase the knowledge on the linear stability problem of the one-dimensional detonation wave as well as to approach more realistic models.

In paper [16] we were able to determine the region of the parameter plan of the reaction heat and forward activation energy that corresponds, giving a specific set of parameters, to unstable solution. Moreover, we tracked the fundamental eigenvalue along different values of the reaction heat. These are two examples of the nice perspectives that we might have on the macroscopic behaviour of a physical phenomenon such as a detonation wave, starting from the microscopic features of the system constituents and their relations.

Stability analysis is a major issue in the study of detonation waves which, in turn, are used to model many relevant engineering processes. As was said before, the model used in this paper is rather simplistic but still faces interesting challenges. We are already working on the development of a bidimensional stability analysis starting from a kinetic level and we expect to obtain new and interesting results on this subject.

Acknowledgements This research was financed by Portuguese Funds through FCT, Fundação para a Ciência e a Tecnologia, within the Project UID/MAT/00013/2013.

References

1. Cercignani, C.: Introduction to Rarefied Gas Dynamics. Oxford University Press, Oxford (1998)
2. Villani, C.: A Review of mathematical topics in collisional kinetic theory. Handbook of Mathematical Fluid Dynamics, vol. 1. North-Holland, Amsterdam (2002)
3. Cercignani, C.: Theory and Application of the Boltzmann Equation. Scottish Academic Press, Edinburg (1975)
4. Cercignani, C.: The Boltzmann Equation and Its Applications. Springer, Berlin (1988)
5. Kac, M.: Foundations of Kinetic Theory, pp. 171–197. University of California Press, Berkeley (1956)
6. Płatkowski, T., Illner, R.: Discrete velocity models of the Boltzmann equation: a survey on the mathematical aspects of the theory. SIAM Rev. **30**(2), 213–255 (1988)
7. Polewczak, J., Soares, A.J.: Kinetic theory of simple reacting spheres I. In: 27th International Symposium on Rarefied Gas Dynamics. AIP Conference Proceedings, vol. 1333, pp. 117–122 (2011)
8. Kremer, G., Soares, A.J.: Effect of reaction heat on Maxwellian distribution functions and rate of reactions. J. Stat. Mech. **P12003**, 1–16 (2007)
9. Groppi, M., Spiga, G.: Kinetic theory of a chemically reacting gas with inelastic transitions. Transp. Theory Stat. Phys. **30**, 305–324 (2001)
10. Bardos, C., Golse, F., Levermore, D.: Fluid dynamic limits of kinetic equations I: formal derivations. J. Stat. Phys. **63**, 323–344 (1991)
11. Bardos, C., Golse, F., Levermore, C.D.: Fluid dynamic limits of kinetic equations II: Convergence proofs for the Boltzmann equation. Commun. Pure Appl. Math. **46**, 667–753 (1993)
12. Fickett, W., Davis, W.C.: Detonation, Theory and Experiment. University of California, Berkeley (1979)
13. Fickett, W.: Introduction to Detonation Theory. University of California, Berkeley (1986)
14. Conforto, F., Monaco, R., Schürrer, F., Ziegler, I.: Steady detonation waves via the Boltzmann equation for a reacting mixture. J. Phys. A: Math. Gen. **36**, 5381–5398 (2003)
15. Conforto, F., Groppi, M., Monaco, R., Spiga, G.: Steady detonation problem for slow and fast chemical reaction. Modelling and Numerics of Kinetics Dissipative Systems, pp. 105–117. Nova Science, New York (2006)
16. Carvalho, F. and Soares, A. J.: On the dynamics and linear stability of one-dimensional steady detonation waves. J. Phys. A: Math. Theor. **45**, 255501 (23pp) (2012)
17. Kremer, G.M.: An Introduction to the Boltzmann Equation and Transport Processes in Gases. Springer, Berlin (2010)
18. Erpenbeck, J.J.: Stability of steady-state equilibrium detonations. Phys. Fluids **5**, 604–614 (1962)
19. Erpenbeck, J.J.: Stability of idealized one-reaction detonations. Phys. Fluids **7**, 684–696 (1964)
20. Lee, H.I., Stewart, D.S.: Calculation of linear detonation stability: one dimensional instability of plane detonation. J. Fluid Mech. **216**, 103–132 (1990)
21. Abouseif, G.E., Toong, T.Y.: Theory of unstable one-dimensional detonations. Combust. Flame **45**, 67–94 (1982)
22. Sharpe, G.J.: Linear stability of pathological detonations. J. Fluid Mech. **401**, 311–338 (1999)
23. Bourlioux, A., Majda, A., Roytburd, V.: Theoretical and numerical structure for unstable one-dimensional detonations. SIAM J. Appl. Math **51**, 303–343 (1991)
24. Stewart, D.S., Kasimov, A.: State of detonation stability theory and its application to propulsion. J. Propuls. Power **22**, 1230–1244 (2006)
25. Short, M., Stewart, D.S.: Cellular detonation stability. Part 1. A normal-mode linear analysis. J. Fluid Mech. **368**, 229–262 (1998)
26. Kasimov, A., Stewart, D.S.: Spinning instability of gaseous detonations. J. Fluid Mech. **466**, 179–203 (2002)

27. Gorchkov, V., Kiyanda, C.B., Short, M., Quirk, J.J.: A detonation stability formulation for arbitrary equations of state and multi-step reaction mechanisms. Proc. Combust. Inst. **31**, 2397–2405 (2007)
28. Buckmaster, J.D., Ludford, G.S.S.: The effect of structure on the stability of detonations I. Role of the induction zone. Proceedings of the XX Symposium (International) on Combustion, pp. 1669–1676 (1986)

De Giorgi Techniques Applied to the Hölder Regularity of Solutions to Hamilton–Jacobi Equations

Chi Hin Chan and Alexis Vasseur

Abstract This article is dedicated to the proof of C^α regularization effects of Hamilton–Jacobi equations. The proof is based on the De Giorgi method. The regularization is independent on the regularity of the Hamilton.

Keywords Hamilton–Jacobi equation · Hölder regularity · De Giorgi method

1 Introduction

This article is dedicated to the proof of C^α regularization effects of Hamilton–Jacobi equations of the form:

$$\partial_t u + H(t, x, \nabla u) = 0, \qquad t \in (0, T), \quad x \in \mathbb{R}^N, \tag{1}$$

with $T > 0$, $N \in \mathbb{Z}^+$, and where the Hamiltonian verifies a uniform, in x and t, coercivity property of the form:

$$\frac{1}{\Lambda}|P|^p - \Lambda \leq H(t, x, P) \leq \Lambda |P|^p + \Lambda, \tag{2}$$

C.H. Chan
Department of Applied Mathematics, National Chiao Tung University,
2F, Science Bld. 1, 1001 Ta Hsueh Road, Hsinchu 30010, Taiwan, ROC
e-mail: cchan@math.nctu.edu.tw

A. Vasseur (✉)
Department of Mathematics, University of Texas at Austin, RLM 8.100,
2515 Speedway Stop C1200, Austin, TX 78712, USA
e-mail: vasseur@math.utexas.edu

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_6

for a $p \in (1, \infty)$ and $\Lambda \geq 1$. More precisely, we consider solutions which verify

$$\partial_t u + \Lambda |\nabla u|^p \geq -\Lambda, \qquad \text{in the sense of viscosity solution,} \tag{3}$$

$$\partial_t u + \frac{1}{\Lambda} |\nabla u|^p \leq \Lambda, \qquad \text{in the sense of distribution.} \tag{4}$$

The main theorem, which is the focus of this article, is the following.

Theorem 1 *Let $N \in \mathbb{Z}^+$, $p \in (1, \infty)$, $T > 0$ and $R > 0$ to be given. Let $u \in L^p(0, T; W^{1,p}(B(R)))$ to be a bounded solution on $(0, T) \times B(R)$ to (3) (4). Then, it follows that, for each $\tau \in (0, T)$ and $r \in (0, R)$, we have $u \in C^\alpha([\tau, T) \times B(r))$, where $\alpha \in (0, 1)$, and $\|u\|_{C^\alpha([\tau,T)\times B(r))}$ depend only on N, T, τ, R, r, Λ, p and $\|u\|_{L^\infty((0,T)\times B(R))}$.*

The result of the paper is not new. However the method of proof, based on the De Giorgi method [6] to study the regularity of elliptic equations with rough coefficients, is pretty unusual for the study of viscosity solutions.

Our proof is inspired by previous applications of the De Giorgi method to integral-differential parabolic equations [1, 2].

The first Hölder regularity result of this kind was obtained by Schwab, in [8], in the case of a convex Hamiltonian. The result was a key ingredient to perform the stochastic homogenization of Hamilton–Jacobi equations in Stationary Ergodic Spatio-Temporal media. This result inspired several generalizations. The non-convex case is technically more challenging. The first proofs, relying on stochastic methods, were obtained by Cardaliaguet [4], and Cannarsa, Cardaliaguet [3]. Cardaliaguet and Silvestre provided a simpler proof in [5]. Their proof is based on the construction of sub-solutions and supersolutions combined with improvement of oscillation techniques. It includes applications to some degenerated parabolic equations (and also includes our case).

Hamilton–Jacobi equations have solutions with breakdown of the C^1-regularity in finite time, due to the formation of so-called caustics. It is quite remarkable that a typical Hamilton–Jacobi equation has some regularization effect on its solutions at a lower level C^α, for some $\alpha \in (0, 1)$.

Our proof uses the coercivity of the Hamilton–Jacobi equation to induce a parabolic-like regularization effect. It is based on De Giorgi techniques which provide C^α-regularization for elliptic equations with rough coefficients. It involves the decrease of the oscillation of the solution from scale to scale. While obtaining improved oscillation of the solution from above, we only use that the solution verifies (1) in the sense of distributions (the "viscosity solution" structure, based on the comparison principle, is not used). When we need to shrink the oscillation of the solution by below, the regularization effect is obtained backward in time. In the backward in time regularization process, the "viscosity structure" of the solution is still irrelevant. However, to be consistent with the regularization by above, the backward in time regularization needs to be pushed back in positive time. This is the only part of the proof which needs the comparison principle.

Note that our definition of solution is unusual. The relation between this kind of solutions and the notion of viscosity solutions is not clear, and should be investigated.

Note that $u + \Lambda t$ verifies

$$\partial_t v + \Lambda |\nabla v|^p \geq 0. \tag{5}$$

This inequality is slightly better in the rescaling process. So, without loss of generality, we will assume that

$$\frac{1}{\Lambda}|P|^p - \Lambda \leq H(P) \leq \Lambda |P|^p, \tag{6}$$

instead of (2).

Remark It would help a lot to have, at every scale,

$$\partial_t v + \Lambda |\nabla v|^p \geq \Lambda.$$

But this inequality will not be preserved via the scaling.

The rest of the paper is structured as follows. In Sect. 2, we derive the first lemma of De Giorgi. In Sect. 3, we prove the second lemma of De Giorgi. In Sect. 4 we show how the oscillation can be reduced locally. The precise scaling leading to the C^α regularity is provided in Sect. 5.

2 The First De-Girogi's Lemma

In this paper, by saying that $u : [0, 2] \times B(1) \to \mathbb{R}$ is a weak solution to (4), we actually mean that $u \in L^p(0, 2; W^{1,p}(B(1)))$ and that the following differential inequality holds for any nonnegative test function $\phi \in C_c^\infty((0, 2) \times B(1))$.

$$\int_{[0,2]\times B(1)} \Big(- u \cdot \partial_t \phi + \frac{1}{\Lambda} |\nabla u|^p \cdot \phi \Big) \leq \Lambda \int_{[0,2]\times B(1)} \phi.$$

The goal of this section is to establish the following De-Giorgi's lemma.

Lemma 2 *There exists an absolute constant $\delta = \delta(N, \Lambda, p) > 0$ such that, for any function $u \in L^p(0, 2; W^{1,p}(B(1)))$ which is a weak solution to (4) on $[0, 2] \times B(1)$, we have the following implication:*

If it happens that

$$\int_{[0,2]\times B(1)} u_+ \leq \delta,$$

then it follows that

$$u \leq 1 \quad \text{on } [1, 2] \times B(1).$$

Here, $B(r)$ stands for the open ball centered at the origin O with radius r in $\mathbb{R}^N$.

Proof Let $u : [0, 2] \times B(1) \to \mathbb{R}$ to be a weak solution to (4) on $[0, 2] \times B(1)$ which satisfies all the hypothesis in Lemma 2. First, we work with a sequence of truncated functions v_k on $[0, 2] \times B(1)$ which is defined as follows.

$$v_k = \big(u - (1 - \frac{1}{2^k})\big)_+.$$

Next, we multiply (4) by $\chi_{\{v_k>0\}}$ to yield the following relation

$$\partial_t v_k + \frac{1}{\Lambda}|\nabla v_k|^p \le \Lambda \chi_{\{v_k>0\}}, \tag{7}$$

which holds in the distributional sense on $[0, 2] \times B(1)$, for each integer $k \ge 1$. Note that this can be obtained rigorously thanks to the fact that $v \in L^p(0, 2; W^{1,p}(B(1)))$. First we obtained an in equation for $\phi(v)$, with $y \to \phi(y)$ a nonnegative regular function. Then we pass to the limit for an approximation of the function $y \to (y - C)_+$.

Inequality (7) is the ground on which we will build up a nonlinear recurrence relation for the following sequence of truncated energies.

$$U_k = \sup_{t\in[T_k,2]} \int_{B(1)} v_k(t, \cdot) + \int_{T_k}^{2} \int_{B(1)} |\nabla v_k(t, y)|^p \mathrm{d}y\mathrm{d}t,$$

where $T_k = 1 - \frac{1}{2^k}$, for each $k \in \mathbb{Z}^+$. Next, take any two real numbers σ, t which satisfy the following constraint.

$$T_{k-1} \le \sigma \le T_k \le t \le 2.$$

By taking the spatial-integral over $B(1)$ and then the time-integral over the interval $[\sigma, t]$ for each term in (7), we yield

$$\int_{B(1)} v_k(t, \cdot) + \frac{1}{\Lambda}\int_\sigma^t \int_{B(1)} |\nabla v_k|^p(\tau, \cdot)d\tau \le \int_{B(1)} v_k(\sigma, \cdot) + \Lambda \int_\sigma^t \int_{B(1)} \chi_{\{v_k>0\}} \tag{8}$$

Next, by taking the time average over $\sigma \in [T_{k-1}, T_k]$ for each term in the above inequality, we easily yield

$$\begin{aligned}&\int_{B(1)} v_k(t, \cdot) + \frac{1}{\Lambda}\int_{T_k}^t \int_{B(1)} |\nabla v_k|^p(\tau, \cdot)d\tau\\ \le& 2^k \int_{T_{k-1}}^{T_k} \int_{B(1)} v_k(\sigma, \cdot) + \Lambda \int_{T_{k-1}}^t \int_{B(1)} \chi_{\{v_k>0\}},\end{aligned}$$

from which it follows, through taking the sup over $t \in [T_k, 2]$, that the following relation holds.

$$\sup_{t\in[T_k,2]}\int_{B(1)} v_k(t,\cdot)+\frac{1}{\Lambda}\int_{T_k}^{2}\int_{B(1)}|\nabla v_k|^p \le \int_{[T_{k-1},2]\times B(1)}\left(2^k v_k+\Lambda\chi_{\{v_k>0\}}\right).$$

The above inequality immediately gives

$$U_k \le \left(2^k\Lambda+\Lambda^2\right)\left\{\int_{[T_{k-1},2]\times B(1)} v_k+\int_{[T_{k-1},2]\times B(1)}\chi_{\{v_k>0\}}\right\}. \tag{9}$$

Now, we choose an exponent $r\in(1,\infty)$ such that r and p satisfy exactly one of the following two possible constraints.

- **Constraint I** $1<p<N$, and $p\le r\le \frac{Np}{N-p}$.
- **Constraint II** $r\ge p\ge N$.

Then, it follows that v_k must satisfy the following very well-known Sobolev's inequality (See for instance Lemma II.3.2 of Sect. 2.3 on p. 52 and Exercise II.3.12 on p. 59 in Sect. 2.3 of [7]).

$$\begin{aligned}\left\|v_k(t)\right\|_{L^r(B(1))} &\le C_{N,p,r}\left\|v_k(t)\right\|^{1-\lambda}_{L^p(B(1))}\cdot\left\|v_k(t)\right\|^{\lambda}_{W^{1,p}(B(1))}\\ &\le C_{N,p,r}\left\{(1-\lambda)\left\|v_k(t)\right\|_{L^p(B(1))}+\lambda\left\|v_k(t)\right\|_{W^{1,p}(B(1))}\right\}\\ &= C_{N,p,r}\left\{\left\|v_k(t)\right\|_{L^p(B(1))}+\lambda\left\|\nabla v_k(t)\right\|_{L^p(B(1))}\right\},\end{aligned} \tag{10}$$

where $\lambda=N(\frac{1}{p}-\frac{1}{r})$. Recall that we have the following very standard apriori estimate.

$$\left\|v_k(t)\right\|_{L^p(B(1))}\le C_{N,p}\left\{\left\|\nabla v_k(t)\right\|_{L^p(B(1))}+\int_{B(1)} v_k(t)\right\}. \tag{11}$$

So, it follows from (10) and (11) that the following apriori estimate holds.

$$\left\|v_k(t)\right\|_{L^r(B(1))}\le C_{N,p,r}\left\{\left\|\nabla v_k(t)\right\|_{L^p(B(1))}+\int_{B(1)} v_k(t)\right\},$$

from which it follows that v_k satisfies the following estimate.

$$\begin{aligned}&\int_{T_k}^{2}\left\|v_k(t)\right\|^p_{L^r(B(1))}\mathrm{d}t\\ \le& C_{N,p,r}\cdot\int_{T_k}^{2}\left\{\left\|\nabla v_k(t)\right\|^p_{L^p(B(1))}+\left(\int_{B(1)} v_k(t)\right)^p\right\}\mathrm{d}t\\ \le& C_{N,p,r}\left\{U_k+U_k^p\right\}.\end{aligned} \tag{12}$$

However, by means of interpolation, we yield the following estimate for v_k.

$$\|v_k\|_{L^s([T_k,2]\times B(1))} \le \|v_k\|^{1-\theta}_{L^\infty(T_k,2;L^1(B(1)))} \cdot \|v_k\|^{\theta}_{L^p(T_k,2;L^r(B(1)))}, \tag{13}$$

where the exponent s and the parameter $\theta \in [0, 1]$ are given by

$$\begin{aligned} s &= 1 + p - \frac{p}{r}, \\ \theta &= \frac{pr}{r + pr - p}. \end{aligned} \tag{14}$$

So, by combining (12) and (13), we yield the following estimate.

$$\begin{aligned} \|v_k\|^s_{L^s([T_k,2]\times B(1))} &\le \|v_k\|^{s-p}_{L^\infty(T_k,2;L^1(B(1)))} \cdot \|v_k\|^{p}_{L^p(T_k,2;L^r(B(1)))} \\ &\le U_k^{s-p} \cdot C_{N,p,r} \cdot \left\{U_k + U_k^p\right\} \\ &= C_{N,p,r}\left\{U_k^{2-\frac{p}{r}} + U_k^{1+p-\frac{p}{r}}\right\}. \end{aligned} \tag{15}$$

Now, by applying estimate (15), we can do the raising up of the index as follows.

$$\begin{aligned} &\int_{[T_{k-1},2]\times B(1)} \left(\chi_{\{v_k>0\}} + v_k\right) \\ \le &\int_{[T_{k-1},2]\times B(1)} \left(\chi_{\{v_{k-1}>\frac{1}{2^k}\}} + v_{k-1}\cdot\chi_{\{v_{k-1}>\frac{1}{2^k}\}}\right) \\ \le &\left(2^{sk} + 2^{(s-1)k}\right) \int_{[T_{k-1},2]\times B(1)} v_{k-1}^s \\ \le &(2^s)^k \cdot C_{N,p,r}\left\{U_{k-1}^{2-\frac{p}{r}} + U_{k-1}^{1+p(1-\frac{1}{r})}\right\}. \end{aligned} \tag{16}$$

So, by combining (9) with (16), we yield the following estimate for v_k, with r and p satisfy either **Constraint I** or **Constraint II** and s to be specified in (14).

$$U_k \le \left\{2^k\Lambda + \Lambda^2\right\} \cdot 2^{sk} \cdot C_{N,p,r}\left\{U_{k-1}^{2-\frac{p}{r}} + U_{k-1}^{1+p(1-\frac{1}{r})}\right\}. \tag{17}$$

In the case of $1 < p < N$, we simply take $r = \frac{Np}{N-p}$. Then, estimate (17) reduces down to the following simple estimate.

$$U_k \le \left(D(N, p, \Lambda)\right)^k \left\{U_{k-1}^{1+\frac{p}{N}} + U_{k-1}^{p(1+\frac{1}{N})}\right\}, \tag{18}$$

where $D(N, p, \Lambda)$ is some absolute constant depending only on N, p, Λ.

In the case of $p \ge N$, we simply take $r = 2p$. Then, (17) reduces down to the following one.

$$U_k \le \Big(D(N,p,\Lambda)\Big)^k \cdot \Big\{U_{k-1}^{\frac{3}{2}} + U_{k-1}^{p+\frac{1}{2}}\Big\}. \tag{19}$$

Now, for some technical purpose, we now need to verify the following relation for all $k \ge 1$.

$$U_k \le 2\Lambda \int_{[0,2]\times B(1)} v_k + \Lambda^2 \int_{[0,2]\times B(1)} \chi_{\{v_k>0\}} \tag{20}$$

In order to verify (20) for each $k \ge 1$, we first recall that relation (8) holds for all variables σ, t which satisfy the constraint $0 \le \sigma \le T_k \le t \le 2$. Thanks to the fact that $T_k \ge \frac{1}{2}$ holds for any $k \ge 1$, by taking the average over $\sigma \in [0, T_k]$ on each term of (8), we easily yield the following estimate

$$\int_{B(1)} v_k(t,\cdot) + \frac{1}{\Lambda}\int_{T_k}^{t}\int_{B(1)} |\nabla v_k|^p \le 2\int_{[0,2]\times B(1)} v_k + \Lambda \int_0^t \int_{B(1)} \chi_{\{v_k>0\}},$$

which holds for $t \in [T_k, 2]$. So, by simply taking sup over $t \in [T_k, 2]$ of each term which appears in the above estimate, we immediately obtain (20) for each $k \ge 1$, as desired. Since it is obvious that the following estimate is valid for each $k \ge 1$

$$\int_{[0,2]\times B(1)} \chi_{\{v_k>0\}} \le 2\int_{[0,2]\times B(1)} u_+,$$

it follows directly from (20) that we have the following estimate for each $k \ge 1$

$$U_k \le 2\big(\Lambda + \Lambda^2\big)\int_{[0,2]\times B(1)} u_+. \tag{21}$$

(21) immediately leads to the following assertion

- If it happens that $\int_{[0,2]\times B(1)} u_+ < \frac{1}{2\Lambda(1+\Lambda)}$, then it follows that $U_k < 1$ holds for all $k \ge 1$.

Due to the above assertion, we can now say that as long as u satisfies $\int_{[0,2]\times B(1)} u_+ < \frac{1}{2\Lambda(1+\Lambda)}$, it follows from either (18) or (19) that the following nonlinear recurrence relation holds for each $k \ge 1$.

$$U_k \le \Big(D(N,p,\Lambda)\Big)^k U_{k-1}^{\beta}, \tag{22}$$

where $\beta = 1 + \frac{p}{N}$ for the case of $1 < p < N$, and that $\beta = \frac{3}{2}$ for the case of $p \ge N$.

In light of (22), it is time to recall the following well-known assertion of the De-Giorig's method.

- **Assertion I** Let $D(N,p,\Lambda) > 0$ and $\beta > 1$ to be the two absolute constants which appear in (22). Then, there exists some $\varepsilon_0 \in (0,1)$, which depends only on $D(N,p,\Lambda) > 0$ and $\beta > 1$, such that for any sequence $\{a_k\}_{k=1}^{\infty}$ of nonnegative

numbers for which $a_1 \le \varepsilon_0$ holds and for which the relation $a_k \le D(N, p, \lambda) a_{k-1}^{\beta}$ holds for all $k \ge 1$, it follows that $\lim_{k\to\infty} a_k = 0$.

In accordance with the above assertion, we now take

$$\delta = \frac{\varepsilon_0}{2\Lambda(1+\Lambda)}.$$

Then, whenever $u : [0, 2] \times B(1) \to \mathbb{R}$ is a solution to (4) which satisfies $\int_{[0,2]\times B(1)} u_+ < \delta$, the associated sequence U_k of truncated energies must satisfy both $U_1 < \varepsilon_0$ and (22) for each $k \ge 1$, and hence it follows from **Assertion I** that $\lim_{k\to\infty} U_k = 0$. This immediately lead to the following conclusion:

- If $u : [0, 2] \times B(1) \to \mathbb{R}$ is a solution to (4) which satisfies $\int_{[0,2]\times B(1)} u_+ < \delta$, it follows that $u_+ \le 1$ holds on $[1, 2] \times B(1)$.

In other words, the proof of Lemma 2 is now completed.

3 The Second De-Giorgi's Lemma

We want to show now the following lemma.

Lemma 3 *Let $N \in \mathbb{Z}^+$, and $p \in (1, \infty)$ to be given. Then there exists some absolute constant $\alpha = \alpha(N, \Lambda, p) > 0$, such that, for any function $u \in L^p(-2, 2; W^{1,p}(B(1)))$ which is a weak solution (in the sense of distribution) to*

$$\partial_t u + \frac{1}{\Lambda}|\nabla u|^p \le \Lambda, \qquad \text{on } [-2, 2] \times B(1),$$

we have the following implication:
If it happens that $u \le 2$ holds on $[-2, 2] \times B(1)$, and that u satisfies the following two properties

$$\Big|\{(t, x) \in [-2, 2] \times B(1) : u(t, x) \le 0\}\Big| \ge \frac{|[-2, 2] \times B(1)|}{2}, \tag{23}$$

$$\Big|\{(t, x) \in [-2, 2] \times B(1) : 0 < u(t, x) < 1\}\Big| \le \alpha, \tag{24}$$

then it follows that

$$\int_{[0,2]\times B(1)} [u - 1]_+ < \frac{\delta}{2},$$

where $\delta = \delta(N, \lambda, p) > 0$ is the absolute constant whose existence is asserted in Lemma 2.

Proof We divide the proof in several parts.
Step 1. For any u verifying (23), we have the following relation for any $s, t \in [-2, 2]$ with $s < t$.

$$\int_{B(1)} u_+(t,x)\,dx \le \int_{B(1)} u_+(s,x)\,dx - \frac{1}{\Lambda}\int_s^t \int_{B(1)} |\nabla u_+|^p\,dx\,ds + (t-s)\Lambda|B(1)|.$$

By taking $t = 2, s = -2$ in the above estimate, we obtain the following estimate.

$$\int_{[-2,2]\times B(1)} |\nabla u_+|^p\,dx\,ds \le C(\Lambda). \tag{25}$$

Note that the following estimate holds

$$\partial_t u_+ \le \Lambda.$$

So obviously, we have

$$\|(\partial_t u_+)_+\|_{\mathcal{M}} \le \Lambda|B(1)|,$$

where $\|\cdot\|_{\mathcal{M}}$ stands for norm of measures in $[-2, 2] \times B(1)$. But

$$\begin{aligned}
&\int_{[-2,2]\times B(1)} |(\partial_t u_+)_-|\,dx\,dt \\
\le& -\int_{[-2,2]\times B(1)} \partial_t u_+\,dx\,dt + \int_{[-2,2]\times B(1)} (\partial_t u_+)_+\,dx\,dt \\
\le& 4|B(1)| + 4\Lambda|B(1)|.
\end{aligned} \tag{26}$$

Hence

$$\|\partial_t u\|_{\mathcal{M}} \le C(\Lambda). \tag{27}$$

Step 2. Assume that Lemma 3 is wrong. Then there exists a sequence of functions $\{u_k\}_{k=1}^{\infty}$ on $[-2, 2] \times B(1)$, with each u_k verifies (25), (27) and the following three estimates.

$$\begin{aligned}
\int_{[0,2]\times B(1)} [u_k - 1]_+ &\ge \frac{\delta}{2}, \\
\left|\{(t,x) \in [-2,2]\times B(1) : u_k(t,x) \le 0\}\right| &\ge \frac{|[-2,2]\times B(1)|}{2}, \\
\left|\{(t,x) \in [-2,2]\times B(1) : 0 < u_k(t,x) < 1\}\right| &\le \frac{1}{k}.
\end{aligned} \tag{28}$$

From **Step 1**, there exists a subsequence $\{u_{k_n}\}_{n=1}^{\infty}$ such that u_{k_n} converges to $\bar{u}$ in $L^1([-2, 2] \times B(1))$, where the limiting function $\bar{u}$ still verifies the following properties.

$$\bar{u} \le 2 \quad \text{in } [-2,2] \times B(1),$$
$$\int_{B(1)} \bar{u}_+(t,x)\,dx \le \int_{B(1)} \bar{u}_+(s,x)\,dx + (t-s)\Lambda|B(1)|, \qquad -2 \le s \le t \le 2,$$
$$\int_{[-2,2]\times B(1)} |\nabla \bar{u}_+|^p\,dx\,ds \le C(\Lambda),$$
$$\int_{[0,2]\times B(1)} [\bar{u}-1]_+ \ge \frac{\delta}{2}.$$

Observe that, the following estimate holds for any $\varepsilon > 0$,

$$\big|\{|u_{k_n} - \bar{u}| \ge \varepsilon\}\big| \le \frac{1}{\varepsilon} \int_{[-2,2]\times B(1)} |u_{k_n} - \bar{u}|,$$

So, it follows that for each fixed $\varepsilon > 0$, the term which appears in the left hand side of the above estimate converges to 0 when k_n goes to infinity. Now we have the following obvious estimate.

$$\begin{aligned}\big|\{\bar{u} \le \varepsilon\}\big| &\ge \big|\{u_{k_n} \le 0\}\big| - \big|\{|u_{k_n} - \bar{u}| \ge \varepsilon\}\big| \\ &\ge \frac{\big|[-2,2]\times B(1)\big|}{2} - \big|\{|u_{k_n} - \bar{u}| \ge \varepsilon\}\big|. \end{aligned} \tag{29}$$

Through passing into the limit on the above estimate as k_n goes to infinity, we get

$$\big|\{\bar{u} \le \varepsilon\}\big| \ge \frac{\big|[-2,2]\times B(1)\big|}{2},$$

which is true for any $\varepsilon > 0$. So, by passing to the limit on the above estimate by letting $\varepsilon \to 0^+$, it follows that we have

$$\big|\{\bar{u} \le 0\}\big| \ge \frac{\big|[-2,2]\times B(1)\big|}{2}.$$

In the same way, we have,

$$\begin{aligned}\big|\{\varepsilon \le \bar{u} \le 1-\varepsilon\}\big| &\le \big|\{0 < u_{k_n} < 1\}\big| + \big|\{|u_{k_n} - \bar{u}| \ge \varepsilon\}\big| \\ &\le \frac{1}{k_n} + \frac{1}{\varepsilon} \int_{[-2,2]\times B(1)} |u_{k_n} - \bar{u}|.\end{aligned}$$

By passing to the limit on the above estimate as $k_n \to \infty$, we can deduce that the following relation holds for every $\varepsilon > 0$

$$\big|\{\varepsilon \le \bar{u} \le 1-\varepsilon\}\big| = 0,$$

from which it follows, through taking $\varepsilon \to 0^+$, that the following relation holds.

$$\left|\{(t,x) \in [-2,2] \times B(1) : 0 < \bar{u}(t,x) < 1\}\right| = 0. \tag{30}$$

Step 3. Now, we observe that, for almost every $t \in [-2,2]$,

$$\int_{B(1)} |\nabla \bar{u}|^p(t,x)\,dx$$

is finite. Also, (30) tells us that the following relation holds for almost every $t \in [-2,2]$

$$\left|\{0 < \bar{u}(t,\cdot) < 1\} \cap B(1)\right| = 0,$$

So, by an application of the isoperimetric lemma of De Girogi (with fixed time t), it follows that, for almost every $t \in [-2,2]$, we have either

$$\begin{aligned} &\bar{u}(t,\cdot) \le 0 \quad \text{in } B(1), \\ &\text{or} \qquad \bar{u}(t,\cdot) \ge 1 \quad \text{in } B(1). \end{aligned}$$

Especially, for almost every $t \in [-2,2]$, we have either

$$\int_{B(1)} \bar{u}_+(t,x)\,dx = 0$$

or

$$\int_{B(1)} \bar{u}_+(t,x)\,dx \ge |B(1)|.$$

Step 4. Since

$$|\{\bar{u} \le 0\}| \ge \frac{\left|[-2,2] \times B(1)\right|}{2},$$

there exists $s \in [-2,0]$ for which the following relation holds

$$\int_{B(1)} \bar{u}_+(s,x)\,dx = 0.$$

Consider $s_0 \in [s,2]$ to be the supremum of all such time $\bar{s} \in [s,2]$ which satisfies the following property

$$\int_{B(1)} \bar{u}_+(\tau,\cdot)\,dx = 0 \qquad \text{for } \tau \in [s,\bar{s}).$$

If it happens that $s_0 < 2$, then for $t \in (s_0, s_0 + \frac{1}{2\Lambda}] \cap [-2,2]$, we have

$$\int_{B(1)} \bar{u}_+(t,x)\,dx \le \Lambda |B(1)|(t-s_0) \le \frac{|B(1)|}{2}.$$

But then the above estimate will lead to the following relation

$$\int_{B(1)} \bar{u}_+(t,x)\,dx = 0,$$

which holds for all $t \in (s_0, s_0 + \frac{1}{2\Lambda}] \cap [-2,2]$. This directly contradicts the definition of s_0. This means that we have no choice but to admit that s_0 must be 2. However, $s_0 = 2$ would mean that $\bar{u} \le 0$ holds on $[s,2] \times B(1)$, which however directly contradicts the fact that we should have

$$\int_{[0,2]\times B(1)} [\bar{u}-1]_+ \ge \frac{\delta}{2}.$$

This ends the proof.

4 Improved Oscillations from Above and Below

Let $\delta = \delta(N,\Lambda,p) > 0$, and $\alpha = \alpha(N,\Lambda,p) > 0$ to be the two absolute constants which are specified in Lemmas 2 and 3 respectively. Now, we consider the integer $K_0 \in \mathbb{Z}^+$ which is defined as follows.

$$K_0 = \left[\frac{\big|[-2,2]\times B(1)\big|}{\alpha} \right] + 1, \tag{31}$$

where the symbol $[x]$ means the largest integer which is less than or equal to x. It is obvious that K_0 is an absolute constant which depends only on N, Λ, and p, since both δ and α do. Now, we consider the following two differential inequalities.

$$\partial_t u + \frac{2^{(K_0+1)(p-1)}}{\Lambda} \big|\nabla u\big|^p \le \frac{\Lambda}{2^{K_0+1}}. \tag{32}$$

$$\partial_t u + 2^{(K_0+1)(p-1)} \Lambda \big|\nabla u\big|^p \ge 0. \tag{33}$$

Now, by applying Lemma 3, and then Lemma 2 successively, we can now obtain the following proposition.

Proposition 1 (Improved oscillation from above) *There exists some absolute constant $\lambda = \lambda(N,\Lambda,p) \in (0,1)$, such that for any weak solution $u \in L^p(-2,2; W^{1,p}(B(1)))$ to (32) on $[-2,2]\times B(1)$, we have the following implication.*

If it happens that $u \leq 2$ holds on $[-2, 2] \times B(1)$ and that u satisfies the following property

$$\Big|\{u(t, x) \leq 0\} \cap ([-2, 2] \times B(1))\Big| \geq \frac{\Big|[-2, 2] \times B(1)\Big|}{2},$$

then it follows that

$$u \leq 2 - \lambda \quad \text{on} \quad [1, 2] \times B(1).$$

Proof Let $u : [-2, 2] \times B(1) \to \mathbb{R}$ to be a weak solution to (32) which satisfies all the hypothesis of Proposition 1. For each integer $k \in \mathbb{Z}^+$ which satisfies $1 \leq k \leq K_0 + 1$, we consider the function $u_k : [-2, 2] \times B(1) \to \mathbb{R}$ which is defined through the following relation in an inductive manner.

$$u_k = 2(u_{k-1} - 1),$$

where the function u_0 is just defined to be u itself (i.e. $u_0 = u$). Then, inductively, it is easy to see that the following identity holds for each $k \in \{1, 2, \ldots K_0 + 1\}$.

$$u_k = 2^k\Big\{u - 2(1 - \frac{1}{2^k})\Big\}.$$

By construction, it is obvious that, for each $1 \leq k \leq K_0 + 1$ the relation $u_k \leq 2$ holds on $[-2, 2] \times B(1)$. Observe that the following relation holds for each $k \in \{1, 2, \ldots, K_0\}$.

$$\begin{aligned} &\Big|\{(t, x) \in [-2, 2] \times B(1) : u_k(t, x) \leq 0\}\Big| \\ \geq &\Big|\{(t, x) \in [-2, 2] \times B(1) : u(t, x) \leq 0\}\Big| \\ \geq &\frac{\Big|[-2, 2] \times B(1)\Big|}{2}. \end{aligned} \tag{34}$$

Inductively, it is apparent that each u_k is a weak solution to the following differential inequality on $[-2, 2] \times B(1)$

$$\partial_t u_k + \frac{2^{(p-1)(K_0+1-k)}}{\Lambda}\Big|\nabla u_k\Big|^p \leq \frac{\Lambda}{2^{K_0+1-k}}. \tag{35}$$

Next, we note that it is *impossible* to have the following relation to be valid for all $k \in \{1, 2, \ldots, K_0\}$

$$\Big|\{(t, x) \in [-2, 2] \times B(1) : 0 < u_k(t, x) < 1\}\Big| > \alpha,$$

since if otherwise, the validity of the above relation for all $1 \leq k \leq K_0$ would lead to

$$\begin{aligned}\big|\{(t,x)\in[-2,2]\times B(1) : u_{K_0}(t,x)\le 0\}\big| &> \frac{\big|[-2,2]\times B(1)\big|}{2} + K_0\alpha \\ &> \frac{3}{2}\big|[-2,2]\times B(1)\big|,\end{aligned}$$

which is absurd. This indicates that there must be some positive integer j_0 which satisfies $1 \le j_0 \le K_0$ for which the following relation holds.

$$\big|\{(t,x)\in[-2,2]\times B(1) : 0 < u_{j_0}(t,x) < 1\}\big| \le \alpha. \tag{36}$$

Since $K_0 - j_0 \ge 0$, it follows that u_{j_0} is also a weak solution to (4) on $[-2,2]\times B(1)$. As a result, (34) and (36) together enable us to apply Lemma 3 directly to u_{j_0} in order to deduce that the following property holds.

$$\int_{[0,2]\times B(1)} (u_{j_0+1})_+ = 2\int_{[0,2]\times B(1)} \big(u_{j_0}-1\big)_+ \le \delta. \tag{37}$$

Since u_{j_0+1} satisfies (4), (37) enables us to apply Lemma 2 directly to u_{j_0+1} to deduce that

$$u_{j_0+1} \le 1 \quad \text{on} \quad [1,2]\times B(1),$$

which gives

$$u \le 2 - \frac{1}{2^{j_0+1}} \le 2 - \frac{1}{2^{K_0+1}} \quad \text{on} \quad [1,2]\times B(1).$$

So, by taking $\lambda = \frac{1}{2^{K_0+1}}$, the proof of Proposition 1 is complete.

Now we want to show the following proposition.

Proposition 2 (Improved oscillation from below) *There exists an absolute constant $\tilde{\lambda} = \tilde{\lambda}(N,\Lambda,p) \in (0,1)$ such that for any function $u \in L^p(-2,2; W^{1,p}(B(1)))$ which is a weak solution to (32) on $[-2,2]\times B(1)$, and which simultaneously is a viscosity solution to (33) on $[-2,2]\times B(1)$, we have the following implication: If u satisfies the following properties*

$$u \ge -2 \quad \text{on} \quad [-2,2]\times B(1). \tag{38}$$

$$\big|\{u(t,x)\ge 0\}\cap([-2,2]\times B(1))\big| \ge \frac{\big|[-2,2]\times B(1)\big|}{2}, \tag{39}$$

then it follows that

$$u \ge -2 + \tilde{\lambda} \quad \text{on} \quad [1,2]\times B(\frac{1}{2}).$$

Proof Let $u : [-2,2]\times B(1) \to \mathbb{R}$ to be a viscosity solution to (33) on $[-2,2]\times B(1)$ which is also a weak solution to (32) on $[-2,2]\times B(1)$ and which satisfies conditions (38) and (39).

The function $v : [-2, 2] \times B(1) \to \mathbb{R}$ defined by

$$v(t, x) = -u(-t, x)$$

is not a viscosity solution to (33) anymore, but it still verifies (32) in the sense of distribution. Condition (39) is equivalent to

$$\big|\{v(t, x) \le 0\} \cap ([-2, 2] \times B(1))\big| \ge \frac{\big|[-2, 2] \times B(1)\big|}{2}. \tag{40}$$

Also, condition (38) gives that $v \le 2$ holds on $[-2, 2] \times B(1)$. Hence, we can directly apply Proposition 1 to deduce that,

$$v \le 2 - \lambda \quad \text{on } [1, 2] \times B(1),$$

which is equivalent to,

$$u \ge -2 + \lambda \quad \text{on } [-2, -1] \times B(1).$$

Here, $\lambda \in (0, 1)$ is the absolute constant which is specified in Proposition 1. With respect to some positive number $\lambda_1 \in (0, \lambda)$ which will be determined later, consider the function $\psi : [-2, 2] \times \mathbb{R}^N \to \mathbb{R}$ which is defined as follows.

$$\psi(t, x) = \inf\left(-2 + \lambda_1; -2 - \frac{\lambda_1}{8}(t + 2) + \left(\frac{\lambda_1}{8\Lambda \cdot 2^{(p-1)(K_0+1)}}\right)^{1/p}\big(1 - |x|\big)\right).$$

The function ψ is a viscosity solution to

$$\partial_t \psi + 2^{(K_0+1)(p-1)}\Lambda|\nabla\psi|^p = 0, \quad \text{on } [-2, 2] \times \mathbb{R}^N.$$

Then, it follows that the relation $\psi(-2, \cdot) \le u(-2, \cdot)$ holds on $B(1)$. It is equally obvious that the following relation also holds for all $(t, x) \in [-2, 2] \times \partial B(1)$.

$$\psi(t, x) = -2 - \frac{\lambda_1}{8}(t + 2) \le -2.$$

This tells us that the relation $\psi \le u$ holds point-wisely on the parabolic boundary $\{-2\} \times B(1) \cup [-2, 2] \times \partial B(1)$ of the cube $[-2, 2] \times B(1)$. This allows us to employ basic comparison principle in the theory of Hamilton–Jacobi equations to deduce that the following relation holds for all $(t, x) \in [-2, 2] \times B(1)$.

$$u(t, x) \ge \psi(t, x),$$

from which it follows that the following relation holds

$$u \geq \inf\left(-2+\lambda_1; -2-\frac{\lambda_1}{2}+\frac{1}{2}\left(\frac{\lambda_1}{8\Lambda\cdot 2^{(p-1)(K_0+1)}}\right)^{1/p}\right) \quad \text{on } [1,2]\times B\left(\frac{1}{2}\right).$$

In the case of $p \in (1,\infty)$, since

$$\lim_{\lambda_1\to 0^+}\left\{\left(\frac{\lambda_1}{8\Lambda\cdot 2^{(p-1)(K_0+1)}}\right)^{1/p}\cdot\frac{1}{\lambda_1}\right\} = +\infty,$$

it follows that we may choose $\lambda_1 \in (0,\lambda)$ to be sufficiently small so that the following relation holds

$$2\lambda_1 < \left(\frac{\lambda_1}{8\Lambda\cdot 2^{(p-1)(K_0+1)}}\right)^{1/p}. \tag{41}$$

With respect to such a $\lambda_1 \in (0,\lambda)$ satisfying (41), we deduce that the following improved oscillation holds, provided u satisfies (38) and (39).

$$u \geq \inf\left(-2+\lambda_1; -2+\frac{\lambda_1}{2}\right) = -2+\frac{\lambda_1}{2} \quad \text{on } [1,2]\times B\left(\frac{1}{2}\right).$$

So, by taking $\tilde{\lambda} = \frac{\lambda_1}{2}$, the proof of proposition 2 is now completed.

5 Final Rescaling

Indeed, by means of a simple re-scaling argument, it is easy to see that Theorem 1 will follow from the following proposition, which we are going to prove in this final section.

Proposition 3 *There exists absolute constants $\varepsilon_1 \in (0,1)$, $\tilde{\lambda}\in(0,1)$ and $\alpha_1 \in (1,p)$, depending only on N, p, Λ, such that for any function $u : L^p(-4,0; W^{1,p}(B(1)))$ which is a weak solution to (4) on $[-4,0]\times B(1)$, and which simultaneously is a viscosity solution to (5) on $[-4,0]\times B(1)$, we have the following implication: If it happens that $|u|\leq 2$ holds on $[-4,0]\times B(1)$, then it follows that u satisfies the following relation for all $m \in \mathbb{Z}^+$.*

$$\operatorname{osc}_{Q_m} u \leq 4\left(\frac{4-\tilde{\lambda}}{4}\right)^m,$$

where $Q_m = [-\varepsilon_1^{m\alpha_1}, 0]\times B(\frac{\varepsilon_1^m}{2})$.

Proof The proof of Proposition 3 will be carried out through several steps as follows.

Step 1: Initial Re-scaling

Let $u : [-4, 0] \times B(1) \to \mathbb{R}$ to be a function which is a weak solution to (4) on $[-4, 0] \times B(1)$, and which simultaneously is a viscosity solution to (5) on $[-4, 0] \times B(1)$. We will assume, without the loss of generality, that $|u| \leq 2$ holds on $[-4, 0] \times B(1)$. In order to use either improved oscillation from above or below by means of Propositions 1 or 2, we need to re-scale our function u so that the re-scaled function will satisfies (32) in the weak sense and (33) in the viscosity sense. Moreover, with respect to some parameters $\varepsilon \in (0, 1)$, and $\alpha \in [1, p)$ which will be determined later, we consider the function $u_1 : [\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon}) \to \mathbb{R}$ defined as follows (This index $\alpha \in [1, p)$ *absolutely has nothing to do* with the previous absolute constant $\alpha(N, \Lambda, p) > 0$ which appears in Lemma 3).

$$u_1(t, x) = u(\varepsilon^\alpha t, \varepsilon x).$$

Notice that u_1 is then a weak solution to the following inequality on $[\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon})$

$$\partial_t u_1 + \frac{1}{\Lambda \varepsilon^{p-\alpha}} \left|\nabla u_1\right|^p \leq \Lambda \varepsilon^\alpha. \tag{42}$$

At the same time u_1 is also a viscosity solution to the following inequality on $[\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon})$.

$$\partial_t u_1 + \frac{\Lambda}{\varepsilon^{p-\alpha}} \left|\nabla u_1\right|^p \geq 0. \tag{43}$$

Now, we simply notice that

$$\left\{ \frac{1}{\varepsilon^{p-\alpha}} : \varepsilon \in (0, 1), \alpha \in [1, p) \right\} = (1, +\infty),$$

which immediately ensures that we can find a suitable pair (ε, α) with $\varepsilon \in (0, 1)$ and $\alpha \in [1, p)$ for which the following required property holds

$$\frac{1}{\varepsilon^{p-\alpha}} = 2^{(K_0+1)(p-1)}. \tag{44}$$

So, with respect to such a pair of (ε, p) satisfying (44), since $\frac{p-1}{p-\alpha} \geq 1$, we clearly have

$$\varepsilon^\alpha = \frac{1}{2^{(K_0+1)(\frac{p-1}{p-\alpha})\alpha}} \leq \frac{1}{2^{K_0+1}}.$$

Hence, it follows from (42) and (43) that u_1 is a weak solution to (32) on $[\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon})$, which simultaneously is also a viscosity solution to (33) on $[\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon})$. Since we still have $|u_1| \leq 2$ on $[\frac{-4}{\varepsilon^\alpha}, 0] \times B(\frac{1}{\varepsilon})$, we can directly apply either Propositions 1 or 2 to yield either $u_1 \leq 2 - \tilde{\lambda}$ on $[-1, 0] \times B(\frac{1}{2})$ or $u_1 \geq -2 + \tilde{\lambda}$ on $[-1, 0] \times B(\frac{1}{2})$.

In either case, we obtain the following improved oscillation

$$\operatorname{osc}_{[-1,0]\times B(\frac{1}{2})} u_1 \leq 4 - \tilde{\lambda}. \tag{45}$$

Step 2: Second Re-scaling and Improved Oscillation at the Scale of ε_1

The situation looks so far so good. However, we have to be more careful in carrying out the second re-scaling. First, due to the above improved oscillation of u_1 on $[-1, 0] \times B(\frac{1}{2})$, we can find some $d_1 \in \mathbb{R}$ with $|d_1| \leq \frac{\tilde{\lambda}}{2}$ such that the following relation $|u_1 - d_1| \leq 2 - \frac{\tilde{\lambda}}{2}$ holds on $[-1, 0] \times B(\frac{1}{2})$. Just as before, with respect to some $\varepsilon_1 \in (0, 1)$ and $\alpha_1 \in [1, p)$ which have to be determined later, we need to consider the re-scaled function u_2 defined as follows

$$u_2(t, x) = \frac{4}{4 - \tilde{\lambda}} \left\{ u_1(\varepsilon_1^{\alpha_1} t, \varepsilon_1 x) - d_1 \right\}. \tag{46}$$

Observe that u_2 is a weak solution to the following inequality on $\left[\frac{-4}{\varepsilon^\alpha \varepsilon_1^{\alpha_1}}, 0\right] \times B(\frac{1}{\varepsilon\varepsilon_1})$.

$$\partial_t u_2 + \frac{2^{(K_0+1)(p-1)}}{\Lambda \varepsilon_1^{p-\alpha_1}} \left(\frac{4 - \tilde{\lambda}}{4}\right)^{p-1} \left|\nabla u_2\right|^p \leq \frac{\Lambda}{2^{K_0+1}} \epsilon_1^{\alpha_1} \left(\frac{4}{4 - \tilde{\lambda}}\right). \tag{47}$$

Also, u_2 is a viscosity solution to the following inequality on $[\frac{-4}{\varepsilon^\alpha \varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{\varepsilon\varepsilon_1})$.

$$\partial_t u_2 + \Lambda 2^{(K_0+1)(p-1)} \frac{1}{\varepsilon_1^{p-\alpha_1}} \left(\frac{4 - \tilde{\lambda}}{4}\right)^{p-1} \left|\nabla u_2\right|^p \geq 0. \tag{48}$$

In light of the structure of (48), it is natural to take ε_1 as follows, with some suitable $\alpha_1 \in (1, p)$ which will be determined later.

$$\varepsilon_1 = \left(\frac{4 - \tilde{\lambda}}{4}\right)^{\frac{p-1}{p-\alpha_1}} \tag{49}$$

With this choice of ϵ_1 as specified in (49), we surely have

$$\frac{1}{\varepsilon_1^{p-\alpha}} \left(\frac{4 - \tilde{\lambda}}{4}\right)^{p-1} = 1.$$

and that

$$\epsilon_1^{\alpha_1} \left(\frac{4}{4 - \tilde{\lambda}}\right) = \left(\frac{4 - \tilde{\lambda}}{4}\right)^{\frac{p(\alpha_1-1)}{p-\alpha_1}} < 1,$$

since $\alpha_1 \in (1, p)$ ensures that $\frac{p(\alpha_1-1)}{p-\alpha_1} > 0$. These observations tell us that, as long as ϵ_1 is given by (49), it follows from (47) and (48) that u_2 is still a weak solution to (32) on $[\frac{-4}{\varepsilon^{\alpha}\varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{\varepsilon\varepsilon_1})$, which simultaneously is also a viscosity solution to (33) on $[\frac{-4}{\varepsilon^{\alpha}\varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{\varepsilon\varepsilon_1})$.

Due to construction, u_2 must satisfies $|u_2| \le 2$ on $[-\frac{1}{\varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{2\varepsilon_1})$. The key point is that we have the freedom to choose $\alpha_1 \in (1, p)$ to be as close to $p > 1$ as possible. Indeed, we observe that

$$\lim_{\alpha_1 \to p^-} \left(\frac{4}{4-\tilde{\lambda}}\right)^{\frac{p-1}{p-\alpha_1}} = +\infty,$$

which allows us to choose some suitable $\alpha_1 \in (1, p)$ to be sufficiently close to p, in a manner which depends only on p, q and $\tilde{\lambda}$ (and hence only on N, p, and λ) such that the following two relations must hold simultaneously.

$$\begin{aligned} &\frac{1}{\varepsilon_1^{\alpha_1}} = \left(\frac{4}{4-\tilde{\lambda}}\right)^{\frac{p-1}{p-\alpha_1}\alpha_1} > 4. \\ &\frac{1}{2\varepsilon_1} > 1. \end{aligned} \tag{50}$$

(50) is here to ensure that $[-4, 0] \times B(1) \subset [-\varepsilon_1^{-\alpha_1}, 0] \times B(\frac{1}{2\varepsilon_1})$ holds.

As a result, with respect to such a $\alpha_1 = \alpha_1(N, p, \Lambda) \in (0, p)$, and ε_1 to be given in (49), we eventually have $|u_2| \le 2$ on $[-4, 0] \times B(1)$. So, we may apply either Propositions 1 or 2 to deduce that we can find some suitable $d_2 \in \mathbb{R}$ with $|d_2| \le \frac{\tilde{\lambda}}{2}$ for which we have

$$\left|u_2 - d_2\right| \le 2 - \frac{\tilde{\lambda}}{2} \quad \text{on} \quad [-1, 0] \times B\left(\frac{1}{2}\right).$$

Step 3: Successive Re-scalings and Improved Oscillations at Finer and Finer Scales

Let $\alpha_1 \in (1, p)$ to be the same absolute constant in **Step 2** which satisfies (50), and let $\varepsilon_1 \in (0, 1)$ to be the one specified in (49). Inductively, suppose that we have done re-scalings on the original solution $u : [-4, 0] \times B(1) \to \mathbb{R}$ for $m-1$-times; so that we have found a list of numbers $d_1, d_2, d_3, \ldots d_{m-1} \in \mathbb{R}$, each of them satisfies $|d_j| \le \frac{\tilde{\lambda}}{2}$, in such a way that the associated list of re-scaled functions u_j as determined by the following recurrence relation (for $2 \le k \le m$)

$$u_k(t, x) = \frac{4}{4-\tilde{\lambda}}\left\{u_{k-1}(\varepsilon_1^{\alpha_1} t, \epsilon_1 x) - d_{k-1}\right\} \tag{51}$$

satisfy the following properties.

- For each $1 \le j \le m$, u_j is a weak solution to (32) on $[-4, 0] \times B(1)$, and simultaneously a viscosity solution to (33) on $[-4, 0] \times B(1)$.
- For each $1 \le j \le m$, the relation $|u_j(t, x)| \le 2$ holds for all $(t, x) \in [-4, 0] \times B(1)$.

Since u_m satisfies the above two properties, we apply either propositions 1 or 2 to deduce that either $u_m \le 2 - \tilde{\lambda}$ or else $u_m \ge -2 + \tilde{\lambda}$ must hold on $[-1, 0] \times B(\frac{1}{2})$. In other words, we can find some suitable $d_m \in \mathbb{R}$ with $|d_m| \le \frac{\tilde{\lambda}}{2}$ for which we have

$$\big|u_m - d_m\big| \le 2 - \frac{\tilde{\lambda}}{2} \quad \text{on} \quad [-1, 0] \times B\big(\frac{1}{2}\big).$$

Now, consider the re-scaled function $u_{m+1} : \big[-\frac{4}{\varepsilon_1^{\alpha_1}}, 0\big] \times B\big(\frac{1}{\varepsilon_1}\big) \to \mathbb{R}$ defined as follows

$$u_{m+1}(t, x) = \frac{4}{4 - \tilde{\lambda}} \Big\{ u_m(\varepsilon_1^{\alpha_1} t, \epsilon_1 x) - d_m \Big\}.$$

u_{m+1} is then a weak solution to (32) on $[-\frac{4}{\varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{\varepsilon_1})$, and simultaneously a viscosity solution to (33) on $[-\frac{4}{\varepsilon_1^{\alpha_1}}, 0] \times B(\frac{1}{\varepsilon_1})$. By construction, we also have

$$\big|u_{m+1}\big| \le 2 \quad \text{on} \quad [-\frac{1}{\varepsilon_1^{\alpha_1}}, 0] \times B\big(\frac{1}{2\varepsilon_1}\big),$$

Since $[-4, 0] \times B(1) \subset [-\varepsilon_1^{-\alpha_1}, 0] \times B\big(\frac{1}{2\varepsilon_1}\big)$, the above estimate ensures that $\big|u_{m+1}\big| \le 2$ holds on $[-4, 0] \times B(1)$. So, we may apply either Propositions 1 or 2 to deduce that we can find some suitable $d_{m+1} \in \mathbb{R}$ with $|d_{m+1}| \le \frac{\tilde{\lambda}}{2}$ for which we have

$$\big|u_{m+1} - d_{m+1}\big| \le 2 - \frac{\tilde{\lambda}}{2} \quad \text{on} \quad [-1, 0] \times B\big(\frac{1}{2}\big).$$

So, inductively, we are able to construct a sequence of numbers $\{d_m\}_{m=1}^{\infty}$ with $|d_m| \le \frac{\tilde{\lambda}}{2}$ such that for the associated sequence of successively re-scaled functions $\{u_m\}_{m=1}^{\infty}$ as defined in (51), we have the following property.

- For each $m \ge 2$, u_m is a weak solution to (32) on $[-4, 0] \times B(1)$, and simultaneously a viscosity solution to (33) on $[-4, 0] \times B(1)$. Moreover, the relation $|u_m(t, x)| \le 2$ holds on $[-4, 0] \times B(1)$,

which allows us to use either Propositions 1 or 2 to obtain the following improved oscillations at finer and finer scales

$$\big|u_m - d_m\big| \le 2 - \frac{\tilde{\lambda}}{2} \quad \text{on} \quad [-1, 0] \times B\big(\frac{1}{2}\big), \tag{52}$$

which holds for every $m \in \mathbb{Z}^+$. Then, it follows directly from (52) that the following property holds for all $m \geq 0$

$$\operatorname{osc}_{Q_m} u_1 \leq 4\left(\frac{4-\tilde{\lambda}}{4}\right)^{m+1},$$

in which $Q_m = [-\varepsilon_1^{m\alpha_1}, 0] \times B\left(\frac{\varepsilon_1^m}{2}\right)$.

Acknowledgements A.F. Vasseur was partially supported by the NSF Grant DMS 1209420. Ch.-H. Chan was partially supported by the grant NSC 103-2115-M-009-010-MY2 from the National Science Council of Taiwan.

References

1. Caffarelli, L., Chan, Ch-H, Vasseur, A.: Regularity theory for parabolic nonlinear integral operators. J. Am. Math. Soc. **24**(3), 849–869 (2011)
2. Caffarelli, L., Vasseur, A.: Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. Ann. Math. (2), **171**(3), 1903–1930 (2010)
3. Cannarsa, P., Cardaliaguet, P.: Hölder estimates in space-time for viscosity solutions of Hamilton-Jacobi equations. Commun. Pure Appl. Math. **63**(5), 590–629 (2010)
4. Cardaliaguet, P.: A note on the regularity of solutions of Hamilton-Jacobi equations with superlinear growth in the gradient variable. ESAIM Control Optim. Calc. Var. **15**(2), 367–376 (2009)
5. Cardaliaguet, P., Silvestre, L.: Hölder continuity to Hamilton-Jacobi equations with superquadratic growth in the gradient and unbounded right-hand side. Commun. Partial Differ. Equ. **37**(9), 1668–1688 (2012)
6. De Giorgi, E.: Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. **3**(3), 25–43 (1957)
7. Galdi, G.P.: An introduction to the mathematical theory of the Navier-Stokes equations. Springer Monographs in Mathematics, 2nd edn. Springer, New York (2011). (Steady-state problems)
8. Schwab, R.: Stochastic homogenization of Hamilton-Jacobi equations in stationary ergodic spatio-temporal media. Indiana Univ. Math. J. **58**(2), 537–581 (2009)

On the Motion of Chemically Reacting Fluids Through Porous Medium

Eduard Feireisl, Jiří Mikyška, Hana Petzeltová and Peter Takáč

Abstract We consider a parabolic-hyperbolic system of nonlinear partial differential equations modeling the motion of a chemically reacting mixture through porous medium. The existence of classical as well as weak solutions is established under several physically relevant choices of the constitutive equations and relevant boundary conditions.

Keywords Chemically reacting fluid · Porous medium · DiPerna, Lions theory

1 Introduction

A simple model of the motion of a mixture of n chemically reacting fluids takes the form (see e.g. Giovangigli [10, Chaps. 2,3]):

$$\partial_t(\rho^i) + \mathrm{div}_x(\rho^i \mathbf{v}) + \mathrm{div}_x \mathscr{F}^i = m_i \omega^i, \tag{1}$$

E. Feireisl (✉) · H. Petzeltová
Institute of Mathematics of the Academy of Sciences of the Czech Republic,
Žitná 25, 115 67 Prague 1, Czech Republic
e-mail: feireisl@math.cas.cz

H. Petzeltová
e-mail: petzelt@math.cas.cz

J. Mikyška
Czech Technical University in Prague, Faculty of Nuclear Sciences
and Physical Engineering, Department of Mathematics, Trojanova 13,
120 00 Prague 2, Czech Republic
e-mail: jiri.mikyska@fjfi.cvut.cz

P. Takáč
Institut Für Mathematik, Universität Rostock, Ulmenstraße 69, Haus 3,
18051 Rostock, Germany
e-mail: peter.takac@uni-rostock.de

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_7

where ρ^i is the mass density of the i-th species, $\mathbf{v}$ is the fluid bulk velocity of the mixture, m_i the molar mass of the i-th species, $\mathscr{F}^i$ the diffusive fluxes, and ω^i represent the molar production, typically given functions of $(\rho^1, \dots, \rho^n)$ and of the temperature. We also denote

$$\rho = \sum_{i=1}^{n} \rho^i,$$

the total density of the mixture and introduce the mass fractions

$$Y^i = \frac{\rho^i}{\rho},\ i = 1, \dots, n.$$

Obviously,

$$Y^i \geq 0,\ \sum_{i=1}^{n} Y^i = 1. \tag{2}$$

We may sum up (1) to deduce the mass conservation (*equation of continuity*):

$$\partial_t \rho + \mathrm{div}_x(\rho \mathbf{v}) = -\mathrm{div}_x \sum_{i=1}^{n} \mathscr{F}^i + \sum_{i=1}^{n} m_i \omega^i = 0, \tag{3}$$

where the last equality should be viewed as a natural constraint to be imposed on $\mathscr{F}^i$, ω^i enforced by the principle of mass conservation.

The diffusion fluxes are typically given through the empirical Fick's law:

$$\mathscr{F}^i = -d_i \nabla_x Y^i,\ d_i > 0,\ i = 1, \dots, n \tag{4}$$

If the motion takes place in the porous medium environment, we may close the system by imposing the standard hypothesis that the velocity $\mathbf{v}$ is given by the pressure gradient, more specifically

$$\mathbf{v} = -\nabla_x p + \rho \mathbf{g}, \tag{5}$$

where $\mathbf{g}$ represents the gravitational force. For the one component compressible flow, the relation (5) has been rigorously identified as a homogenization limit of the compressible Navier–Stokes system, see Masmoudi [11]. The result has been extended to a more general class of pressure laws and also to the full Navier–Stokes–Fourier system in [8]. The Eq. (1) has been considered in many special situations, namely in connection with the Navier–Stokes system, see, e.g., [9] and references there. Here, the main point consists in considering mixed boundary conditions for a simplified model, and an application of the DiPerna, Lions theory to the general problem.

1.1 A Parabolic-Hyperbolic System

We consider the problem of solvability of the system described above for $(\rho^i, \mathbf{v})$, or alternatively, for (Y^i, p), in the set $\{(t, x) \in [0, \tau] \times \Omega\}$, where $\Omega \subset R^3$ is a bounded smooth domain, under the following simplifying assumptions:

- The diffusion coefficients d_i vanish for all $i = 1, 2, \ldots$.
- The process is isothermal, the temperature $T > 0$ is constant.
- The effect of the gravitational force is neglected, $\mathbf{g} = 0$.
- The production rates $\omega^i = \omega^i(\rho^1, \ldots, \rho^n)$ are given smooth functions of species densities. More specifically,

$$\omega^i = \mathscr{C}_i - \rho^i \mathscr{D}_i, \tag{6}$$

where $\mathscr{C}_i \geq 0$, $\mathscr{D}_i \geq 0$,

$$\mathscr{C}_i = \sum_{j=1}^{m} \left[\nu^b_{i,j} K^f_j(T) \Pi^n_{l=1} \left(\frac{\rho^l}{m_l} \right)^{\nu^f_{l,j}} + \nu^f_{i,j} K^b_j(T) \Pi^n_{l=1} \left(\frac{\rho^l}{m_l} \right)^{\nu^b_{l,j}} \right], \tag{7}$$

$$\mathscr{D}_i = \frac{1}{m_i} \left[\sum_{j=1, \nu^f_{i,j} \geq 1}^{m} \nu^f_{i,j} K^f_j(T) \left(\frac{\rho^i}{m_i} \right)^{\nu^f_{i,j}-1} \Pi^n_{l=1, l\neq i} \left(\frac{\rho^l}{m_l} \right)^{\nu^f_{l,j}} \right. \tag{8}$$

$$\left. + \sum_{j=1, \nu^b_{i,j} \geq 1}^{m} \nu^b_{i,j} K^b_j(T) \left(\frac{\rho^i}{m_i} \right)^{\nu^b_{i,j}-1} \Pi^n_{l=1, l\neq i} \left(\frac{\rho^l}{m_l} \right)^{\nu^b_{l,j}} \right],$$

where m is the number of chemical reactions, K^f_j, K^b_j are positive functions of the temperature, and $\nu^f_{i,j}$, $\nu^b_{i,j}$ are non-negative integers (stoichiometric coefficients), see [10, Sect. 6.4.6].
- The pressure of the mixture is given by the perfect gas law,

$$p = \sum_{i=1}^{n} \frac{1}{m_i} \rho^i RT, \tag{9}$$

where R is the perfect gas constant.

Remark 1 It is interesting to note that (9) with *equal* molar masses $m_i = m$ is the only choice of the pressure compatible with the *Second Law of Thermodynamics* as soon as Fick's law is imposed, cf. [9].

Our goal in the present paper is to discuss the solvability and a proper choice of boundary conditions for system (1) under the simplifying conditions stated above.

In Sect. 2, we study the case when the pressure p satisfies a parabolic equation of porous medium type independent of the species densities ρ^i. The standard parabolic theory yields a regular pressure p that can be subsequently substituted in (1) to determine uniquely ρ^i, $i = 1, \dots, n$ by the method of characteristics. Relevant boundary conditions are easy to discuss in this context.

In Sect. 3, we address the general situation when all equations in (1) are strongly coupled. The resulting system is of mixed parabolic-hyperbolic type. We derive a priori bounds and show weak sequential stability of the family of solutions. To this end, a variant of DiPerna, Lions [7] theory for the transport equation is used.

2 The Case of "Independent" Pressure

We start with the simple situation of equal molar masses $m_i = m > 0$ for all $i = 1, \dots, n$. In accordance with (9) and (3), (5), we may sum up the Eq. (1) to obtain

$$\partial_t p - \mathrm{div}_x (p \nabla_x p) = 0. \tag{10}$$

Thus the pressure satisfies a parabolic type differential equation that may be solved separately and independently of the other quantities. Note that the same situation occurs in the absence of chemical reactions, meaning $\omega_i = 0$ for all $i = 1, \dots, n$. Most generally, we have (10) whenever

$$\sum_{j \in S_j} \omega_j = 0,\ m_j = m_{S_j} > 0 \text{ for all } j \in S_j,\ S_i \cap S_j = \emptyset \text{ if } i \neq j,\ \cup_j S_j = \{1, \dots, n\}. \tag{11}$$

2.1 Boundary Value Problem for the Pressure Equation

Equation (10) represents the standard porous medium equation studied frequently in the literature, see e.g. Di Benedetto [6]. Here, in addition, we avoid the "vacuum" problem by imposing positive initial and boundary conditions on p.

2.1.1 Mixed Neumann–Dirichlet Boundary Conditions

We suppose the boundary $\partial\Omega$ can be decomposed as

$$\partial\Omega = \Gamma_D \cup \Gamma_N,\ \Gamma_D,\ \Gamma_N \text{ smooth and compact with } \Gamma_D \cap \Gamma_N = \emptyset\,. \tag{12}$$

We impose the (non-homogeneous) Dirichlet boundary condition

$$p = p_b \text{ on } \Gamma_D, \quad p_b \text{ is a positive constant,} \tag{13}$$

together with the (homogenous) Neumann boundary condition

$$\nabla_x p \cdot \mathbf{n} = 0 \text{ on } \Gamma_N . \tag{14}$$

As usual, $\mathbf{n}$ denotes the *outer* unit normal vector to the boundary $\partial\Omega$ of Ω.

Remark 2 This choice of boundary conditions corresponds to the presence of a "well" in the container Ω on the boundary of which a (constant) pressure is maintained, with the rest of $\partial\Omega$ being an impermeable wall.

In order to deal with a well-posed problem, we prescribe the initial pressure distribution

$$p(0, x) = p_0(x) \text{ in } \Omega. \tag{15}$$

(i) Consider the situation

$$p_0(x) \geq p_b > 0 \text{ for all } x \in \Omega, \ p_0 \in W^{2,\infty}(\Omega), \ p_0 \not\equiv p_b.$$

By virtue of the standard parabolic theory, problem (10), (12–15) admits a unique solution

$$p(t, x) \geq p_b \text{ for any } (t, x) \in (0, \tau) \times \Omega .$$

Moreover, the solution is smooth in the *open* set $(0, \tau) \times \Omega$ and, by virtue of the strong maximum principle (Hopf's boundary point lemma),

$$\nabla_x p \cdot \mathbf{n} < 0 \text{ on } \Gamma_D . \tag{16}$$

Now, equation (1) reduces to the transport problem

$$\partial_t(\rho^i) - \mathrm{div}_x(\rho^i \nabla_x p) = m_i \omega^i(\rho^1, \dots, \rho^n), \ i = 1, \dots n, \tag{17}$$

with a given (regular) velocity field $\mathbf{v} = -\nabla_x p$. Keeping (14), (16) in mind, the method of characteristics yields that Eq. (17) admits a unique solution for any initial data

$$\rho^i(0, \cdot) = \rho_0^i , \quad \text{in } \Omega \tag{18}$$

satisfying the obvious compatibility condition

$$\sum_{i=1}^{n} \frac{1}{m_i} \rho_0^i RT = p_0 \quad \text{in } \Omega . \tag{19}$$

(ii) Now, we examine the complementary situation

$$0 < p_0(x) \leq p_b \quad \text{for all } x \in \overline{\Omega},\ p_0 \in W^{2,\infty}(\Omega),\ p_0 \not\equiv p_b .$$

It is easy to check, by means of the same arguments as above, that

$$\nabla_x p \cdot \mathbf{n} > 0 \text{ on } \Gamma_D. \tag{20}$$

Consequently, for the transport problem (17), (18) to be uniquely solvable, we have to prescribe the boundary conditions

$$\rho^i|_{\Gamma_D} = \rho_b^i,\ i = i, \ldots, n,$$

with the compatibility condition

$$\sum_{i=1}^{n} \frac{1}{m_i} \rho_b^i RT = p_b.$$

(iii) In general, the sign of the normal component of the velocity $-\nabla_x p \cdot \mathbf{n}$ on Γ_D is determined by the pressure. In particular, the relevant boundary conditions for ρ^i must be prescribed a posteriori, after having solved problem (10), (12–15).

2.2 Other Boundary Conditions

More general boundary conditions can be handled in a similar fashion. One should always keep in mind that the boundary conditions for the species densities ρ_b^i must be determined after having identified the sign of $\nabla_x p \cdot \mathbf{n}$ together with p on $\partial\Omega$.

3 General System

We focus on the general case in which the equations for the pressure and the species densities are coupled. It turns out that it is more convenient to consider p, together with the mass fractions Y^i, as independent variables. Taking into account (3) with $\mathbf{v} = -\nabla_x p$, the resulting system of equations reads:

$$\partial_t p - \mathrm{div}_x(p\nabla_x p) = RT\sum_{i=1}^{n}\omega^i, \tag{21}$$

$$\partial_t Y^i - \nabla_x p\cdot\nabla_x Y^i = \frac{m_i}{\rho}\omega^i,\ i=1,\dots,n. \tag{22}$$

Recalling the pressure-density relation

$$\rho = p\left(\sum_{i=1}^{n}\frac{1}{m_i}Y^i RT\right)^{-1},\ \rho^i = Y^i\rho, \tag{23}$$

and using the specific form of ω^i stated in (6–8), we view the right-hand sides of the above equations as functions of p and $Y^1,\dots,Y^n$.

System (21–23) is nonlinear of parabolic-hyperbolic type. To avoid unnecessary technicalities, we impose the homogeneous Neumann boundary conditions for the pressure,

$$\nabla_x p\cdot\mathbf{n}|_{\partial\Omega} = 0\,. \tag{24}$$

Accordingly, only the *initial conditions* for Y^i are necessary to make the problem, at least formally, well-posed.

3.1 A Priori Estimates

We start by deriving suitable a priori estimates for (smooth) solutions of problem (21), (22), (24).

3.1.1 Uniform Bounds on the Pressure

Uniform bounds on the pressure are usually derived by application of some form of the maximum principle. A short inspection of the pressure Eq. (21) and the structure (6) of the functions ω^i reveals that

$$\sum_{i=1}^{n}\omega_i = \sum_{i=1}^{n}\mathscr{C}_i - \rho^i\mathscr{D}_i \overset{<}{\sim} \sum_{j=1}^{n}\left(p^{\sum_{l=1}^{m}\nu_{l,j}^f} + p^{\sum_{l=1}^{m}\nu_{l,j}^b}\right).$$

Consequently, in view of the standard maximum principle estimates, we get a uniform bound

$$0 \le p(t,x) \le \overline{p} \text{ on the time interval } (0,\tau), \tag{25}$$

where $\tau > 0$ depends, in general, on $\|p(0,\cdot)\|_{L^\infty(\Omega)}$. Moreover, the estimate is uniform, meaning extendable to any positive τ if at least one of the following situations occurs:

- $$\sum_{i=1}^{n} \mathscr{C}_i - \rho^i \mathscr{D}_i \stackrel{<}{\sim} (p+1),$$

  for specific examples see [10, Sect. 3.2.3];
- $$\|p(0,\cdot)\|_{L^\infty(\Omega)} \text{ is sufficiently small,}$$

  where "small" means in terms of τ and the structural constants appearing in (7), (8).

Accordingly, in the remaining part of this section, we assume the validity of the bound (25). Note that, in view of the structure of ω^i stated in (6), relation (25) implies that

$$p(t,\cdot) \geq \underline{p} > 0 \text{ for any } t \in (0,\tau) \text{ as soon as } \inf_{x\in\Omega} p(0,x) > 0, \tag{26}$$

where the lower bound $\underline{p}$ may depend on τ.

3.1.2 Maximal Regularity Estimates

In view of (25), (26) we may use the maximal regularity estimates for (non-degenerate) parabolic equations, see Denk, Hieber, and Pruess [4] or Ashyralyev and Sobolevskii [5], to deduce the bounds

$$\partial_t p,\ \nabla_x^l p\,;\ l = 0,1,2\,,\quad \text{bounded in } L^q((0,\tau)\times\Omega) \text{ for any finite } 1 < q < \infty\,. \tag{27}$$

Unfortunately, the bounds (27) are still not sufficient for the transport Eq. (22) to be well-posed. The available DiPerna, Lions theory [7] (see also Ambrosio [2], Crippa and De Lellis [3]) require that, at least,

$$\mathrm{div}_x \nabla_x p = \Delta_x p \in L^1(0,\tau; L^\infty(\Omega))\,. \tag{28}$$

In order to guarantee (28), higher order regularity estimates are needed that will be established in the next section.

3.1.3 Higher Order Regularity

Taking the time derivative of (21) with respect to t and denoting $P = \partial_t p$, we obtain

$$\partial_t P - \mathrm{div}_x(p\nabla_x P) = \mathrm{div}_x(\partial_t p \nabla_x p) + RT\sum_{i=1}^{n}\partial_t \omega^i. \tag{29}$$

To evaluate $\partial_t \omega^i$ we realize that, thanks to (6–8),

$$\omega_i = \sum_{k=1}^{k_i} \rho^k G_{k,i}(\rho^1, \ldots, \rho^n),\ i = 1, \ldots, n,$$

where $G_{k,i}$ are continuously differentiable functions. Using (17) we compute

$$\partial_t\left(\rho^k G_{k,i}(\rho^1, \ldots, \rho^n)\right) = \partial_t \rho^k G_{k,i}(\rho^1, \ldots, \rho^n) + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\partial_t \rho^j$$

$$= \mathrm{div}_x(\rho^k \nabla_x p) G_{k,i}(\rho^1, \ldots, \rho^n) + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\mathrm{div}_x(\rho^j \nabla_x p)$$

$$+ m_i \omega_i G_{k,i}(\rho^1, \ldots, \rho^n) + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j} m_j \omega_j.$$

Furthermore,

$$\mathrm{div}_x(\rho^k \nabla_x p) G_{k,i}(\rho^1, \ldots, \rho^n) + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\mathrm{div}_x(\rho^j \nabla_x p)$$

$$= \mathrm{div}_x\left[\rho^k \nabla_x p G_{k,i}(\rho^1, \ldots, \rho^n)\right] - \rho_k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\nabla_x p \cdot \nabla_x \rho^j$$

$$+ \rho_k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\nabla_x p \cdot \nabla_x \rho^j + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\rho^j \Delta_x p$$

$$= \mathrm{div}_x\left[\rho^k \nabla_x p G_{k,i}(\rho^1, \ldots, \rho^n)\right] + \rho^k \sum_{j=1}^{n} \frac{G_{k,i}(\rho^1, \ldots, \rho^n)}{\partial \rho^j}\rho^j \Delta_x p.$$

Summing up the previous observations and going back to (29) we infer that

$$\partial_t P - \mathrm{div}_x(p\nabla_x P) = \mathrm{div}_x(\mathbf{F}) + G\,,$$

with

$$\mathbf{F},\ G \quad \text{bounded in } L^q((0,\tau)\times\Omega) \text{ for any finite } 1 < q < \infty,\ \mathbf{F}\cdot\mathbf{n}|_{\partial\Omega} = 0\,.$$

Thus, applying the (weak) maximal regularity theory for parabolic equations (see Amann [1]), we conclude that

$$\partial_t p = P \quad \text{is bounded in } L^q(0,\tau; W^{1,q}(\Omega)) \text{ for any } 1 < q < \infty\,. \tag{30}$$

Note that this step requires higher regularity of the initial data (at $t = 0$), specifically,

$$\partial_t p(0,\ \cdot\ ) = P(0,\ \cdot\ ) \in B^{1-(2/q);q,q}(\Omega)\,,$$

see Amann [1, Theorem 2.1]. This kind of initial regularity hypothesis is not unusual for a parabolic problem.

Finally, embedding $W^{1,q}(\Omega)$ into $L^\infty(\Omega)$ for $q > 3$, together with boundedness of the right hand side of (21), (see (25)) yields the desired conclusion

$$\mathrm{div}_x\nabla_x p = \Delta_x p \in L^q(0,\tau; L^\infty(\Omega)) \quad \text{for any } 1 < q < \infty\,. \tag{31}$$

3.2 Weak Sequential Stability

Our goal is to establish the following result:

Theorem 1 *Let $\{p_\varepsilon\}_{\varepsilon>0}$, $\{Y^i_\varepsilon\}_{\varepsilon>0}$; $i = 1,\dots,n$, be a family of (smooth) solutions of problem (21), (22) such that:*

$$\begin{gathered} p_\varepsilon \to p,\ \nabla_x p_\varepsilon \to \nabla_x p \text{ in } C([0,\tau]\times\Omega), \\ \Delta_x p_\varepsilon \to \Delta_x p \text{ weakly-(*) in } L^q(0,\tau; L^\infty(\Omega)),\ 1 < q < \infty, \end{gathered} \tag{32}$$

$$Y^i_\varepsilon \to Y^i \text{ weakly-(*) in } L^\infty((0,\tau)\times\Omega), \tag{33}$$

$$Y^i_\varepsilon(0,\cdot) \to Y^i_0 \text{ in } L^1(\Omega). \tag{34}$$

Then

$$Y^i_\varepsilon \to Y^i \text{ a.e. in } (0,\tau)\times\Omega\,, \tag{35}$$

where p and $Y^1, \dots, Y^n$ satisfy (22), specifically,

$$\partial_t Y^i - \mathrm{div}_x(Y^i \nabla_x p) + Y^i \Delta_x p = \frac{1}{p}\omega_i(p, Y^1, \dots, Y^n) \sum_{j=1}^{n} \frac{m_i}{m_j} Y^j RT, \ i = 1, \dots, n. \tag{36}$$

The rest of the paper is devoted to the proof of Theorem 1. We use the approach proposed in the seminal paper by DiPerna and Lions [7].

3.2.1 Existence for the Limit Problem

We show that the limit problem (36) admits a weak solution $Y^1, \dots, Y^n$ such that

$$Y^i \geq 0 \text{ for any } i = 1, \dots, n, \ \sum_{i=1}^{n} Y^i = 1\,,$$

provided the initial data satisfy

$$Y_0^i \geq 0\,, \quad \sum_{i=1}^{n} Y_0^i = 1\,.$$

Step 1
We approximate the pressure p by a family of smooth functions $\{p_\delta\}_{\delta>0}$,

$$p_\delta \to p, \ \nabla_x p_\delta \to \nabla_x p \text{ uniformly in } [0, \tau] \times \overline{\Omega},$$

$$\Delta_x p_\delta \to \Delta_x p \text{ a.e. in } (0, \tau) \times \Omega\,, \quad \|\Delta_x p_\delta\|_{L^q(0,\tau;L^\infty(\Omega))} \overset{<}{\sim} 1 \quad \text{for any } 1 < q < \infty\,.$$

as $\delta \to 0$. Using the standard method of characteristics, we find a unique solution $Y_\delta^1, \dots, Y_\delta^n$ emanating from the initial data $Y_0^1, \dots, Y_0^n$.

Thanks to hypothesis (6),

$$Y_\delta^i \geq 0 \quad \text{for all } i = 1, \dots, n\,,$$

and, by virtue of (3),

$$\sum_{i=1}^{n} Y_\delta^i = 1\,.$$

Consequently, passing to a suitable subsequence if necessary, we may assume that

$$Y_\delta^i \to Y^i \text{ weakly-(*) in } L^\infty((0, \tau) \times \Omega) \cap C_{\mathrm{weak}}([0, \tau]; L^1(\Omega)) \text{ as } \delta \to 0,$$

where

$$\partial_t Y^i - \mathrm{div}_x(Y^i \nabla_x p) + Y^i \Delta_x p = \frac{1}{p} \overline{\omega_i(p, Y^1, \ldots, Y^n) \sum_{j=1}^{n} \frac{m_i}{m_j} Y^j} RT, \ i = 1, \ldots, n. \tag{37}$$

$$Y^i(0, \cdot) = Y_0^i. \tag{38}$$

Here and hereafter, the upper bar denotes a weak limit of compositions of smooth functions applied to weakly convergent sequences.

Step 2

In order to complete the proof, we have to show strong convergence

$$Y_\delta^i \to Y^i \text{ a.a. in } (0, \tau) \times \Omega \text{ as } \delta \to 0. \tag{39}$$

To this end, we write down a renormalized formulation of the $\delta-$problem in the form:

$$\partial_t |Y_\delta|^2 - \mathrm{div}_x(|Y_\delta|^2 \nabla_x p_\delta) + |Y_\delta|^2 \Delta_x p_\delta = \frac{2RT}{p_\delta} \sum_{i,j=1}^{n} \frac{m_i}{m_j} \omega_i(p_\delta, Y_\delta^1, \ldots, Y_\delta^n) Y_\delta^i Y_\delta^j.$$

Letting $\delta \to 0$ we obtain

$$\partial_t \overline{|Y|^2} - \mathrm{div}_x(\overline{|Y|^2} \nabla_x p) + \overline{|Y|^2} \Delta_x p = \frac{2RT}{p} \sum_{i,j=1}^{n} \frac{m_i}{m_j} \overline{\omega_i(p, Y^1, \ldots, Y^n) Y^i Y^j}. \tag{40}$$

Now, applying the regularization procedure of DiPerna and Lions [7] to (37) we deduce that

$$\partial_t |Y|^2 - \mathrm{div}_x(|Y|^2 \nabla_x p) + |Y|^2 \Delta_x p = \frac{2RT}{p} \sum_{i,j=1}^{n} \frac{m_i}{m_j} \overline{\omega_i(p, Y^1, \ldots, Y^n) Y^j} Y^i. \tag{41}$$

Step 3

Finally, we integrate the difference of (40), (41) over Ω:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \left(\overline{|Y|^2} - |Y|^2 \right) \mathrm{d}x = - \int_\Omega \Delta_x p \left(\overline{|Y|^2} - |Y|^2 \right) \mathrm{d}x$$

$$+ \int_\Omega \frac{2RT}{p} \sum_{i,j=1}^{n} \frac{m_i}{m_j} \left[\overline{\omega_i(p, Y^1, \ldots, Y^n) Y^i Y^j} - \overline{\omega_i(p, Y^1, \ldots, Y^n) Y^j} Y^i \right] \mathrm{d}x,$$

where

$$\int_\Omega \left[\overline{\omega_i(p, Y^1, \dots, Y^n)Y^iY^j} - \overline{\omega_i(p, Y^1, \dots, Y^n)Y^j}Y^i\right] \mathrm{d}x$$

$$= \lim_{\delta\to 0}\int_\Omega \left[\omega_i(p_\delta, Y_\delta^1, \dots, Y_\delta^n)Y_\delta^i - \omega_i(p_\delta, Y^1, \dots, Y^n)Y^i\right](Y_\delta^j - Y^j)\,\mathrm{d}x$$

$$\overset{<}{\sim} \lim_{\delta\to 0}\int_\Omega |Y_\delta - Y|^2\,\mathrm{d}x = \int_\Omega \left(\overline{|Y|^2} - |Y|^2\right)\,\mathrm{d}x.$$

Thus, applying Gronwall's lemma and using the fact that the initial values converge strongly, we conclude

$$\overline{|Y|^2} = |Y|^2$$

yielding (39).

3.2.2 Compactness

Our ultimate goal is to show (35), (36). As Y_ε are smooth, we may rewrite (36) as

$$\partial_t Y_\varepsilon^i - \nabla_x Y_\varepsilon^i \cdot \nabla_x p_\varepsilon = \frac{RT}{p_\varepsilon}\sum_{j=1}^n \frac{m_i}{m_j}\omega_i(p_\varepsilon, Y_\varepsilon^1, \dots, Y_\varepsilon^n)Y_\varepsilon^j,\ i = 1, \dots, n. \quad (42)$$

At this stage, we employ once more the regularization procedure of DiPerna, Lions [7] to Eq. (36):

$$\partial_t Y_r^i - \nabla_x Y_r^i \nabla_x p = \frac{RT}{p}\sum_{j=1}^n \frac{m_i}{m_j}\omega_i(p, Y_r^1, \dots, Y_r^n)Y_r^j + e_r,\ i = 1, \dots, n, \quad (43)$$

where

$$e_r \to 0 \text{ in } L^1((0,\tau)\times\Omega) \text{ as } r \to 0.$$

Similarly to the above, we subtract (42), (43), multiply the resulting expression by $Y_\varepsilon^i - Y_r^i$, and integrate over Ω obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |Y_\varepsilon - Y_r|^2\,\mathrm{d}x + \int_\Omega \Delta_x p_\varepsilon |Y_\varepsilon - Y_r|^2\,\mathrm{d}x = \int_\Omega \sum_{i=1}^n (\nabla_x p_\varepsilon - \nabla_x p)\cdot\nabla_x Y_r^i (Y_\varepsilon^i - Y_r^i)\,\mathrm{d}x$$

$$= \int_\Omega \frac{RT}{p_\varepsilon} \sum_{i,j=1}^{n} \frac{m_i}{m_j} \left[\omega_i(p_\varepsilon, Y_\varepsilon^1, \dots, Y_\varepsilon^n) Y_\varepsilon^j - \omega_i(p_\varepsilon, Y_r^1, \dots, Y_r^n) Y_r^j \right] \mathrm{d}x + e_\varepsilon(r) + e_r,$$

where

$$e_\varepsilon(r) \to 0 \text{ in } L^1((0,\tau) \times \Omega) \text{ as } \varepsilon \to 0 \text{ for any fixed } r.$$

Finally, letting first $\varepsilon \to 0$, then $r \to 0$, and realizing that

$$Y_r^i \to Y^i \text{ in } C([0,\tau]; L^2(\Omega)),$$

we get the desired conclusion (35), (36).

Acknowledgements The research of E.F. leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ ERC Grant Agreement 320078. The Institute of Mathematics of the Academy of Sciences of the Czech Republic is supported by RVO:67985840. J.M. was supported by the project Computational methods in thermodynamics of multicomponent mixtures, KONTAKT LH12064, 2012–2015 of the Czech Ministry of Education, Youth, and Sports. H.P. was supported by the Institute of Mathematics of the Academy of Sciences of the Czech Republic, RVO:67985840. P.T. was partially supported by the Deutsche Forschungsgemeinschft (D.F.G., Germany) under Grant # TA 213 / 16-1.

References

1. Amann, H.: Maximal regularity and quasilinear parabolic boundary value problems. In: Recent advances in elliptic and parabolic problems, pp. 1–17. World Scientific Publishing, Hackensack, NJ, (2005)
2. Ambrosio, L.: Transport equation and Cauchy problem for BV vector fields. Invent. Math. **158**, 227–260 (2004)
3. Crippa, G., De Lellis, C.: Estimates and regularity results for the DiPerna-Lions flow. J. Reine Angew. Math. **616**, 15–46 (2008)
4. Denk, R., Hieber, M., Prüss, J.: R-boundedness, Fourier multipliers, and problems of elliptic and parabolic type, Mem. Amer. Math. Soc., **166**(788) (2003)
5. Ashyralyev, A., Sobolevskii, P.E.: Well-posedness of Parabolic Difference Equations, in Operator Theory: Advances and Applications, vol. 69. Birkhäuser Verlag, BaselBostonBerlin (1994)
6. Di Benedetto, E.: Continuity of weak solutions to a general porous media equations. Indiana Univ. Math. J. **32**, 83–118 (1983)
7. DiPerna, R.J., Lions, P.-L.: Ordinary differential equations, transport theory and Sobolev spaces. Invent. Math. **98**, 511–547 (1989)
8. E. Feireisl, A. Novotný, and T. Takahashi.: Homogenization and singular limits for the complete Navier-Stokes-Fourier system. J. Math. Pures Appl. (9), **94**(1):33–57, (2010)
9. Feireisl, E., Petzeltová, H., Trivisa, K.: Multicomponent reactive flows: global-in-time existence for large data. Commun. Pure Appl. Anal. **7**, 1017–1047 (2008)
10. Giovangigli, V.: Multicomponent Flow Modeling. Birkhäuser, Basel (1999)
11. Masmoudi, N.: Homogenization of the compressible Navier-Stokes equations in a porous medium. ESAIM: Control. Optim. Calc. Var. **8**, 885–906 (2002)

Entropy Methods and Convergence to Equilibrium for Volume-Surface Reaction-Diffusion Systems

Klemens Fellner and Bao Quoc Tang

Abstract We consider two volume-surface reaction-diffusion systems arising from cell biology. The first system describes the localisation of the protein Lgl in the asymmetric division of Drosophila SOP stem cells, while the second system models the JAK2/STAT5 signalling pathway. Both model systems have in common that (i) different species are located in *different spatial compartments*, (ii) the involved chemical reaction kinetics between the species satisfies a *complex balance condition* and (iii) that the associated complex balance equilibrium is *spatially inhomogeneous*. By using recent advances on the entropy method for complex balanced reaction-diffusion systems, we show for both systems exponential convergence to the equilibrium with constants and rates, which can be explicitly estimated.

Keywords Volume-surface reaction-diffusion systems · Surface diffusion hfillbreak Asymmetric stem cell division · JAK/STAT signalling pathway hfillbreak Convergence to equilibrium · Entropy method

1 Introduction and Main Results

Various physical, chemical, biological or ecological systems involve processes in different spatial compartments. A particular important example is given by considering quantities on a domain and on its surrounding boundary. Reactions taking place in these situations result in a class of PDE models called *Volume-Surface Reaction-Diffusion* systems (hereafter, we will use the abbreviation VSRD systems). The intrinsic volume-surface coupling of VSRD systems introduces new difficulties

K. Fellner (✉) · B.Q. Tang
Institute of Mathematics and Scientific Computing, University of Graz,
Heinrichstrasse 36, 8010 Graz, Austria
e-mail: klemens.fellner@uni-graz.at

B.Q. Tang
e-mail: quoc.tang@uni-graz.at

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_8

in both the mathematical and the numerical analysis compared to classical reaction-diffusion systems supported on only one spatial domain.

Recently, a rapidly increasing amount of attention has been devoted to the mathematical theory of VSRD systems arising from such different applications as fluid mechanics [16, 21], ecology [1–3], crystal growth [19] or, especially, cell biology, see e.g. [9, 11, 13, 17].

This paper aims to investigate the large time behaviour of two particular linear VSRD systems arising from two different application backgrounds in cell biology: the first being a model for the localisation of the key-protein Lgl during the asymmetric stem cell division of SOP precursor cells in Drosophila (see e.g. [11]), the second one being a model on the so-called JAK2/STAT5 signalling pathway, see [13]. In both examples, the cell cytoplasm constitutes the volume domain and the surrounding cell-membrane/cortex constitutes the surrounding boundary. In the JAK2/STAT5 model, also the volume of cell nucleus and its boundary are considered. Moreover, reactions occur within and between the volume- and surface-compartments, which do not satisfy the so called detailed balance condition, but the more general *complex balance condition*, see e.g. [18].

The main results of this paper prove exponential convergence to the complex balance equilibrium with explicitly computable rates by using the so-called *entropy method*. Moreover, we extend the entropy method to apply to spatially inhomogeneous equilibria. This paper may serve as a proof of concept for the applicability of the entropy method to a wide class of VSRD systems including mixed ODE/PDE systems. The presented proofs, however, rely on positive lower and upper bounds of the equilibria, which are difficult to obtain for general systems.

The entropy method applied in this work has recently become a very powerful tool in proving exponential convergence to equilibrium with explicit rates for reaction-diffusion systems, but mostly under the assumption of the detailed balance condition and on a single spatial domain, see e.g. [5, 6, 12, 15, 22]. The entropy method for complex balance reaction-diffusion networks was so far only considered on a single domain for linear, respectively, nonlinear systems in [7, 10].

The current paper constitutes the first results of the entropy method for volume-surface systems with spatially inhomogeneous complex balance equilibria. In the next two sections, we detail the considered VSRD systems and state the main results.

1.1 A VSRD Model for the Localisation of Lgl During Asymmetric Stem Cell Division

In stem cells undergoing asymmetric cell division, particular proteins (so-called cell-fate determinants) are localised at the cortex of only one of the two daughter cells during mitosis. These cell-fate determinants trigger in the following the differentiation of one daughter cell into specific tissue while the other daughter cell remains a stem cell.

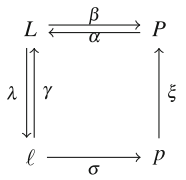


Fig. 1 The reaction dynamics between L, P, l and p

In Drosophila, SOP precursor stem cells provide a well-studied biological example model of asymmetric stem cell division, see e.g. [4, 20, 25] and the references therein. In particular, asymmetric cell division of SOP cells is driven by the asymmetric localisation of the key protein Lgl (Lethal giant larvae), which exists in two conformational states: a non-phosphorylated form which regulates the localisation of the cell-fate-determinants in the membrane of one daughter cell, and a phosphorylated form which is inactive.

The asymmetric localisation of Lgl during mitosis is the result of the activation of the kinase aPKC, which phosphorylates Lgl (as part of a highly evolutionary conserved protein complex) only on a subpart of the cortex, as well as the results of the weakly reversible reaction/sorption dynamics of the two conformations of Lgl between cortex and cytoplasm. In particular, it is the irreversible release of phosphorylated Lgl from the cortex, which initiates the asymmetric localisation of Lgl upon the activation of aPKC.

Let $\Omega \subset \mathbb{R}^n, n = 2, 3$ describe the cell cytoplasm as a connected, bounded domain with sufficiently smooth boundary $\partial\Omega$ (e.g. $\partial\Omega \in C^{2+\varepsilon}, \varepsilon > 0$). Denote by $\Gamma = \partial\Omega$ the surrounding cell cortex. Moreover, the cell cortex is divided into two disjoint, connected subsets Γ_1 and Γ_2 with $\Gamma = \Gamma_1 \cup \Gamma_2$ and where Γ_2 is the active part of the cell cortex, where phosphorylation takes place.

Concerning the various conformations of Lgl, we denote by L and P the concentrations of non-phosphorylated Lgl and phosphorylated Lgl within the volume domain Ω. Moreover, the concentrations of non-phosphorylated Lgl and phosphorylated Lgl on the cell cortex are denoted by ℓ and p, respectively. Note that ℓ is supported on Γ while p is supported only on the sub-domain Γ_2, since phosphorylation only occurs on Γ_2.

Schematically, we consider the following reactions between the four different conformations of Lgl with positive reaction rate constants $\alpha, \beta, \lambda, \gamma, \sigma$ and ξ as depicted in Fig. 1.

Moreover, we assume positive diffusion coefficients of L, P on Ω, of ℓ on Γ and of p on Γ_2, i.e. $d_L, d_P, d_\ell, d_p > 0$, respectively. Then, applying the mass action law, the resulted VSRD system consists of two volume equations

$$\begin{cases} L_t - d_L \Delta L = -\beta L + \alpha P, & x \in \Omega,\ t > 0, \\ P_t - d_P \Delta P = \beta L - \alpha P, & x \in \Omega,\ t > 0, \end{cases} \tag{1a}$$

and two surface equations

$$\begin{cases} \ell_t - d_\ell \Delta_\Gamma \ell = \lambda L - (\gamma + \sigma \chi_{\Gamma_2})\ell, & x \in \Gamma,\ t > 0, \\ p_t - d_p \Delta_{\Gamma_2} p = \sigma \ell - \xi p, & x \in \Gamma_2,\ t > 0, \\ \partial_{\nu_2} p = 0, & x \in \partial\Gamma_2,\ t > 0, \end{cases} \tag{1b}$$

which are connected via mixed Neumann/Robin boundary conditions

$$\begin{cases} d_L \partial_\nu L = -\lambda L + \gamma \ell, & x \in \Gamma,\ t > 0, \\ d_P \partial_\nu P = \chi_{\Gamma_2} \xi p, & x \in \Gamma,\ t > 0, \end{cases} \tag{1c}$$

and subject to nonnegative initial data

$$\begin{cases} L(x,0) = L_0(x), \quad P(x,0) = P_0(x), & x \in \Omega, \\ \ell(x,0) = \ell_0(x), \quad x \in \Gamma, & p(x,0) = p_0(x), \quad x \in \Gamma_2, \end{cases} \tag{1d}$$

where ν and ν_2 are the outward unit normal vectors of Γ and $\partial\Gamma_2$, respectively, and Δ_Γ and Δ_{Γ_2} are Laplace-Beltrami operators on Γ and Γ_2, respectively. Moreover, χ_{Γ_2} denotes the characteristic function of the boundary part Γ_2.

Note that the above system (1) conserves the total mass of Lgl, which is expressed in the following conservation law:

$$\begin{aligned} &\int_\Omega (L(t,x) + P(t,x))\,dx + \int_\Gamma \ell(t,x)\,dS + \int_{\Gamma_2} p(t,x)\,dS \\ &\quad = \int_\Omega (L_0(x) + P_0(x))\,dx + \int_\Gamma \ell_0(x)\,dS + \int_{\Gamma_2} p_0(x)\,dS > 0, \qquad \forall t > 0 \end{aligned} \tag{2}$$

Concerning the existence of global weak solutions of (1), we refer to [11], where also the quasi-steady-state approximation in the limit $\xi \to +\infty$ and numerical simulations were carried out.

The first main result of this paper is stated in the following theorem.

Theorem 1 (Exponential convergence to equilibrium of Lgl system (1))
Assume that $\Omega \subset \mathbb{R}^n$, $n = 2, 3$ is a bounded domain with sufficiently smooth boundary $\Gamma = \partial\Omega$. Moreover, $\Gamma = \Gamma_1 \cup \Gamma_2$ is the union of two disjoint, connected subsets and Γ_2 has a smooth boundary $\partial\Gamma_2$.

Then, for any positive initial mass $M > 0$, the system (1) possesses a unique positive equilibrium $(L_\infty, P_\infty, \ell_\infty, p_\infty)$ satisfying the mass conservation law

$$\int_\Omega (L_\infty(x) + P_\infty(x))dx + \int_\Gamma \ell_\infty(x)dS + \int_{\Gamma_2} p_\infty(x)dS = M.$$

Moreover, $L_\infty \in C(\overline{\Omega}) \cap H^2(\Omega)$, $P_\infty \in L^\infty(\Omega) \cap H^{3/2}(\Omega)$, $\ell_\infty \in C(\Gamma) \cap H^2(\Gamma_2)$ and $p_\infty \in C(\overline{\Gamma_2}) \cap H^2(\Gamma_2)$, and there are $0 < a \le A < +\infty$ such that

$$a \le L_\infty(x), P_\infty(x), \ell_\infty(x), p_\infty(x) \le A$$

where the bounds hold for x in $\overline{\Omega}$, a.e. in Ω, in Γ and in $\overline{\Gamma_2}$ respectively.

Finally, every global weak solution to (1) with positive initial mass M (as constructed in [11]) converges exponentially to $(L_\infty, P_\infty, \ell_\infty, p_\infty)$ in the following sense

$$\int_\Omega \frac{|L(t)-L_\infty|^2}{L_\infty} + \int_\Omega \frac{|P(t)-P_\infty|^2}{P_\infty} + \int_\Gamma \frac{|\ell(t)-\ell_\infty|^2}{\ell_\infty} + \int_{\Gamma_2} \frac{|p(t)-p_\infty|^2}{p_\infty} \leq C_0 e^{-\lambda_0 t}$$

for all $t > 0$, where λ_0 is as in Lemma 3 and C_0 and λ_0 can be computed explicitly.

1.2 A PDE/ODE System Modelling the JAK2/STAT5 Signalling Pathway

The communication between cells in multicellular organisms is often mediated by signalling molecules secreted to the extracellular space, which then bind to cell surface receptors, see [13]. However, the modalities of the transport from the site of signal transducer and activator of transcription (STAT) phosphorylation at the plasma membrane to the site of action in the nucleus is still unclear. In [13], Friedmann, Neumann and Rannacher introduced a mixed PDE/ODE model system to analyse the influence of the cell shape on the regulatory response to the activated pathway.

By following the notations in [13], we denote by u_0 and u_1 the unphosporylated and phosphorylated STAT5 in the cytoplasm, while u_2 and u_3 denote the unphosphorylated and phosphorylated STAT5 in the nucleus, respectively. Moreover, we denote by $u_4, \ldots, u_7$ so-called "fictitious concentrations", which describe processes in the nucleus via linear equations yielding a delayed response. The reaction dynamics of the eight species u_i, $i = 0, 1, \ldots, 7$ are depicted by the diagram in Fig. 2.

The JAK2/STAT5 model considers a smooth, bounded domain $\Omega_0 \subset \mathbb{R}^n$ and distinguishes $\Omega_{\text{nuc}} \subsetneq \Omega_0$ the domain of the cell nucleus and $\Omega_{\text{cyt}} = \Omega_0 \backslash \Omega_{\text{nuc}}$ the cell cytoplasm. With a little abuse of notation, we denote by $\partial\Omega_{\text{cyt}} = \partial\Omega_0$ the membrane of the cell, while $\partial\Omega_{\text{nuc}}$ is the boundary of the nucleus.

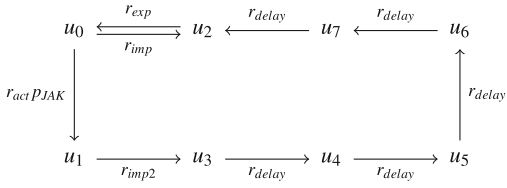


Fig. 2 Reaction network of the JAK2/STAT5 signalling pathway

The following mixed PDE/ODE model was considered in [13]: The two PDEs

$$\begin{cases} \partial_t u_0(t,x) = D\Delta u_0(t,x), & t>0, \quad x\in\Omega_{\rm cyt},\\ D\partial_{n_1} u_0(t,y) = -\frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|}p_{JAK}u_0(t,y), & t>0, \quad y\in\partial\Omega_{\rm cyt},\\ D\partial_{n_2} u_0(t,y) = -\frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|}u_0(t,y)+\frac{r_{\rm exp}}{|\partial\Omega_{\rm nuc}|}u_2(t), & t>0, \quad y\in\partial\Omega_{\rm nuc}, \end{cases} \tag{3a}$$

$$\begin{cases} \partial_t u_1(t,x) = D\Delta u_1(t,x), & t>0, \quad x\in\Omega_{\rm cyt},\\ D\partial_{n_1} u_1(t,y) = \frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|}p_{JAK}u_0(t,y), & t>0, \quad y\in\partial\Omega_{\rm cyt},\\ D\partial_{n_2} u_1(t,y) = -\frac{r_{\rm imp2}}{|\partial\Omega_{\rm nuc}|}u_1(t,y), & t>0, \quad y\in\partial\Omega_{\rm nuc}, \end{cases} \tag{3b}$$

and six ODEs

$$\begin{cases} (u_2)'(t)+\frac{r_{\rm exp}}{|\Omega_{\rm nuc}|}u_2(t) = \frac{r_{\rm delay}}{|\Omega_{\rm nuc}|}u_7(t)+\frac{r_{\rm imp}}{|\Omega_{\rm nuc}||\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}}u_0(t,y)dS,\\ (u_3)'(t)+\frac{r_{\rm delay}}{|\Omega_{\rm nuc}|}u_3(t) = \frac{r_{\rm imp2}}{|\Omega_{\rm nuc}||\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}}u_1(t,y)dS,\\ (u_i)'(t)+\frac{r_{\rm delay}}{|\Omega_{\rm nuc}|}u_i(t) = \frac{r_{\rm delay}}{|\Omega_{\rm nuc}|}u_{i-1}(t), \quad i=4,5,6,7, \end{cases} \tag{3c}$$

subject to nonnegative initial data $u_0(x,0)=u_0^{in}(x), u_1(x,0)=u_1^{in}(x),\ x\in\Omega$, $u_i(0)=u_i^{in}$ for $i=2,3,\dots,7$, where ν_1 and ν_2 are outward normals of $\partial\Omega_{\rm cyt}$ and $\partial\Omega_{\rm nuc}$ respectively. Note that the system (3) satisfies the mass conservation law

$$\int\limits_{\Omega_{\rm cyt}}(u_0(t,x)+u_1(t,x))dx+|\Omega_{\rm nuc}|\sum_{i=2}^{7}u_i(t)=\int\limits_{\Omega_{\rm cyt}}(u_0^{in}(x)+u_1^{in}(x))dx+|\Omega_{\rm nuc}|\sum_{i=2}^{7}u_i^{in}. \tag{4}$$

The well-posedness of the mixed PDE/ODE model (3) was shown in [13]. Moreover, in the pure ODE case, the authors proved exponential convergence to equilibrium by extensively studying the structure of the reaction matrix. This approach, however, doesn't apply to the PDE/ODE case and the authors were only able to prove Lyapunov stability of the stationary states of (3).

The second main result of this paper proves exponential convergence to equilibrium of the mixed PDE/ODE JAK2/STAT5 model (3).

Theorem 2 (Exponential convergence to equilibrium of the JAK2/STAT5 model (3)) *Let $\Omega_0, \Omega_{\rm nuc}, \Omega_{\rm cyt}\subset\mathbb{R}^n$, $n\geq 2$ be smooth, bounded domains with $\Omega_{\rm nuc}\subsetneq\Omega_0$, $\Omega_{\rm cyt}=\Omega_0\backslash\Omega_{\rm nuc}$ and $\partial\Omega_{\rm cyt}=\partial\Omega_0$.*

Then, for any positive initial mass $M>0$, system (3) possesses a unique equilibrium $(u_{0,\infty},\dots,u_{7,\infty})$ satisfying the mass conservation law

$$\int\limits_{\Omega_{\rm cyt}}(u_{0,\infty}(x)+u_{1,\infty}(x))dx+|\Omega_{\rm nuc}|\sum_{i=2}^{7}u_{i,\infty}=M>0. \tag{5}$$

Moreover, $u_{2,\infty}, \dots, u_{7,\infty}$ *are positive and* $u_{0,\infty}, u_{1,\infty} \in C(\overline{\Omega}_{\mathrm{cyt}}) \cap C^2(\Omega_{\mathrm{cyt}})$ *satisfy* $b \le u_{0,\infty}(x), u_{1,\infty}(x) \le B$ *for all* $x \in \overline{\Omega}_{\mathrm{cyt}}$ *for some constants* $0 < b \le B \le +\infty$.

Finally, any global weak solution $(u_i)_{i=0,1,\dots,7}$ *to (3) with positive initial mass M (as constructed in [13]) converges exponentially to the equilibrium* $(u_{i,\infty})_{i=0,1,\dots,7}$ *in the sense that*

$$\int_{\Omega_{\mathrm{cyt}}} \left(\frac{|u_0(t,x) - u_{0,\infty}(x)|^2}{u_{0,\infty}(x)} + \frac{|u_1(t,x) - u_{1,\infty}(x)|^2}{u_{1,\infty}(x)} \right) dx + |\Omega_{\mathrm{nuc}}| \sum_{i=2}^{7} \frac{|u_i(t) - u_{i,\infty}|^2}{u_{i,\infty}} \le C_1 e^{-\lambda_1 t},$$

for all $t > 0$, *where* λ_1 *is as in Lemma 5 and* C_1, λ_1 *are constants, which can be computed explicitly in terms of the domains, parameters and initial mass M.*

Remark 1 In system (3), the diffusion coefficients of u_0 and u_1 are taken the same as in [13]. However, the proof of Theorem 2 holds equally for different diffusion coefficients for u_0 and u_1, e.g. u_0 diffuses with $D_0 > 0$ and u_1 with $D_1 > 0$.

2 Preliminaries

The proofs of Theorems 1 and 2 use some previous results about linear complex balance reaction-diffusion networks proven in [10], which we shall briefly recall in the following. The interested reader is referred to [10] for more details.

We consider a first order (i.e. linear) reaction network of the form

$$S_i \underset{a_{ji}}{\overset{a_{ij}}{\rightleftharpoons}} S_j \qquad i \ne j = 1, 2, \dots, N, \qquad (\mathscr{N})$$

where $S_i, i = 1, 2, \dots, N$ are different chemical substances (or species) and a_{ij}, $a_{ji} \ge 0$ are constant reaction rates. In particular, a_{ij} denotes the reaction rates from the species S_j to S_i. The considered reaction network is contained in a bounded vessel (or reactor) $\Omega \subset \mathbb{R}^n$, where Ω is a smooth, bounded domain with outer unit normal ν. The substances S_i are described by spatio-temporal concentrations $u_i(t, x)$ at position $x \in \Omega$ and time $t \ge 0$. In addition, each substance S_i is assumed to diffuse with a diffusion coefficient $d_i \ge 0$, $i = 1, 2, \dots, N$. Finally, using the mass action law as model for the reaction rates leads to the following linear reaction-diffusion system:

$$\begin{cases} \boldsymbol{c}_t = \mathbb{D}\Delta \boldsymbol{c} + A\boldsymbol{c}, & x \in \Omega, \quad t > 0, \\ \partial_\nu \boldsymbol{c} = 0, & x \in \partial\Omega, \quad t > 0, \\ \boldsymbol{c}(x, 0) = \boldsymbol{c}_0(x) \ge 0, & x \in \Omega, \end{cases} \qquad (6)$$

where $\boldsymbol{c}(t,x) = [u_1(t,x), \dots, u_N(t,x)]^T$ denotes the vector of concentrations subject to non-negative initial conditions $\boldsymbol{c}_0(x) = [u_{1,0}(x) \geq 0, \dots, u_{N,0}(x) \geq 0]^T$, $\mathbb{D} = \mathrm{diag}(d_1, d_2, \dots, d_N)$ denotes the diagonal diffusion matrix and the reaction matrix $A = (a_{ij}) \in \mathbb{R}^{N\times N}$ satisfies the following conditions:

$$\begin{cases} a_{ij} \geq 0, & \text{for all } i \neq j, \quad i,j = 1,2,\dots,N, \\ a_{jj} = -\sum_{i=1, i\neq j}^{N} a_{ij}, & \text{for all } j = 1,2,\dots,N. \end{cases} \tag{7}$$

The solution to system (6) satisfies the following mass conservation law

$$\sum_{i=1}^{N} \int_\Omega u_i(t,x)dx = M := \sum_{i=1}^{N} \int_\Omega u_{i,0}(x)dx > 0 \qquad \text{for all } t > 0. \tag{8}$$

To study the convergence to equilibrium we consider the following quadratic relative entropy functional

$$\mathscr{E}(\boldsymbol{c}_1|\boldsymbol{c}_2)(t) = \sum_{i=1}^{N} \int_\Omega \frac{|u_i|^2}{v_i} dx \tag{9}$$

between two solutions $\boldsymbol{c}_1 = [u_1, \dots, u_N]^T$ and $\boldsymbol{c}_2 = [v_1, \dots, v_N]^T$ (with respect to possibly different initial data) and its entropy dissipation $\mathscr{D}(\boldsymbol{c}_1|\boldsymbol{c}_2) = -\frac{d}{dt}\mathscr{E}(\boldsymbol{c}_1|\boldsymbol{c}_2)$:

$$\mathscr{D}(\boldsymbol{c}_1|\boldsymbol{c}_2) = 2\sum_{i=1}^{N} d_i \int_\Omega v_i \left|\nabla \frac{u_i}{v_i}\right|^2 dx + \sum_{i,j=1; i<j}^{N} \int_\Omega (a_{ij}v_j + a_{ji}v_i)\left(\frac{u_i}{v_i} - \frac{u_j}{v_j}\right)^2 dx. \tag{10}$$

The network $\mathscr{N}$ is called *weakly reversible* if for any reaction $S_i \xrightarrow{a_{ji}} S_j$ with $a_{ji} > 0$, there exist $S_{k_1}, \dots, S_{k_r}$ such that $S_j \equiv S_{k_0} \xrightarrow{a_{k_1 k_0}} S_{k_1} \xrightarrow{a_{k_2 k_1}} S_{k_2} \to \dots \to S_{k_r} \xrightarrow{a_{k_{r+1} k_r}} S_i \equiv S_{k_{r+1}}$ with $a_{k_{i+1}k_i} > 0$ for all $i = 0, 1, \dots, r$. Intuitively, a network $\mathscr{N}$ is weakly reversible if for any reaction from S_i to S_j we can find a returning chain of reactions, which starts from S_j and finishes at S_i.

If $\mathscr{N}$ is weakly reversible, then the associated reaction graph can be composed of multiple disjoint strongly connected components, see [10]. However, these components are entirely independent since every node of a first order reaction network represents only one species. Therefore, all such disjoint components can be treated separately from one another and w.l.o.g. we *shall consider in the following only weakly reversible networks consisting of one strongly connect component*. Finally, all weakly reversible first order reaction networks satisfy the complex balance condition, see [10]. The following theorem is one of the main results in [10].

Theorem 3 (Exponential convergence to equilibrium for first order reaction networks, [10]) *Assume that the reaction network $\mathscr{N}$ is weakly reversible and consists w.l.o.g. of only one strongly connect component. Moreover, assume that there is at least one positive diffusion coefficient, that is, there exists i_0 such that $d_{i_0} > 0$. Let*

$\Omega \subset \mathbb{R}^n$, $n \geq 1$ be a bounded domain with sufficiently smooth boundary $\partial\Omega$ (e.g. $\partial\Omega \in C^{2+\varepsilon}$, $\varepsilon > 0$).

Then, for any given positive initial mass $M > 0$, system (6) possesses a unique positive complex balance equilibrium $\boldsymbol{c}_\infty = (u_{1,\infty}, \ldots, u_{n,\infty})$ satisfying the mass conservation law (8). Moreover, each solution to (6) with positive initial mass M converges exponentially to this equilibrium with computable rates, i.e.

$$\sum_{i=1}^{N} \|u_i(t) - u_{i,\infty}\|^2_{L^2(\Omega)} \leq Ce^{-\lambda t} \qquad \text{for all } t > 0,$$

where C, λ are constants, which can be computed explicitly.

The following elementary inequality, which was proved in [10] (in a variant), will be useful in the following sections.

Lemma 1 (A finite dimensional inequality, see [10]) *Assume that the network $\mathscr{N}$ is weakly reversible and consists of one strongly connected component.*

Then, for all $\boldsymbol{c} = (c_1, \ldots, c_N)$ satisfying $\sum_{i=1}^N \alpha_i c_i = 0$ with $\alpha_1, \ldots, \alpha_N$ being positive constants, there holds

$$\sum_{i,j=1;i<j}^{N} (a_{ij} + a_{ji})(c_i - c_j)^2 \geq \eta \sum_{i=1}^{N} c_i^2$$

for an explicit *constant $\eta > 0$ depending only on α_i, a_{ij} and N.*

Proof First, thanks to the weak reversibility of $\mathscr{N}$, and since $\mathscr{N}$ contains only one strongly connected component, we observe for every S_i and S_j in $\mathscr{N}$ that there exist a chain of nontrivial reactions starting from S_i and finishing at S_j and vice versa. Hence, by using iteratively the triangle inequality along such chains of reactions in cases where $a_{ij} = a_{ji} = 0$, we can estimate (see [10, Lemma 2.4] for the details)

$$\sum_{i,j=1;i<j}^{N} (a_{ij} + a_{ji})(c_i - c_j)^2 \geq \zeta \sum_{i,j=1;i<j}^{N} (c_i - c_j)^2 \tag{11}$$

for some explicit constant $\zeta(N) > 0$ only depending on N. Now, for any $1 \leq i_0 \leq N$

$$\begin{aligned}\sum_{i,j=1;i<j}^{N} (c_i - c_j)^2 &\geq \sum_{j=1;j\neq i_0}^{N} (c_{i_0} - c_j)^2 \geq \sum_{j=1;j\neq i_0}^{N} \frac{1}{\alpha_j^2}(\alpha_j c_{i_0} - \alpha_j c_j)^2 \\ &\geq \frac{\left(c_{i_0}\sum_{j=1;j\neq i_0}^N \alpha_j - \sum_{j=1;j\neq i_0}^N \alpha_j c_j\right)^2}{(N-1)\max_{i=1,\ldots,N}\{\alpha_i^2\}} = \frac{\left(\sum_{j=1}^N \alpha_j\right)^2}{(N-1)\max_{i=1,\ldots,N}\{\alpha_i^2\}}\, c_{i_0}^2.\end{aligned} \tag{12}$$

Since $1 \leq i_0 \leq N$ is arbitrary, by combining (11) and (12) we get the desired result.

3 Proof of Theorem 1

In this section, we denote by $\boldsymbol{c} = (L, P, \ell, p)$ the vector of concentrations of system (1). Following the previous section, we introduce the relative entropy between two solution trajectories $\boldsymbol{c}_1 = (L_1, P_1, \ell_1, p_1)$ and $\boldsymbol{c}_2 = (L_2, P_2, \ell_2, p_2)$ as follow

$$\mathscr{E}(\boldsymbol{c}_1|\boldsymbol{c}_2) = \int_\Omega \left(\frac{L_1^2}{L_2} + \frac{P_1^2}{P_2}\right) dx + \int_\Gamma \frac{\ell_1^2}{\ell_2} dS + \int_{\Gamma_2} \frac{p_1^2}{p_2} dS. \tag{13}$$

Corresponding to (10), we can compute the entropy dissipation functional corresponding to (13), i.e. $\mathscr{D}(\boldsymbol{c}_1|\boldsymbol{c}_2) = -\frac{d}{dt}\mathscr{E}(\boldsymbol{c}_1|\boldsymbol{c}_2)$ as

$$\begin{aligned}
\mathscr{D}(\boldsymbol{c}_1|\boldsymbol{c}_2) &= 2d_L \int_\Omega L_2 \left|\nabla \frac{L_1}{L_2}\right|^2 dx + 2d_P \int_\Omega P_2 \left|\nabla \frac{P_1}{P_2}\right|^2 dx + 2d_\ell \int_\Gamma \ell_2 \left|\nabla_\Gamma \frac{\ell_1}{\ell_2}\right|^2 dS \\
&+ 2d_p \int_{\Gamma_2} p_2 \left|\nabla \frac{p_1}{p_2}\right|^2 dS + \int_\Omega (\alpha P_2 + \beta L_2) \left|\frac{L_1}{L_2} - \frac{P_1}{P_2}\right|^2 dx \\
&+ \int_\Gamma (\lambda L_2 + \gamma \ell_2) \left|\frac{L_1}{L_2} - \frac{\ell_1}{\ell_2}\right|^2 dS + \int_{\Gamma_2} \xi p_2 \left|\frac{P_1}{P_2} - \frac{p_1}{p_2}\right|^2 dS + \int_{\Gamma_2} \sigma \ell_2 \left|\frac{\ell_1}{\ell_2} - \frac{p_1}{p_2}\right|^2 dS.
\end{aligned} \tag{14}$$

The following Lemma 2 proves the existence of a unique positive equilibrium provided positive initial mass $M > 0$. The main difficulties are the complex balance structure of the equilibrium (with the associated mass conservation law preventing standard coercivity arguments) and the mixed boundary conditions impeding classical solutions and thus, the direct use of classical maximum principles.

Lemma 2 (Existence of a unique positive equilibrium)
Let Ω be a bounded domain of $\mathbb{R}^n$, $n = 2, 3$ with smooth boundary $\Gamma = \partial\Omega$.

Then, for any positive initial mass $M > 0$, the system (1) possesses a unique positive equilibrium $\boldsymbol{c}_\infty = (L_\infty, P_\infty, \ell_\infty, p_\infty)$ satisfying the mass conservation

$$\int_\Omega (L_\infty(x) + P_\infty(x))dx + \int_\Gamma \ell_\infty(x) dS + \int_{\Gamma_2} p_\infty(x) dS = M. \tag{15}$$

Moreover, $L_\infty \in C(\overline{\Omega}) \cap H^2(\Omega)$, $P_\infty \in L^\infty(\Omega) \cap H^{3/2}(\Omega)$, $\ell_\infty \in C(\Gamma) \cap H^2(\Gamma)$ and $p_\infty \in C(\overline{\Gamma_2}) \cap H^2(\Gamma_2)$, and satisfy for some constants $0 < a \le A < +\infty$

$$0 < a \le L_\infty(x) \le A \quad \textit{for all } x \in \overline{\Omega},$$

$$0 < a \le P_\infty(x) \le A \quad \textit{for almost all } x \in \Omega,$$

$$0 < a \le \ell_\infty(x) \le A \quad \textit{for all } x \in \Gamma,$$

$$0 < a \le p_\infty(x) \le A \quad \textit{for all } x \in \overline{\Gamma_2}.$$

Proof We will first prove that the equilibrium system has non-negative solutions and then show that equilibria are indeed bounded and strictly positive. Finally, the uniqueness follows from the vanishing of entropy-dissipation functional (14).

In order to prove the existence of nonnegative equilibria via a fixed point argument, we consider the following auxiliary system

$$\begin{cases} -d_L \Delta L + \beta L = \alpha P_0, \quad -d_P \Delta P + \alpha P = \beta L_0, & x \in \Omega, \\ \quad d_L \partial_\nu L + \lambda L = \gamma \ell_0, \quad d_P \partial_\nu P = \chi_{\Gamma_2} \xi p_0, & x \in \Gamma, \\ -d_\ell \Delta_\Gamma \ell + (\lambda + \sigma \chi_{\Gamma_2}) \ell = \lambda L_0|_\Gamma, & x \in \Gamma, \\ -d_p \Delta_{\Gamma_2} p + \xi p = \sigma \ell_0, & x \in \Gamma_2, \\ \quad d_p \partial_{\nu_{\Gamma_2}} p = 0, & x \in \partial\Gamma_2 \end{cases} \tag{16}$$

where $(L_0, P_0, \ell_0, p_0) \in \mathscr{Y}$ are given in the space

$$\mathscr{Y} = \{(U, V, u, v) \in H^1(\Omega) \times L^2(\Omega) \times L^2(\Gamma) \times L^2(\Gamma_2) : U, V, u, v \ge 0\}.$$

By standard linear elliptic equation theory, there exists a unique weak solution $(L, P, \ell, p) \in H^1(\Omega) \times H^1(\Omega) \times H^1(\Gamma) \times H^1(\Gamma_2)$ for (16). Thanks to the non-negativity of $(L_0, P_0, \ell_0, p_0) \in \mathscr{Y}$ and the weak maximum principle (cf. e.g. [14]), this solution is also nonnegative: Indeed, by testing, for instance, the equation for P by the negative part $P_- = -\min\{P, 0\}$, we calculate with $P P_- = -(P_-)^2$

$$-\int_\Gamma \chi_{\Gamma_2} \xi p_0 P_- \, dS - d_P \int_\Omega \chi_{\{P \le 0\}} |\nabla P|^2 \, dx = +\alpha \int_\Omega (P_-)^2 + \beta \int_\Omega L_0 P_- \, dx, \tag{17}$$

and observe that the left hand side is nonpositive while the right hand side is nonnegative provided that p_0 and L_0 are nonnegative. Thus, both sides have to equal zero and, as a consequence, $\int_\Omega (P_-)^2 = 0$, which implies the nonnegativity of P.

Moreover, the smoothness of the boundary $\partial\Omega$ allows to deduce higher regularity for L, namely $L \in H^2(\Omega)$ thanks to $P_0 \in L^2(\Omega)$ and $\ell_0 \in L^2(\Gamma)$. In particular, the following *a prior* estimate holds

$$\begin{aligned} &\|L\|_{H^2(\Omega)} + \|P\|_{H^1(\Omega)} + \|\ell\|_{H^1(\Gamma)} + \|p\|_{H^1(\Gamma_2)} \\ &\le C \left(\|L_0\|_{H^1(\Omega)} + \|P_0\|_{L^2(\Omega)} + \|\ell_0\|_{L^2(\Gamma)} + \|p_0\|_{L^2(\Gamma_2)} \right). \end{aligned}$$

By defining $\mathfrak{T} : \mathscr{Y} \to \mathscr{Y}$ by $\mathfrak{T}(L_0, P_0, \ell_0, p_0) = (L, P, \ell, p)$, we obtain from the previous *a priori* estimate that $\mathfrak{T}$ is a compact operator. Hence, it follows from the Schauder fixed point theorem that there exist fixed points $(L_\infty, P_\infty, \ell_\infty, p_\infty)$ of $\mathfrak{T}$, and these fixed points are thus nonnegative solutions to the equilibrium system. Note

that uniqueness of the fixed points $(L_\infty, P_\infty, \ell_\infty, p_\infty)$ can not be supposed as such, since we expect equilibria to exist for any given mass M.

Next, again by maximal regularity for linear elliptic equations, we obtain $L_\infty \in H^2(\Omega)$, $\ell_\infty \in H^2(\Gamma)$ and $p_\infty \in H^2(\Gamma_2)$, which implies $L_\infty \in C(\overline{\Omega})$, $\ell_\infty \in C(\Gamma)$ and $p_\infty \in C(\overline{\Gamma_2})$ thanks to Sobolev embeddings and $\Omega \subset \mathbb{R}^n$ with $n = 2, 3$. The continuity of L_∞, ℓ_∞ and p_∞ and the compactness of $\overline{\Omega}$, Γ and $\overline{\Gamma_2}$ imply also the upper bounds $L_\infty, \ell_\infty, p_\infty \le A < +\infty$ for a constant A.

For P_∞ satisfying the mixed Neumann boundary condition $d_P \partial_\nu P_\infty = \chi_{\Gamma_2} p_\infty$ with $\chi_{\Gamma_2} \in L^\infty(\Gamma)$ being discontinuous, maximal elliptic regularity only yields $P_\infty \in H^{3/2}(\Omega)$, which is insufficient to conclude boundedness in three space dimensions, see e.g. [8]. However, we are able to construct supersolutions $\hat{P}$ as the solutions of

$$\begin{cases} -d_P \Delta \hat{P} + \alpha \hat{P} = \beta \|L_\infty\|_\infty, & x \in \Omega, \\ d_P \partial_\nu \hat{P} = \xi \|p_\infty\|_\infty, & x \in \partial\Gamma. \end{cases}$$

For the supersolutions $\hat{P}$, standard elliptic theory implies the maximal regularity $\hat{P} \in H^2(\Omega)$ and thus continuity and boundedness. Moreover, the same weak maximum principle argument as in (17) yields $\hat{P} \ge P_\infty$ and thus the upper bound $P_\infty \le A < +\infty$ for a constant A.

We will show now that $\ell_\infty \equiv 0$ implies $L_\infty = P_\infty = p_\infty = 0$. Indeed, with $\ell_\infty \equiv 0$ it follows readily that $p_\infty \equiv 0$. By multiplying the equation for L_∞ by βL_∞, the equation for P_∞ by αP_∞, and by summing the two equations, we calculate with $p_\infty \equiv 0$

$$d_L \|\nabla L_\infty\|^2_{L^2(\Omega)} + \lambda \|L_\infty\|^2_{L^2(\Gamma)} + d_P \|\nabla P_\infty\|^2_{L^2(\Omega)} + \int_\Omega (\beta L_\infty - \alpha P_\infty)^2 \, dx = 0,$$

which implies $L_\infty \equiv 0$ and eventually $P_\infty \equiv 0$. Therefore, whenever a positive mass $M > 0$ is considered, the corresponding equilibrium state has to satisfy $\ell_\infty \not\equiv 0$, and consequently $p_\infty \not\equiv 0$ and $P_\infty \not\equiv 0$.

From the continuity of ℓ_∞, we obtain that p_∞ is the unique classical solution to

$$\begin{cases} -d_p \Delta_{\Gamma_2} p_\infty + \xi p_\infty = \sigma \ell_\infty, & x \in \Gamma_2, \\ d_p \partial_{\nu_{\Gamma_2}} p_\infty = 0, & x \in \partial\Gamma_2 \end{cases}$$

Since ℓ_∞ is nonnegative and not identically zero, we can apply the classical maximum principle to conclude $p_\infty(x) \ge a > 0$ for $x \in \overline{\Gamma_2}$ and a constant $a > 0$.

Next, by considering the auxiliary equation

$$-d_\ell \Delta_\Gamma \ell^* + (\gamma + \sigma)\ell^* = \lambda L_\infty|_\Gamma, \quad x \in \Gamma,$$

and by recalling the continuity and nonnegativity of $L_\infty \not\equiv 0$, the strong maximum applied to the unique classical solution ℓ^* implies that $\ell^*(x) \ge a > 0$ for all $x \in \Gamma$ and a constant $a > 0$. Moreover, by a weak maximum principle argument analog

to (17), we have that ℓ^* is a subsolution to ℓ_∞, i.e. $\ell_\infty(x) \geq \ell^*(x) \geq a > 0$ for all $x \in \Gamma$.

Moreover, we consider the unique classical solutions L^* of the auxiliary system

$$-d_L \Delta L^* + \beta L^* = 0, \quad x \in \Omega, \qquad d_L \partial_\nu L^* + \lambda L^* = \gamma \ell_\infty, \quad x \in \Gamma,$$

for which the classical maximum principle and the lower bound $\ell_\infty(x) \geq a > 0$ implies $L^*(x) \geq a > 0$ for $x \in \Omega$ and a constant $a > 0$. Furthermore, by the weak maximum principle, L^* is a subsolution to L_∞, i.e. $L_\infty(x) \geq L^* \geq a > 0$ for all $x \in \Omega$.

Finally, by considering the unique classical solution P^* of the auxiliary system

$$-d_P \Delta P^* + \alpha P^* = \beta L_\infty, \quad x \in \Omega, \qquad d_P \partial_\nu P^* = 0, \quad x \in \partial\Gamma,$$

the weak maximum principle shows P^* to be subsolutions, which is bounded below by a positive constants due to the strong maximum principle applied to P^*, i.e. $P_\infty(x) \geq P^*(x) \geq a > 0$ for $x \in \Omega$ and a constant $a > 0$. This finishes the proof of the lower and upper bounds.

To prove the uniqueness of the equilibrium for a given positive mass $M > 0$, we suppose two different equilibria $\boldsymbol{c}_\infty^{(1)} = (L_\infty^{(1)}, P_\infty^{(1)}, \ell_\infty^{(1)}, p_\infty^{(1)})$, $\boldsymbol{c}_\infty = (L_\infty, P_\infty, \ell_\infty, p_\infty)$ as constructed above. Then, obviously the entropy-dissipation of the relative entropy between $\boldsymbol{c}^{(1)}$ and $\boldsymbol{c}_\infty$ vanishes, i.e. $\mathscr{D}(\boldsymbol{c}_\infty^{(1)}, \boldsymbol{c}_\infty) = 0$. Thanks to (14), this implies

$$\frac{L_\infty^{(1)}}{L_\infty} \equiv \frac{P_\infty^{(1)}}{P_\infty} \equiv \frac{\ell_\infty^{(1)}}{\ell_\infty} \equiv \frac{p_\infty^{(1)}}{p_\infty} \equiv k$$

for some constant $k \in \mathbb{R}\backslash\{0\}$. Hence, the conservation law (15) implies $\boldsymbol{c}_\infty^{(1)} \equiv \boldsymbol{c}_\infty$ provided a fixed positive mass $M > 0$.

Remark 2 Note that the existence of a nonnegative equilibrium is proved independently of the space dimension. The positive lower and upper bounds, however, are based on classical maximum principles arguments. Due to the discontinuity of the characteristic function χ_{Γ_2}, we do not get classical solutions but only weak solutions with higher regularity (e.g. $L_\infty \in H^2(\Omega)$, $P_\infty \in H^{3/2}(\Omega)$), which restricts our proof to dimensions $n \leq 3$. The case of higher spatial dimensions remains open.

The proof of Theorem 1 is based on the following crucial Lemma 3, which establishes a so called *entropy entropy-dissipation estimate* and constitutes the key idea of the entropy method, which aims to quantify the entropy dissipation in terms of the relative entropy towards the equilibrium via a functional inequality independent of the flow of a PDE model, see e.g. [5, 6].

In order to prove the entropy entropy-dissipation estimate in Lemma 3 and in particular as a consequence of having to prove an entropy method for the space inhomogeneous equilibria $\boldsymbol{c}_\infty = (L_\infty, P_\infty, \ell_\infty, p_\infty)$, it will be highly convenient to introduce the following abbreviations, weighted quantities and inequalities:

- Norms: $\|\cdot\|_\Omega, \|\cdot\|_\Gamma, \|\cdot\|_{\Gamma_2}$ are the norms in $L^2(\Omega), L^2(\Gamma), L^2(\Gamma_2)$, respectively;
- New variables (weighted deviations around equilibrium values):

$$U = \frac{L - L_\infty}{L_\infty}, \quad V = \frac{P - P_\infty}{P_\infty}, \quad u = \frac{\ell - \ell_\infty}{\ell_\infty}, \quad v = \frac{p - p_\infty}{p_\infty};$$

- New measures: $dL_\infty = L_\infty dx, \quad dP_\infty = P_\infty dx,$

$$dS_{L_\infty} = L_\infty|_\Gamma\, dS, \quad dS_{P_\infty} = P_\infty|_\Gamma\, dS, \quad d\ell_\infty = \ell_\infty dS, \quad dp_\infty = p_\infty dS.$$

- Weighted averages: $\overline{U} = \frac{1}{\int_\Omega dL_\infty} \int_\Omega U dL_\infty, \quad \overline{V} = \frac{1}{\int_\Omega dP_\infty} \int_\Omega V dP_\infty,$

$$\overline{u} = \frac{1}{\int_\Gamma d\ell_\infty} \int_\Gamma \ell d\ell_\infty, \quad \overline{v} = \frac{1}{\int_{\Gamma_2} dp_\infty} \int_{\Gamma_2} p dp_\infty.$$

- Weighted Poincaré Inequalities: The following weighted inequalities hold thanks to the upper and lower bounds of L_∞, P_∞, ℓ_∞ and p_∞ in Lemma 2

$$\int_\Omega |\nabla U|^2 dL_\infty \geq P_L \int_\Omega |U - \overline{U}|^2 dL_\infty, \quad \int_\Omega |\nabla V|^2 dP_\infty \geq P_P \int_\Omega |V - \overline{V}|^2 dP_\infty, \tag{18}$$

$$\int_\Gamma |\nabla_\Gamma u|^2 d\ell_\infty \geq P_\ell \int_\Gamma |u - \overline{u}|^2 d\ell_\infty, \quad \int_{\Gamma_2} |\nabla_{\Gamma_2} v|^2 dp_\infty \geq P_p \int_{\Gamma_2} |v - \overline{v}|^2 dp_\infty. \tag{19}$$

- Weighted Trace Inequalities: Thanks to the lower and upper bounds of $L_\infty \in C(\overline{\Omega})$ and the usual Trace inequality, we have

$$\int_\Omega |\nabla U|^2 dL_\infty \geq T_L \int_\Gamma |U|_\Gamma - \overline{U}|^2 dS_{L_\infty} \tag{20}$$

With respect to the new notations, note that the relative entropy (13), in particular the relative entropy w.r.t. the equilibrium $\boldsymbol{c}_\infty$, i.e. $\mathscr{E}(\boldsymbol{c}|\boldsymbol{c}_\infty)$ can be rewritten as

$$\mathscr{E}(\boldsymbol{c}|\boldsymbol{c}_\infty) = \mathscr{E}(\boldsymbol{c} - \boldsymbol{c}_\infty|\boldsymbol{c}_\infty) + M, \qquad \text{where} \quad \mathscr{E}(\boldsymbol{c}_\infty|\boldsymbol{c}_\infty) = M, \tag{21}$$

and that $\mathscr{E}(\boldsymbol{c} - \boldsymbol{c}_\infty|\boldsymbol{c}_\infty) = \int_\Omega U^2 dL_\infty + \int_\Omega V^2 dP_\infty + \int_\Gamma u^2 d\ell_\infty + \int_{\Gamma_2} v^2 dp_\infty$. Moreover, the entropy dissipation law (14) rewrite as

$$\begin{aligned}
\mathscr{D}(\boldsymbol{c} - \boldsymbol{c}_\infty|\boldsymbol{c}_\infty) &= -\frac{d}{dt}\mathscr{E}(\boldsymbol{c} - \boldsymbol{c}_\infty|\boldsymbol{c}_\infty) \\
&= 2d_L \int_\Omega |\nabla U|^2 dL_\infty + 2d_P \int_\Omega |\nabla V|^2 dP_\infty + 2d_\ell \int_\Gamma |\nabla_\Gamma u|^2 d\ell_\infty + 2d_p \int_{\Gamma_2} |\nabla_{\Gamma_2} v|^2 dp_\infty \\
&\quad + \alpha \int_\Omega |U - V|^2 dP_\infty + \beta \int_\Omega |U - V|^2 dL_\infty + \lambda \int_\Gamma |U|_\Gamma - u|^2 dS_{L_\infty} \\
&\quad + \gamma \int_\Gamma |U|_\Gamma - u|^2 d\ell_\infty + \xi \int_{\Gamma_2} |V|_\Gamma - v|^2 dp_\infty + \sigma \int_{\Gamma_2} |u - v|^2 d\ell_\infty,
\end{aligned} \tag{22}$$

and from (21) it follows readily that $\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)=\mathscr{D}(\boldsymbol{c}|\boldsymbol{c}_\infty)$.

Proof (*Proof of Theorem* 1) The proof of Theorem 1 is a direct consequence of the key functional inequality (23) in the following Lemma 3 and a classical Gronwall argument applied to the entropy dissipation law (22), which is weakly satisfied (in the sense of being integrated in time) by the weak global solutions to system (1) constructed in [11].

Lemma 3 (Entropy entropy-dissipation estimate for system (1)) *Fix a positive initial mass* $M>0$. *Then, for any non-negative measurable functions* $\boldsymbol{c}=(L,P,\ell,p)$ *satisfying the mass conservation*

$$\int_\Omega (L(x)+P(x))dx+\int_\Gamma \ell(x)dS+\int_{\Gamma_2} p(x)dS=M,$$

the entropy entropy-dissipation estimate

$$\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\geq \lambda_0\,\mathscr{E}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty) \tag{23}$$

holds, where $\boldsymbol{c}_\infty$ *is as in Lemma 2 and the constant* $\lambda_0>0$ *can be estimated explicitly.*

Proof Note that $\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)=0$ for all constant states satisfying $U=V=u=v$ while $\mathscr{E}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)=0$ if and only if $U=V=u=v=0$. Hence, the constraint provided by mass conservation law, i.e.

$$\int_\Omega U(x)dL_\infty+\int_\Omega V(x)dP_\infty+\int_\Gamma u(x)d\ell_\infty+\int_{\Gamma_2} v(x)dp_\infty=0,$$

plays a crucial role in inequality (23), which can not hold otherwise.

The proof of this lemma is therefore divided into two steps, where the mass conservation law enters the proof in the first step. At first, we remark that the relative entropy enjoys to following additivity property w.r.t. $\overline{\boldsymbol{c}}=(\overline{U},\overline{V},\overline{u},\overline{v})$

$$\mathscr{E}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)=\mathscr{E}(\boldsymbol{c}-\overline{\boldsymbol{c}}|\boldsymbol{c}_\infty)+\mathscr{E}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)$$

and that the second term on the right hand side is controlled in terms of the entropy dissipation in Step 1, while the first term is controlled in Step 2:

Step 1. First, we prove that there exists an explicit constant $K_0>0$ such that

$$\mathscr{D}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\geq K_0\,\mathscr{E}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty) \tag{24}$$

Indeed, (24) writes explicitly as

$$\alpha|\overline{U}-\overline{V}|^2\int_\Omega dP_\infty+\beta|\overline{U}-\overline{V}|^2\int_\Omega dL_\infty+\lambda|\overline{U}-\overline{u}|^2\int_\Gamma dS_{L_\infty}$$
$$+\gamma|\overline{U}-\overline{u}|^2\int_\Gamma d\ell_\infty+\xi|\overline{V}-\overline{v}|^2\int_{\Gamma_2}dp_\infty+\sigma|\overline{u}-\overline{v}|^2\int_{\Gamma_2}d\ell_\infty$$
$$\geq K_0\Big(\overline{U}^2\int_\Omega dL_\infty+\overline{V}^2\int_\Omega dP_\infty+\overline{u}^2\int_\Gamma d\ell_\infty+\overline{v}^2\int_{\Gamma_2}dp_\infty\Big)$$

under the mass constrain $\overline{U}\int_\Omega dL_\infty+\overline{V}\int_\Omega dP_\infty+\overline{u}\int_\Gamma d\ell_\infty+\overline{v}\int_{\Gamma_2}dp_\infty=0$ and where α, β, λ, γ, σ and ξ denote the positive reaction rate constants of the network Fig. 1.

However, the above inequality is a direct consequence of Lemma 1 applied to the vector of averaged concentrations $(\overline{U},\overline{V},\overline{u},\overline{v})$ after noting that the network of reactions between L, P, ℓ and p (see Fig. 1) is weakly reversible with one strongly connected component. Thus, the constants K_0 can be taken as the corresponding constant $\eta>0$ of Lemma 1.

Step 2. We introduce the following deviations from the averaged values by $\delta_U=U-\overline{U}$, $\delta_V=V-\overline{V}$, $\delta_u=u-\overline{u}$ and $\delta_v=v-\overline{v}$. Note that $\int_\Omega\delta_U dL_\infty=\int_\Omega\delta_V dP_\infty=\int_\Gamma\delta_u d\ell_\infty=\int_{\Gamma_2}\delta_v dp_\infty=0$. Moreover, we can rewrite

$$\mathscr{E}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)=\int_\Omega\delta_U^2dL_\infty+\int_\Omega\delta_v^2dP_\infty+\int_\Gamma\delta_u^2d\ell_\infty+\int_{\Gamma_2}\delta_v^2dp_\infty+\mathscr{E}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty).\tag{25}$$

By using the weighted Poincaré and Trace inequalities (18)–(20), we estimate $\mathscr{D}$ as

$$\begin{aligned}\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\geq{}&d_LP_L\int_\Omega\delta_U^2dL_\infty+d_PP_P\int_\Omega\delta_V^2dP_\infty+d_LT_L\int_\Gamma\delta_U^2dS_{L_\infty}\\&+2d_\ell P_\ell\int_\Gamma\delta_u^2d\ell_\infty+2d_pP_p\int_{\Gamma_2}\delta_v^2dp_\infty\\&+\Big[\alpha\int_\Omega|U-V|^2dP_\infty+\beta\int_\Omega|U-V|^2dL_\infty+\lambda\int_\Gamma|U|_\Gamma-u|^2dS_{L_\infty}\\&+\gamma\int_\Gamma|U|_\Gamma-u|^2d\ell_\infty+\xi\int_{\Gamma_2}|V|_{\Gamma_2}-v|^2dp_\infty+\sigma\int_{\Gamma_2}|u-v|^2d\ell_\infty\Big]\end{aligned}\tag{26}$$

We denote by J_i, $i=1,2,\dots,6$ the last six terms on the right hand side of (26). We have

$$\begin{aligned}J_1=\alpha\int_\Omega|U-V|^2dP_\infty&=\alpha\int_\Omega|\overline{U}-\overline{V}+\delta_U-\delta_V|^2dP_\infty\\&\geq\varepsilon_1\alpha\int_\Omega|\overline{U}-\overline{V}|^2dP_\infty-\frac{2\alpha\varepsilon_1}{1-\varepsilon_1}\Big(\int_\Omega\delta_U^2dP_\infty+\int_\Omega\delta_V^2dP_\infty\Big)\\&\geq\varepsilon_1\alpha\int_\Omega|\overline{U}-\overline{V}|^2dP_\infty-\frac{2\alpha\varepsilon_1}{1-\varepsilon_1}\Big(\Big\|\frac{P_\infty}{L_\infty}\Big\|_{L^\infty(\Omega)}\int_\Omega\delta_U^2dL_\infty+\int_\Omega\delta_V^2dP_\infty\Big)\end{aligned}\tag{27}$$

for all $\varepsilon_1 \in (0,1)$. Similarly, we get

$$
\begin{aligned}
J_2 &\geq \varepsilon_2\beta \int_\Omega |\overline{U}-\overline{V}|^2 dL_\infty - \frac{2\beta\varepsilon_2}{1-\varepsilon_2}\left(\int_\Omega \delta_U^2 dL_\infty + \left\|\frac{L_\infty}{P_\infty}\right\|_{L^\infty(\Omega)} \int_\Omega \delta_V^2 dP_\infty\right),\\
J_3 &\geq \varepsilon_3\lambda \int_\Gamma |\overline{U}-\overline{u}|^2 dS_{L_\infty} - \frac{2\lambda\varepsilon_3}{1-\varepsilon_3}\left(\int_\Gamma (\delta_U|_\Gamma)^2\, dS_{L_\infty} + \left\|\frac{L_\infty}{\ell_\infty}\right\|_{L^\infty(\Gamma)} \int_\Gamma \delta_u^2 d\ell_\infty\right),\\
J_4 &\geq \varepsilon_4\gamma \int_\Gamma |\overline{U}-\overline{u}|^2 d\ell_\infty - \frac{2\gamma\varepsilon_4}{1-\varepsilon_4}\left(\left\|\frac{\ell_\infty}{L_\infty}\right\|_{L^\infty(\Gamma)} \int_\Gamma (\delta_U|_\Gamma)^2\, dS_{L_\infty} + \int_\Gamma \delta_u^2 d\ell_\infty\right),\\
J_6 &\geq \varepsilon_6\sigma \int_{\Gamma_2} |\overline{u}-\overline{v}|^2 d\ell_\infty - \frac{2\sigma\varepsilon_6}{1-\varepsilon_6}\left(\int_{\Gamma_2} \delta_u^2 d\ell_\infty + \left\|\frac{\ell_\infty}{p_\infty}\right\|_{L^\infty(\Gamma_2)} \int_{\Gamma_2} \delta_v^2 dp_\infty\right),
\end{aligned}
\tag{28}
$$

with $\varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_6 \in (0,1)$. For J_5, the lack of continuity of P_∞ at the boundary Γ_2 prevents a similar estimate as above, since it is unclear how to control the term $\|\frac{p_\infty}{P_\infty}\|_{L^\infty(\Gamma_2)}$. However, the weak reversibility of system (1) allows first to estimate $J_5 \geq 0$ and then use the triangle inequality to have

$$
\begin{aligned}
&\frac{1}{2}\left(\varepsilon_2\beta \int_\Omega |\overline{U}-\overline{V}|^2 dL_\infty + \varepsilon_3\lambda \int_\Gamma |\overline{U}-\overline{u}|^2 dS_{L_\infty} + \varepsilon_6\sigma \int_{\Gamma_2} |\overline{u}-\overline{v}|^2 d\ell_\infty\right)\\
&\geq \frac{1}{6}\min\left\{\varepsilon_2\beta \int_\Omega dL_\infty; \varepsilon_3\lambda \int_\Gamma dS_{L_\infty}; \varepsilon_6\sigma \int_{\Gamma_2} d\ell_\infty\right\} |\overline{V}-\overline{v}|^2 =: \omega|\overline{V}-\overline{v}|^2.
\end{aligned}
\tag{29}
$$

By combining (26)–(28) and by choosing $\varepsilon_1, \ldots, \varepsilon_6$ small enough (for instance in order to ensure that for some $\eta_1 > 0$

$$
d_L P_L - \frac{2\alpha\varepsilon_1}{1-\varepsilon_1}\left\|\frac{P_\infty}{L_\infty}\right\|_{L^\infty(\Omega)} - \frac{2\beta\varepsilon_2}{1-\varepsilon_2} \geq \eta_1 > 0
$$

with $\|\frac{P_\infty}{L_\infty}\|_{L^\infty(\Omega)} \leq \frac{A}{a}$ by Lemma 2), we can estimate $\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)$ below as

$$
\begin{aligned}
\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty) &\geq \frac{1}{2}\min\{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,2\omega,\varepsilon_6\}\mathscr{D}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\\
&\quad + \eta_1\int_\Omega \delta_U^2 dL_\infty + \eta_2\int_\Omega \delta_V^2 dP_\infty + \eta_3\int_\Gamma \delta_u^2 d\ell_\infty + \eta_4\int_{\Gamma_2} \delta_v^2 dp_\infty
\end{aligned}
\tag{30}
$$

where ω is defined in (29). Hence, by using (24) and (25), we have

$$\begin{aligned}\mathscr{D}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty) &\geq \frac{1}{2}K_0\min\{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,2\omega,\varepsilon_6\}\mathscr{E}(\overline{\boldsymbol{c}}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\\ &\quad+\min_{i=1..4}\{\eta_i\}\left(\int_\Omega \delta_U^2 dL_\infty+\int_\Omega \delta_V^2 dP_\infty+\int_\Gamma \delta_u^2 d\ell_\infty+\int_{\Gamma_2}\delta_v^2 dp_\infty\right)\\ &\geq \lambda_0\,\mathcal{E}(\boldsymbol{c}-\boldsymbol{c}_\infty|\boldsymbol{c}_\infty)\end{aligned}$$

with $\lambda_0=\frac{1}{2}\min\{2K_0\varepsilon_1,2K_0\varepsilon_2,2K_0\varepsilon_3,2K_0\varepsilon_4,4K_0\omega,2K_0\varepsilon_6,\eta_1,\eta_2,\eta_3,\eta_4\}$.

4 Proof of Theorem 2

In this section, we denote by $\mathbf{u}=(u_0,u_1,\dots,u_7)$ and $\mathbf{w}=(w_0,w_1,\dots,w_7)$. Moreover, we define (in the spirit of the relative entropy (9) of first-order reaction networks) the relative entropy functional associated to (3):

$$\mathscr{E}(\mathbf{u}|\mathbf{w})=\int_{\Omega_{\rm cyt}}\left(\frac{|u_0|^2}{w_0}+\frac{|u_1|^2}{w_1}\right)dx+|\Omega_{\rm nuc}|\sum_{i=2}^{7}\frac{|u_i|^2}{w_i},\tag{31}$$

which dissipates (analog to the entropy dissipation (10)) due to the following entropy dissipation functional:

$$\begin{aligned}\mathscr{D}(\mathbf{u}|\mathbf{w})&=-\frac{d}{dt}\mathscr{E}(\mathbf{u}|\mathbf{w})=D\int_{\Omega_{\rm cyt}}w_0\left|\nabla\frac{u_0}{w_0}\right|^2dx+D\int_{\Omega_{\rm cyt}}w_1\left|\nabla\frac{u_1}{w_1}\right|^2dx\\ &+r_{\rm act}\,p_{JAK}\int_{\partial\Omega_{\rm cyt}}w_0\left[\frac{u_0}{w_0}\Big|_{\partial\Omega_{\rm cyt}}-\frac{u_1}{w_1}\Big|_{\partial\Omega_{\rm cyt}}\right]^2dS+r_{\rm imp2}\int_{\partial\Omega_{\rm nuc}}w_1\left[\frac{u_1}{w_1}\Big|_{\partial\Omega_{\rm nuc}}-\frac{u_3}{w_3}\right]^2dS\\ &+r_{\rm imp}\int_{\partial\Omega_{\rm nuc}}w_0\left[\frac{u_0}{w_0}\Big|_{\partial\Omega_{\rm nuc}}-\frac{u_2}{w_2}\right]^2dS+r_{\rm exp}w_2\int_{\partial\Omega_{\rm nuc}}\left[\frac{u_0}{w_0}\Big|_{\partial\Omega_{\rm nuc}}-\frac{u_2}{w_2}\right]^2dS\\ &+r_{\rm delay}\,w_7\left[\frac{u_7}{w_7}-\frac{u_2}{w_2}\right]^2+\sum_{i=3}^{6}r_{\rm delay}\,w_i\left[\frac{u_i}{w_i}-\frac{u_{i+1}}{w_{i+1}}\right]^2.\end{aligned}\tag{32}$$

Lemma 4 (Existence of a unique positive equilibrium of (3)) *For any positive initial mass $M>0$, system (3) possesses a unique equilibrium $\mathbf{u}_\infty=(u_{0,\infty},\dots,u_{7,\infty})$ satisfying the mass conservation (5), i.e.*

$$\int_{\Omega_{\rm cyt}}(u_{0,\infty}(x)+u_{1,\infty}(x))dx+|\Omega_{\rm nuc}|\sum_{i=2}^{7}u_{i,\infty}=M>0.$$

Moreover, $u_{2,\infty}, \ldots, u_{7,\infty}$ *are positive and* $u_{0,\infty}, u_{1,\infty} \in C(\overline{\Omega}_{\rm cyt}) \cap C^2(\Omega_{\rm cyt})$ *satisfy*

$$0 < b \le u_{0,\infty}(x), u_{1,\infty}(x) \le B < +\infty, \qquad \text{for all } x \in \Omega_{\rm cyt}$$

for some constants $0 < b \le B \le +\infty$.

Proof From (3c), we easily see that

$$u_{3,\infty} = u_{4,\infty} = u_{5,\infty} = u_{6,\infty} = u_{7,\infty} = \frac{r_{\rm imp2}}{r_{\rm delay}|\partial\Omega_{\rm nuc}|}\int_{\partial\Omega_{\rm nuc}} u_{1,\infty}(y)dS, \tag{33}$$

$$\begin{aligned} u_{2,\infty} &= \frac{r_{\rm delay}}{r_{\rm exp}}u_{7,\infty} + \frac{r_{\rm imp}}{r_{\rm exp}|\partial\Omega_{\rm nuc}|}\int_{\partial\Omega_{\rm nuc}} u_{0,\infty}(y)dS \\ &= \frac{r_{\rm imp2}}{r_{\rm exp}|\partial\Omega_{\rm nuc}|}\int_{\partial\Omega_{\rm nuc}} u_{1,\infty}(y)dS + \frac{r_{\rm imp}}{r_{\rm exp}|\partial\Omega_{\rm nuc}|}\int_{\partial\Omega_{\rm nuc}} u_{0,\infty}(y)dS. \end{aligned} \tag{34}$$

It thus remains to solve the following non-local elliptic system for $u_{0,\infty}$ and $u_{1,\infty}$,

$$\begin{cases} D\Delta u_{0,\infty}(x) = 0, & x \in \Omega_{\rm cyt}, \\ D\partial_{n_1}u_{0,\infty}(y) = -\frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|}p_{JAK}u_{0,\infty}(y), & y \in \partial\Omega_{\rm cyt}, \\ D\partial_{n_2}u_{0,\infty}(y) = -\frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|}u_{0,\infty}(y) \\ \quad + \frac{1}{|\partial\Omega_{\rm nuc}|}\left(\frac{r_{\rm imp2}}{|\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}} u_{1,\infty}\,dS + \frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}} u_{0,\infty}\,dS\right), & y \in \partial\Omega_{\rm nuc}, \end{cases} \tag{35a}$$

$$\begin{cases} D\Delta u_{1,\infty}(x) = 0, & x \in \Omega_{\rm cyt}, \\ D\partial_{n_1}u_{1,\infty}(y) = \frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|}p_{JAK}u_{0,\infty}(y), & y \in \partial\Omega_{\rm cyt}, \\ D\partial_{n_2}u_{1,\infty}(y) = -\frac{r_{\rm imp}2}{|\partial\Omega_{\rm nuc}|}u_{1,\infty}(y), & y \in \partial\Omega_{\rm nuc}. \end{cases} \tag{35b}$$

subject to the constraint, which follows from the mass conservation, (33) and (34),

$$\begin{aligned} &\int_{\Omega_{\rm cyt}} (u_{0,\infty} + u_{1,\infty})dx + \frac{r_{\rm imp}}{r_{\rm exp}|\partial\Omega_{\rm nuc}|}\int_{\partial\Omega_{\rm nuc}} u_{0,\infty}dS \\ &+ \left(5\frac{r_{\rm imp2}}{r_{\rm delay}|\partial\Omega_{\rm nuc}|}\frac{r_{\rm imp2}}{r_{\rm exp}|\partial\Omega_{\rm nuc}|}\right)\int_{\partial\Omega_{\rm nuc}} u_{1,\infty}dS = M. \end{aligned} \tag{36}$$

By considering an auxiliary system as follows

$$\begin{cases} D\Delta u_0 = 0, & x \in \Omega_{\rm cyt}, \\ D\partial_{n_1} u_0 + \frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|} p_{JAK} u_0 = 0, & y \in \partial\Omega_{\rm cyt}, \\ D\partial_{n_2} u_0 + \frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|} u_0 = \frac{1}{|\partial\Omega_{\rm nuc}|}\left(\frac{r_{\rm imp2}}{|\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}} \widehat{u_1}\, dS + \frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|}\int\limits_{\partial\Omega_{\rm nuc}} \widehat{u_0}\, dS\right), & y \in \partial\Omega_{\rm nuc}, \\ D\Delta u_1 = 0, & x \in \Omega_{\rm cyt}, \\ D\partial_{n_1} u_1 = \frac{r_{\rm act}}{|\partial\Omega_{\rm cyt}|} p_{JAK} \widehat{u_0}(y), & y \in \partial\Omega_{\rm cyt}, \\ D\partial_{n_2} u_1 + \frac{r_{\rm imp}2}{|\partial\Omega_{\rm nuc}|} u_1 = 0, & y \in \partial\Omega_{\rm nuc} \end{cases} \tag{37}$$

we can use similar arguments in Lemma 2, namely define for non-negative $(\widehat{u_0}, \widehat{u_1})$ an operator $\mathfrak{L} : (\widehat{u_0}, \widehat{u_1}) \mapsto (u_0, u_1)$ as the solution to (37). Then, the existence of a nonnegative weak solution $(u_{0,\infty}, u_{1,\infty}) \in H^1(\Omega_{\rm cyt}) \times H^1(\Omega_{\rm cyt})$ to (35a) and (35b) follows from the Schauder fixed point theorem applied to $\mathfrak{L}$. In return, this implies nonnegative equilibria $(u_{0,\infty}, \ldots, u_{7,\infty})$ to system (3).

By applying standard bootstrap arguments to (35a) and (35b), we obtain that $(u_{0,\infty}, u_{1,\infty})$ is in fact a classical solution, namely $u_{0,\infty}, u_{1,\infty} \in C(\overline{\Omega}_{\rm cyt}) \cap C^2(\Omega_{\rm cyt})$. Hence, $u_{0,\infty}$ and $u_{1,\infty}$ are uniformly bounded above by a positive constant thanks to the compactness of $\overline{\Omega}_{\rm cyt}$. We now prove the strict positivity of $u_{0,\infty}$ and $u_{1,\infty}$. First we show that $u_{0,\infty}$ is not identical to zero on $\partial\Omega_{\rm cyt}$. Indeed, assume the reserve is true, we then obtain from (35b) that $u_{1,\infty} \equiv 0$. It follows then from (35a) and the strong maximum principle that $u_{0,\infty} \equiv 0$. This violates the mass conservation (36). Therefore, $u_{0,\infty} \geq 0$ is not identical to zero on $\partial\Omega_{\rm cyt}$ and we get from (35b) and the maximum principle that $u_{1,\infty}(x) > 0$ for all $x \in \overline{\Omega}_{\rm cyt}$ which leads to a strictly positive lower bound of $u_{1,\infty}$. The strict positivity of $u_{0,\infty}$ also follows from $u_{0,\infty} \in C(\overline{\Omega}_{\rm cyt})$ and strong maximum principle since the second boundary condition in (35a) implies

$$D\partial_n u_{0,\infty}(y) + \frac{r_{\rm imp}}{|\partial\Omega_{\rm nuc}|} u_{0,\infty}(y) \geq \frac{r_{\rm imp2}}{|\partial\Omega_{\rm nuc}|^2} \int_{\partial\Omega_{\rm nuc}} u_{1,\infty} dS$$

and $u_{1,\infty}$ is strictly positive. The uniqueness of the equilibrium can be proved similar to Lemma 2 thanks to the entropy structures (31) and (32). We omit here the proof.

As in the previous section, we introduce the following short notations

- New variables: $v_i = \frac{u_i - u_{i,\infty}}{u_{i,\infty}}$, for $i = 0, 1, \ldots, 7$.
- New measures: $du_{0,\infty} = u_{0,\infty} dx$, $du_{1,\infty} = u_{1,\infty} dx$, $d\sigma_{u_{0,\infty}} = u_{0,\infty}|_{\partial\Omega_{\rm cyt}}\, dS$,

$$d\sigma_{u_{1,\infty}} = u_{1,\infty}|_{\partial\Omega_{\rm cyt}}\, dS, \quad dS_{u_{0,\infty}} = u_{0,\infty}|_{\partial\Omega_{\rm nuc}}\, dS, \quad dS_{u_{1,\infty}} = u_{1,\infty}|_{\partial\Omega_{\rm nuc}}\, dS.$$

- New parameters: $\alpha = r_{\rm act}$, $\beta = p_{JAK}$, $\gamma = r_{\rm imp}$, $\sigma = r_{\rm imp2}$, $\kappa = r_{\rm exp}$, $\xi = r_{\rm delay}$.
- Weighted averages:

$$\overline{v}_0 = \frac{1}{\int_{\Omega_{\rm nuc}} du_{0,\infty}} \int_{\Omega_{\rm nuc}} v_0(t,x) du_{0,\infty}, \qquad \overline{v}_1 = \frac{1}{\int_{\Omega_{\rm nuc}} du_{1,\infty}} \int_{\Omega_{\rm nuc}} v_1(t,x) du_{1,\infty}.$$

Then, analog to (21)–(22), we can rewrite the relative entropy

$$\mathscr{E}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) = \int_{\Omega_{\rm cyt}} |v_0|^2 du_{0,\infty} + \int_{\Omega_{\rm cyt}} |v_1|^2 du_{1,\infty} + |\Omega_{\rm nuc}| \sum_{i=2}^{7} u_{i,\infty} |v_i|^2, \quad (38)$$

and its entropy dissipation functional $\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) = -\frac{d}{dt}\mathscr{E}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty)$ as

$$\begin{aligned}
\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) &= D\int_{\Omega_{\rm cyt}} |\nabla v_0|^2 \, du_{0,\infty} + D\int_{\Omega_{\rm cyt}} |\nabla v_1|^2 \, du_{1,\infty} + \sum_{i=3}^{6} \xi u_{i,\infty} \left[v_i - v_{i+1}\right]^2 \\
&+ \alpha\beta \int_{\partial\Omega_{\rm cyt}} \left[v_0|_{\partial\Omega_{\rm cyt}} - v_1|_{\partial\Omega_{\rm cyt}}\right]^2 d\sigma_{u_{0,\infty}} + \sigma \int_{\partial\Omega_{\rm nuc}} \left[v_1|_{\partial\Omega_{\rm nuc}} - v_3\right]^2 dS_{u_{1,\infty}} \\
&+ \gamma \int_{\partial\Omega_{\rm nuc}} \left[v_0|_{\partial\Omega_{\rm nuc}} - v_2\right]^2 dS_{u_{0,\infty}} + \kappa u_{2,\infty} \int_{\partial\Omega_{\rm nuc}} [v_0|_{\partial\Omega_{\rm nuc}} - v_2]^2 dS + \xi u_{7,\infty} [v_7 - v_2]^2.
\end{aligned} \quad (39)$$

Finally, the mass conservation law rewrite as

$$\overline{v}_0 \int_{\Omega_{\rm cyt}} du_{0,\infty} + \overline{v}_1 \int_{\Omega_{\rm cyt}} du_{1,\infty} + |\Omega_{\rm nuc}| \sum_{i=2}^{7} u_{i,\infty} v_i = 0. \quad (40)$$

Proof (Proof of Theorem 2) The proof of Theorem 2 is a direct consequence of the functional inequality (41) in the following Lemma 5 and a Gronwall argument applied to the entropy dissipation law (39), which is weakly satisfied by the global solutions constructed in [13].

Lemma 5 (Entropy entropy-dissipation estimate for system (3)) *Fix a positive initial mass $M > 0$. Then, for any measurable state $\mathbf{u} = (u_0, \ldots, u_7)$ satisfying the mass conservation $\int_{\Omega_{\rm cyt}} (u_0(x) + u_1(x))dx + |\Omega_{\rm nuc}| \sum_{i=2}^{7} u_i = M$ the entropy entropy-dissipation inequality*

$$\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) \geq \lambda_1 \, \mathscr{E}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) \quad (41)$$

holds where $\mathbf{u}_\infty$ is as in Lemma 4 and the constant $\lambda_1 > 0$ can be estimated explicitly.

Proof The proof of this lemma is similar to that of Lemma 3. First, it is straightforward to verify that the relative entropy satisfies to following additivity property:

$$\mathscr{E}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) = \mathscr{E}(\mathbf{u}-\overline{\mathbf{u}}|\mathbf{u}_\infty) + \mathscr{E}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty) \quad (42)$$

where

$$\mathscr{E}(\mathbf{u}-\overline{\mathbf{u}}|\mathbf{u}_\infty) = \int_{\Omega_{\rm cyt}} |v_0-\overline{v}_0|^2 du_{0,\infty} + \int_{\Omega_{\rm cyt}} |v_1-\overline{v}_1|^2 du_{1,\infty},$$

$$\mathscr{E}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty) = \overline{v}_0^2 \int_{\Omega_{\rm cyt}} du_{0,\infty} + \overline{v}_1^2 \int_{\Omega_{\rm cyt}} du_{1,\infty} + |\Omega_{\rm nuc}| \sum_{i=2}^{7} u_{i,\infty}|v_i|^2.$$

Step 1. Thanks to the mass conservation law (40), it follows from Lemma 1 and the weak reversibility of the reaction network Fig. 2 that for an explicit constant $L_0 > 0$

$$\mathscr{D}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty) \geq L_0\, \mathscr{E}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty). \tag{43}$$

Step 2. The term $\mathscr{E}(\mathbf{u}-\overline{\mathbf{u}}|\mathbf{u}_\infty)$ can be controlled in terms of the entropy dissipation by using the weighted Poincaré inequalities as follow

$$\begin{aligned}\frac{1}{2}\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) &\geq D\int_{\Omega_{\rm cyt}} |\nabla v_0|^2 du_{0,\infty} + D\int_{\Omega_{\rm cyt}} |\nabla v_1|^2 du_{1,\infty} \\ &\geq L_1\Big(\int_{\Omega_{\rm cyt}} |v_0-\overline{v}_0|^2 du_{0,\infty} + \int_{\Omega_{\rm cyt}} |v_1-\overline{v}_1|^2 du_{1,\infty}\Big) = L_1\,\mathscr{E}(\mathbf{u}-\overline{\mathbf{u}}|\mathbf{u}_\infty),\end{aligned} \tag{44}$$

where $L_1 = D\min\{P_{u_{0,\infty}}, P_{u_{1,\infty}}\}$ depends on D and the corresponding weighted Poincaré constants. Next, we prove for some constant $L_2 > 0$ that

$$\frac{1}{2}\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) \geq L_2\,\mathscr{D}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty), \tag{45}$$

which yields the desired estimate (41) with $\lambda_1 = \min\{L_1, L_0L_2\}$ by combining (42), (43), (44) and (45). To prove (45), we will use the μ-weighted trace inequalities

$$\int_{\Omega_{\rm cyt}} |\nabla f|^2 d\mu \geq T_{\sigma_\mu}\int_{\partial\Omega_{\rm cyt}} |f-\overline{f}|^2\, d\sigma_\mu, \quad \int_{\Omega_{\rm nuc}} |\nabla f|^2 d\mu \geq T_{S_\mu}\int_{\partial\Omega_{\rm nuc}} |f-\overline{f}|^2\, dS_\mu,$$

where $\overline{f} = \frac{1}{\int_{\Omega_{\rm cyt}} d\mu}\int_{\Omega_{\rm cyt}} f\, d\mu$ to estimate, using the triangle inequality,

$$\begin{aligned}\mathscr{D}(\mathbf{u}-\mathbf{u}_\infty|\mathbf{u}_\infty) &\geq DT_\mu \fint_{\partial\Omega_{\rm cyt}} |v_0-\overline{v}_0|^2 d\sigma_{u_{0,\infty}} + DT_\mu \fint_{\partial\Omega_{\rm cyt}} |v_1-\overline{v}_1|^2 d\sigma_{u_{1,\infty}} + \xi u_{7,\infty}|v_7-v_2|^2 \\ &+ \alpha\beta\int_{\partial\Omega_{\rm cyt}} |v_0-v_1|^2 d\sigma_{u_{0,\infty}} + DT_\mu\int_{\partial\Omega_{\rm nuc}} |v_0-\overline{v}_0|^2 dS_{u_{0,\infty}} + \gamma\int_{\partial\Omega_{\rm nuc}} |v_0-v_2|^2 dS_{u_{0,\infty}} \\ &+ DT_\mu\int_{\partial\Omega_{\rm nuc}} |v_1-\overline{v}_1|^2 dS_{u_{1,\infty}} + \sigma\int_{\partial\Omega_{\rm nuc}} |v_1-v_3|^2 dS_{u_{1,\infty}} + \sum_{i=3}^{6}\xi u_{i,\infty}\big[v_i-v_{i+1}\big]^2 \\ &\geq C_1|\overline{v}_0-\overline{v}_1|^2\int_{\partial\Omega_{\rm cyt}} d\sigma_{u_{0,\infty}} + C_2|\overline{v}_0-v_2|^2\int_{\partial\Omega_{\rm nuc}} dS_{u_{0,\infty}} + C_3|\overline{v}_1-v_3|^2\int_{\partial\Omega_{\rm nuc}} dS_{u_{1,\infty}} \\ &+ \sum_{i=3}^{6}\xi u_{i,\infty}|v_i-v_{i+1}|^2 \geq 2L_2\,\mathscr{D}(\overline{\mathbf{u}}-\mathbf{u}_\infty|\mathbf{u}_\infty)\end{aligned}$$

where $T_\mu = \min\{T_{\sigma_{u_{0,\infty}}}, T_{\sigma_{u_{1,\infty}}}, T_{S_{u_{0,\infty}}}, T_{S_{u_{1,\infty}}}\}, C_1 = \frac{DT_\mu}{3} \min\left\{1; \alpha\beta; \min_{y\in\partial\Omega_{\text{cyt}}} \frac{u_{1,\infty}(y)}{u_{0,\infty}(y)}\right\}$, $C_2 = \frac{\min\{DT_\mu;\gamma\}}{2}$, $C_3 = \frac{\min\{DT_\mu;\sigma\}}{2}$ and $L_2 = \frac{1}{2}\min\left\{\frac{C_1}{\alpha\beta}; \frac{C_2}{2\gamma}; \frac{C_2\int_{\partial\Omega_{\text{nuc}}} dS_{u_{0,\infty}}}{2\kappa u_{2,\infty}\int_{\partial\Omega_{\text{nuc}}} dS}; \frac{C_3}{\sigma}; 1\right\}$ which can be computed explicitly, for instance

$$
\begin{aligned}
&DT_\mu \int_{\partial\Omega_{\text{cyt}}} |v_0 - \overline{v}_0|^2 d\sigma_{u_{0,\infty}} + DT_\mu \int_{\partial\Omega_{\text{cyt}}} |v_1 - \overline{v}_1|^2 d\sigma_{u_{1,\infty}} + \alpha\beta \int_{\partial\Omega_{\text{cyt}}} |v_0 - v_1|^2 d\sigma_{u_{0,\infty}} \\
&\geq \frac{1}{3} DT_\mu \min\left\{1; \min_{y\in\partial\Omega_{\text{cyt}}} \frac{u_{1,\infty}(y)}{u_{0,\infty}(y)}; \alpha\beta\right\} \int_{\partial\Omega_{\text{cyt}}} (v_0 - \overline{v}_0 + \overline{v}_1 - v_1 - v_0 + v_1)^2 d\sigma_{u_{0,\infty}} \\
&= C_1 |\overline{v}_0 - \overline{v}_1|^2 \int_{\partial\Omega_{\text{cyt}}} d\sigma_{u_{0,\infty}}.
\end{aligned}
$$

Acknowledgements The second author was supported by International Research Training Group IGDK 1754. This work has partially been supported by NAWI Graz.

References

1. Berestycki, H., Coulon, A.C., Roquejoffre, J.M., Rossi, L.: The effect of a line with nonlocal diffusion on Fisher-KPP propagation. Math. Models Methods Appl. Sci. **25**(13), 2519–2562 (2015)
2. Berestycki, H., Roquejoffre, J.M., Rossi, L.: The influence of a line with fast diffusion on Fisher-KPP propagation. J. Math. Biol. **66**(4-5), 743–766 (2013)
3. Berestycki, H., Roquejoffre, J.M., Rossi, L.: Fisher–KPP propagation in the presence of a line: further effects. Nonlinearity **26**(9), 2623 (2013)
4. Betschinger, J., Mechtler, K., Knoblich, J.A.: The Par complex directs asymmetric cell division by phosphorylating the cytoskeletal protein Lgl. Nature **422**, 326–329 (2003)
5. Desvillettes, L., Fellner, K.: Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations. J. Math. Anal. Appl. **319**(1), 157–176 (2006)
6. Desvillettes, L. Fellner, K.: Entropy methods for reaction-diffusion systems: degenerate diffusion. Discrete Contin. Dyn. Syst. (Supplements Special). Dynamical Systems and Differential Equations. Proceedings of the 6th AIMS International Conference, 304–312 (2007)
7. Desvillettes, L., Fellner, K., Tang, B.Q.: Trend to equilibrium for reaction-diffusion systems arising from complex balanced chemical reaction networks. SIAM J. Math. Anal. **49**(4), 2666–2709 (2017). arXiv:1604.04536
8. Evans, L.C.: American Mathematical Society, Partial Differential Equations (2010)
9. Fellner, K., Latos, E., Tang, B.Q.: Well-posedness and exponential equilibration of a volume-surface reaction-diffusion system with nonlinear boundary coupling. to appear in Ann. Inst. H. Poincar Anal. Non Linéaire
10. Fellner, K., Prager, W., Tang, B.Q.: The entropy method for reaction-diffusion systems without detailed balance: first order chemical reaction works. Kinet. Relat. Models **10**(4), 1055–1087 (2017). arXiv:1504.08221
11. Fellner, K., Rosenberger, S., Tang, B.Q.: Quasi-steady-state approximation and numerical simulation for a volume-surface reaction-diffusion system. Commun. Math. Sci. **14**(6), 1553–1580 (2016)
12. Fellner, K., Tang, B.Q.: Explicit exponential convergence to equilibrium for nonlinear reaction-diffusion systems with detailed balance condition. Nonlinear Anal. **159**, 145–180 (2017). arXiv:1601.05992
13. Friedmann, E., Neumann, R., Rannacher, R.: Well-posedness of a linear spatio-temporal model of the JAK2/STAT5 signaling pathway. Comm. Math. Anal. **15**, 76–102 (2013)

14. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, Berlin (1977)
15. Glitzky, A., Gröger, K., Hünlich, R.: Free energy and dissipation rate for reaction-diffusion processes of electrically charged species. Appl. Anal. **60**, 201–217 (1996)
16. Glitzky, A., Mielke, A.: A gradient structure for systems coupling reaction-diffusion effects in bulk and interfaces. Z. Angew. Math. Phys. **64**(1), 29–52 (2013)
17. Henneke, F., Tang, Bao Q.: Fast reaction limit of a volume-surface reaction-diffusion system towards a heat equation with dynamical boundary conditions. Asymptot. Anal. **98**(4), 325–339 (2016)
18. Horn, F.J.M., Jackson, R.: General mass action kinetics. Arch. Ration. Mech. Anal. **47**, 81–116 (1972)
19. Kwon, Y.-I., Derby, J.J.: Modeling the coupled effects of interfacial and bulk phenomena during solution crystal growth. J. Cryst. Growth **230**, 328–335 (2001)
20. Mayer, B., Emery, G., Berdnik, D., Wirtz-Peitz, F., Knoblich, J.: Quantitative analysis of protein dynamics during asymmetric cell division. Curr. Biol. **15**, 1847–54 (2005)
21. Mielke, A.: Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. Discret. Contin. Dyn. Syst. Ser. S **6**(2), 479–499 (2013)
22. Mielke, A., Haskovec, J., Markowich, P.A.: On uniform decay of the entropy for reaction-diffusion systems. J. Dyn. Differ. Equ. **27**, 897–928 (2015)
23. Pao, C.V.: Nonlinear Parabolic and Elliptic Equations. Plenum Press, New York (1992)
24. Taylor, M.: Partial Differential Equations I: Basic Theory. Applied Mathematical Sciences. Springer, New York (1996)
25. Wirtz-Peitz, F., Nashimura, T., Knoblich, J.: Linking cell cycle to asymmetric division: Aurora-A phosphorylates the Par complex to regulate numb localization. Cell **135**, 161–173 (2008)

Equilibrium Fluctuations for the Slow Boundary Exclusion Process

Tertuliano Franco, Patrícia Gonçalves and Adriana Neumann

Abstract We prove that the equilibrium density fluctuations of the symmetric simple exclusion process in contact with slow boundaries is given by an Ornstein–Uhlenbeck process with Dirichlet, Robin or Neumann boundary conditions depending on the range of the parameter that rules the slowness of the boundaries.

Keywords Equilibrium density fluctuations · Exclusion process
Slow boundaries · Ornstein–Uhlenbeck process

1 Introduction

The study of nonequilibrium behavior of interacting particle systems is one of the most challenging problems in the field and it has only been completely solved in very particular cases. The toy model for the study of a system in a nonequilibrium scenario is the symmetric simple exclusion process (SSEP) whose dynamics is rather simple to explain and it already captures many features of more complicated systems.

T. Franco
UFBA, Instituto de Matemática, Campus de Ondina, Av. Adhemar de Barros, S/N. CEP, 40170-110 Salvador, Brazil
e-mail: tertu@ufba.br

P. Gonçalves (✉)
Center for Mathematical Analysis, Geometry and Dynamical Systems, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
e-mail: patricia.goncalves@math.tecnico.ulisboa.pt

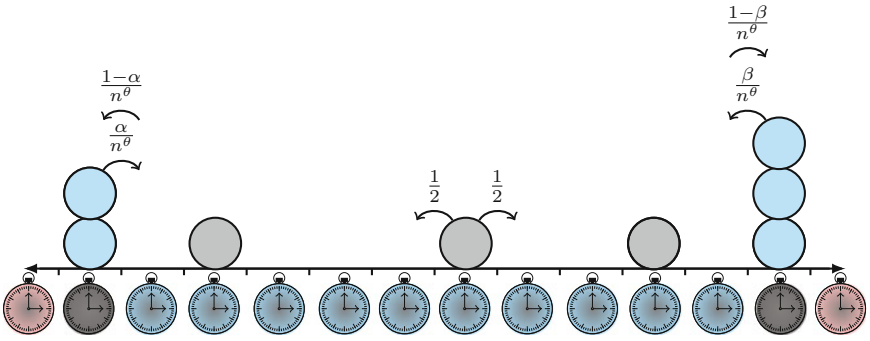
A. Neumann
UFRGS, Instituto de Matemática, Campus do Vale, Av. Bento Gonçalves, 9500, 91509-900 Porto Alegre, Brazil
e-mail: aneumann@mat.ufrgs.br

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_9

The dynamics of this model can be described as follows. We fix a scaling parameter n and we consider the SSEP evolving on the discrete space $\Sigma_n = \{1, \dots, n-1\}$ to which we call the bulk. To each pair of bonds $\{x, x+1\}$ with $x = 1, \dots, n-2$ we associate a Poisson process $N_{x,x+1}(t)$ of rate 1. Now we artificially add two end points at the bulk, namely, we add the sites $x = 0$ and $x = n$ and we superpose the exclusion dynamics with a Glauber dynamics which has only effect at the boundary points of the bulk, namely at the sites $x = 1$ and $x = n-1$. For that purpose, we add extra Poisson processes at the bonds $\{0, 1\}$ and $\{n-1, n\}$. In each one of these bonds there are two Poisson processes: $N_{0,1}(t)$ with parameter $\alpha n^{-\theta}$, $N_{1,0}(t)$ with parameter $(1-\alpha)n^{-\theta}$, $N_{n-1,n}(t)$ with parameter $\beta n^{-\theta}$ and $N_{n,n-1}(t)$ with parameter $(1-\beta)n^{-\theta}$. All the Poisson processes are independent. Above $\alpha, \beta \in (0, 1)$ and $\theta \geq 0$ is a parameter that rules the slowness of the boundary dynamics. Below in the figure we colored the Poisson clocks associated to the bonds in the bulk in the blue color, while the Poisson clocks associated to the bonds at the boundary are colored in the gray and pink colors, to emphasize that they have different rates.

Now that the clocks are fixed we can explain the dynamics. For that purpose, initially we place particles in the bulk according to some probability measure and we denote this configuration of particles and holes by $\eta = (\eta(1), \dots, \eta(n-1))$, so that for $x \in \Sigma_n$, $\eta(x) = 1$ if there is a particle at the site x and $\eta(x) = 0$ if the site x is empty. Now, if a clock rings for a bond $\{x, x+1\}$ in the bulk, then we exchange the coordinates x and $x+1$ of η, that is we exchange $\eta(x)$ with $\eta(x+1)$ at rate 1. If the clock rings for the bond at the boundary as, for example, from the Poisson process $N_{0,1}(t)$ then a particle gets into the bulk through the site 1 at rate $\alpha n^{-\theta}$ if and only if there is no particle at the site 1, otherwise nothing happens. If the clock rings from the Poisson process $N_{1,0}(t)$ and there is a particle at the site 1, then it exits the bulk from the site 1 at rate $(1-\alpha)n^{-\theta}$. Note that the higher the value of θ the slower is the dynamics at the boundaries. For a display of the description above, see the figure below.



The dynamics just described is Markovian and can be completely characterized in terms of its infinitesimal generator given below in (1). We note that the space state of this Markov process is $\Omega_n := \{0, 1\}^{\Sigma_n}$. Observe that the bulk dynamics preserves the number of particles and our interest is to describe the space-time evolution of

this conserved quantity as a solution of some partial differential equation called the *hydrodynamic equation*.

Note that for the choice $\alpha = \beta = \rho$ a simple computation shows that the Bernoulli product measure of parameter ρ given by: $\nu_\rho^n(\eta \in \Omega_n : \eta(x) = 1) = \rho$ is invariant under the dynamics. For this choice of the parameters the boundary reservoirs have the same intensity and we do not see any induced current on the system. Nevertheless, in the case $\alpha \neq \beta$, let us say for example $\alpha < \beta$, there is a tendency to have more particles entering into the bulk from the right reservoir and leaving the system from the left reservoir. This is a current which is induced by the difference of the density at the boundary reservoirs. Note that in the bulk the dynamics is symmetric. In the case $\alpha \neq \beta$, since we have a finite state Markov process, there is only one stationary measure that we denote by μ_n^{ss} which is no longer a product measure as in the case $\alpha = \beta$. By using the matrix ansatz method developed by [3, 10, 11] and references therein, it is possible to compute the correlation function in the stationary state and an important problem is to analyze the behavior of the system starting from this non-equilibrium stationary state.

We note that the hydrodynamic limit of this model was studied in [1] and the hydrodynamic equations consist in the heat equation with different types of boundary conditions depending on the range of the parameter θ. More precisely, for $0 \leq \theta < 1$ the heat equation has Dirichlet boundary conditions which fix the value of the density profile at the points 0 and 1 to be α and β, respectively. In this case we do not see any difference at the macroscopic level with respect to the case $\theta = 0$. Nevertheless, for $\theta = 1$ the boundary dynamics is slowed enough in such a way that macroscopically the Dirichlet boundary conditions are replaced by a type of Robin boundary conditions. These Robin boundary conditions state that the rate at which particles are injected into the system through the boundary points, is given by the difference of the density at the bulk and the boundary. Finally for $\theta > 1$, the boundaries are sufficiently slowed so that the Robin boundary conditions are replaced by Neumann boundary conditions stating that macroscopically there is no flux of particles from the boundary reservoirs.

We emphasize that there are many similar models to the one studied in these notes which we summarize as follows. In [7–9], the authors consider a model where removal of particles can only occur at an interval around the left boundary and the entrance of particles is allowed only at an interval around the right boundary. Their model presents a current exchange between the two reservoirs and shows some similarities with our model for the choice $\theta = 1$. Another case already studied in the literature (see [12, 19]) is when the boundary is not slowed, that corresponds to our model for the choice $\theta = 0$. As mentioned above, the hydrodynamic equation of this model has Dirichlet boundary conditions, see [12] or the Eq. (7). A similar model, whose hydrodynamic equation has both Dirichlet boundary conditions and Neumann boundary conditions, was studied in [6]. The main difference, at the macroscopic level, is that the end points of the boundary conditions vary with time. The microscopic dynamics there is given by the SSEP evolving on $\mathbb{Z}$ with additional births and deaths restricted to a subset of configurations where there is a leftmost hole and a rightmost particle. In this situation, at a fixed rate j birth of particles

occur at the position of the leftmost hole and at the same rate, independently, the rightmost particle dies. Another model which has a current is considered in [4]. The dynamics evolves on the discrete torus $\mathbb{Z}/n\mathbb{Z}$ without reservoirs, but has a surprising phenomenon: a "battery effect". This effect produces a current of particles through the system and is due to a single abnormal bond, where the rates to cross from left to right and from right to left are different. Finally, another model which has similarities with the model we consider in these notes is the SSEP with a slow bond, which was studied in [13–15]. The dynamics evolves on the discrete torus $\mathbb{Z}/n\mathbb{Z}$, and particles exchange positions between nearest neighbor bonds at rate 1, except at one particular bond, where the exchange occurs at rate $n^{-\beta}$. In this case $\beta > 0$ is a parameter that rules the slowness of the bond and for that reason the bond is called the *slow bond*. The similarity between the slow bond model and the slow boundary model considered in these notes is that if we "open" the discrete torus exactly at the position of the slow bond, then the slow bond gives rise to a slow boundary. In [13, 15] different hydrodynamic behaviors were obtained, depending on the range of the parameter β, more precisely, the hydrodynamic equation is, in all cases, the heat equation but the boundary conditions vary with the value of β, exhibiting three different regimes as for the slow boundary model, see [1].

Our interest in these notes is to go beyond the hydrodynamical behavior, analyzing the fluctuations around the hydrodynamical profile. To accomplish this, we restrict ourselves to the case $\alpha = \beta = \rho$ and starting from the stationary measure ν_ρ^n defined above. Our result states that the fluctuations starting from ν_ρ^n are given by an Ornstein–Uhlenbeck process solution of

$$d\mathcal{Y}_t = \Delta_\theta \mathcal{Y}_t dt + \sqrt{2\chi(\rho)t}\, \nabla_\theta\, d\mathcal{W}_t\,,$$

where $\chi(\rho)$ is the variance of $\eta(x)$ with respect to ν_ρ^n, $\mathcal{W}_t$ is a space-time white noise of unit variance and Δ_θ and ∇_θ are, respectively, the Laplacian and derivative operators defined on a space of test functions with different types of boundary conditions depending on the value of θ. We note that the case $\theta = 0$ was studied in [19] and the case $\theta = 1$ was studied in [16]. In those articles, the nonequilibrium fluctuations were obtained starting from general initial measures, which include the equilibrium case ν_ρ^n treated here. We note however, that the case $\theta \neq 1$ is quite difficult to attack at the nonequilibrium scenario since we need to establish a local replacement (see Lemma 3) in order to close the martingale problem, which we can only prove starting the system from the equilibrium state. In a future work, we will dedicate to extending this result to the nonequilibrium situation as, for example, starting the system from the steady state when $\alpha \neq \beta$.

Here follows an outline of these notes. In Sect. 2 we give the definition of the model, we recall from [1] the hydrodynamic limit and in Sect. 3 we state our main result, namely, Theorem 3. In Sect. 4 we characterize the limit process by means of a martingale problem. Tightness is proved in Sect. 5 and in Sect. 6 we prove the Replacement Lemma which is the most technical part of these notes.

2 Statement of Results

2.1 The Model

For $n \geq 1$, we denote by Σ_n the set $\{1, \ldots, n-1\}$, which will be referred by the expression *bulk*. The symmetric simple exclusion process with slow boundaries is a Markov process $\{\eta_t : t \geq 0\}$ with state space $\Omega_n := \{0,1\}^{\Sigma_n}$. The slowness of the boundaries is ruled by a parameter that we denote by $\theta \geq 0$. If η is a configuration of the state space Ω, then for $x \in \Sigma_n$, the random variable $\eta(x)$ can take only two values, namely 0 or 1. If $\eta(x) = 0$, it means that the site x is vacant, while $\eta(x) = 1$ means that the site x is occupied. The dynamics of this model can be described as follows. In the bulk particles move according to continuous time random walks, but whenever a particle wants to jump to an occupied site, the jump is suppressed. At the left boundary, particles can be created (resp. removed) at rate $\alpha n^{-\theta}$ (resp. $(1-\alpha)n^{-\theta}$). At the right boundary, particles can be created (resp. removed) at rate $\beta n^{-\theta}$ (resp. $(1-\beta)n^{-\theta}$).

Fix now a finite time horizon T. The Markov process $\{\eta_t(x) : x \in \Sigma_n; t \in [0,T]\}$ can be characterized in terms of its infinitesimal generator that we denote by $\mathcal{L}_n^\theta$ and is defined as follows. For a function $f : \Omega_n \to \mathbb{R}$, we have that

$$\begin{aligned}(\mathcal{L}_n^\theta f)(\eta) \;=\; & \left[\frac{\alpha}{n^\theta}(1-\eta(1)) + \frac{(1-\alpha)}{n^\theta}\eta(1)\right]\big(f(\eta^1) - f(\eta)\big) \\ & + \left[\frac{\beta}{n^\theta}(1-\eta(n-1)) + \frac{(1-\beta)}{n^\theta}\eta(n-1)\right]\big(f(\eta^{n-1}) - f(\eta)\big) \\ & + \sum_{x=1}^{n-2}\big(f(\sigma^{x,x+1}\eta) - f(\eta)\big)\,,\end{aligned} \tag{1}$$

where $\sigma^{x,x+1}\eta$ is the configuration obtained from η by exchanging the occupation variables $\eta(x)$ and $\eta(x+1)$, that is,

$$(\sigma^{x,x+1}\eta)(y) = \begin{cases} \eta(x+1), & \text{if } \; y = x\,, \\ \eta(x), & \text{if } \; y = x+1\,, \\ \eta(y), & \text{otherwise.}\end{cases} \tag{2}$$

and for $x = 1, n-1$ η^x is the configuration obtained from η by flipping the occupation variable $\eta(x)$:

$$(\eta^x)(y) = \begin{cases} 1-\eta(y), & \text{if } \; y = x\,, \\ \eta(y), & \text{otherwise.}\end{cases} \tag{3}$$

Let $\mathcal{D}([0,T], \Omega_n)$ be the space of trajectories which are right continuous and with left limits, taking values in Ω_n. Denote by $\mathbb{P}_{\mu_n}^{\theta,n}$ the probability on $\mathcal{D}([0,T], \Omega_n)$

induced by the Markov process with generator $n^2\mathcal{L}_n^\theta$ and the initial measure μ_n and denote by $\mathbb{E}_{\mu_n}^{\theta,n}$ the expectation with respect to $\mathbb{P}_{\mu_n}^{\theta,n}$.

2.2 Stationary Measures

The stationary measure μ_n^{ss} for this model when $\alpha = \beta = \rho \in (0, 1)$ is the Bernoulli product measure given by

$$\nu_\rho^n\Big(\eta \in \Omega_n : \eta(x) = 1\Big) = \rho\,.$$

But in the general case, where $\alpha \neq \beta$, the stationary measure μ_n^{ss} does not have independent marginals, see [10]. What we can say about the stationary behavior of this model is that the density of particles has a behavior very close to a linear profile, which depends on the range of θ in the sense of the following definition:

Definition 1 Let $\gamma : [0, 1] \to [0, 1]$ be a measurable profile. A sequence $\{\mu_n\}_{n\in\mathbb{N}}$ is said to be *associated* to γ if, for any $\delta > 0$ and any continuous function $f : [0, 1] \to \mathbb{R}$ the following limit holds:

$$\lim_{n\to\infty} \mu_n\left(\eta : \left|\frac{1}{n}\sum_{x=1}^{n-1} f\left(\tfrac{x}{n}\right)\eta(x) - \int f(u)\,\gamma(u)\,du\right| > \delta\right) = 0\,.$$

For μ_n equal to the stationary measure μ_n^{ss}, the limit above is called the *hydrostatic limit*.

Theorem 1 (Hydrostatic Limit, [1]) *Let μ_n^{ss} be the stationary probability measure in Ω_n wrt the Markov process with infinitesimal generator $n^2\mathcal{L}_n^\theta$, defined in (1). The sequence $\{\mu_n^{ss}\}_{n\in\mathbb{N}}$ is associated (in the sense of Definition 1) to the profile $\overline{\rho} : [0, 1] \to \mathbb{R}$ given by*

$$\overline{\rho}(u) = \begin{cases} (\beta-\alpha)u + \alpha, & \text{if } \theta \in [0, 1), \\ \frac{\beta-\alpha}{3}u + \alpha + \frac{\beta-\alpha}{3}, & \text{if } \theta = 1, \\ \frac{\beta+\alpha}{2}, & \text{if } \theta \in (1, \infty), \end{cases} \tag{4}$$

for all $u \in [0, 1]$.

Another feature that we can say about the stationary state of the model studied in this paper is that the profiles in (4) are very close to the mean of $\eta(x)$ taken with respect to the stationary measure μ_n^{ss}. To state this result properly, we start by defining for an initial measure μ_n in Ω_n, for $x \in \Sigma_n$ and for $t \geq 0$ the empirical mean given by

$$\rho_t^n(x) := \mathbb{E}_{\mu_n}^{\theta,n}[\eta_t(x)]\,. \tag{5}$$

If in the expression above $\mu_n = \mu_n^{ss}$, then $\rho_t^n(x)$ does not depend on t, so that $\rho_t^n(x) = \rho^n(x)$. From [1], we have that $\rho^n(x)$ satisfies the following recurrence relations:

$$\begin{cases} 0 = [\rho^n(x+1) - \rho^n(x)] + [\rho^n(x-1) - \rho^n(x)], & \text{if} \quad x \in \{2, \dots, n-2\}, \\ 0 = [\rho^n(2) - \rho^n(1)] + n^{-\theta}[\alpha - \rho^n(1)], \\ 0 = n^{-\theta}[\beta - \rho^n(n-1)] + [\rho^n(n-2) - \rho^n(n-1)]. \end{cases}$$

A simple computation shows that $\rho^n(x)$ is given by $\rho^n(x) = a_n x + b_n$, for all $x \in \Sigma_n$, where $a_n = \frac{\beta-\alpha}{2n^\theta+n-2}$ and $b_n = \alpha + a_n(n^\theta - 1)$. Moreover, we conclude that

$$\lim_{n\to\infty} \left(\max_{x\in\Sigma_n} \left| \rho^n(x) - \overline{\rho}(\tfrac{x}{n}) \right| \right) = 0.$$

2.3 Hydrodynamic Limit

In [1] it was established the hydrodynamic limit of the model for any $\theta \geq 0$. For completeness we recall that result now. Fix a measurable density profile $\rho_0 : [0,1] \to [0,1]$ and for each $n \in \mathbb{N}$, let μ_n be a probability measure on Ω_n.

Theorem 2 (Hydrodynamic Limit, [1]) *Suppose that the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ is associated to a profile $\rho_0(\cdot)$ in the sense of Definition 1. Then, for each $t \in [0,T]$, for any $\delta > 0$ and any continuous function $f : [0,1] \to \mathbb{R}$,*

$$\lim_{n\to+\infty} \mathbb{P}_{\mu_n}^{\theta,n} \left[\eta_\cdot : \left| \frac{1}{n} \sum_{x=1}^{n-1} f(\tfrac{x}{n})\, \eta_{tn^2}(x) - \int f(u)\, \rho(t,u)\, du \right| > \delta \right] = 0,$$

where $\rho(t,\cdot)$ is the unique weak solution of the heat equation

$$\begin{cases} \partial_t \rho(t,u) = \partial_u^2 \rho(t,u)\,, & \text{for } t > 0\,,\ u \in (0,1), \\ \rho(0,u) = \rho_0(u)\,, & u \in [0,1]. \end{cases} \tag{6}$$

with boundary conditions that depend on the range of θ, which are given by:

$$\text{For } \theta < 1,\ \partial_u \rho(t,0) = \alpha \text{ and } \partial_u \rho(t,1) = \beta, \text{ for } t > 0. \tag{7}$$

$$\text{For } \theta = 1,\ \partial_u \rho(t,0) = \rho(t,0) - \alpha \text{ and } \partial_u \rho(t,1) = \beta - \rho(t,1), \text{ for } t > 0. \tag{8}$$

$$\text{For } \theta > 1,\ \partial_u \rho(t,0) = \partial_u \rho(t,1) = 0, \text{ for } t > 0. \tag{9}$$

Remark 1 We note that the profiles in (4) are stationary solutions of the heat equation with the corresponding boundary conditions given above.

3 Density Fluctuations

3.1 The Space of Test Functions

The space $C^\infty([0,1])$ is the space of functions $f : [0,1] \to \mathbb{R}$ such that f is continuous in $[0,1]$ as well as all its derivatives.

Definition 2 Let $\mathcal{S}_\theta$ denote the set of functions $f \in C^\infty([0,1])$ such that for any $k \in \mathbb{N} \cup \{0\}$ it holds that

(1) for $\theta < 1$: $\partial_u^{2k} f(0) = \partial_u^{2k} f(1) = 0$.
(2) for $\theta = 1$: $\partial_u^{2k+1} f(0) = \partial_u^{2k} f(0)$ and $\partial_u^{2k+1} f(1) = -\partial_u^{2k} f(1)$.
(3) for $\theta > 1$: $\partial_u^{2k+1} f(0) = \partial_u^{2k+1} f(1) = 0$.

Definition 3 For $\theta \geq 0$, let $-\Delta_\theta$ be the positive operator, self-adjoint on $L^2[0,1]$, defined on $f \in \mathcal{S}_\theta$ by

$$\Delta_\theta f(u) \;=\; \begin{cases} \partial_u^2 f(u)\,, & \text{if } u \in (0,1), \\ \partial_u^2 f(0^+)\,, & \text{if } u = 0, \\ \partial_u^2 f(1^-)\,, & \text{if } u = 1. \end{cases} \tag{10}$$

Above, $\partial_u^2 f(a^\pm)$ denotes the side limits at the point a. Analogously, let $\nabla_\theta : \mathcal{S}_\theta \to C^\infty([0,1])$ be the operator given by

$$\nabla_\theta f(u) \;=\; \begin{cases} \partial_u f(u)\,, & \text{if } u \in (0,1), \\ \partial_u f(0^+)\,, & \text{if } u = 0, \\ \partial_u f(1^-)\,, & \text{if } u = 1. \end{cases} \tag{11}$$

Definition 4 Let $T_t^\theta : \mathcal{S}_\theta \to \mathcal{S}_\theta$ be the semigroup associated to (6) with the corresponding boundary conditions for the case $\alpha = \beta = 0$. That is, given $f \in \mathcal{S}_\theta$, by $T_t^\theta f$ we mean the solution of the homogeneous version of (6) with initial condition f.

Definition 5 Let $\mathcal{S}'_\theta$ be the topological dual of $\mathcal{S}_\theta$ with respect to the topology generated by the seminorms

$$\|f\|_k = \sup_{u \in [0,1]} |\partial_u^k f(u)|\,, \tag{12}$$

where $k \in \mathbb{N} \cup \{0\}$. In other words, $\mathcal{S}'_\theta$ consists of all linear functionals $f : \mathcal{S}_\theta \to \mathbb{R}$ which are continuous with respect to all the seminorms $\|\cdot\|_k$.

Let $\mathcal{D}([0,T], \mathcal{S}'_\theta)$ (resp. $\mathcal{C}([0,T], \mathcal{S}'_\theta)$) be the space of trajectories which are right continuous and with left limits (resp. continuous), taking values in $\mathcal{S}'_\theta$.

The expression for T_t^θ, $\theta \geq 0$, is presented in the next proposition:

Proposition 1 *Let $\theta \geq 0$. Suppose that $\rho_0 \in L^2[0,1]$. Then*

$$(T_t^\theta \rho_0)(u) \;:=\; \sum_{n=1}^{\infty} a_n \, e^{-\lambda_n t} \, \Psi_n(u)\,, \tag{13}$$

where $\{\Psi_n\}_{n\in\mathbb{N}}$ is an orthonormal basis of $L^2[0,1]$ constituted by eigenfunctions of the associated Regular Sturm-Liouville Problem (concerning the operator Δ_θ) and a_n are the Fourier coefficients of ρ_0 in the basis $\{\Psi_n\}_{n\in\mathbb{N}}$.

- *For $\theta < 1$, the corresponding orthonormal basis of $L^2[0,1]$ is*

$$\begin{cases} \Psi_n(u) \;=\; \sqrt{2}\sin(n\pi u)\,, \quad \text{for } n \in \mathbb{N}\,, \\ \Psi_0(u) \;\equiv\; 1\,. \end{cases}$$

  The eigenvalues of the associated Regular Sturm-Liouville Problem (concerning the operator Δ_θ) are given by $\lambda_n = n^2\pi^2$.
- *For $\theta = 1$, the corresponding orthonormal basis of $L^2[0,1]$ is a linear combination of sines and cosines, namely,*

$$\Psi_n(u) \;=\; A_n \sin(\sqrt{\lambda_n}u) + A_n\sqrt{\lambda_n}\cos(\sqrt{\lambda_n}u)\,, \quad \text{for } n \in \mathbb{N}\cup\{0\}\,,$$

  where A_n is a normalizing constant. The eigenvalues λ_n do not have an explicit formula, but it can verified that $\lambda_n \sim n^2\pi^2$.
- *For $\theta > 1$, the corresponding orthonormal basis of $L^2[0,1]$ is*

$$\begin{cases} \Psi_n(u) \;=\; \sqrt{2}\cos(n\pi u)\,, \quad \text{for } n \in \mathbb{N}\,, \\ \Psi_0(u) \;\equiv\; 1\,. \end{cases}$$

The eigenvalues of the associated Regular Sturm-Liouville Problem (concerning the operator Δ_θ) are given by $\lambda_n = n^2\pi^2$.

Proof For $\theta = 1$ the expression for T_t^θ has been obtained in [16]. For the case $\theta \neq 1$, as in [16], we state the associated *Regular Sturm-Liouville Problem* (for details on this subject we refer to [2], for instance):

$$\text{For } \theta < 1: \quad \begin{cases} \Psi''(u) + \lambda\Psi(u) \;=\; 0\,, \quad u \in (0,1)\,, \\ \Psi(0) \;=\; 0\,, \quad \Psi(1) \;=\; 0\,; \end{cases}$$

$$\text{For } \theta \geq 1: \quad \begin{cases} \Psi''(u) + \lambda\Psi(u) \;=\; 0\,, \quad u \in (0,1)\,, \\ \Psi'(0) \;=\; 0\,, \quad \Psi'(1) \;=\; 0\,. \end{cases}$$

The solution of each one of the problems above (the eigenvalues λ_n and the eigenfunctions Ψ_n) can be found in Chap. 10 of [5].

As a consequence, the series (13) converges exponentially fast, implying that $(T_t^\theta \rho_0)(u)$ is smooth in space and time for any $t > 0$. This observation implies a property of $T_t^\theta : \mathcal{S}_\theta \to \mathcal{S}_\theta$ stated in the next corollary.

Corollary 1 *If $f \in \mathcal{S}_\theta$, then for any $t > 0$, $T_t^\theta f \in \mathcal{S}_\theta$ and $\Delta_\theta T_t^\theta f \in \mathcal{S}_\theta$.*

We observe that the previous result is needed in the proof of uniqueness of the corresponding Ornstein–Uhlenbeck process (which is defined in the next section). Its proof is a consequence of the formula (13), see [16] for more details.

3.2 Ornstein–Uhlenbeck Process

Fix $\rho \in (0, 1)$. Based on [17, 18], we give here a characterization of the generalized Ornstein–Uhlenbeck process which is a solution of

$$d\mathcal{Y}_t = \Delta_\theta \mathcal{Y}_t dt + \sqrt{2\chi(\rho)t}\, \nabla_\theta\, d\mathcal{W}_t\,, \tag{14}$$

where $\mathcal{W}_t$ is a space-time white noise of unit variance and $\chi(\rho) = \int (\eta(x) - \rho)^2 d\nu_\rho^n = \rho(1-\rho)$, in terms of a martingale problem. We will see below that this process governs the equilibrium fluctuations of the density of particles of our model. In spite of having a dependence of $\mathcal{Y}_t$ on θ, we do not index on it to not overload notation. Denote by $\mathbf{Q}_\rho^\theta$ the distribution of $\mathcal{Y}_\cdot$ and $\mathbf{E}_{\mathbf{Q}_\rho^\theta}$ the expectation with respect to $\mathbf{Q}_\rho^\theta$.

Define the inner product between the functions $f, g : [0, 1] \to \mathbb{R}$ by

$$\langle f,\, g\rangle_{L_\rho^{2,\theta}} = 2\chi(\rho)\Big[\int_0^1 f(u)\, g(u)\, du + \Big(f(0)g(0) + f(1)g(1)\Big)\mathbf{1}_{\theta=1}\Big],$$

where $\mathbf{1}_\cdot$ is the indicator function. Then, $L_\rho^{2,\theta}([0, 1])$ is the space of functions $f : [0, 1] \to \mathbb{R}$ with $\|f\|_{L_\rho^{2,\theta}} < \infty$, where

$$\|f\|^2_{L_\rho^{2,\theta}} = \langle f,\, f\rangle_{L_\rho^{2,\theta}}. \tag{15}$$

Proposition 2 *There exists an unique random element $\mathcal{Y}$ taking values in the space $\mathcal{C}([0, T], \mathcal{S}'_\theta)$ such that:*

(i) For every function $f \in \mathcal{S}_\theta$, $\mathcal{M}_t(f)$ and $\mathcal{N}_t(f)$ given by

$$\mathcal{M}_t(f) = \mathcal{Y}_t(f) - \mathcal{Y}_0(f) - \int_0^t \mathcal{Y}_s(\Delta_\theta f)ds\,,$$
$$\mathcal{N}_t(f) = \big(\mathcal{M}_t(f)\big)^2 - 2\chi(\rho)\, t\, \|\nabla_\theta f\|^2_{L^{2,\theta}_\rho} \tag{16}$$

are $\mathcal{F}_t$-martingales, where for each $t \in [0,T]$, $\mathcal{F}_t := \sigma(\mathcal{Y}_s(f); s \le t,\ f \in \mathcal{S}_\theta)$.

(ii) $\mathcal{Y}_0$ is a Gaussian field of mean zero and covariance given on $f, g \in \mathcal{S}_\theta$ by

$$\mathbf{E}_{\mathbf{Q}^\theta_\rho}\big[\mathcal{Y}_0(f)\mathcal{Y}_0(g)\big] = \langle f,\ g\rangle_{L^{2,\theta}_\rho} \tag{17}$$

Moreover, for each $f \in \mathcal{S}_\theta$, the stochastic process $\{\mathcal{Y}_t(f)\,;\ t \ge 0\}$ is Gaussian, being the distribution of $\mathcal{Y}_t(f)$ conditionally to $\mathcal{F}_s$, for $s < t$, normal of mean $\mathcal{Y}_s(T^\theta_{t-s} f)$ and variance $\int_0^{t-s} \|\nabla_\theta T^\theta_r f\|^2_{L^{2,\theta}_\rho}\, dr$, where T^θ_t was given in Definition 4.

The random element $\mathcal{Y}_\cdot$ is called the generalized Ornstein–Uhlenbeck process of characteristics Δ_θ and ∇_θ. From the second equation in (16) and Lévy's Theorem on the martingale characterization of Brownian motion, for each $f \in \mathcal{S}_\theta$, the process

$$\mathcal{M}_t(f)\Big(2\chi(\rho)t\|\nabla_\theta f\|^2_{L^{2,\theta}_\rho}\Big)^{-1/2} \tag{18}$$

is a standard Brownian motion. Therefore, in view of Proposition 2, it makes sense to say that $\mathcal{Y}_\cdot$ is the formal solution of (14).

3.3 The Density Fluctuation Field

We define the density fluctuation field $\mathcal{Y}^n_\cdot$ as time-trajectory of the linear functional acting on functions $f \in \mathcal{S}_\theta$ as

$$\mathcal{Y}^n_t(f) \;=\; \frac{1}{\sqrt{n}} \sum_{x=1}^{n-1} f\Big(\frac{x}{n}\Big)\Big(\eta_{tn^2}(x) - \rho^n_t(x)\Big), \quad \text{for all } t \ge 0, \tag{19}$$

where ρ^n_t was defined in (5). Our results are given for the case $\alpha = \beta = \rho$ and for μ_n being equal to ν^n_ρ, that is, the Bernoulli product measure with parameter $\rho \in (0,1)$, so that $\rho^n_t(x) = \rho$, for all $x \in \Sigma_n$ and $t \ge 0$. Let $\mathbb{Q}^{\theta,n}_\rho$ be the probability measure on $\mathcal{D}([0,T], \mathcal{S}'_\theta)$ induced by the density fluctuation field $\mathcal{Y}^n_\cdot$ and ν^n_ρ. We note that since we will consider only the initial measure μ_n as ν^n_ρ, we will simplify the notations $\mathbb{P}^{\theta,n}_{\nu^n_\rho}$ and $\mathbb{E}^{\theta,n}_{\nu^n_\rho}$ as $\mathbb{P}^{\theta,n}_\rho$ and $\mathbb{E}^{\theta,n}_\rho$, respectively. Our main result is the following theorem.

Theorem 3 (Ornstein–Uhlenbeck limit) *For $\alpha = \beta = \rho \in (0,1)$, if we take the initial measure to be ν^n_ρ, namely, the Bernoulli product measure with parameter ρ, then, the sequence $\{\mathbb{Q}^{\theta,n}_\rho\}_{n\in\mathbb{N}}$ converges, as $n \to \infty$, to a generalized Ornstein–Uhlenbeck*

(O.U.) process, which is the formal solution of equation (14)*. As a consequence, the variance of the limit field* $\mathcal{Y}_t$ *is given on* $f \in \mathcal{S}_\theta$ *by*

$$\mathbf{E}_{\mathbf{Q}_\rho^\theta}[\mathcal{Y}_t(f)\mathcal{Y}_s(f)] = \chi(\rho)\int_0^1 (f(u))^2\,du + \int_0^s \|T_{t-r}^\theta f\|^2_{L^{2,\theta}_\rho} dr\,, \tag{20}$$

where $\|\cdot\|^2_{L^{2,\theta}_\rho}$ *was defined in* (15).

4 Proof of Theorem 3

4.1 Characterization of Limit Points

Fix a test function f. By Dynkin's formula, we have that

$$M_t^n(f) = \mathcal{Y}_t^n(f) - \mathcal{Y}_0(f) - \int_0^t (\partial_s + n^2\mathcal{L}_n^\theta)\mathcal{Y}_s^n(f)ds, \tag{21}$$

$$N_t^n(f) = (M_t^n(f))^2 - \int_0^t n^2\mathcal{L}_n^\theta\mathcal{Y}_s^n(f)^2 - 2\mathcal{Y}_s^n(f)n^2\mathcal{L}_n^\theta\mathcal{Y}_s^n(f)ds \tag{22}$$

are martingales with respect to the natural filtration $\mathcal{F}_t := \sigma(\eta_s : s \le t)$. To simplify notation we denote $\Gamma_s^n(f) := (\partial_s + n^2\mathcal{L}_n^\theta)\mathcal{Y}_s^n(f)$. A long but elementary computation shows that

$$\begin{aligned}\Gamma_s^n(f) =& \frac{1}{\sqrt{n}}\sum_{x=1}^{n-1}\Delta_n f\left(\tfrac{x}{n}\right)(\eta_s(x)-\rho)\\ &+\sqrt{n}\nabla_n^+ f(0)(\eta_s(1)-\rho) - \sqrt{n}\nabla_n^- f(n)(\eta_s(n-1)-\rho)\\ &-\frac{n^{3/2}}{n^\theta} f\left(\tfrac{1}{n}\right)(\eta_s(1)-\rho) - \frac{n^{3/2}}{n^\theta} f\left(\tfrac{n-1}{n}\right)(\eta_s(n-1)-\rho).\end{aligned} \tag{23}$$

Above

$$\Delta_n f(x) := n^2\Big[f\left(\tfrac{x+1}{n}\right) + f\left(\tfrac{x-1}{n}\right) - 2f\left(\tfrac{x}{n}\right)\Big],$$

$$\nabla_n^+ f(x) := n\Big[f\left(\tfrac{x+1}{n}\right) - f\left(\tfrac{x}{n}\right)\Big]$$

and

$$\nabla_n^- f(x) := n\Big[f\left(\tfrac{x}{n}\right) - f\left(\tfrac{x-1}{n}\right)\Big].$$

We note that for the choice $\theta = 0$, using the fact that $f(0) = 0 = f(1)$, the expression (23) reduces to

$$\Gamma_s^n(f) = \frac{1}{\sqrt{n}} \sum_{x=1}^{n-1} \Delta_n f\left(\tfrac{x}{n}\right)(\eta_s(x) - \rho)\,, \tag{24}$$

which is $\mathcal{Y}_s^n(\Delta_n f)$.

Now, we close the Eq. (23) for each regime of θ. The goal is to show that we can rewrite (23) as (24) plus a term which vanishes as $n \to \infty$.

• The case $\theta < 1$: we note that since $f \in \mathcal{S}_\theta$ we can write $\Gamma_s^n(f)$ as

$$\mathcal{Y}_s^n(\Delta_n f) + \sqrt{n}(1 - n^{-\theta})\Big\{\nabla_n^+ f(0)(\eta_s(1) - \rho) - \nabla_n^- f(n)(\eta_s(n-1) - \rho)\Big\}.$$

In order to close the equation for the martingale we need to show that:

$$\lim_{n\to\infty} \mathbb{E}_\rho^{\theta,n}\left[\left(\int_0^t \sqrt{n}(\eta_s(x) - \rho)\,ds\right)^2\right] = 0, \quad \text{for } x = 1, n-1, \tag{25}$$

which is a consequence of Lemma 3, see Remark 2.

• The case $\theta = 1$: we can write $\Gamma_s^n(f)$ as

$$\begin{aligned}&\mathcal{Y}_s^n(\Delta_n f) + \sqrt{n}\Big(\partial_u f(0) - f(0)\Big)(\eta_s(1) - \rho)\\ &+ \sqrt{n}\Big(\partial_u f(1) + f(1)\Big)(\eta_s(n-1) - \rho) + O\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Since $f \in \mathcal{S}_\theta$ the last expression equals to $\mathcal{Y}_s^n(\Delta_n f)$.

• The case $\theta > 1$: we can repeat the computations above and since $f \in \mathcal{S}_\theta$, $\Gamma_s^n(f)$ can be rewritten as

$$\mathcal{Y}_s^n(\Delta_n f) - \frac{n^{3/2}}{n^\theta} f\left(\tfrac{1}{n}\right)(\eta_s(1) - \rho) - \frac{n^{3/2}}{n^\theta} f\left(\tfrac{n-1}{n}\right)(\eta_s(n-1) - \rho) + O\left(\frac{1}{\sqrt{n}}\right).$$

Then, in order to close the equation for the martingale term we need to show that

$$\lim_{n\to\infty} \mathbb{E}_\rho^{\theta,n}\left[\left(\int_0^t \frac{n^{3/2}}{n^\theta}(\eta_s(x) - \rho)\,ds\right)^2\right] = 0, \quad \text{for } x = 1, n-1\,, \tag{26}$$

which is a consequence of Lemma 3, see Remark 2.

From the previous observations, for each regime of θ we can rewrite (23) as (24) plus a negligible term.

Lemma 1 *For all $\theta \geq 0$, $t > 0$ and $f \in \mathcal{S}_\theta$ it holds that*

$$\lim_{n\to\infty} \mathbb{E}_\rho^{\theta,n}[|M_t^n(f)|^2] = t\|\nabla_\theta f\|_{L^{2,\theta}_\rho},$$

where the norm above was defined in (15).

Proof A simple computation shows that the integral part of the martingale $N_t^n(f)$ can be written as

$$\begin{aligned} n^2\mathcal{L}_n^\theta \mathcal{Y}_s^n(f)^2 - 2\mathcal{Y}_s^n(f)n^2\mathcal{L}_n^\theta \mathcal{Y}_s^n(f) &= \frac{1}{n}\sum_{x=1}^{n-2}\left(\nabla_n^+ f\left(\tfrac{x}{n}\right)\right)^2\Big(\eta_s(x)-\eta_s(x+1)\Big)^2 \\ &+\frac{n}{n^\theta}\left(f\left(\tfrac{1}{n}\right)\right)^2\Big(\rho - 2\rho\eta_s(1)+\eta_s(1)\Big) \\ &+\frac{n}{n^\theta}\left(f\left(\tfrac{n-1}{n}\right)\right)^2\Big(\rho - 2\rho\eta_s(n-1)+\eta_s(n-1)\Big), \end{aligned}$$

from where we get that

$$\begin{aligned} &\mathbb{E}_\rho^{\theta,n}\big[|M_t^n(f)|^2\big] \\ &= 2\chi(\rho)\, t\left\{\frac{1}{n}\sum_{x=1}^{n-2}\left(\nabla_n^+ f\left(\tfrac{x}{n}\right)\right)^2 + \frac{n}{n^\theta}\left(\left(f\left(\tfrac{1}{n}\right)\right)^2 + \left(f\left(\tfrac{n-1}{n}\right)\right)^2\right)\right\}. \end{aligned} \tag{27}$$

Let $f \in \mathcal{S}_\theta$. The first term at the right hand side of the previous expression converges to $2\chi(\rho)\int_0^1 \left(\nabla_\theta f(u)\right)^2 du$, for all $\theta \geq 0$. The second term at the right hand side of last expression has to be analyzed for each case of θ separately:

• The case $\theta < 1$: since $f(0) = 0 = f(1)$, the second term at the right hand side of (27) can be rewritten as $2\chi(\rho)\, t$ times

$$\frac{n}{n^\theta}\left(\left(f\left(\tfrac{1}{n}\right)\right)^2 + \left(f\left(\tfrac{n-1}{n}\right)\right)^2\right) = \frac{1}{n^{1+\theta}}\left(\left(\nabla_n^+ f(0)\right)^2 + \left(\nabla_n^- f(n)\right)^2\right),$$

which goes to zero as $n \to \infty$.

• The case $\theta = 1$: the second term at the right hand side of (27) converges, as $n \to \infty$, to

$$2\chi(\rho)\, t\left(f^2(0) + f^2(1)\right).$$

Recalling that $f(0) = \partial_u f(0)$ and $f(1) = -\partial_u f(1)$, the proof ends.

• The case $\theta > 1$: since $f \in \mathcal{S}_\theta$ and $\frac{n}{n^\theta} \to 0$, as $n \to \infty$, the second term at the right hand side of (27) converges to zero when $n \to \infty$.

We have just proved that the quadratic variation of the martingale converges in mean. In the next Lemma we state the stronger convergence of the martingales to a Brownian motion.

Lemma 2 *For $f \in \mathcal{S}_\theta$, the sequence of martingales $\{M_t^n(f); t \in [0,T]\}_{n\in\mathbb{N}}$ converges in the topology of $\mathcal{D}([0,T],\mathbb{R})$, as $n \to \infty$, towards a Brownian motion $\mathcal{W}_t(f)$ of quadratic variation given by $t\|\nabla_\theta f\|_{L^{2,\theta}_\rho}$ where $\|\cdot\|_{L^{2,\theta}_\rho}$ was defined in* (15).

Proof We can repeat here the same proof of [14, p. 4170], which is based on Lemma 1 and the fact that a limit in distribution of a uniformly integrable sequence of martingales is a martingale. We leave the details to the interested reader.

4.2 Convergence at Initial Time

Proposition 3 *The sequence $\{\mathcal{Y}_0^n\}_{n\in\mathbb{N}}$ converges in distribution to $\mathcal{Y}_0$, where $\mathcal{Y}_0$ is a Gaussian field with mean zero and covariance given by* (17).

Proof We first claim that, for every $f \in \mathcal{S}_\theta$ and every $t > 0$,

$$\lim_{n\to+\infty} \log \mathbb{E}_\rho^{\theta,n}\Big[\exp\{i\lambda\mathcal{Y}_0^n(f)\}\Big] = -\frac{\lambda^2}{2}\chi(\rho)\int_0^1 f^2(u)\,du\,.$$

Since ν_ρ^n is a Bernoulli product measure,

$$\begin{aligned}\log \mathbb{E}_\rho^{\theta,n}[\exp\{i\lambda\mathcal{Y}_0^n(f)\}] &= \log\int\left[\exp\left\{\frac{i\lambda}{\sqrt{n}}\sum_{x\in\Sigma_n}(\eta_0(x)-\rho)\,f\left(\frac{x}{n}\right)\right\}\right]d\nu_\rho^n \\ &= \sum_{x\in\Sigma_n}\log\int\left[\exp\left\{\frac{i\lambda}{\sqrt{n}}(\eta_0(x)-\rho)\,f\left(\frac{x}{n}\right)\right\}\right]d\nu_\rho^n\,.\end{aligned}$$

Since f is smooth and using Taylor's expansion, the right hand side of last expression is equal to

$$-\frac{\lambda^2}{2n}\sum_{x\in\Sigma_n} f^2\left(\frac{x}{n}\right)\chi(\rho) + O\left(\frac{1}{\sqrt{n}}\right)\,.$$

Taking the limit as $n \to +\infty$ and using the continuity of f, the proof of the claim ends. Replacing f by a linear combination of functions and recalling the Cramér-Wold device, the proof finishes.

5 Tightness

Now we prove that the sequence of processes $\{\mathcal{Y}_t^n; t \in [0,T]\}_{n\in\mathbb{N}}$ is tight. Recall that we have defined the density fluctuation field on test functions $f \in \mathcal{S}_\theta$. Since we want to use Mitoma's criterium [20] for tightness, we need the following property from the space $\mathcal{S}_\theta$.

Proposition 4 *The space $\mathcal{S}_\theta$ endowed with the semi-norms given in (12) is a Fréchet space.*

Proof The definition of a Fréchet space can be found, for instance, in [21]. Since $C^\infty([0,1])$ endowed with the semi-norms (12) is a Fréchet space, and a closed subspace of a Fréchet space is also a Fréchet space, it is enough to show that $\mathcal{S}_\theta$ is a closed subspace of $C^\infty([0,1])$, which is a consequence of the fact that uniform convergence implies point-wise convergence.

As a consequence of Mitoma's criterium [20] and Proposition 4, the proof of tightness of the $\mathcal{S}'_\theta$ valued processes $\{\mathcal{Y}_t^n; t \in [0,T]\}_{n\in\mathbb{N}}$ follows from tightness of the sequence of real-valued processes $\{\mathcal{Y}_t^n(f); t \in [0,T]\}_{n\in\mathbb{N}}$, for $f \in \mathcal{S}_\theta$.

Proposition 5 (Mitoma's criterium, [20]) *A sequence of processes $\{x_t; t \in [0,T]\}_{n\in\mathbb{N}}$ in $\mathcal{D}([0,T],\mathcal{S}_\theta')$ is tight with respect to the Skorohod topology if, and only if, the sequence $\{x_t(f); t \in [0,T]\}_{n\in\mathbb{N}}$ of real-valued processes is tight with respect to the Skorohod topology of $\mathcal{D}([0,T],\mathbb{R})$, for any $f \in \mathcal{S}_\theta$.*

Now, to show tightness of the real-valued process we use the Aldous' criterium:

Proposition 6 *A sequence $\{x_t; t \in [0,T]\}_{n\in\mathbb{N}}$ of real-valued processes is tight with respect to the Skorohod topology of $\mathcal{D}([0,T],\mathbb{R})$ if:*

(i) $\lim_{A\to+\infty} \limsup_{n\to+\infty} \mathbb{P}_{\mu_n}\Big(\sup_{0\le t\le T} |x_t| > A\Big) = 0,$

(ii) for any $\varepsilon > 0$, $\lim_{\delta\to 0} \limsup_{n\to+\infty} \sup_{\lambda\le\delta} \sup_{\tau\in\mathcal{T}_T} \mathbb{P}_{\mu_n}(|x_{\tau+\lambda} - x_\tau| > \varepsilon) = 0,$

where $\mathcal{T}_T$ is the set of stopping times bounded by T.

Fix $f \in \mathcal{S}_\theta$. By (21), it is enough to prove tightness of $\{\mathcal{Y}_0^n(f)\}_{n\in\mathbb{N}}$, $\{\int_0^t \Gamma_s^n(f)\,ds; t \in [0,T]\}_{n\in\mathbb{N}}$, and $\{\mathcal{M}_t^n(f); t \in [0,T]\}_{n\in\mathbb{N}}$.

5.1 Tightness at the Initial Time

This follows from Proposition 3.

5.2 Tightness of the Martingales

By Lemma 2, the sequence of martingales converges. In particular, it is tight.

5.3 Tightness of the Integral Terms

The first claim of Aldous' criterium can be easily checked for the integral term $\int_0^t \Gamma_s^n(f)\,ds$, where the expression for $\Gamma_s^n(f)$ can be found in (23). Let $f \in \mathcal{S}_\theta$.

- The case $\theta < 1$: by Young's inequality and Cauchy-Schwarz's inequality we have that

$$\begin{aligned}
&\mathbb{E}_\rho^{\theta,n}\left[\sup_{t\le T}\left(\int_0^t \Gamma_s^n(f)\,ds\right)^2\right] \\
&\qquad \le CT\int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\frac{1}{\sqrt{n}}\sum_{x=1}^{n-1}\Delta_n f\left(\tfrac{x}{n}\right)(\eta_{sn^2}(x)-\rho)\right)^2\right]ds \\
&\qquad + C\,(\nabla_n^+ f(0))^2\,T\int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\sqrt{n}(\eta_{sn^2}(1)-\rho)\right)^2\right]ds \\
&\qquad + C\,(\nabla_n^- f(1))^2\,T\int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\sqrt{n}(\eta_{sn^2}(n-1)-\rho)\right)^2\right]ds.
\end{aligned}$$

  Since $f \in \mathcal{S}_\theta$ and by (25), the second and third terms at the right hand side of the previous expression go to zero, as $n \to \infty$. Then there exists $C > 0$ such that these two terms are bounded from above by CT. The first term at the right hand side of last expression is bounded from above by T^2 times

$$\frac{1}{n}\sum_{x=1}^{n-1}\left(\Delta_n f\left(\tfrac{x}{n}\right)\right)^2\chi(\rho)\,. \tag{28}$$

  Now, since $f \in \mathcal{S}_\theta$ last expression is bounded from above by a constant. Now we need to check the second claim of Aldous' criterium. For that purpose, fix a stopping time $\tau \in \mathcal{T}_T$. By Chebyshev's inequality together with (28), we get that

$$\mathbb{P}_\rho^{\theta,n}\left(\left|\int_\tau^{\tau+\lambda}\Gamma_s^n(f)\,ds\,\right| > \varepsilon\right) \le \frac{1}{\varepsilon^2}\mathbb{E}_\rho^{\theta,n}\left[\left(\int_\tau^{\tau+\lambda}\Gamma_s^n(f)\,ds\,\right)^2\right] \le \frac{\delta C}{\varepsilon^2}\,,$$

  which vanishes as $\delta \to 0$.
- The case $\theta = 1$: we note that it was treated in [16].
- The case $\theta > 1$: as in the case $\theta < 1$, we have that

$$\begin{aligned}
&\mathbb{E}_\rho^{\theta,n}\left[\sup_{t\le T}\left(\int_0^t \Gamma_s^n(f)\,ds\right)^2\right] \\
&\qquad \le CT\int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\frac{1}{\sqrt{n}}\sum_{x=1}^{n-1}\Delta_n f\left(\tfrac{x}{n}\right)(\eta_{sn^2}(x)-\rho)\right)^2\right]ds
\end{aligned}$$

$$+ C\, f^2\left(\tfrac{1}{n}\right) T \int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\frac{n^{3/2}}{n^\theta}(\eta_{sn^2}(1)-\rho)\right)^2\right] ds$$
$$+ C\, f^2\left(\tfrac{n-1}{n}\right) T \int_0^T \mathbb{E}_\rho^{\theta,n}\left[\left(\frac{n^{3/2}}{n^\theta}(\eta_{sn^2}(n-1)-\rho)\right)^2\right] ds\,,$$

plus a term of order $\frac{1}{\sqrt{n}}$. To bound the first term at the right hand side of the previous inequality we repeat the same computations as in the case $\theta < 1$. In order to bound the second and the third terms at the right hand side of the previous inequality, we use (26) and the proof follows as in the case $\theta < 1$.

6 Replacement Lemma

This section is devoted to estimate the expectations (25) and (26). In order to do this we start introducing some notations. Let μ be an initial measure. For $x = 0, 1, \ldots, n-1$, define

$$I_{x,x+1}(f,\mu) := \int r_{x,x+1}(\eta)\left(f(\sigma^{x,x+1}\eta) - f(\eta)\right)^2 d\mu\,,$$

where $\sigma^{x,x+1}\eta$ was defined in (2), for $x = 1, \ldots, n-2$, $\sigma^{0,1}\eta := \eta^1$, $\sigma^{n-1,n}\eta := \eta^{n-1}$ (the configurations η^1 and η^{n-1} were defined in (3)), and the rates are given by

$$\begin{aligned} r_{0,1}(\eta) &:= r_\alpha(\eta) := \frac{\alpha}{n^\theta}(1-\eta(1)) + \frac{1-\alpha}{n^\theta}\eta(1)\,,\\ r_{n-1,n}(\eta) &:= r_\beta(\eta) := \frac{\beta}{n^\theta}(1-\eta(n-1)) + \frac{1-\beta}{n^\theta}\eta(n-1)\,,\\ r_{x,x+1}(\eta) &:= 1, \quad \text{if } x = 1, \ldots, n-2\,. \end{aligned}$$

Define the quantity:

$$\mathcal{D}_n(f,\mu) := \sum_{x=0}^{n-1} I_{x,x+1}(f,\mu) = \sum_{x=0}^{n-1}\int r_{x,x+1}(\eta)\left(f(\sigma^{x,x+1}\eta) - f(\eta)\right)^2 d\mu\,. \tag{29}$$

The Dirichlet form is defined by $\langle -\mathcal{L}_n^\theta f, f\rangle_\mu$, where we can rewrite for short the infinitesimal generator as

$$\mathcal{L}_n^\theta f(\eta) := \sum_{x=0}^{n-1} L_{x,x+1} f(\eta) := \sum_{x=0}^{n-1} r_{x,x+1}(\eta)(f(\sigma^{x,x+1}\eta) - f(\eta))\,.$$

Now, we recall that we consider the case $\alpha = \beta = \rho \in (0,1)$, so that the measure ν_ρ^n (the Bernoulli product measure) is invariant for this process and it satisfies

$$r_{x,x+1}(\eta)\,\nu_\rho^n(\eta) = r_{x,x+1}(\sigma^{x,x+1}\eta)\,\nu_\rho^n(\sigma^{x,x+1}\eta)\,, \tag{30}$$

for all $x \in \{0,1,\ldots,n-1\}$. Let us check this equality in the case $x = 0$, the case $x = n-1$ is similar and the others are also very simple to check. Note that

$$r_{0,1}(\sigma^{0,1}\eta)\,\frac{\nu_\rho^n(\sigma^{0,1}\eta)}{\nu_\rho^n(\eta)} = \left[\frac{\rho}{n^\theta}(1-\eta^1(1)) + \frac{1-\rho}{n^\theta}\eta^1(1)\right]\frac{\nu_\rho^n(\eta^1)}{\nu_\rho^n(\eta)}. \tag{31}$$

Since

$$\frac{\nu_\rho^n(\eta^1)}{\nu_\rho^n(\eta)} = \mathbf{1}_{\eta(1)=1}\frac{1-\rho}{\rho} + \mathbf{1}_{\eta(1)=0}\frac{\rho}{1-\rho}\,, \tag{32}$$

then (31) becomes

$$\mathbf{1}_{\eta(1)=1}\left[\frac{\rho}{n^\theta}\right]\frac{1-\rho}{\rho} + \mathbf{1}_{\eta(1)=0}\left[\frac{1-\rho}{n^\theta}\right]\frac{\rho}{1-\rho} = r_{0,1}(\eta)\,.$$

Thus, using (30), we get

$$\langle -\mathcal{L}_n^\theta f,\, f\rangle_{\nu_\rho^n} = \frac{1}{2}\mathcal{D}_n(f,\nu_\rho^n)\,. \tag{33}$$

Lemma 3 (Replacement Lemma) *Let $x = 1, n-1$ and $t > 0$ fixed. Then*

$$\mathbb{E}_\rho^{\theta,n}\left[\left(\int_0^t c_n\big(\eta_s(x)-\rho\big)\,ds\right)^2\right] \le C\frac{c_n^2 n^\theta}{n^2}.$$

Remark 2 Recall that for $\theta < 1$ we have in (25) $c_n = \sqrt{n}$, so that the error above becomes n^θ/n, which vanishes as $n \to \infty$. Recall that for $\theta > 1$ we have in (26) $c_n = n^{3/2}/n^\theta$, so that the error above becomes n/n^θ, which vanishes as $n \to \infty$.

Proof The proof follows by a classical argument combining both the Kipnis–Varadhan's inequality (see [18, p. 333, Lemma 6.1]) with Young's inequality. For that purpose let $x = 1$ (the other case is completely analogous) and note that the expectation in the statement of the lemma can be bounded from above by

$$\sup_{f\in L^2_{\nu_\rho^n}}\left\{\int c_n(\eta(1)-\rho)\,f(\eta)\,d\nu_\rho^n + n^2\langle\mathcal{L}_n^\theta f,\, f\rangle_{\nu_\rho^n}\right\}, \tag{34}$$

where $L^2_{\nu_\rho^n}$ is the space of functions f such that $\int f^2(\eta)\,d\nu_\rho^n < +\infty$. We start by writing the integral $\int c_n(\eta(1)-\rho)\,f(\eta)d\nu_\rho^n$ as twice its half and in one of the terms

we make the exchange $\eta \to \eta^1$ to have

$$\frac{1}{2}\int c_n(\eta(1)-\rho)f(\eta)\,d\nu_\rho^n + \frac{1}{2}\int c_n(1-\eta(1)-\rho)f(\eta^1)\frac{\nu_\rho^n(\eta^1)}{\nu_\rho^n(\eta)}\,d\nu_\rho^n\,,$$

see (32) to get the expression of $\frac{\nu_\rho^n(\eta^1)}{\nu_\rho^n(\eta)}$. A simple computation shows that the integral at the right hand side of last expression is equal to

$$-\frac{1}{2}\int c_n(\eta(1)-\rho)f(\eta^1)\,d\nu_\rho^n\,,$$

so that the display above is equal to

$$\frac{1}{2}\int c_n(\eta(1)-\rho)(f(\eta)-f(\eta^1))\,d\nu_\rho^n.$$

By Young's inequality we can bound the previous expression by

$$B\int c_n^2(\eta(1)-\rho)^2 d\nu_\rho^n + \frac{1}{4B}\int (f(\eta)-f(\eta^1))^2\,d\nu_\rho^n\,.$$

Now, remember the notation $\eta^1 = \sigma^{0,1}\eta$ and multiply and divide by $r_{0,1}(\eta)$ the integrand function inside the second integral above. We can do it, because there exists $\tilde{C}_\rho$ such that $\frac{\tilde{C}_\rho}{n^\theta} \le r_{0,1}(\eta) \le \frac{C_\rho}{n^\theta}$. Then we can bound the previous expression from above by

$$B\int c_n^2(\eta(1)-\rho)^2\,d\nu_\rho^n + \frac{n^\theta}{4B\tilde{C}_\rho}\int r_{0,1}(\eta)\,(f(\sigma^{0,1}\eta)-f(\eta))^2\,d\nu_\rho^n\,.$$

Using (29) the second integral in the last expression is bounded from above by $\mathcal{D}_n(f,\nu_\rho^n)$. Recalling (33), we get

$$\int c_n(\eta(1)-\rho)f(\eta)\,d\nu_\rho^n \le B\,c_n^2\int(\eta(1)-\rho)^2\,d\nu_\rho^n + \frac{n^\theta}{2B\tilde{C}_\rho}\langle -\mathcal{L}_n^\theta f, f\rangle_{\nu_\rho^n}\,.$$

Putting this inequality in (34) and choosing $B = n^{\theta-2}/2\tilde{C}_\rho$, the term at the right hand side of the last expression cancels with $n^2\langle\mathcal{L}_n^\theta f, f\rangle_{\nu_\rho^n}$. Therefore, the expectation appearing in the statement of the lemma is bounded from above by

$$\frac{c_n^2 n^\theta}{2\tilde{C}_\rho n^2}\int(\eta(1)-\rho)^2\,d\nu_\rho^n\,.$$

Since η is bounded the proof ends.

Acknowledgements A. N. was supported through a grant "L'ORÉAL - ABC - UNESCO Para Mulheres na Ciência". P. G. thanks FCT/Portugal for support through the project UID/MAT/04459/2013. T. F. was supported by FAPESB through the project Jovem Cientista-9922/2015. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovative programme (grant agreement No 715734).

References

1. Baldasso, R., Menezes, O., Neumann, A., Souza, R.R.: Exclusion process with slow boundary. J. Stat. Phys. **167**, 1112–1142 (2017)
2. Birkhoff, G., Rota, G.-C.: Ordinary Differential Equations, 4th edn. Wiley, New York (1989)
3. Blythe, R.A., Evans, M.R.: Nonequilibrium steady states of matrix-product form: a solver's guide. J. Phys. A: Mathe. Theor. **40**(46), R333 (2007)
4. Bodineau, T., Derrida, B., Lebowitz, J.L.: A diffusive system driven by a battery or by a smoothly varying field. J. Stat. Phys. **140**(4), 648–675 (2010)
5. Boyce, W., DiPrima, R.: Elementary Differential Equations and Boundary Value Problems, 9th edn. Wiley, New York (2009)
6. De Masi, A., Presutti, E., Tsagkarogiannis, D., Vares, M.E.: Current reservoirs in the simple exclusion process. J. Stat. Phys. **144**(6), 1151–1170 (2011)
7. De Masi, A., Presutti, E., Tsagkarogiannis, D., Vares, M.E.: Non equilibrium stationary state for the sep with births and deaths. J. Stat. Phys. **147**(3), 519–528 (2012)
8. De Masi, A., Presutti, E., Tsagkarogiannis, D., Vares, M.E.: Truncated correlations in the stirring process with births and deaths. Electron. J. Probab. **17**(6), 1–35 (2012)
9. De Masi, A., Ferrari, P., Presutti, E.: Symmetric simple exclusion process with free boundaries. Probab. Theory Relat. Fields **161**(1), 155–193 (2015)
10. Derrida, B.: Non-equilibrium steady states: fluctuations and large deviations of the density and of the current. J. Stat. Mech. **2007**(7), P07023 (2007)
11. Derrida, B., Janowsky, S.A., Lebowitz, J.L., Speer, E.R.: Exact solution of the totally asymmetric simple exclusion process: shock profiles. J. Stat. Phys. **73**(5), 813–842 (1993)
12. Farfán, J.: Hydrostatics, statical and dynamical large deviations of boundary driven gradient symmetric exclusion processes. Ph.D. thesis, 2008
13. Franco, T., Gonçalves, P., Neumann, A.: Hydrodynamical behavior of symmetric exclusion with slow bonds. Ann. Inst. H. Poincaré Probab. Stat. **49**(2), 402–427 (2013)
14. Franco, T., Gonçalves, P., Neumann, A.: Phase transition in equilibrium fluctuations of symmetric slowed exclusion. Stoch. Process. Appl. **123**(12), 4156–4185 (2013)
15. Franco, T., Gonçalves, P., Neumann, A.: Phase transition of a heat equation with Robin's boundary conditions and exclusion process. Trans. Am. Math. Soc. **367**(9), 6131–6158 (2015)
16. Franco, T., Gonçalves, P., Neumann, A.: Non-equilibrium and stationary fluctuations of a slowed boundary symmetric exclusion, arXiv e-prints (2016)
17. Holley, R.A., Stroock, D.W.: Generalized Ornstein-Uhlenbeck processes and infinite particle branching brownian motions. Publ. Res. Inst. Math. Sci. **14**(3), 741–788 (1978)
18. Kipnis, C., Landim, C.: Scaling limits of interacting particle systems. Grundlehren der Mathematischen Wissenschaften. [Fundamental Principles of Mathematical Sciences] vol. 320. Springer, Berlin (1999)
19. Landim, C., Milanés, A., Olla, S.: Stationary and nonequilibrium fluctuations in boundary driven exclusion processes. Markov Process Relat. Fields **14**(2), 165–184 (2008)
20. Mitoma, I.: Tightness of probabilities on $C([0, 1]; Y')$ and $D([0, 1]; Y')$. Ann. Probab. **11**(4), 989–999 (1983)
21. Reed, M., Simon, B.: Methods of Modern Mathematical Physics I: Functional Analysis, 1st edn. Academic Press, New York (1981)

From the N-Body Schrödinger Equation to the Vlasov Equation

François Golse

Abstract This paper describes a method for obtaining an estimate of the convergence rate for the joint mean-field and semiclassical limit of the N-particle Schrödinger equation leading to the Vlasov equation. The interaction force is assumed to be Lipschitz continuous. This is an account of a recent work in collaboration with T. Paul [Arch. Rational Mech. Anal. https://doi.org/10.1007/s00205-016-1031-x].

Keywords Mean-field limit · Vlasov equation · Schrödinger equation Wasserstein distance · Töplitz operator

1 Introduction

The Vlasov equation is a mean-field model in the kinetic theory of charged/massive particles. Vlasov equations are used in plasma physics to describe the dynamics of charged particles, or in cosmology to describe the collective motion of massive celestial bodies.

In the case where the interaction force between elementary constituents (ions or electrons in plasma physics, for instance) are Lipschitz continuous, the Vlasov equation has been derived from the N-body problem of classical mechanics in the large N, small coupling constant limit (see the works of Neunzert–Wick [26], Braun–Hepp [8] and Dobrushin [11]).

In the present paper, we shall investigate the following natural problem.

Problem: Is it possible to derive the Vlasov equation from the *quantum* N-body problem by a *joint semiclassical* ($\hbar \to 0$) *and mean-field* ($N \to \infty$) limit[1]?

F. Golse (✉)
CMLS, École Polytechnique and CNRS, Université Paris-Saclay, 91128 Palaiseau Cedex, France
e-mail: francois.golse@polytechnique.edu

[1] It is customary among mathematicians to consider the reduced Planck constant $\hbar$ as a vanishingly small parameter in the semiclassical limit. This is obviously improper since $\hbar$ is a constant. Strictly speaking, the semiclassical limit holds for mechanical systems whose typical action is of an order of magnitude $\gg \hbar$.

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_10

This problem has received the attention of several authors: see for instance the works of Graffi–Martinez–Pulvirenti [16], Pezzotti–Pulvirenti [28], and the more recent paper by Benedikter–Porta–Saffirio–Schlein [7], which discusses the Vlasov limit for a gas of fermions described by the Hartree–Fock equations. (The mean-field limit in the case of fermions involves an equivalent Planck constant of order $N^{-1/3}$, and therefore shares some features with the semiclassical limit.)

The interest for the mean-field limit in the quantum N-body problem in quantum mechanics can be explained as follows. In many applications, the particle number N is very large (typically, from 10^2 to 10^{23} or more…)

Now, the N-body problem in *classical mechanics* involves a phase space of dimension $6N$ (i.e. 3 degrees of freedom for the position and 3 degrees of freedom for the momentum for each point particle). On the other hand, the N-body problem in *quantum mechanics* involves wave functions defined on a configuration space of dimension $3N$. This explains the need for a reduced description, in the single-particle phase-space $\mathbf{R}^6$ in the case of classical mechanics, and in the single-particle configuration space $\mathbf{R}^3$ in the case of quantum mechanics.

This situation can be summarized by the following diagram.

$$\begin{array}{ccc} \boxed{\textbf{Schrödinger}} & \overset{N\to\infty}{\longrightarrow} & \boxed{\textbf{Hartree}} \\ \downarrow & & \downarrow \\ \hbar \to 0 & \searrow & \hbar \to 0 \\ \downarrow & & \downarrow \\ \boxed{\textbf{Liouville}} & \overset{N\to\infty}{\longrightarrow} & \boxed{\textbf{Vlasov}} \end{array}$$

In this diagram, the horizontal arrows correspond to the mean-field limit (i.e. the limit as $N \to \infty$). Thus, the Hartree equation is the mean-field limit of the N-body Schrödinger equation in quantum mechanics, in the same way as the Vlasov equation is the mean-field limit of the N-body Liouville equation in classical mechanics. The vertical arrows correspond to the semiclassical limit (i.e. the limit as $\hbar \to 0$). The Liouville equation can be obtained as the semiclassical limit of the Schrödinger equation, while the Vlasov equation can be likewise obtained as the semiclassical limit of the Hartree equation.

The semiclassical limit of quantum mechanics is an old and distinguished subject, on which there is a huge body of literature. The reader is referred to Chap. VII of [19] for an introduction to the subject aimed at physicists, and Appendix 11 of [2] or Theorem 5.1 in [1] for a mathematical statement. Both [1, 2] refer to Maslov's treatise [25] for a proof of the semiclassical limit; an alternate, perhaps simpler proof of the same result, based on the Laptev–Sigal parametrix [20], can be found in [5].

The limit discussed in the present paper corresponds to the diagonal arrow — i.e. to the *simultaneous* mean-field ($N \to \infty$) and semiclassical ($\hbar \to 0$) limits. Obviously, the validity of this limit is a consequence of the uniformity as $\hbar \to 0$ of the mean-field

limit in quantum mechanics (the upper horizontal arrow), and of the semiclassical limit of the Hartree equation, leading to the Vlasov equation. The semiclassical limit of the Hartree equation has been studied for instance by Lions–Paul [24] in terms of Wigner measures (see Theorem IV.2 in [24]). Interestingly, their result applies to singular potentials, including the Coulomb potential, which is of great interest for applications (see Theorem IV.5 in [24]). The uniformity as $\hbar \to 0$ of the mean-field limit in quantum mechanics has been recently obtained in [13], and formulated in terms of some quantum analogue of the Monge–Kantorovich (or Kantorovich-Rubinshtein, or Vasershtein) distances, following the work of Dobrushin [11] in classical mechanics. The main result in [13] applies to the case of Lipschitz continuous interaction forces.

Independently of this question of uniformity, there is an important body of literature on the mean-field limit in quantum mechanics. Most of these results use the formalism of BBGKY hierarchies: [3, 31] in the case of bounded potentials, [12] in the case of the Coulomb potential. The problem of estimating the convergence rate for this limit has been studied in [30] (notice that this reference does not use the formalism of BBGKY hierarchies). More recently, a much simpler method, also avoiding the use of BBGKY hierarchies has been proposed in [29]. This new method applies to singular interaction forces, with significant restrictions on the limiting solution (which has to be a pure state). All these estimates involve either the trace norm or the operator norm, and therefore cannot be uniform as $\hbar \to 0$.

The results described in this paper have been obtained in collaboration with T. Paul [14].

2 Quantum Versus Classical Dynamics

2.1 The Vlasov Equation

The unknown of the Vlasov equation is $f \equiv f(t, x, \xi)$, the particle distribution function, which is a time-dependent probability density on the single particle phase space $\mathbf{R}^d \times \mathbf{R}^d$: in other words, f is measurable and

$$f \geq 0 \text{ a.e. on } \mathbf{R}_+ \times \mathbf{R}^d \times \mathbf{R}^d \,, \quad \iint_{\mathbf{R}^d \times \mathbf{R}^d} f(t, x, \xi) dx d\xi = 1 \,.$$

The Vlasov equation is a mean-field kinetic equation governing the dynamics of a gas of identical particles interacting via the potential V:

$$\partial_t f + \xi \cdot \nabla_x f - \nabla V_f(t, x) \cdot \nabla_\xi f = 0 \,, \quad x, \xi \in \mathbf{R}^d \,, \quad t \geq 0 \,.$$

The potential V is an even, real-valued function whose regularity will be discussed later.

This equation is obviously Hamiltonian: the left hand side is indeed of the form

$$(\partial_t + \xi \cdot \nabla_x - \nabla V_f(t,x) \cdot \nabla_\xi) f(t,x,\xi) = \partial_t f(t,x,\xi) + \{H_f(t,x,\xi),\, f(t,x,\xi)\}\,,$$

where the mean-field, self-consistent potential and Hamiltonian are given by the formulas

$$\begin{cases} V_f(t,x) := \displaystyle\iint_{\mathbf{R}^d \times \mathbf{R}^d} V(x-y) f(t,y,\xi) dy d\xi\,, \\ \\ H_f(t,x,\xi) := \frac{1}{2}|\xi|^2 + V_f(t,x)\,, \end{cases}$$

while the Poisson bracket is defined by the prescription

$$\{x_m, x_n\} = \{\xi_m, \xi_n\} = 0\,, \qquad \{\xi_m, x_n\} = \delta_{mn}\,, \quad m,n = 1,\dots,d\,.$$

2.2 The N-Body Schrödinger Equation

The unknown of the N-body Schrödinger equation is the N-body wave function $\Psi_N \equiv \Psi_N(t,x_1,\dots,x_N) \in \mathbf{C}$, where x_j is the position of the j-th particle for each $j = 1,\dots,N$. The N-body configuration space is $\mathfrak{H}_N := \mathfrak{H}^{\otimes N} \simeq L^2((\mathbf{R}^d)^N)$, with $\mathfrak{H} := L^2(\mathbf{R}^d)$.

The Schrödinger equation is written in terms of the quantum N-body Hamiltonian

$$\mathscr{H}_N := \sum_{j=1}^N -\tfrac{1}{2}\hbar^2 \Delta_{x_j} + \frac{1}{N} \sum_{1 \le j < k \le N} V(x_j - x_k)\,,$$

which involves the real-valued, even interaction potential V. The quantum Hamiltonian is viewed as an unbounded operator acting on $\mathfrak{H}_N$, and the Schrödinger equation takes the form

$$i\hbar \partial_t \Psi_N = \mathscr{H}_N \Psi_N := \sum_{j=1}^N -\tfrac{1}{2}\hbar^2 \Delta_{x_j} \Psi_N + \frac{1}{N} \sum_{1 \le j < k \le N} V(x_j - x_k) \Psi_N\,, \qquad x_1,\dots,x_N \in \mathbf{R}^d\,. \tag{1}$$

Observe the $1/N$ coupling constant in front of the potential. It is chosen so that the kinetic energy of the N-particle system, i.e.

$$\sum_{j=1}^N -\tfrac{1}{2}\hbar^2 \Delta_{x_j}$$

which involves a N-term summation, is of the same order of magnitude as its potential energy

$$\frac{1}{N}\sum_{1\le j<k\le N} V(x_j - x_k)$$

involving a N^2-term summation. The $1/N$ coupling constant is typical of the mean-field limit for particles which do not satisfy the Pauli exclusion principle — in other words, one should keep in mind that the mean-field scaling for fermions is completely different (see [6] on p. 1089).

2.3 The N-Body Heisenberg Equation

Let $\rho(t)$ be the $\mathfrak{H}_N$-orthogonal projection on the complex line $\mathbf{C}\Psi_N(t,\cdot) \subset \mathfrak{H}_N$. Adopting Dirac's notation involving "bras" and "kets" (see Chap. I.6 in [10])

$$\rho_N(t) := |\Psi_N(t,\cdot)\rangle\langle\Psi_N(t,\cdot)|\,.$$

If Ψ_N satisfies the Schrödinger equation (1), then $\rho_N(t)$ satisfies the Heisenberg equation

$$i\hbar\partial_t\rho_N(t) = \mathscr{H}_N\rho_N(t) - \rho_N(t)\mathscr{H}_N =: [\mathscr{H}_N, \rho_N(t)]\,,$$

since the N-body quantum Hamiltonian $\mathscr{H}_N$ is self-adjoint on $\mathfrak{H}_N$.

More generally, the Heisenberg equation defines a dynamics on the set of *density operators* on $\mathfrak{H}_N$.

A density operator on $\mathfrak{H}_N$ is a linear operator ρ on $\mathfrak{H}_N$ satisfying

$$\rho = \rho^* \ge 0\,, \quad \mathrm{tr}_{\mathfrak{H}_N}(\rho) = 1\,.$$

The set of density operators on $\mathfrak{H}_N$ is henceforth denoted $\mathscr{D}(\mathfrak{H}_N)$.

3 Comparing Quantum and Classical Densities

3.1 Monge–Kantorovich(-Rubinshtein)/Vasershtein Distance

Let μ, ν be Borel probability measures on $\mathbf{R}^d$ with bounded p-th order moments, where $p \ge 1$.

A *coupling* of μ, ν is a Borel probability measure π on $\mathbf{R}^d \times \mathbf{R}^d$ with 1st and 2nd marginals μ and ν respectively. In other words, for all $\phi, \psi \in C_b(\mathbf{R}^d)$, one assumes that

$$\iint_{\mathbf{R}^d\times\mathbf{R}^d} (\phi(x)+\psi(y))\pi(dxdy) = \int_{\mathbf{R}^d} \phi(x)\mu(dx) + \int_{\mathbf{R}^d} \psi(y)\nu(dy)\,.$$

The set of couplings of μ, ν is denoted $\Pi(\mu,\nu)$.

The Monge–Kantorovich distance of exponent p between μ and ν is defined as

$$\mathrm{dist}_{MK,p}(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \left(\iint_{\mathbf{R}^d\times\mathbf{R}^d} |x-y|^p \pi(dxdy) \right)^{1/p}.$$

This distance is often referred to as the "Vasershtein (or Wasserstein) distance" and sometimes as the "Kantorovich–Rubinshtein distance" (the latter terminology being used mostly in the case $p=1$). An excellent reference for this class of distances is Chap. 7 of [32]. (Even the fact that the expression above satisfies the triangle inequality is far from obvious: see Sect. 1 in Chap. 7 of [32].)

Perhaps the most important result on Monge–Kantorovich distances for the purpose of our study is that the Monge–Kantorovich distance of exponent p metrizes the weak topology of probability measures with bounded moments of order p (see Sect. 2 in Chap. 7 of [32]).

The following elementary computations show that the Monge–Kantorovich distance between probability measures is well suited for measuring the proximity/stability of particle trajectories — at variance with the distance defined by the total variation.

Example 1 Let $a, b \in \mathbf{R}^d$; then

$$\mathrm{dist}_{MK,p}(\delta_a,\delta_b) = |a-b|\,,$$

whereas

$$\|\delta_a-\delta_b\|_{TV} = \begin{cases} 2 & \text{if } a\neq b\,,\\ 0 & \text{if } a=b\,. \end{cases}$$

The essence of the semiclassical limit is that the dynamics of the quantum density operators concentrates on particle trajectories in phase space. Therefore, it is natural to expect that a uniform in $\hbar$ convergence rate estimate for the mean-field limit in quantum mechanics should involve some analogue of a Monge–Kantorovich distance. Indeed, Monge–Kantorovich distances metrize the weak topology of probability measures (with additional moment bounds). These distances are particularly well adapted to estimating the difference between a smooth density and a probability measure concentrated on a lower dimensional set, such as a Dirac measure. This was the rationale for introducing the "pseudo-distance" $MK_2^\hbar$ in [13]. The reader is referred to (4) below for a precise definition of this quantity, which is obviously analogous to the quadratic Monge–Kantorovich distance $\mathrm{dist}_{MK,2}$ on the set of Borel probability measures in phase space.

3.2 Coupling Quantum and Classical Densities

Next we seek to define a notion of "pseudo-distance" between a quantum and a classical density. This pseudo-distance is constructed by analogy with the Monge–Kantorovich distance of exponent 2, following Dobrushin's 1979 derivation of the Vlasov equation in [11].

As a first step we need to define a notion of coupling of a quantum and of a classical density.

Definition 1 Let $\rho \in \mathscr{D}(\mathfrak{H})$ and let p be a Borel probability density on $\mathbf{R}^d \times \mathbf{R}^d$. A coupling of ρ and p is an operator-valued function $(x, \xi) \mapsto Q(x, \xi)$ defined on $\mathbf{R}^d \times \mathbf{R}^d$ such that

$$\begin{cases} (x, \xi) \mapsto Q(x, \xi) = Q(x, \xi)^* \in \mathscr{L}(\mathfrak{H}) \quad \text{s.t. } Q(x, \xi) \geq 0\,, \\ \operatorname{tr}(Q(x, \xi)) = p(x, \xi)\,, \qquad \displaystyle\iint_{\mathbf{R}^d \times \mathbf{R}^d} Q(x, \xi) dx d\xi = \rho\,. \end{cases}$$

The set of all couplings of the densities ρ and p is denoted $\mathscr{C}(p, \rho)$.

Example 2 The set $\mathscr{C}(p, \rho)$ is obviously nonempty, since the map

$$p \otimes \rho : (x, \xi) \mapsto p(x, \xi)\rho \qquad \text{belongs to } \mathscr{C}(p, \rho)\,.$$

3.3 Pseudo-distance Between Quantum and Classical Densities

Next we introduce a cost function comparing classical and quantum "coordinates" (i.e. position and momentum). This cost function is defined as follows:

$$c_\hbar(x, \xi) := |x - y|^2 + |\xi + i\hbar\nabla_y|^2\,.$$

The reason for considering a quadratic cost function is that quadratic polynomials are quantized exactly as second order differential operators. In other words, $c_\hbar$ is the quantization of the cost-function $|x - y|^2 + |\xi - \eta|^2$ defined on the 2-body phase space $\mathbf{R}^6_{x,\xi} \times \mathbf{R}^6_{y,\eta}$.

Definition 2 Let $\rho \in \mathscr{D}(\mathfrak{H})$ and let p be a Borel probability density on $\mathbf{R}^d \times \mathbf{R}^d$. We define

$$E_\hbar(p, \rho) := \inf_{Q \in \mathscr{C}(p,\rho)} \sqrt{\iint_{\mathbf{R}^d \times \mathbf{R}^d} \operatorname{tr}(c_\hbar(x, \xi) Q(x, \xi)) dx d\xi}\,.$$

This pseudo-distance[2] is obviously analogous to the Monge–Kantorovich distance with exponent 2.

Remark 1 The integrand in the expression defining $E_\hbar(p,\rho)$ may not be defined, since $c_\hbar(x,\xi)Q(x,\xi)$ may fail in general to be a trace-class operator on $\mathfrak{H}$. In fact this integrand should be thought of as being defined by the following formula:

$$\mathrm{tr}(c_\hbar(x,\xi)Q(x,\xi)) := \mathrm{tr}(Q(x,\xi)^{1/2}c_\hbar(x,\xi)Q(x,\xi)^{1/2}) \in [0,\infty]\,.$$

Indeed, the right hand side above always defines an element of $[0,+\infty]$ since $Q(x,\xi)^{1/2}c_\hbar(x,\xi)Q(x,\xi)^{1/2}$ is a self-adjoint nonnegative operator for a.e. $x,\xi \in \mathbf{R}^d$. (This is analogous to the integral of a nonnegative measurable function, which always defines an element of $[0,+\infty]$.)

Unfortunately, the pseudo-distance defined by the expression above is a slightly mysterious object. In the sequel, we seek to compare it to more familiar quantities.

3.4 Husimi Transform and Lower Bound for $E_\hbar$

First we recall the notions of Wigner and Husimi transforms of a density operator on a Hilbert space.

Definition 3 Let $\rho \in \mathscr{D}(\mathfrak{H})$. The Wigner transform at scale $\hbar$ of the density operator ρ is the function defined on the phase space $\mathbf{R}^d_x \times \mathbf{R}^d_\xi$ by the formula

$$W_\hbar[\rho](x,\xi) := \tfrac{1}{(2\pi)^d}\int_{\mathbf{R}^d} e^{-i\xi\cdot y}\rho(x+\tfrac{1}{2}\hbar y, x-\tfrac{1}{2}\hbar y)dy\,.$$

The Husimi transform at scale $\hbar$ of the density operator ρ is the function on the phase space $\mathbf{R}^d_x \times \mathbf{R}^d_\xi$ defined in terms of the Wigner function by the formula

$$\widetilde{W}_\hbar[\rho] := e^{\hbar\Delta_{x,\xi}/4}W_\hbar[\rho]\,.$$

The Wigner function of a density operator $\rho \in \mathscr{D}(\mathfrak{H})$ satisfies

$$W_\hbar[\rho](x,\xi) \in \mathbf{R} \quad \text{for a.e. } x,\xi \in \mathbf{R}^d\,.$$

However, $W_\hbar[\rho]$ is in general not a.e. nonnegative. In other words, the Wigner transform does not map the set $\mathscr{D}(\mathfrak{H})$ of (quantum) density operators on configuration space into the set of (classical) probability densities on phase space.

---

[2] We shall refer to $E_\hbar$ as a "pseudo-distance" without further apology, although there is a well-defined notion of "pseudometric" in mathematics. Usually, generalizations of the notion of metric measure the proximity between elements of a same set, while $E_\hbar$ is designed to measure the proximity between objects of a different nature (specifically, between a classical and a quantum density).

On the other hand, the Husimi transform satisfies

$$\widetilde{W}_\hbar[\rho](x,\xi) \geq 0 \quad \text{for a.e. } x,\xi \in \mathbf{R}^d\,.$$

The next result establishes a lower bound for the pseudo-distance $E_\hbar$ in terms of the quadratic Monge–Kantorovich distance between the classical probability density and the Husimi transform of the quantum density.

Theorem 1 *Let p be a probability density on $\mathbf{R}^d \times \mathbf{R}^d$ s.t.*

$$\iint_{\mathbf{R}^d\times\mathbf{R}^d} (|x|^2+|\xi|^2)p(x,\xi)dxd\xi < \infty\,.$$

For each $\rho \in \mathscr{D}(\mathfrak{H})$, one has

$$E_\hbar(p,\rho)^2 \geq \max\left(d\hbar\,,\ \mathrm{dist}_{\mathrm{MK},2}(p,\widetilde{W}_\hbar[\rho])^2 - d\hbar\right)\,.$$

See Theorem 2.4 (2) in [14] for a proof of this result. Notice in particular that $E_\hbar(p,\rho) > 0$ for all p and ρ.

3.5 Töplitz Quantization and Upper Bound for $E_\hbar$

In this section, we first define the notion of Töplitz quantization, which can be thought of as the reciprocal of the Husimi transform, up to an error term of order $O(\hbar)$.

Given $q,\,p \in \mathbf{R}^d$, we define the *coherent state* $|q+ip,\hbar\rangle$, which is the wave function given by the formula

$$|q+ip,\hbar\rangle(x) := (\pi\hbar)^{-d/4}e^{-|x-q|^2/2\hbar}e^{ip\cdot x/\hbar}\,.$$

This wave function consists of a plane wave in the direction of p, modulated with a Gaussian amplitude centered at q. See Figs. 1 and 2 for a graphic representation of the modulus of $|q+ip,\hbar\rangle$.

The coherent state $|q,\,p\rangle$ is the quantum analogue of the perfectly localized phase space density $\delta_{q,p}$ on $\mathbf{R}^6_{x,\xi}$. Of course, this state is not perfectly localized in position and momentum because of Heisenberg's uncertainty principle. The Gaussian localization is known to saturate the Heisenberg uncertainty principle: see Theorem 226 in [17], which is quoted as "Weyl's inequality".

Definition 4 For each positive Borel measure μ on $\mathbf{C}^d$, the Töplitz operator with symbol μ is

$$\mathrm{OP}^T_\hbar(\mu) := \tfrac{1}{(2\pi\hbar)^d}\int_{\mathbf{C}^d} |z,\hbar\rangle\langle z,\hbar|\,\mu(dz) \geq 0\,.$$

For instance, an elementary computation shows that $\mathrm{OP}^T_\hbar(1) = I_{\mathfrak{H}}$.

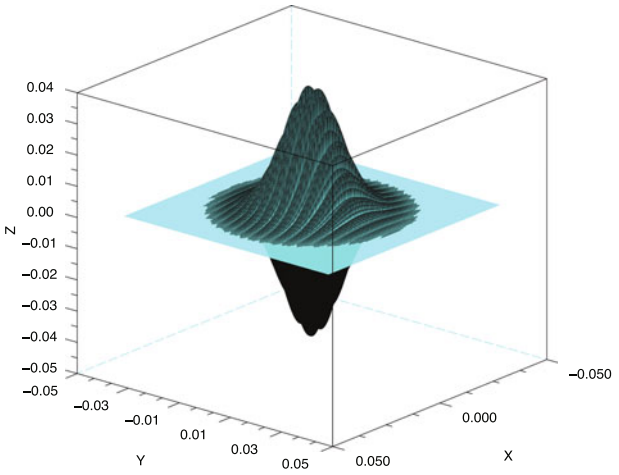


Fig. 1 With $\hbar = 8 \cdot 10^{-5}$, $Z =$ real part of coherent state centered at $q = (0, 0)$ with momentum $p = (1, 0)$ with space variable $(X, Y) \in \mathbf{R}^2$

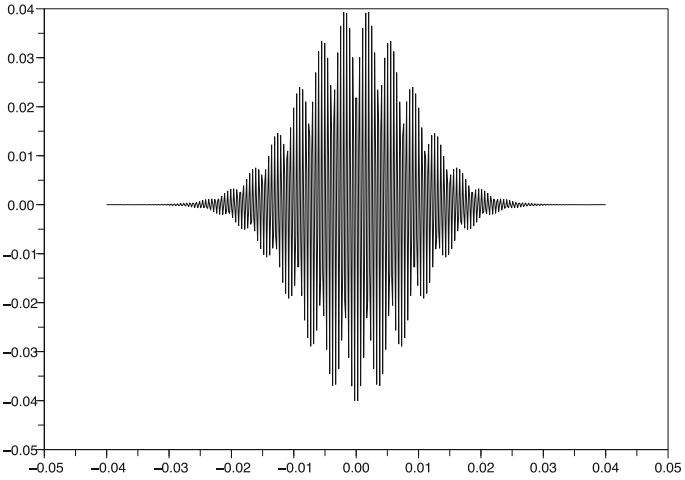


Fig. 2 Oscillating structure of a Gaussian coherent state. This is the section of the graph in Fig. 1 in the plane of equation $x = 0$

The relation between the Töplitz quantization and the Husimi transform is explained by the following formula:

$$\tilde{W}_\hbar[\mathrm{OP}_\hbar^T(\mu)] = \tfrac{1}{(2\pi\hbar)^d} e^{\hbar\Delta_{x,\xi}/2}\mu\,.$$

(See formula (51) in [13].) In other words, the Husimi transform of a Töplitz operator is equal to its symbol up to a smoothing operator which is an $O(\hbar)$ perturbation of the identity.

Theorem 2 *Under the same assumptions as in Theorem 1, let μ be a Borel probability measure on $\mathbf{C}^d$. Then $\mathrm{OP}^T_\hbar((2\pi\hbar)^d\mu) \in \mathscr{D}(\mathfrak{H})$ and*

$$E_\hbar(p, \mathrm{OP}^T_\hbar((2\pi\hbar)^d\mu))^2 \leq \mathrm{dist}_{\mathrm{MK},2}(p,\mu)^2 + d\hbar\,.$$

See Theorem 2.4 (1) in [14] for a proof of this result.

4 From the N-Body Heisenberg Equation to the Vlasov Equation

4.1 Indistinguishable Particles and Symmetries

Henceforth, we use the following notation for N-tuples of positions or momenta:

$$X_N := (x_1, \ldots, x_N)\,, \qquad \Xi_N := (\xi_1, \ldots, \xi_N)\,.$$

We also need the following representation of the symmetric group $\mathfrak{S}_N$ in $\mathfrak{H}_N$. For each $\sigma \in \mathfrak{S}_N$ and each $\Psi \in \mathfrak{H}_N$, we define $U_\sigma \in \mathscr{L}(\mathfrak{H}_N)$ by the following formula:

$$U_\sigma \Psi(X_N) := \Psi(\sigma \cdot X_N)\,, \qquad \text{where } \sigma \cdot X_N := (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(N)})\,.$$

Obviously

$$U_\sigma^* = U_\sigma^{-1} (= U_{\sigma^{-1}})\,;$$

in other words, U_σ is a unitary operator on $\mathfrak{H}_N$.

With this, we define a notion of symmetric quantum N-body density operator. A density operator $\rho \in \mathscr{D}(\mathfrak{H}_N)$ is said to be symmetric if

$$U_\sigma \rho_N U_\sigma^* = \rho_N\,, \qquad \text{for all } \sigma \in \mathfrak{S}_N\,.$$

The set of symmetric density operators on $\mathfrak{H}_N$ is denoted $\mathscr{D}^s(\mathfrak{H}_N)$. Since the N particles under consideration are indistinguishable, their density operator is obviously symmetric.

This symmetry property is propagated by Heisenberg's equation, in the following sense. If $\rho_N(t)$ is a solution of Heisenberg's equation, then

$$\rho_N(0) \in \mathscr{D}^s(\mathfrak{H}_N) \Rightarrow \rho_N(t) \in \mathscr{D}^s(\mathfrak{H}_N)\,, \quad \text{for all } t \in \mathbf{R}\,.$$

4.2 Symmetric Densities and k−Particle Marginals

For each symmetric N-particle density operator $\rho_N \in \mathscr{D}^s(\mathfrak{H}_N)$, its k-particle marginal is the symmetric density operator $\rho_N^{\mathbf{k}} \in \mathscr{D}^s(\mathfrak{H}_k)$ such that

$$\mathrm{tr}_{\mathfrak{H}_k}(A\rho_N^{\mathbf{k}}) = \mathrm{tr}_{\mathfrak{H}_N}((A \otimes I_{\mathfrak{H}_{N-k}})\rho_N)$$

for each $A \in \mathscr{L}(\mathfrak{H}_k)$.

Example 3 For each $\rho \in \mathscr{D}(\mathfrak{H})$ and all $N \geq k \geq 1$, one has

$$\rho_N := \rho^{\otimes N} \Rightarrow \rho_N^{\mathbf{k}} = \rho^{\otimes k}\,.$$

4.3 From the N-Body Heisenberg Equation to the Vlasov Equation

The main result in this paper is the following theorem (which is Theorem 2.6 in [14]).

Theorem 3 *Let $f^{in} \equiv f^{in}(x,\xi) \in L^1((|x|^2+|\xi|^2)dxd\xi)$ be a probability density on $\mathbf{R}^d \times \mathbf{R}^d$, and let $\rho_{N,\hbar}^{in} \in \mathscr{D}^s(\mathfrak{H}_N)$. Let f and $\rho_{N,\hbar}$ be the solutions of the Vlasov and the Heisenberg equation with initial data f^{in} and $\rho_{N,\hbar}^{in}$ respectively. Let Γ be given by the formula*

$$\Gamma := 2 + 4\max(1, \mathrm{Lip}(\nabla(V))^2\,.$$

(1) Then, for each $t \geq 0$ one has

$$E_\hbar(f(t), \rho_{\hbar,N}^{\mathbf{1}}(t))^2 \leq \frac{1}{N} E_\hbar((f^{in})^{\otimes N}, \rho_{\hbar,N}^{in})^2 e^{\Gamma t} + \frac{8\|\nabla V\|_{L^\infty}^2}{N-1}\frac{e^{\Gamma t}-1}{\Gamma}\,.$$

(2) If moreover $\rho_{\hbar,N}^{in} = \mathrm{OP}_\hbar^T[(2\pi\hbar)^{dN}(f^{in})^{\otimes N}]$, then

$$\mathrm{dist}_{\mathrm{MK},2}(f(t), \widetilde{W}_\hbar[\rho_{\hbar,N}^{\mathbf{1}}(t)])^2 \leq d\hbar(1+e^{\Gamma t}) + \frac{8\|\nabla V\|_{L^\infty}^2}{N-1}\frac{e^{\Gamma t}-1}{\Gamma}\,.$$

An example of initial bosonic state satisfying the assumption of Theorem 3 (2) is

$$\rho_{\hbar,N}^{in} = \mathrm{OP}_\hbar^T[(2\pi\hbar)^{dN}\delta_{(q,p)}^{\otimes N}] \quad \text{for some } q,\, p \in \mathbf{R}^d\,.$$

In other words, $\rho_{\hbar,N}^{in}$ is the density operator associated to the factorized N-body wave function

$$\Psi_{\hbar,N}^{in}(x_1,\dots,x_N) = \prod_{k=1}^N |q+ip,\hbar\rangle(x_k)\,.$$

Actually, one can replace the Gaussian profile in the definition of the wave packet $|q + ip, \hbar\rangle$ with any other real-valued function $a \in \mathscr{S}(\mathbf{R}^d)$. This would lead to an initial state defined in terms of the factorized N-body wave function

$$\Psi^{in}_{\hbar,N}(x_1, \ldots, x_N) = \prod_{k=1}^{N} a\left(\frac{x_k - q}{\sqrt{\hbar}}\right) e^{ip\cdot x_k/\hbar}.$$

With this modification, the inequality in Theorem 3 (2) remains unchanged, except for the first term on the right hand side, which should be replaced with $\hbar(d + H[a]e^{\Gamma t})$, where $H[a]$ is a positive constant depending on the function a only. The interested reader is referred to the forthcoming publication [15] for a detailed discussion on this matter.

We shall not insist further on this issue, since we are concerned with the joint mean-field and semiclassical limit of the quantum dynamics of N-particle systems. Indeed, the influence of the quantum statistics (i.e. the difference between bosons and fermions) disappears in the semiclassical limit.

5 Proof of Theorem 3

The core of the proof is based on an Eulerian analogue of Dobrushin's argument in [11], which was based on following particle trajectories. In addition, Dobrushin used the following essential feature of the N-particle dynamics in classical mechanics: the phase-space empirical measure built on a solution of Newton's system of motion equations is an exact (weak) solution of the Vlasov equation. No analogue of this property is known in quantum mechanics to this date. Our analysis uses instead the N-particle Heisenberg equation, which is the quantum analogue of the Liouville equation in classical mechanics.

5.1 Dynamics of Couplings

Our first task is to write an equation defining a dynamics on the set of couplings of the N-fold tensor product of the Vlasov solution and of the N-particle Heisenberg solution. This idea follows [13] — see also [9, 27] which used a similar idea in a rather different context (specifically, for nonlinear diffusion equations viewed as gradient flows in the sense of Otto). However, the procedure used in [13] and in [14] — described below — for propagating couplings of solutions of the Heisenberg and of the Vlasov–Poisson equations differs significantly from the method used in [9, 27]. We shall not insist further on this point, which is rather subtle; suffice it to say that the argument in [9, 27] uses tools from optimal transport — specifically, the Benamou–Brenier variational formula for the quadratic Monge–Kantorovich dis-

tance (formula (7.34) in [33]), specialized to gradient fields. There does not seem to be any obvious analogue of this method in the case considered below and in [14], which uses a completely different procedure.

Let $Q^{in}_{N,\hbar} \in \mathscr{C}((f^{in})^{\otimes N}, \rho^{in}_{N,\hbar})$; solve the classical+quantum transport equation

$$\begin{cases} \partial_t Q_{N,\hbar}(t, X_N, \Xi_N) + \left\{ \sum_{j=1}^N H_f(t, x_j, \xi_j), Q_{N,\hbar}(t, X_N, \Xi_N) \right\}_N \\ \qquad + \frac{i}{\hbar}[\mathscr{H}_N, Q_{N,\hbar}(t, X_N, \Xi_N)] = 0 , \\ Q_{N,\hbar}\big|_{t=0} = Q^{in}_{N,\hbar} , \end{cases} \tag{2}$$

where we recall that

$$\begin{cases} \mathscr{H}_N := \sum_{j=1}^N -\frac{1}{2}\hbar^2 \Delta_{y_j} + \frac{1}{N} \sum_{1 \le j < k \le N} V(y_j - y_k) , \\ H_f(t, x, \xi) := \frac{1}{2}|\xi|^2 + \iint_{\mathbf{R}^d \times \mathbf{R}^d} V(x - z) f(t, z, \zeta) dz d\zeta . \end{cases}$$

The notation $\{\cdot, \cdot\}_N$ designates the N-body Poisson bracket, defined by

$$\{x_{j,m}, x_{k,n}\}_N = \{\xi_{j,m}, \xi_{k,n}\}_N = 0 , \quad \{\xi_{j,m}, x_{k,n}\}_N = \delta_{jk}\delta_{mn} ,$$

with $j, k = 1, \dots, N$ and $m, n = 1, \dots, d$. In other words, the Eq. (2) describes the joint dynamics of N quantum, indistinguishable particles and of N associated classical, independent particles.

This mixed system of particles satisfies the following symmetry property.

Definition 5 Let $\rho_N \in \mathscr{D}^s(\mathfrak{H}_N)$. A coupling Q_N of $f^{\otimes N}$ and ρ_N is said to be a symmetric coupling if

$$U_\sigma Q_N(\sigma \cdot X_N, \sigma \cdot \Xi_N) U_\sigma^* = Q_N(X_N, \Xi_N) \qquad \text{for all } \sigma \in \mathfrak{S}_N ,$$

for a.e. $X_N, \Xi_N \in (\mathbf{R}^d)^N$.

In other words, the classical particles are exchanged by the *same* permutation as their quantum associates.

Not surprisingly, this symmetry property is propagated by (2).

Lemma 1 *Let $Q_{N,\hbar}$ be the solution of (2).*
(1) For each $t \ge 0$, one has

$$Q_{N,\hbar}(t) \in \mathscr{C}(f(t)^{\otimes N}, \rho_{N,\hbar}(t)) .$$

(2) If $Q^{in}_{N,\hbar}$ is a symmetric coupling, then, for all $t \geq 0$, the coupling $Q_{N,\hbar}(t)$ is symmetric.

This is Lemma 5.1 in [14], to which we refer for a proof. The key idea in the proof of statement (1) is the following: elementary computations show that the $\mathrm{tr}_{\mathfrak{H}_N}(Q(t, X_N, \Xi_N))$ is a classical N-body probability density which satisfies the same equation as

$$\prod_{j=1}^{n} f(t, x_j, \xi_j) .$$

Likewise

$$\int_{(\mathbf{R}^d \times \mathbf{R}^d)^N} Q(t, X_N, \Xi_N) dX_N d\Xi_N \in \mathscr{D}(\mathfrak{H}_N)$$

and satisfies the N-body Heisenberg equation. By uniqueness of the solution of the Vlasov and of the Heisenberg equation, this implies statement (1). Statement (2) follows from observing that Q_N and

$$(t, X_N, \Xi_N) \mapsto U_\sigma Q_N(t, \sigma \cdot X_N, \sigma \cdot \Xi_N) U_\sigma^*$$

are both solutions of the Cauchy problem (2) with the same initia data.

5.2 The Functional $D(t)$

For each symmetric coupling that is a solution of (2), define

$$D(t) := \frac{1}{N} \iint_{(\mathbf{R}^d \times \mathbf{R}^d)^N} \sum_{k=1}^{N} \mathrm{tr}_{\mathfrak{H}_N}(c_\hbar(x_j, \xi_j, y_j, \nabla_{y_j}) Q_{N,\hbar}(t, X_N, \Xi_N)) dX_N d\Xi_N .$$

Lemma 2 *For each $t \geq 0$, one has*

$$D(t) \geq E_\hbar(f(t), \rho^{\mathbf{1}}_{N,\hbar}(t))^2 .$$

Proof First $Q^{\mathbf{1}}_{N,\hbar}(t) \in \mathscr{C}(f(t), \rho^{\mathbf{1}}_{N,\hbar}(t))$, and by symmetry of $Q_{N,\hbar}$

$$\begin{aligned} D(t) &= \iint_{(\mathbf{R}^d \times \mathbf{R}^d)^N} \mathrm{tr}_{\mathfrak{H}_N}(c_\hbar(x_1, \xi_1, y_1, \nabla_{y_1}) Q_{N,\hbar}(t, X_N, \Xi_N)) dX_N d\Xi_N \\ &= \iint_{(\mathbf{R}^d \times \mathbf{R}^d)} \mathrm{tr}_{\mathfrak{H}}(c_\hbar(x_1, \xi_1, y_1, \nabla_{y_1}) Q^{\mathbf{1}}_{N,\hbar}(t, x_1, \xi_1)) dx_1 d\xi_1 \\ &\geq E_\hbar(f(t), \rho^{\mathbf{1}}_{N,\hbar}(t))^2 \end{aligned}$$

by definition of $E_\hbar$.

5.3 The Evolution of D

Apply the cost operator $c_\hbar(x_1, \xi_1, y_1, \nabla y_1)$ to both sides of the equation for $Q_{N,\hbar}$, integrate by parts in the classical variables and use the identity

$$\mathrm{tr}([A, S]T) = -\mathrm{tr}([A, T]S)\,,$$

which is in some sense analogous to "integration by parts in the quantum variables".

We thus arrive at the differential equality

$$\begin{aligned}\dot{D} = &\iint \mathrm{tr}_{\mathfrak{H}} \left((\{\tfrac{1}{2}|\xi_1|^2, c_\hbar\} - [\tfrac{i\hbar}{2}\Delta_{y_1}, c_\hbar])(x_1, \xi_1, y_1, \nabla y_1) Q^{\mathbf{1}}_{N,\hbar}(t, x_1, \xi_1) \right) dx_1 d\xi_1 \\ &+ \frac{i}{\hbar} \iint \mathrm{tr}_{\mathfrak{H}_2} \left(\tfrac{N-1}{N} [V(y_1 - y_2), c_\hbar(x_1, \xi_1, y_1, \nabla y_1)] Q^{\mathbf{2}}_{N,\hbar}(t, X_2, \Xi_2) \right) dX_2 d\Xi_2 \\ &+ \iint \mathrm{tr}_{\mathfrak{H}} \left(\{V_f, c_\hbar\}(x_1, \xi_1, y_1, \nabla y_1) Q^{\mathbf{1}}_{N,\hbar}(t, x_1, \xi_1) \right) dx_1 d\xi_1\,.\end{aligned}$$

At this point, we denote the symmetric operator product as follows

$$A \vee B := AB + BA\,,$$

and use the inequality

$$A^*B + B^*A \leq A^*A + B^*B\,, \quad \text{for all } A, B \in \mathscr{L}(\mathfrak{H}_N)\,. \tag{3}$$

(Indeed $A^*A + B^*B - A^*B - B^*A = (A^* - B^*)(A - B) \geq 0$.)

Then, one has

$$\{\tfrac{1}{2}|\xi_1|^2, c_\hbar\} - [\tfrac{i\hbar}{2}\Delta_{y_1}, c_\hbar] = (\xi_1 + i\hbar\nabla_{y_1}) \vee (x_1 - y_1) \leq c_\hbar\,.$$

Since $Q_{N,\hbar} \geq 0$, this inequality and the equality above for $\dot{D}$ imply that

$$\begin{aligned}\dot{D} \leq D &- \tfrac{N-1}{N} \int \mathrm{tr}_{\mathfrak{H}_2}(Q^{\mathbf{2}}_{N,\hbar}(\xi_1 + i\hbar\nabla_{y_1}) \vee \mathscr{W}(X_2, Y_2)) dX_2 d\Xi_2 \\ &- \int \mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}(\xi_1 + i\hbar\nabla_{y_1}) \vee \tfrac{1}{N-1} \sum_{j=2}^{N} \mathscr{V}(t, x_1, x_j)) dX_N d\Xi_N\,,\end{aligned}$$

using again the symmetry of $Q_{N,\hbar}$ in the last term, and denoting

$$\mathscr{V}(t,x_1,x_j) := \nabla V \star f(t,x_1) - \tfrac{N-1}{N}\nabla V(x_1-x_j)\,,$$
$$\mathscr{W}(X_2,Y_2) := \nabla V(x_1-x_2) - \nabla V(y_1-y_2)\,.$$

Indeed

$$\mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}(A^*B+B^*A)) \le \mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}(A^*A+B^*B))\,,$$

since

$$\mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}(A^*A+B^*B)) - \mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}(A^*B+B^*A))$$
$$= \mathrm{tr}_{\mathfrak{H}_N}(Q_{N,\hbar}^{1/2}(A^*-B^*)(A-B)Q_{N,\hbar}^{1/2}) \ge 0\,.$$

Therefore

$$\dot{D} \le (3+2L^2)D + \int_{(\mathbf{R}^d\times\mathbf{R}^d)^N} \left|\tfrac{1}{N-1}\sum_{k=2}^N \mathscr{V}(t,x_1,x_k)\right|^2 f^{\otimes N} dX_N d\Xi_N\,.$$

Expand the square in the last term on the right hand side, and observe that

$$j\neq k \Rightarrow \int \mathscr{V}(t,x_1,x_j)\cdot\mathscr{V}(t,x_1,x_k) f^{\otimes N} dX_N d\Xi_N = 0\,.$$

Therefore, the last term on the right hand side of the last inequality is equal to

$$\int \tfrac{1}{(N-1)^2}\sum_{j=2}^N |\mathscr{V}(t,x_1,x_k)|^2 f^{\otimes N} dX_N d\Xi_N$$
$$= \int \tfrac{1}{N-1}|\mathscr{V}(t,x_1,x_2)|^2 f^{\otimes 2} dX_2 d\Xi_2 = O(1/N)\,.$$

One concludes with the Gronwall inequality.

6 Final Remarks and Perspectives

The core of the convergence rate estimate in the proof of Theorem 3 involves a stability inequality and a consistency estimate, as in Lax's equivalence theorem [21] in numerical analysis.

The stability part of the analysis (leading to the exponential amplification by Gronwall's inequality) can be seen at the level of the first equation in the BBGKY hierarchy. Because the cost function in D is a *sum* of quantities depending on x_j, y_j, ξ_j, there is a "*localization in degree*" effect in the BBGKY hierarchy. In particular, there is no bound *à la* Cauchy–Kowalevsky when estimating D.

The consistency part of the analysis requires distributing the interaction term $\mathscr{V}$ on *all* the particles. Since the $\mathscr{V}$ term depends on the X_N variables only, and the X_N

marginal of $Q_{N,\hbar}$ is the N-fold tensor power of the Vlasov solution, one concludes by (a trivial quantitative variant of the) law of large numbers.

The quantity $E_\hbar$ introduced in [14] and Definition 2 can also be used to prove quantitative estimate of the convergence rate for the semiclassical limit of the Hartree equation, leading to the Vlasov equation. Likewise, one can also prove the semiclassical limit of the N-body Heisenberg equation leading to the N-body Liouville equation, and obtain a uniform in N estimate for the convergence rate in terms of the $E_\hbar$ pseudo-distance. These results are stated as Theorems 2.5 and 2.7 respectively in [14]. Both proofs assume that ∇V is Lipschitz continuous on $\mathbf{R}^d$, as in the derivation of the Vlasov equation from the N-body Heisenberg equation discussed here. While the semiclassical limit has been known for a long time, it is perhaps not without interest to compare the existing results in the literature with Theorems 2.5 and 2.7 in [14].

The traditional method for describing the semiclassical limit is the WKB ansatz (see Chap. VII in [19], especially Sect. 46). Controling the propagation of the WKB ansatz puts severe regularity requirements on the potential (which should be C^∞), and on the initial phase function (assumed to be C^m with $m > 6d + 5$): see [5]. The formulation of the semiclassical limit in [24] requires much less on the potential (which should be of class $C^{1,1}$, as in [14] and the present work), and essentially nothing on the family of initial density operators indexed by $\hbar$. The quantitative result in [14] only requires that the phase space moments of order 2 of the initial family of density operators should be bounded as $\hbar \to 0$. However, obtaining a precise description of the *structure* of the propagated Wigner measure in the semiclassical limit seems to require additional regularity properties on the initial condition. The propagation of the Wigner measure associated to a WKB ansatz in the semiclassical limit has been studied in detail in [4] with the tools of geometric measure theory. One of the conclusions to be drawn from the discussion in [4] is that the basic structure of the propagated WKB ansatz, especially as regards the notion of "caustic", may differ dramatically if the regularity of the initial phase function falls below some regularity threshold. However, this does not affect the validity of the semiclassical limit and the convergence rate for this limit in terms of the pseudo-distance $E_\hbar$. In other words, one can argue that the pseudo-distance $E_\hbar$ considered in [14] and in the present paper is probably the most appropriate tool for propagating convergence rate estimates in the semiclassical limit under minimal regularity assumptions.

We conclude this paper with a (very incomplete) list of open problems related to the ones considered in [14].

Problem 1. Can one generalize Theorem 3 to the case of the Coulomb potential $V = 1/r$? This would lead to a rigorous justification of the Vlasov–Poisson system from the quantum mechanics of a large number of identical particles. Deriving the Vlasov–Poisson system from the classical N-point dynamics with Coulomb interaction is still an outstanding open problem. There has been some recent progress on this problem. The best result with point particles at the time of this writing is due tu Hauray and Jabin [18]; it can handle singular interaction forces of order $O(r^{-\alpha})$ with $\alpha < 1$ in

terms of the inter-particle distance r. The Hauray–Jabin method falls short of treating the Coulomb interaction even in the most favorable space dimension 2.

Problem 2. Another class of results assumes that the interacting particles have a positive radius, vanishing as the number of particles $N \to \infty$. This suggests considering a mollified potential V, with a regularizing effect that is gradually removed as $N \to \infty$: see [22, 23] for the best results in that direction, leading ultimately to the Vlasov–Poisson system. Is it possible to obtain a similar result starting from the quantum N-body problem, and viewing $\hbar$ as a regularization parameter analogous to the particle radius in the classical case, with some asymptotic ordering of $\hbar$ and $1/N$?

Problem 3. It could be interesting to explore more thoroughly the properties of $E_\hbar$ and of the pseudo-distance $MK_2^\hbar$ used in [13]. Let us briefly recall the definition of $MK_2^\hbar$. Let $\mathfrak{H} := L^2(\mathbf{R}^d)$ and let $\rho_1, \rho_2 \in \mathscr{D}(\mathfrak{H})$. A coupling of ρ_1 and ρ_2 is a density operator $R \in \mathscr{D}(\mathfrak{H} \otimes \mathfrak{H})$ such that

$$\mathrm{tr}_{\mathfrak{H}_2}(R(A \otimes I)) = \mathrm{tr}_{\mathfrak{H}}(\rho_1 A)\,, \quad \mathrm{tr}_{\mathfrak{H}_2}(R(I \otimes A)) = \mathrm{tr}_{\mathfrak{H}}(\rho_2 A)\,.$$

Denoting by $\Pi(\rho_1, \rho_2)$ the set of couplings of ρ_1, ρ_2, one defines

$$MK_2^\hbar(\rho_1, \rho_2) := \inf_{R \in \Pi(\rho_1, \rho_2)} \sqrt{\mathrm{tr}_{\mathfrak{H}_2}((|x_1 - x_2|^2 - |\nabla_{x_1} - \nabla_{x_2}|^2)R)}\,. \tag{4}$$

We do not know whether $MK_2^\hbar$ satisfies the triangle inequality — at least none of the two proofs of the triangle inequality for the Monge–Kantorovich distances proposed in Chap. 7, Sect. 1 of [32], seems to have any obvious analogue for $MK_2^\hbar$.

Similarly, one can study the following question: let p be a probability density on $\mathbf{R}^d \times \mathbf{R}^d$ such that

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} (|x|^2 + |\xi|^2) p(x, \xi) dx d\xi < \infty\,,$$

and let $\rho_1, \rho_2 \in \mathscr{D}(\mathfrak{H})$. Does one have

$$E_\hbar(p, \rho_2) \le E_\hbar(p, \rho_1) + MK_2^\hbar(\rho_1, \rho_2)?$$

If such an inequality was known, one could immediately deduce Theorem 3 from Theorem 2.4 in [13] and from the convergence rate for the semiclassical limit obtained in Theorem 2.5 in [14].

Acknowledgements The author is indebted to Luigi Ambrosio and Thierry Paul for various suggestions and remarks during the preparation of this paper.

References

1. Arnold, V.I.: Characteristic class entering in quantization condition. Func. Anal. Appl. **1**, 1–14 (1967)
2. Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, New York (1989)
3. Bardos, C., Golse, F., Mauser, N.: Weak coupling limit of the N-particle Schrödinger equation. Methods Appl. Anal. **7**, 275–293 (2000)
4. Bardos, C., Golse, F., Markowich, P., Paul, T.: Hamiltonian evolution of monokinetic measures with rough momentum profile. Arch. Rational Mech. Anal. **217**, 71–111 (2015)
5. Bardos, C., Golse, F., Markowich, P., Paul, T.: On the classical limit of the Schrödinger equation. Discrete Contin. Dyn. Syst. **35**, 5689–5709 (2015)
6. Benedikter, N., Porta, M., Schlein, B.: Mean field evolution of fermionic systems. Commun. Math. Phys. **331**, 1087–1131 (2014)
7. Benedikter, N., Porta, M., Saffirio, C., Schlein, B.: Form the Hartree dynamics to the Vlasov equation. Arch. Rational Mech. Anal. **221**, 273–334 (2016)
8. Braun, W., Hepp, K.: The Vlasov dynamics and its fluctuations in the 1/N limit of interacting classical particles. Commun. Math. Phys. **56**, 101–113 (1977)
9. Daneri, S., Savaré, G.: Eulerian calculus for the displacement convexity in the Wasserstein distance. SIAM J. Math. Anal. **40**, 1104–1122 (2008)
10. Dirac, P.A.M.: The Principles of Quantum Mechanics. Oxford University Press (1958)
11. Dobrushin, R.: Vlasov equations. Funct. Anal. Appl. **13**, 115–123 (1979)
12. Erdös, L., Yau, H.-T.: Derivation of the nonlinear Schrödinger equation from a many body Coulomb system. Adv. Theor. Math. Phys. **5**, 1169–1205 (2001)
13. Golse, F., Mouhot, C., Paul, T.: On the mean-field and classical limits of quantum mechanics. Commun. Math. Phys. **343**, 165–205 (2016)
14. Golse, F., Paul, T.: The Schrödinger equation in the mean-field and semiclassical regime. Arch. Rational Mech. Anal. doi:10.1007/s00205-016-1031-x
15. Golse, F., Paul, T.: Wave packets in quantum mechanics and the quadratic Monge-Kantorovich distance. In preparation
16. Graffi, S., Martinez, A., Pulvirenti, M.: Mean-field approximation of quantum systems and classical limit. Math. Models Methods Appl. Sci. **13**, 59–73 (2003)
17. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities. Cambridge University Press, London (1934)
18. Hauray, M., Jabin, P.-E.: Particle approximations of Vlasov equations with singular forces. Ann. Sci. Ecol. Norm. Sup. **48**, 891–940 (2015)
19. Landau, L.D., Lifshitz, E.M.: Quantum Mechanics. Course of Theoretical Physics, vol. 3. Pergamon Press (1977)
20. Laptev, A., Sigal, I.: Global Fourier integral operators and semiclassical asymptotics. Rev. Math. Phys. **12**, 749–766 (2000)
21. Lax, P.D., Richtmeyer, R.: Survey of the stability of linear finite difference equations. Comm. Pure Appl. Math. **9**, 267–293 (1956)
22. Lazarovici, D.: The Vlasov-Poisson dynamics as the mean-field limit of extended charges. Preprint arXiv:1502.07047, to appear in Commun. Math. Phys
23. Lazarovici, D., Pickl, P.: A mean-field limit for the Vlasov-Poisson system (preprint). arXiv:1502.04608
24. Lions, P.-L., Paul, T.: Sur les mesures de Wigner. Rev. Math. Iber. **9**, 553–618 (1993)
25. Maslov, V.P.: Théorie des perturbations et méthodes asymptotiques. Dunod, Paris (1972)
26. Neunzert, H., Wick, J.: Die Approximation der Lösung von IntegroDifferentialgleichungen durch endliche Punktmengen. Lecture Notes in Math., vol. 395, pp. 275–290. Springer, Berlin (1974)
27. Otto, F., Westdickenberg, M.: Eulerian calculus for the contraction in the Wasserstein distance. SIAM J. Math. Anal. **37**, 1227–1255 (2005)
28. Pezzotti, F., Pulvirenti, M.: Mean-field limit and semiclassical expansion of a quantum particle system. Ann. Henri Poincaré **10**, 145–187 (2009)

29. Pickl, P.: A simple derivation of mean field limits for quantum systems. Lett. Math. Phys. **97**, 151–164 (2011)
30. Rodnianski, I., Schlein, B.: Quantum fluctuations and rate of convergence towards mean field dynamics. Commun. Math. Phys. **291**, 31–61 (2009)
31. Spohn, H.: Kinetic equations from Hamiltonian dynamics. Rev. Mod. Phys. **52**, 600–640 (1980)
32. Villani, C.: Topics in Optimal Transportation. American Mathematical Society, Providence (RI) (2003)
33. Villani, C.: Optimal Transport. Old and New. Springer, Berlin (2009)

Matrix Product Ansatz for Non-equilibrium Quantum Steady States

D. Karevski, V. Popkov and G.M. Schütz

Abstract We present a general construction of matrix product states for stationary density matrices of one-dimensional quantum spin systems kept out of equilibrium through boundary Lindblad dynamics. As an application we review the isotropic Heisenberg quantum spin chain which is closely related to the generator of the simple symmetric exclusion process. Exact and heuristic results as well as numerical evidence suggest a local quantum equilibrium and long-range correlations reminiscent of similar large-scale properties in classical stochastic interacting particle systems that can be understood in terms of fluctuating hydrodynamics.

Keywords Dissipative quantum spin chains · Lindblad equation
Matrix product ansatz · Nonequilibrium stationary states · Exact solution

1 The Quantum Master Equation

This article is concerned with stationary states of non-equilibrium quantum spin systems, addressing a mathematically minded readership. We spent some effort on recalling – in mathematical terms – relevant basic quantum mechanical notions as well as providing motivations from physics as to why quantum spin systems are of great current interest. Among them is, we feel, a striking analogy with some properties

D. Karevski
Institut Jean Lamour, Département P2M, Groupe de Physique Statistique,
Université de Lorraine, CNRS UMR 7198, B.P. 70239, 54506
Vandoeuvre Les Nancy Cedex, France
e-mail: dragi.karevski@univ-lorraine.fr

V. Popkov
Helmholtz-Institut Für Strahlen-und Kernphysik, Universität Bonn,
Nussallee 14-16, 53119 Bonn, Germany
e-mail: popkov@uni-bonn.de

G.M. Schütz (✉)
Institute of Complex Systems II, Forschungszentrum Jülich, 52425 Jülich, Germany
e-mail: g.schuetz@fz-juelich.de

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_11

of *classical* stochastic interacting particle systems [4, 10, 31] that we point out in the hope of stimulating further mathematically rigorous work.

Let $\mathfrak{H}$ be a separable complex Hilbert space. A concrete physical quantum system is mathematically defined by a specific self-adjoint (not necessarily bounded) linear operator H on $\mathfrak{H}$, called quantum Hamiltonian (in the following simply Hamiltonian). Vectors in $\mathfrak{H}$ are denoted by the ket-symbol $|\cdot\rangle$ and vectors in the dual space $\mathfrak{H}^*$ are denoted by the bra-symbol $\langle\cdot|$. The scalar product of two vectors $|\Psi\rangle = \sum_n c_n |n\rangle \in \mathfrak{H}$ and $|\Phi\rangle = \sum_n b_n |n\rangle \in \mathfrak{H}$ with coordinates $b_n, c_n \in \mathbb{C}$ in some orthonormal basis $|n\rangle$, $\langle n|$ of $\mathfrak{H}$ and its dual resp. is denoted $\langle\Phi|\Psi\rangle$ and defined to be linear in the *second* argument, i.e., $\langle\Phi|\Psi\rangle := \sum_n \bar{b}_n c_n$ where the bar denotes complex conjugation. We denote the unit operator on $\mathfrak{H}$ by $\mathbf{1}$. The Kronecker symbol $\delta_{a,b}$ is defined by $\delta_{a,b} = 1$ if $a = b$ and $\delta_{a,b} = 0$ else for a and b from any set.

The eigenvalues E_n of the Hamiltonian H are the physical energies measured in an experiment when the physical system is in an eigenstate n of H, defined by the corresponding eigenvector $|\Psi_n\rangle$. One normalizes these eigenvectors, which span the Hilbert space $\mathfrak{H}$, to satisfy the orthogonality relation $\langle\Psi_n|\Psi_m\rangle = \delta_{n,m}$. A spectral ray $|\Psi\rangle \in \mathfrak{H}$ normalized such that $||\Psi||^2 := \langle\Psi|\Psi\rangle = 1$ (i.e. a vector defined up to an arbitrary phase) is called a state vector. It represents the full information that one can have about a quantum system under the idealizing assumption that it is isolated (and has always been isolated) from its physical environment.[1] The modulus $|\psi_n|^2$ of the components of $|\Psi\rangle$ are the probabilities to find the physical system in eigenstate n.

In general, physically observable properties of a quantum system (e.g. particle positions, momenta and so on) are represented by self-adjoint linear operators O_i on $\mathfrak{H}$ which we call *observables*. The "fuzzy" and non-deterministic nature of quantum mechanics is reflected by the fact that the O_i are not all diagonal in some fixed basis of $\mathfrak{H}$ and that only the mean outcome of a large number (mathematically speaking, an infinite number) of measurements of such an observable is predictable. By the mean (or expected) value of an observable O in a general state vector $|\Psi\rangle$ we mean the scalar product $\langle O\rangle \equiv \langle\Psi|O|\Psi\rangle = \sum_{m,n} \bar{c}_m c_n \langle\Psi_m|O|\Psi_n\rangle$.

A self-adjoint positive definite linear operator on $\mathfrak{H}$ with unit trace is called a *density matrix* or *state* (not eigenstate!) of a physical system. Therefore a density matrix ρ with eigenvalues $\rho_n \in \mathbb{R}$ has the properties

$$\rho^\dagger = \rho, \quad \rho_n \geq 0, \quad \mathrm{Tr}(\rho) = 1 \tag{1}$$

where the dagger-symbol † denotes hermitian conjugation. For a given Hilbert space we denote the set of all density matrices by $\mathfrak{S}(\mathfrak{H})$. The mean value of an observable O_i in a state ρ is given by the Frobenius scalar product $\langle O_i\rangle := \mathrm{Tr}(O_i^\dagger \rho)$.

[1] Due to the quantum mechanical phenomenon of entanglement, a quantum subsystem that has interacted with its environment in the past (until some time t_0) cannot be considered isolated for $t \geq t_0$ even when there are no interactions from t_0 onwards.

Unlike a state vector describing a *single and isolated* quantum system, a density matrix contains the full information about a quantum system in either of the following three scenarios:

(1) A density matrix of the specific form

$$\rho = |\,\Psi\,\rangle\langle\,\Psi\,| \tag{2}$$

may describe a single isolated system.[2] In this case we say that ρ is a *pure* state. If a density matrix is not a pure state then there is no state vector $|\,\Psi\,\rangle$ such that $\mathrm{Tr}(O_i^\dagger \rho) = \langle\,\Psi\,|O|\,\Psi\,\rangle$ for all observables O_i.

(2) One describes an *ensemble* of identical isolated quantum systems. In particular, if for some $\beta \in \mathbb{R}_0^+$ the density matrix is of the form

$$\rho = \frac{1}{Z}\mathrm{e}^{-\beta H} \tag{3}$$

where $Z = \mathrm{Tr}\,(\exp(-\beta H))$ we say that the physical system defined by the Hamiltonian H is in thermal equilibrium at temperature $T = 1/\beta$ and the normalization factor Z is called the partition function. In this case the probability to find the system in an eigenstate n of H is proportional to the Boltzmann weight $\exp(-\beta E_n)$ analogous to classical thermodynamics.

(3) ρ describes a subsystem (or an ensemble thereof) of a larger physical system with which *it interacts* (or has interacted in the past).[3]

Pure states and equilibrium states have in common that they remain so when the physical system is isolated from its environment or becomes isolated from some time $t \geq t_0$ onwards. This follows from the time-evolution equation for the density matrix ρ_t of an isolated quantum system with quantum Hamiltonian H

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = -i[H, \rho_t] \tag{4}$$

where the commutator is defined by $[A, B] := AB - BA$. Therefore an equilibrium state is stationary. A pure state $\rho_0 = |\,\Psi(0)\,\rangle\langle\,\Psi(0)\,|$ is only stationary if $|\,\Psi(0)\,\rangle$ is an eigenstate of H, but generally remains a pure state since the evolution equation (4) is solved by the unitary transformation $\rho_t = \exp(-iHt)\rho_0 \exp(iHt)$ which gives $\rho_t = |\,\Psi(t)\,\rangle\langle\,\Psi(t)\,|$ with $|\,\Psi(t)\,\rangle = \exp(-iHt)|\,\Psi(0)\,\rangle$.

We are interested in open systems that are in contact with an environment. In the Markovian approach to open quantum systems [2, 7] the time evolution

---

[2]Following quantum mechanical convention we use the short hand $|\cdot\rangle\langle\cdot| \equiv |\cdot\rangle \otimes \langle\cdot|$ for the Kronecker product $\otimes$ of a state vector $|\cdot\rangle \in \mathfrak{H}$ and some dual state vector $\langle\cdot| \in \mathfrak{H}^*$. We stress that by the rules of tensor calculus one has $\langle\,\Psi\,| \otimes |\,\Phi\,\rangle = |\,\Psi\,\rangle \otimes \langle\,\Phi\,| \equiv |\,\Psi\,\rangle\langle\,\Phi\,|$ but $\langle\,\Psi\,| \otimes |\,\Phi\,\rangle \neq \langle\,\Psi\,|\,\Phi\,\rangle$ since $\langle\,\Psi\,|\,\Phi\,\rangle$ represents the scalar product.

[3]For this scenario, which we have in mind for applications, one often calls ρ the *reduced density matrix*, but we shall refrain doing so here.

$$\rho_t = \Lambda_t \rho_0 \tag{5}$$

is given by a one-parameter semigroup Λ_t of linear endomorphisms on the space $\mathfrak{S}(\mathfrak{H})$ of all density matrices [19]. Under some continuity conditions and for bounded H the Lindblad theorem [12, 21] asserts that the infinitesimal generator $\mathscr{L}$ of the semigroup Λ_t that preserves self-adjointness, positivity and unit trace is of the form

$$\mathscr{L}(\rho) = -i[H, \rho] + \mathscr{D}(\rho). \tag{6}$$

The commutator describes the unitary part of the time evolution (as in an isolated quantum system) and the dissipative part $\mathscr{D}(\rho) \in \mathfrak{End}(\mathfrak{S}(\mathfrak{H}))$, which encodes the physical properties of the coupling to the environment, is of the form

$$\mathscr{D}(\rho) = \sum_j \mathscr{D}_j(\rho), \quad \mathscr{D}_j(\rho) = D_j \rho D_j^\dagger - \frac{1}{2}\{\rho, D_j^\dagger D_j\} \tag{7}$$

with bounded operators $D_j \in \mathfrak{End}(\mathfrak{H})$ and the anticommutator $\{A, B\} := AB + BA$. The evolution equation (6) with dissipators (7) is called *quantum master equation*. The operators D_j that specify an individual dissipator are called Lindblad operators. In an open system a state that is initially pure or in equilibrium does not in general remain so as would be the case in the absence of dissipators in (6). This raises the question of stationary states in open systems.

In order to address existence we introduce the adjoint generator $\mathscr{L}^\dagger$ which is defined as follows [19]. Consider the Banach space $L^1(\mathfrak{H})$ over $\mathbb{R}$ of self-adjoint trace class linear operators $\sigma \in \mathfrak{H}$ with norm given by $||\sigma||_1 = \sup \sum_n |(x_n, \sigma y_n)|$ where the supremum is taken over all orthonormal and complete bases $\{x_n\}$ and $\{y_n\}$ of $\mathfrak{H}$. Then all linear, real and continuous functionals F on $L^1(\mathfrak{H})$ are of the form $\langle F, \sigma\rangle = \mathrm{Tr}(F^\dagger \sigma)$ where F is a bounded self-adjoint linear operator on $\mathfrak{H}$. The set of all such bounded observables F defines the space $L^\infty(\mathfrak{H})$ dual to $L^1(\mathfrak{H})$. Its norm is given by $||F||_\infty = \sup_{||\sigma||_1=1} |\langle F, \sigma\rangle| = \sup_{\Psi\in\mathfrak{H}} ||F\Psi||/||\Psi||$. Then the adjoint generator is given by

$$\mathscr{L}^\dagger(F) = -i[H, F] + \sum_j \left(D_j^\dagger F D_j - \frac{1}{2}\{F, D_j^\dagger D_j\}\right) \tag{8}$$

and one sees that $\mathscr{L}^\dagger(\mathbf{1}) = 0$. If $\mathfrak{H}$ is finite-dimensional then this guarantees the existence of a density matrix ρ such that

$$\mathscr{L}(\rho) = 0. \tag{9}$$

We call a density matrix satisfying (9) a *stationary state*, and, in particular, when $\rho \neq \mathrm{e}^{-\beta H}/Z$ for any $\beta \in \mathbb{R}_0^+$, we call ρ a *non-equilibrium steady state* (NESS) of the open quantum system with Hamiltonian H. For ergodicity and approach to stationarity, which are not our concern, we refer to [11]. For Lindblad operators of

the form $D_j = \Gamma L_j$ with a common coupling constant Γ the strong coupling limit $\Gamma \to \infty$ is called the *Zeno limit*.

Finally we remark that shifting the Lindblad operators by (in general complex) constants c_j generates an additional unitary term in the quantum master equation. More precisely, defining for some $c_j \in \mathbb{C}$ the self-adjoint operators

$$G_j = \frac{i}{2}\left(c_j D_j^\dagger - \bar{c}_j D_j\right), \quad \tilde{H} = H - \sum_j G_j, \tag{10}$$

one has

$$\mathscr{L}(\rho) = \tilde{\mathscr{L}}(\rho) \tag{11}$$

where $\tilde{\mathscr{L}}$ is defined by the modified Hamiltonian $\tilde{H}$ and shifted Lindblad operators

$$\tilde{D}_j := D_j - c_j. \tag{12}$$

Notice that $\tilde{G}_j = G_j$.

This paper deals with the construction of non-equilibrium stationary states ρ defined by (9) for a specific family of physical systems of great interest, viz. quantum spin chains coupled to environment at their boundaries, defined in Sect. 2. In Sect. 3 we generalize in mathematically rigorous form the matrix product ansatz (MPA) of Prosen [27, 29] with local divergence condition introduced by us in [16]. As an application (Sect. 4) we summarize recent progress that we made for the stationary non-equilibrium magnetization profiles in the isotropic spin-1/2 Heisenberg quantum spin chain [16, 17, 25] and discuss it in the light of very recent results [8] on correlation functions for this quantum system. The upshot is that there are substantial and perhaps somewhat unexpected similarities between quantum and classical stationary states of boundary-driven non-equilibrium systems.

2 Quantum Spin Chains

2.1 Why Quantum Spin Chains?

The prototypical model for the quantum mechanical description of magnetism in linear chains of atoms is the so-called Heisenberg quantum spin chain, proposed first in 1928 [13] as an improvement over the classical Ising model which was introduced a few years earlier by Lenz and solved by his student Ernst Ising in 1925 [15]. The simplest version of the Heisenberg model, the spin-1/2 chain defined below, is exactly solvable in the sense of quantum integrability [3]. Hence the equilibrium properties of the system, which were derived in the past decades in a vast body of literature, are rather well understood from a theoretical perspective and to some extent also

experimentally for various spin-chain materials which exhibit quasi one-dimensional interactions between neighbouring atoms.

In recent years, novel experimental Laser techniques involving single cold atoms in optical traps have made the investigation of spin chains *far from thermal equilibrium* feasible. The unique possibilities that the study of individual interacting atoms offers has triggered an immense experimental research activity. On the theoretical side, however, not much is known about non-equilibrium steady states of spin chains which are of particular interest in the case of *boundary driving*, since in this way one obtains information about anomalous transport properties. By boundary driving we mean a scenario where the two ends of a chain are forced into different states by some boundary interaction with the physical environment of the chain, thus inducing stationary currents of locally conserved quantities along the chain. The bulk of the system is considered to be effectively isolated from its physical environment, i.e., described by some quantum Hamiltonian H. The boundary interaction is described by Lindblad dissipators.

Exact results are scarce for chains with more than just a few atoms and there are, to our knowledge, no exact concrete results for specific quantum chains of arbitrary length kept far from thermal equilibrium by some kind of Lindblad boundary-drive. This state of affairs is in stark contrast to classical stochastic interacting particle systems whose Markov generators can be expressed in terms of (non-Hermitian) quantum spin chains [30] and for which many exact and rigorous results exist [6, 9, 18, 20, 30] and which are also amenable to generally applicable analytical approaches such as macroscopic fluctuation theory [5] and non-linear fluctuating hydrodynamics [32].

Nevertheless, a breakthrough in the study of quantum systems far from thermal equilibrium came a few years ago through the work of Prosen [26, 27] who devised a matrix product ansatz (MPA) somewhat reminiscent of the matrix product ansatz for classical stochastic interacting particle systems [6]. This MPA was subsequently developed by us, using a local divergence technique that reveals a link to quantum integrability and symmetries of the quantum system [16]. The MPA allowed for the derivation of recursion relations for mean values of physical observables from which stationary currents, magnetization profiles and correlations could be computed numerically exactly for *large* finite chains and analytically from a continuum approximation to these recursion relations [8, 17, 27]. As pointed out below, these results point to an interesting analogy with a well-known result in classical stochastic interacting particle systems [4, 10, 31].

2.2 Definitions and Notation

The set of integers $\{0, \ldots, n-1\}$ is denoted $\mathbb{S}_n$. We denote the canonical basis vectors of the n-dimensional complex vector space $\mathbb{C}^n$ by the symbol $|\alpha)$ with $\alpha \in \mathbb{S}_n$. Complex conjugation of some $z \in \mathbb{C}$ is denoted by $\bar{z}$. The canonical basis vectors of the dual space are denoted by $(\alpha|$. With the scalar product $(w|v) := \sum_\alpha \bar{w}_\alpha v_\alpha$ and

norm $||v|| = \sqrt{\sum_\alpha |v_\alpha|^2}$ the vector space $\mathbb{C}^n$ becomes a finite-dimensional Hilbert space which we shall call the *local physical space* and denote by $\mathfrak{p}$.

From the canonical basis vectors of $\mathbb{C}^n$ we construct the canonical basis of the space $\mathfrak{End}(\mathbb{C}^n)$ of endomorphisms $\mathbb{C}^n \to \mathbb{C}^n$ by the Kronecker products $E^{\alpha\beta} := |\alpha)(\beta| \equiv |\alpha) \otimes (\beta|$. Generally we shall somewhat loosely identify endomorphisms on some vector space with their matrix representation and sometimes call them operators. The n-dimensional matrices $E^{\alpha\beta}$ have matrix elements $(E^{\alpha\beta})_{jk} = \delta_{\alpha,j}\delta_{\beta,k}$ and they satisfy

$$E^{\alpha\beta}E^{\gamma\delta} = \delta_{\beta,\gamma}E^{\alpha\delta} \tag{13}$$

$$\mathrm{Tr}(E^{\alpha\beta}) = \delta_{\alpha,\beta}. \tag{14}$$

The n-dimensional unit matrix is denoted by $\mathbb{1}$. If a complex number appears as one term in any equation for matrices, then this complex number is understood to be a multiple of the unit matrix.

We construct a canonical basis of $\mathbb{C}^{n^N}$ by the tensor product $|\,\mathbf{a}\,\rangle = |\alpha_1) \otimes \cdots \otimes |\alpha_N)$ with the N-tuple $\mathbf{a} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{S}_n^N$. A general vector in $\mathbb{C}^{n^N}$ with components $v_\mathbf{a}$ is then denoted by $|\,v\,\rangle$. We also define basis vectors $\langle\,\mathbf{a}\,|$ of the dual space $\mathbb{C}^{2^N *}$ (isomorphic to $\mathbb{C}^{2^N}$) and the scalar product $\langle\,w\,|\,v\,\rangle := \sum_{\mathbf{a}\in\mathbb{S}^N} \bar{w}_\mathbf{a} v_\mathbf{a}$ and norm $||v|| = \sqrt{\sum_\mathbf{a} |v_\mathbf{a}|^2}$. With these definitions $\mathbb{C}^{n^N}$ becomes a finite-dimensional Hilbert space which we shall call the *physical space* and denote by $\mathfrak{P}$. Here and below

$$\sum_\mathbf{a} := \sum_{\alpha_1\in\mathbb{S}} \cdots \sum_{\alpha_N\in\mathbb{S}} \tag{15}$$

is the N-fold sum over all indices in $\mathbb{S}$.

From arbitrary matrices $Q \in \mathfrak{End}(\mathbb{C}^n)$ we construct the local tensor operators

$$Q_k = \mathbb{1}^{\otimes(k-1)} \otimes Q \otimes \mathbb{1}^{\otimes(N-k)} \in \mathfrak{End}(\mathfrak{P}). \tag{16}$$

By convention $Q^{\otimes 0} := 1$ and $Q^{\otimes 1} := Q$ for any matrix Q. We denote the unit matrix acting on $\mathfrak{P}$ by $\mathbf{1}$, i.e., $\mathbf{1} = \mathbb{1}^{\otimes N}$. The set of products

$$\{E^{\mathbf{a},\mathbf{a}'}\} = \left\{ \prod_{j=1}^N E_j^{\alpha_j\alpha'_j} \right\} \tag{17}$$

for $\mathbf{a}, \mathbf{a}' \in \mathbb{S}_n^N$ forms a complete basis of $\mathfrak{End}(\mathfrak{P})$. Transposition of a matrix A is denoted by A^T. The adjoint of an operator is denoted $A^\dagger$ which in matrix form means $A^\dagger = \bar{A}^T$. Self-adjoint operators are called Hermitian. It is convenient to represent ket-vectors $|\,v\,\rangle$ as column vectors with components $v_\mathbf{a}$. Then $\langle\,v\,|$ is represented by a row vector with components $\bar{v}_\mathbf{a}$. Elements of a generic vector space $\mathfrak{V}$ (not Hilbert) over $\mathbb{C}$ are denoted by the double-ket symbol $|\cdot\rangle\rangle$ and elements of its dual $\mathfrak{V}^*$ by the double-bra symbol $\langle\langle\cdot|$. A linear form $\phi_W : \mathfrak{V} \to \mathbb{C}$ is denoted by $\langle\langle\,W\,|\cdot\rangle\rangle$.

With these conventions we are now in a position to define the objects of our investigation.

Definition 1 Let $h \in \mathfrak{End}(\mathbb{C}^{n^2})$ and $b^L, b^R \in \mathfrak{End}(\mathbb{C}^n)$ be self-adjoint and $b_1^L = b^L \otimes \mathbb{1}^{\otimes(N-1)}$, $b_N^R = \mathbb{1}^{\otimes(N-1)} \otimes b^R$, $h_{k,k+1} = \mathbb{1}^{\otimes(k-1)} \otimes h \otimes \mathbb{1}^{\otimes(N-k-1)}$. Then a homogeneous quantum spin chain with $N \geq 2$ sites with nearest-neighbour interaction h and boundary fields $b^{L,R}$ is defined by the Hamiltonian

$$H = b_1^L + b_N^R + \sum_{k=1}^{N-1} h_{k,k+1}. \tag{18}$$

A quantum spin system with one site is defined by a self-adjoint operator $b \in \mathfrak{End}(\mathbb{C}^n)$.

Definition 2 For $D^{\chi_k} \in \mathfrak{End}(\mathbb{C}^n)$ and a density matrix $\rho \in \mathfrak{S}(\mathfrak{P})$ the operator

$$\mathscr{D}_k(\rho) := D_k^{\chi_k} \rho D_k^{\chi_k\dagger} - \frac{1}{2}\left(\rho D_k^{\chi_k\dagger} D_k^{\chi_k} + D_k^{\chi_k\dagger} D_k^{\chi_k} \rho\right), \quad 1 \leq k \leq N, \quad N \geq 1 \tag{19}$$

is called dissipator at site k with local Lindblad operator D^{χ_k}, indexed by a symbol χ_k. For $N = 1$ the lower index $k = 1$ is dropped.

Definition 3 Let H be a quantum spin Hamiltonian with N sites according to Definition 1, $\mathscr{D}_1$ and $\mathscr{D}_N$ be dissipators with local Lindblad operators D^L and D^R resp. according to Definition 2 and let $\rho \in \mathfrak{S}(\mathfrak{P})$ be the solution of the equation

$$-i[H, \rho] + \mathscr{D}_1(\rho) + \mathscr{D}_N(\rho) = 0. \tag{20}$$

Then ρ is called a non-equilibrium stationary state of the boundary-driven quantum spin system defined by H.

We remark that the construction of matrix product states given below is straightforwardly generalized to more than one boundary dissipator at each edge of the chain.

3 Construction of Stationary Matrix Product States

3.1 Matrix Product Ansatz

In order to construct a solution of the stationary Lindblad equation of the form (20) we first make the following observations:

(a) For any density matrix $\rho \in \mathfrak{S}(\mathfrak{P})$ one can find a matrix $M \in \mathfrak{End}(\mathfrak{P})$ such that

$$\rho = MM^\dagger / Z \tag{21}$$

with the *partition function*

$$Z := \mathrm{Tr}(MM^\dagger). \tag{22}$$

Thus, given M one knows ρ.[4]

(b) One can expand M in the basis (17) of $\mathfrak{End}(\mathfrak{P})$ as

$$M = \sum_{\mathbf{a},\mathbf{a}'} M_{\mathbf{a},\mathbf{a}'} E_1^{\alpha_1,\alpha_1'} \dots E_N^{\alpha_N,\alpha_N'}. \tag{23}$$

The idea of the matrix product ansatz (MPA) is to write the matrix elements $M_{\mathbf{a},\mathbf{a}'}$ as the linear form [27, 29]

$$M_{\mathbf{a},\mathbf{a}'} = \langle\langle\, W \,|\Omega^{\alpha_1,\alpha_1'} \dots \Omega^{\alpha_N,\alpha_N'}|\, V \,\rangle\rangle \tag{24}$$

where $|\, V \,\rangle\rangle$ is a vector in some (generally infinite-dimensional) auxiliary space $\mathfrak{A}$, the n^2 matrices $\Omega^{\alpha,\alpha'}$ are suitably chosen endomorphisms of $\mathfrak{A}$ and $\langle\langle\, W \,|$ is a suitably chosen vector from the dual space $\mathfrak{A}^*$.

In order to use this MPA in applications we need to add some more structure. We define $\bar{\Omega}^{\alpha,\alpha'} \in \mathfrak{End}(\mathfrak{A})$ by complex conjugation of the matrix representation of $\Omega^{\alpha,\alpha'}$. Next we construct

$$\Omega := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \Omega^{\alpha\alpha'}, \quad \Omega^\star := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \bar{\Omega}^{\alpha'\alpha} \quad \in \mathfrak{End}(\mathbb{C}^n \otimes \mathfrak{A}) \tag{25}$$

$$\Omega^{\otimes_p N} := \sum_{\mathbf{a},\mathbf{a}'} E^{\alpha_1\alpha_1'} \otimes \dots \otimes E^{\alpha_N\alpha_N'} \otimes \Omega^{\alpha_1\alpha_1'} \dots \Omega^{\alpha_N\alpha_N'} \in \mathfrak{End}(\mathfrak{P} \otimes \mathfrak{A}) \tag{26}$$

and analogously $(\Omega^\star)^{\otimes_p N} = \left(\Omega^{\otimes_p N}\right)^\star$. The subscript p at the tensor symbol indicates that the tensor product is only taken over the local physical space $\mathfrak{p}$, i.e., the term $\Omega^{\alpha_1\alpha_1'} \dots \Omega^{\alpha_N\alpha_N'} \in \mathfrak{End}(\mathfrak{A})$ in (26) is the usual matrix product. The star $\star$ denotes the adjoint operation on the physical space $\mathfrak{P}$ only, not on the auxiliary space. This means that the matrix $\left(\Omega^{\otimes_p N}\right)^\star$ is obtained from the matrix $\Omega^{\otimes_p N}$ by transposition and complex conjugation of its components $\Omega^{\otimes_p N}_{\mathbf{a},\mathbf{a}'} = \Omega^{\alpha_1\alpha_1'} \dots \Omega^{\alpha_N\alpha_N'} \mapsto \bar{\Omega}^{\alpha_1'\alpha_1} \dots \bar{\Omega}^{\alpha_N'\alpha_N} = \left(\Omega^{\otimes_p N}\right)^\star_{\mathbf{a},\mathbf{a}'}$ as in the second definition in (25) *without* reversing the order of the matrix products and *without* transposing the matrices $\Omega^{\alpha_j\alpha_j'}$.

This construction immediately leads to the following lemma:

Lemma 1 *Let $\mathfrak{A}$ be a vector space, $\Omega^{\alpha,\alpha'} \in \mathfrak{End}(\mathfrak{A})$ for $\alpha, \alpha' \in \mathbb{S}_n$, $|\, V \,\rangle\rangle$, $|\, \overline{V} \,\rangle\rangle \in \mathfrak{A}$ and $\langle\langle\, W \,|$, $\langle\langle\, \overline{W} \,| \in \mathfrak{A}^*$ where the bar denotes complex conjugation of each vector component. Then M, $M^\dagger \in \mathfrak{End}(\mathfrak{P})$ defined by (23) and (24) can be written*

[4] M is not uniquely defined. For a given M and arbitrary unitary U the product MU gives the same ρ. This non-uniqueness seems to be exactly the point that makes M easier to treat than ρ.

$$M = \langle\langle\, W \,|\Omega^{\otimes_p N}|\, V \,\rangle\rangle, \quad M^\dagger = \langle\langle\, \overline{W} \,|\, \left(\Omega^\star\right)^{\otimes_p N} |\, \overline{V} \,\rangle\rangle \tag{27}$$

where the linear form $\langle\langle\, W \,|\cdot|\, V \,\rangle\rangle$ *on* $\mathfrak{A}$ *is taken on each component* $\Omega^{\alpha_1\alpha'_1} \ldots \Omega^{\alpha_N\alpha'_N}$ *of the endomorphism* $\Omega^{\otimes_p N}$ *on* $\mathfrak{P} \otimes \mathfrak{A}$.

Lemma 1 follows immediately from the expansion $\Omega = \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \Omega^{\alpha\alpha'}$ and the multilinearity of the Kronecker product. It expresses the fact that $\Omega^{\otimes N}$ can be thought of as a matrix of dimension n^N (the dimension of the physical space $\mathfrak{P}$) whose matrix elements $(\mathbf{a}, \mathbf{a}')$ are the products $\Omega^{\alpha_1\alpha'_1} \ldots \Omega^{\alpha_N\alpha'_N}$ of (generally infinite-dimensional) matrices acting on the auxiliary space $\mathfrak{A}$. The linear form $\langle\langle\, W \,|\cdot|\, V \,\rangle\rangle$ maps each of these matrix products onto $\mathbb{C}$ so that M is indeed a usual matrix of dimension n^N.

The next technical idea is to double the auxiliary space. To this end we denote the unit operator on $\mathfrak{A}$ by I and define $\Omega_1^{\alpha\alpha'} := \Omega^{\alpha\alpha'} \otimes I$ and $\Omega_2^{\alpha\alpha'} := I \otimes \Omega^{\alpha\alpha'}$ which are endomorphisms of $\mathfrak{A}^2$. The multilinearity of the tensor product allows us to write $\Omega_1^{\alpha\alpha'}\Omega_2^{\beta\beta'} = \Omega^{\alpha\alpha'} \otimes \Omega^{\beta\beta'}$ for any $\alpha, \alpha', \beta, \beta' \in \mathbb{S}$. We also define in analogy to (25) the following endomorphisms of $\mathfrak{P} \otimes \mathfrak{A}^2$

$$\Omega_1 := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \Omega^{\alpha\alpha'} \otimes I, \quad \Omega_2 := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes I \otimes \Omega^{\alpha\alpha'} \tag{28}$$

$$\Omega_1^\star := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \bar{\Omega}^{\alpha'\alpha} \otimes I, \quad \Omega_2^\star := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes I \otimes \bar{\Omega}^{\alpha'\alpha}. \tag{29}$$

Lemma 2 *Let* $\mathfrak{A}$ *be a vector space and* $|\, V \,\rangle\rangle, |\, \overline{V} \,\rangle\rangle \in \mathfrak{A}$ *and* $\Omega^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{A})$ *for* $\alpha, \alpha' \in \mathbb{S}_n$ *and* $\langle\langle\, W \,|, \langle\langle\, \overline{W} \,| \in \mathfrak{A}^*$. *For some* $\Theta^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{A}^2)$ *define*

$$\Theta := \sum_{\alpha,\alpha'} E^{\alpha\alpha'} \otimes \Theta^{\alpha\alpha'} \in \mathfrak{End}(\mathbb{C}^2 \otimes \mathfrak{A}^2) \tag{30}$$

$$\Theta^{\otimes_p N} := \sum_{\mathbf{a},\mathbf{a}'} E^{\alpha_1\alpha'_1} \otimes \cdots \otimes E^{\alpha_N\alpha'_N} \otimes \Theta^{\alpha_1\alpha'_1} \ldots \Theta^{\alpha_N\alpha'_N} \in \mathfrak{End}(\mathfrak{P} \otimes \mathfrak{A}^2). \tag{31}$$

Then for

$$\Theta = \Omega_1\Omega_2^\star \tag{32}$$

a density matrix $\rho \in \mathfrak{S}(\mathfrak{P})$ *has the matrix product representation*

$$\rho = \langle\langle\, W, \overline{W} \,|\Theta^{\otimes_p N}|\, V, \overline{V} \,\rangle\rangle. \tag{33}$$

where the tensor products

$$|\, V, \overline{V} \,\rangle\rangle := |\, V \,\rangle\rangle \otimes |\, \overline{V} \,\rangle\rangle \in \mathfrak{A} \otimes \mathfrak{A}, \quad \langle\langle\, W, \overline{W} \,| := \langle\langle\, W \,| \otimes \langle\langle\, \overline{W} \,| \in \mathfrak{A}^* \otimes \mathfrak{A}^* \tag{34}$$

define a bilinear form $\phi_{W,\overline{W}} : \mathfrak{A} \otimes \mathfrak{A} \to \mathbb{C}$.

Proof We first note that for the scalar product on the physical space $\mathfrak{P}$ we have

$$\langle \mathbf{a} \,|\Theta^{\otimes_p N}|\, \mathbf{a}' \rangle = \Theta^{\alpha_1\alpha_1'} \ldots \Theta^{\alpha_N\alpha_N'}. \tag{35}$$

Furthermore, by the construction (32) for a single site and definition (28), one finds

$$\Theta^{\alpha\alpha'} = \sum_\beta (\alpha|\Omega_1|\beta)(\beta|\Omega_2^\star|\alpha') = \sum_\beta \Omega_1^{\alpha\beta}(\Omega_2^\star)^{\beta\alpha'} = \sum_\beta \Omega_1^{\alpha\beta}\bar{\Omega}_2^{\alpha'\beta} \tag{36}$$

and therefore with $\mathbf{b} := (\beta_1, \ldots, \beta_N) \in \mathbb{S}_n^N$

$$\begin{aligned}\Theta^{\alpha_1\alpha_1'} \ldots \Theta^{\alpha_N\alpha_N'} &= \sum_{\mathbf{b}} \Omega_1^{\alpha_1\beta_1}\bar{\Omega}_2^{\alpha_1'\beta_1} \ldots \Omega_1^{\alpha_N\beta_N}\bar{\Omega}_2^{\alpha_N'\beta_N} \\ &= \sum_{\mathbf{b}} \Omega^{\alpha_1\beta_1} \ldots \Omega^{\alpha_N\beta_N} \otimes \bar{\Omega}^{\alpha_1'\beta_1} \ldots \bar{\Omega}^{\alpha_N'\beta_N}.\end{aligned} \tag{37}$$

This shows that $\Theta^{\alpha_1\alpha_1'} \ldots \Theta^{\alpha_N\alpha_N'} \in \mathfrak{End}(\mathfrak{A}^2)$ is decomposable into a finite sum of endomorphisms of $\mathfrak{A} \otimes \mathfrak{A}$. Then the factorization property of the scalar product involving the tensor vectors (34) and the tensor operators (37) and Lemma 1 give

$$\begin{aligned}\langle\langle W, \overline{W} \,|\Theta^{\otimes_p N}|\, V, \overline{V} \rangle\rangle_{\mathbf{a}\mathbf{a}'} &= \langle\langle W, \overline{W} \,|\Theta^{\alpha_1\alpha_1'} \ldots \Theta^{\alpha_N\alpha_N'}|\, V, \overline{V} \rangle\rangle \\ &= \sum_{\mathbf{b}} \langle\langle W \,|\Omega^{\alpha_1\beta_1} \ldots \Omega^{\alpha_N\beta_N}|\, V \rangle\rangle \overline{\langle\langle W \,|\Omega^{\alpha_1'\beta_1} \ldots \Omega^{\alpha_N'\beta_N}|\, V \rangle\rangle} \\ &= \sum_{\mathbf{b}} M_{\mathbf{a}\mathbf{b}}\bar{M}_{\mathbf{a}'\mathbf{b}}\end{aligned} \tag{38}$$

The l.h.s. of the first equation is the matrix element ρ. Observing that $\bar{M}_{\mathbf{a}'\mathbf{b}} = M^\dagger_{\mathbf{b}\mathbf{a}'}$ and completeness of the basis (17) shows that the r.h.s. of the last equation is equal to $(MM^\dagger)_{\mathbf{a}\mathbf{a}'}$. Thus (33) is proved for each matrix element of ρ. □

The point of this lemma is the fact that a matrix product form of M induces a matrix product form for ρ which allows for a computation of physical observables in terms of the matrices $\Omega^{ss'}$. This is the content of the following proposition.

Proposition 1 *Let $\rho \in \mathfrak{S}(\mathfrak{P})$ be a density matrix with partition function Z (22) and $|\, V, \overline{V} \,\rangle\rangle$, $\langle\langle\, W, \overline{W} \,|$ as defined in (34). With*

$$\Theta_0 := \sum_\alpha \Theta^{\alpha\alpha} = \sum_{\alpha\beta} \Omega_1^{\alpha\beta}\bar{\Omega}_2^{\alpha\beta} \quad \in \mathfrak{End}\left(\mathfrak{A}^2\right) \tag{39}$$

one has

$$Z = \langle\langle\, W, \overline{W}\,|\Theta_0^N|\, V, \overline{V}\,\rangle\rangle \tag{40}$$

$$\begin{aligned}&\langle\, E_{k_1}^{\alpha_1\alpha_1'} E_{k_2}^{\alpha_2\alpha_2'} \dots E_{k_n}^{\alpha_n\alpha_n'}\,\rangle \\ &= \langle\langle\, W, \overline{W}\,|\Theta_0^{k_1-1}\Theta^{\alpha_1'\alpha_1}\Theta_0^{k_2-k_1-1}\Theta^{\alpha_2'\alpha_2}\dots\Theta^{\alpha_n'\alpha_n}\Theta_0^{N-k_n}|\, V, \overline{V}\,\rangle\rangle/Z\end{aligned} \tag{41}$$

Proof The equality following the definition in (39) follows from (36). By construction we have for the partition function (40)

$$\begin{aligned}Z &= \sum_{\mathbf{a},\mathbf{a}'} \mathrm{Tr}\left(E_1^{\alpha_1\alpha_1'}\dots E_N^{\alpha_N\alpha_N'}\right)\langle\langle\, W, \overline{W}\,|\Theta^{\alpha_1\alpha_1'}\dots\Theta^{\alpha_N\alpha_N'}|\, V, \overline{V}\,\rangle\rangle \\ &= \sum_{\mathbf{a},\mathbf{a}'}\left(\prod_{j=1}^{N}\mathrm{Tr}\left(E^{\alpha_j\alpha_j'}\right)\right)\langle\langle\, W, \overline{W}\,|\Theta^{\alpha_1\alpha_1'}\dots\Theta^{\alpha_N\alpha_N'}|\, V, \overline{V}\,\rangle\rangle\end{aligned} \tag{42}$$

where in the second equality we have used the factorization property of the trace for tensor products. The trace property (14) yields the expression (40) for the partition function Z. The expression (41) follows in similar fashion by noting that due to (13) one has $\mathrm{Tr}(E^{\alpha\alpha'}E^{\beta\beta'}) = \delta_{\alpha,\beta'}\delta_{\alpha',\beta}$. □

Remark 1 Since an observable $O_k \in \mathfrak{End}(\mathfrak{P})$ can be expanded $O_k = \sum_{\alpha\alpha'} O_k^{\alpha\alpha'} E_k^{\alpha\alpha'}$ with numerical coefficients of the form $O_k^{\alpha\alpha'} = \bar{O}_k^{\alpha'\alpha} \in \mathbb{C}$, Proposition 1 allows for computing averages of products of local observables in terms of matrix products involving the matrices $\Theta^{\alpha\alpha'}$ and Θ_0.

We note two useful corollaries of Lemma 2 which follow directly from (36).

Corollary 1 *Let ρ be a density matrix according to Lemma 2 and $D_k = \mathbb{1}^{\otimes(k-1)} \otimes D \otimes \mathbb{1}^{\otimes(N-k)}$ be a Lindblad operator acting non-trivially only on site k with some local Lindblad operator $D \in \mathfrak{End}(\mathbb{C}^n)$. Then for the local dissipator $\mathscr{D}_k$ with Lindblad operator D_k one has*

$$\mathscr{D}_k(\rho) = \frac{1}{Z}\langle\langle\, W, \overline{W}\,|\Theta^{\otimes(k-1)} \otimes \Delta \otimes \Theta^{\otimes(N-k)}|\, V, \overline{V}\,\rangle\rangle \tag{43}$$

with Z of Proposition 1 and

$$\Delta = \sum_{\beta}\sum_{\alpha\alpha'} \mathscr{D}\left(E^{\alpha\alpha'}\right) \otimes \Omega_1^{\alpha\beta}\bar{\Omega}_2^{\alpha'\beta} \tag{44}$$

where $\mathscr{D}$ is the dissipator with the local Lindblad operator D.

Corollary 2 *Let ρ be a density matrix according to Lemma 2 and $b_k = \mathbb{1}^{\otimes(k-1)} \otimes b \otimes \mathbb{1}^{\otimes(N-k)} \in \mathfrak{End}(\mathfrak{P})$ be a self-adjoint operator acting non-trivially only on site k with some local self-adjoint operator $b \in \mathfrak{End}(\mathbb{C}^n)$. Then for the unitary part of the time-evolution of the density matrix under b_k one has*

$$-i[b_k, \rho] = \frac{1}{Z}\langle\langle W, \overline{W}\,|\Theta^{\otimes(k-1)} \otimes \Gamma \otimes \Theta^{\otimes(N-k)}|\, V, \overline{V}\,\rangle\rangle \tag{45}$$

with Z of Proposition 1 and

$$\Gamma = -i \sum_{\beta} \sum_{\alpha\alpha'} \left[b, E^{\alpha\alpha'}\right] \otimes \Omega_1^{\alpha\beta} \bar{\Omega}_2^{\alpha'\beta}. \tag{46}$$

For $D = \sum_{\alpha\alpha'} D_{\alpha\alpha'} E^{\alpha\alpha'}$, $b = b^\dagger = \sum_{\alpha\alpha'} b_{\alpha\alpha'} E^{\alpha\alpha'}$ we note

$$\mathscr{D}(E^{\alpha\alpha'}) = \sum_{\beta\beta'} \left(D_{\beta\alpha} \bar{D}_{\beta'\alpha'} E^{\beta\beta'} - \frac{1}{2} D_{\beta\beta'} \bar{D}_{\beta\alpha'} E^{\alpha\beta'} - \frac{1}{2} D_{\beta'\alpha} \bar{D}_{\beta'\beta} E^{\beta\alpha'} \right) \tag{47}$$

$$\left[b, E^{\alpha\alpha'}\right] = \sum_{\beta} \left(b_{\beta\alpha} E^{\beta\alpha'} - \bar{b}_{\beta\alpha'} E^{\alpha\beta} \right) \tag{48}$$

which follows from (13) by straightforward computation and $b = b^\dagger$.

3.2 Main Result

The previous discussion is "abstract nonsense" in so far as we have provided no information about the matrices $\Omega^{\alpha\alpha'}$ and the vectors $\langle\langle W \,|$ and $|\, V \rangle\rangle$ from which a stationary density matrix ρ solving (20) could be constructed. In order to state a sufficient property of the $\Omega^{\alpha\alpha'}$ we define the local divergence condition which was first introduced for $n = 2$ in [16].

Definition 4 (*Local divergence condition*) Let H be a quantum spin Hamiltonian according to Definition 1 and with finite local physical space $\mathfrak{p}$ and let $\mathfrak{A}$ be a vector space with unit operator denoted by I. For $\Omega^{\alpha\alpha'}, \Xi^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{A})$ define $\Omega := \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \Omega^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{p} \otimes \mathfrak{A})$, $\Xi := \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \Xi^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{p} \otimes \mathfrak{A})$, and $\hat{h} := h \otimes I \in \mathfrak{End}(\mathfrak{p}^2 \otimes \mathfrak{A})$. We say that H satisfies a local divergence condition w.r.t. some non-zero Ω and Ξ if

$$\left[\hat{h}, \Omega \otimes_p \Omega\right] = \Xi \otimes_p \Omega - \Omega \otimes_p \Xi \tag{49}$$

where the tensor product $\otimes_p$ over the physical space is defined by $\Xi \otimes_p \Omega := \sum_{\alpha\alpha'} \sum_{\beta\beta'} E^{\alpha\alpha'} \otimes E^{\beta\beta'} \otimes (\Xi^{\alpha\alpha'} \Omega^{\beta\beta'})$.

Remark 2 The local divergence condition (49) defines a quadratic algebra [22] for $2n^2$ generators $\Omega^{\alpha\alpha'}$ and $\Xi^{\alpha\alpha'}$. Quadratic algebras arise e.g. as universal enveloping algebras of Lie algebras and also play an important role in the theory of quantum groups. They also arise in the study of invariant measures of stochastic interacting particle systems [1, 6]. The local divergence condition can be generalized to include

a term $\hat{T}\Omega \otimes_p \Omega - \hat{T}\Omega \otimes_p \Omega$ where $\hat{T} = \mathbb{1} \otimes \mathbb{1} \otimes T$ and $T \in \mathfrak{End}(\mathfrak{A})$ [24]. This extension gives rise to a cubic algebra.

Next we define the *Lindblad boundary matching condition* which underlies in some shape or form many concrete applications of the MPA [29], but which to our knowledge has never been stated as such and in full generality.

Definition 5 (*Lindblad boundary matching condition*) Let $\mathfrak{A}$ be a vector space. For $|\, V \,\rangle\rangle \in \mathfrak{A}$ and $\langle\langle\, W \,| \in \mathfrak{A}^*$ define the vectors $|\, V, \overline{V} \,\rangle\rangle := |\, V \,\rangle\rangle \otimes |\, \overline{V} \,\rangle\rangle$ and $\langle\langle\, W, \overline{W} \,| := \langle\langle\, W \,| \otimes \langle\langle\, \overline{W} \,|$ and for $\Omega^{\alpha\alpha'}, \bar{\Omega}^{\alpha\alpha'}, \Xi^{\alpha\alpha'}, \bar{\Xi}^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{A})$ define the endomorphisms

$$\Lambda^{\alpha\alpha'} := i \sum_{\beta} \left(\Omega^{\alpha\beta} \otimes \bar{\Xi}^{\alpha'\beta} - \Xi^{\alpha\beta} \otimes \bar{\Omega}^{\alpha'\beta} \right) \tag{50}$$

and for $B \in \{L, R\}$ with $b^B_{\alpha\alpha'} = \bar{b}^B_{\alpha'\alpha} \in \mathbb{C}$, $D^B_{\alpha\alpha'} \in \mathbb{C}$

$$\Gamma^{\alpha\alpha'}_B := -i \sum_{\beta\beta'} \left(b^B_{\alpha\beta} \Omega^{\beta\beta'} \otimes \bar{\Omega}^{\alpha'\beta'} - \bar{b}^B_{\alpha'\beta} \Omega^{\alpha\beta'} \otimes \bar{\Omega}^{\beta\beta'} \right) \tag{51}$$

$$\begin{aligned} \Delta^{\alpha\alpha'}_B := \sum_{\beta\beta'} \sum_{\gamma} \Big(& D^B_{\alpha\beta} \bar{D}^B_{\alpha'\beta'} \Omega^{\beta\gamma} \otimes \bar{\Omega}^{\beta'\gamma} \\ & - \frac{1}{2} D^B_{\beta\alpha'} \bar{D}^B_{\beta\beta'} \Omega^{\alpha\gamma} \otimes \bar{\Omega}^{\beta'\gamma} - \frac{1}{2} D^B_{\beta'\beta} \bar{D}^B_{\beta'\alpha} \Omega^{\beta\gamma} \otimes \bar{\Omega}^{\alpha'\gamma} \Big) . \end{aligned} \tag{52}$$

We say that vectors $|\, V \,\rangle\rangle \in \mathfrak{A}$ and $\langle\langle\, W \,| \in \mathfrak{A}^*$ satisfy the *Lindblad boundary matching condition* w.r.t. Ω and Ξ if for all $\alpha, \alpha' \in \mathbb{S}_n$

$$0 = \langle\langle\, X \,| \left(\Gamma^{\alpha\alpha'}_R + \Delta^{\alpha\alpha'}_R - \Lambda^{\alpha\alpha'} \right) |\, V, \overline{V} \,\rangle\rangle = \langle\langle\, W, \overline{W} \,| \left(\Gamma^{\alpha\alpha'}_L + \Delta^{\alpha\alpha'}_L + \Lambda^{\alpha\alpha'} \right) |\, Y \,\rangle\rangle \tag{53}$$

for all $\langle\langle\, X \,| \in \mathfrak{A}^{2*}$ and all $|\, Y \,\rangle\rangle \in \mathrm{span}(\Theta^{\alpha_2,\alpha'_2} \ldots \Theta^{\alpha_N,\alpha'_N} |\, V, \overline{V} \,\rangle\rangle)$ for $N \geq 2$.

Remark 3 Define $\Lambda_0 := \sum_\alpha \Lambda^{\alpha\alpha}$. It is easy to see that $0 = \sum_\alpha \Delta^{\alpha\alpha}_B = \sum_\alpha \Gamma^{\alpha\alpha}_B$. Hence (53) implies $0 = \langle\langle\, X \,| \Lambda_0 |\, V, \overline{V} \,\rangle\rangle = \langle\langle\, W, \overline{W} \,| \Lambda_0 |\, Y \,\rangle\rangle$. For the extended local divergence condition with operator T the Lindblad boundary matching condition acquires an extra term $\{T, \Omega^{\alpha\alpha'}\}$ in both brackets in (53).

With these preparations we are in a position to state the main result in terms of the original local divergence condition (49). The adaptation to the extended local divergence condition is trivial.

Theorem 1 *Given a quantum spin Hamiltonian*

$$H = H_b + H_s \tag{54}$$

according to Definition 1 with bulk part $H_b = \sum_{k=1}^{N-1} h_{k,k+1}$ *and surface part* $H_s = b_1^L + b_N^R$*, and given a vector space* $\mathfrak{A}$*, let* $\Omega^{\alpha\alpha'}$, $\Xi^{\alpha\alpha'} \in \mathfrak{End}(\mathfrak{A})$ *be representation matrices of the quadratic algebra* (49) *defined by h, and let* $|\, V \,\rangle$ *and* $|\, W \,\rangle$ *be vectors satisfying the Lindblad boundary matching condition* (53) *with coefficients* $L^B_{\alpha\alpha'} \in \mathbb{C}$ *and* $b^B_{\alpha\alpha'} = (\alpha|b^B|\alpha')$ *for* $B \in \{L, R\}$*. Then a density matrix* ρ *in the matrix product form* (33) *is a stationary solution of the quantum master equation* (20) *with Lindblad operators* L^B *given by* $L^B_{\alpha\alpha'} = (\alpha|L^B|\alpha')$*.*

This theorem breathes life into the matrix product form (33) of the stationary density matrix by providing sufficient (but not necessary!) conditions on the matrices $\Omega^{\alpha\alpha'}$, vectors $\langle\langle\, W\,|, |\, V\,\rangle\rangle$ and the auxiliary matrices $\Xi^{\alpha\alpha'}$. The basic idea of the proof is to split the quantum master equation into a bulk part and a boundary part. The bulk part comes from the unitary part of the evolution under the action of H_b and leads through the local divergence condition (49) to a quadratic algebra for the matrices $\Omega^{\alpha\alpha'}$, $\Xi^{\alpha\alpha'}$ plus some boundary terms. The boundary part, which involves (i) these boundary terms, (ii) the unitary evolution under the boundary fields, and (iii) the Lindblad dissipators then becomes a set of equations for the vectors $\langle\, W\,|$ and $|\, V\,\rangle$. Choosing a representation for the quadratic algebra and fixing these vectors to satisfy the Lindblad boundary matching condition then guarantees stationarity.

Proof We decompose $\rho = MM^\dagger/Z$ where $Z = \mathrm{Tr}(MM^\dagger) < \infty$ since $\dim(\mathfrak{P}) < \infty$. Hence it suffices to prove

$$\mathscr{L}(MM^\dagger) := -i\big[H, MM^\dagger\big] + \mathscr{D}_1(MM^\dagger) + \mathscr{D}_N(MM^\dagger) = 0 \tag{55}$$

for M and $M^\dagger$ given by Lemma 1.

We consider first the bulk part of the unitary evolution. By definition of the commutator one has $\big[H, MM^\dagger\big] = [H, M]M^\dagger + M\big[H, M^\dagger\big]$. The quadratic algebra (49) ensures validity of the local divergence condition according to Definition 4. The telescopic property of the sum in H_b then implies for $\hat{H}_b := H_b \otimes I \in \mathfrak{End}(\mathfrak{P} \otimes \mathfrak{A})$ the commutation relation

$$\Big[\hat{H}_b, \Omega^{\otimes_p N}\Big] = \Xi \otimes_p \Omega^{\otimes_p (N-1)} - \Omega^{\otimes_p (N-1)} \otimes_p \Xi, \tag{56}$$

and by transposition and complex conjugation in the physical space $\mathfrak{P}$

$$\Big[\hat{H}_b, (\Omega^\star)^{\otimes_p N}\Big] = (\Omega^\star)^{\otimes_p (N-1)} \otimes_p \Xi^\star - \Xi^\star \otimes_p (\Omega^\star)^{\otimes_p (N-1)} \tag{57}$$

where

$$\Xi^\star = \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \bar{\Xi}^{\alpha'\alpha}. \tag{58}$$

Therefore, with

$$N_L := \langle\langle\, W\,|\Xi \otimes_p \Omega^{\otimes_p (N-1)}|\, V\,\rangle\rangle, \quad N_R := \langle\langle\, W\,|\Omega^{\otimes_p (N-1)} \otimes_p \Xi|\, V\,\rangle\rangle \tag{59}$$

and consequently

$$N_L^\dagger = \langle\langle\, \overline{W}\,|\Xi^\star \otimes_p (\Omega^\star)^{\otimes_p(N-1)}|\, \overline{V}\,\rangle\rangle, \quad N_R^\dagger = \langle\langle\, \overline{W}\,|(\Omega^\star)^{\otimes_p(N-1)} \otimes_p \Xi^\star|\, \overline{V}\,\rangle\rangle \tag{60}$$

one has

$$[H_b, M] = N_L - N_R, \quad \left[H_b, M^\dagger\right] = N_R^\dagger - N_L^\dagger. \tag{61}$$

This yields

$$-i\left[H_b, MM^\dagger\right] = iM(N_L^\dagger - N_R^\dagger) - i(N_L - N_R)M^\dagger. \tag{62}$$

Now notice that

$$MN_L^\dagger = \langle\langle\, W, \overline{W}\,|\Omega_1\Xi_2^\star \otimes_p \Theta^{\otimes_p(N-1)}|\, V, \overline{V}\,\rangle\rangle \tag{63}$$

$$N_L M^\dagger = \langle\langle\, W, \overline{W}\,|\Xi_1\Omega_2^\star \otimes_p \Theta^{\otimes_p(N-1)}|\, V, \overline{V}\,\rangle\rangle \tag{64}$$

$$N_R M^\dagger = \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p \Xi_1\Omega_2^\star|\, V, \overline{V}\,\rangle\rangle \tag{65}$$

$$MN_R^\dagger = \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p \Omega_1\Xi_2^\star|\, V, \overline{V}\,\rangle\rangle. \tag{66}$$

Hence

$$-i\left[H_b, MM^\dagger\right] = \langle\langle\, W, \overline{W}\,|\Lambda \otimes_p \Theta^{\otimes_p(N-1)}|\, V, \overline{V}\,\rangle\rangle - \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p \Lambda|\, V, \overline{V}\,\rangle\rangle \tag{67}$$

with $\Lambda = i\left(\Omega_1\Xi_2^\star - \Xi_1\Omega_2^\star\right)$. Expanding Λ using (13) yields

$$\Lambda = \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \Lambda^{\alpha\alpha'} \tag{68}$$

with $\Lambda^{\alpha\alpha'}$ given by (50).

Next we consider the surface part of the unitary evolution. For the boundary fields we obtain from Corollary 2

$$-i\left[b_1^L, MM^\dagger\right] = \langle\langle\, W, \overline{W}\,|\Gamma_L \otimes_p \Theta^{\otimes_p(N-1)}|\, V, \overline{V}\,\rangle\rangle \tag{69}$$

$$-i\left[b_N^R, MM^\dagger\right] = \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p \Gamma_R|\, V, \overline{V}\,\rangle\rangle \tag{70}$$

with

$$\Gamma_B = -i\sum_\beta \sum_{\alpha\alpha'} \left[b^L, E^{\alpha\alpha'}\right] \otimes \Omega_1^{\alpha\beta}\bar{\Omega}_2^{\alpha'\beta}, \quad B \in \{L, R\}. \tag{71}$$

With (48) this yields

$$\Gamma_B = \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \Gamma_B^{\alpha\alpha'} \tag{72}$$

with $\Gamma_B^{\alpha\alpha'}$ defined by (51). Putting together the bulk and the surface contribution thus yields

$$-i\big[H, MM^\dagger\big] = \langle\langle\, W, \overline{W}\,|\,(\Gamma_L + \Lambda) \otimes_p \Theta^{\otimes_p(N-1)}\,|\, V, \overline{V}\,\rangle\rangle + \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p (\Gamma_R - \Lambda)\,|\, V, \overline{V}\,\rangle\rangle. \tag{73}$$

For the dissipator part of the generator $\mathscr{L}$ (55) we have from Corollary 1

$$\mathscr{D}_1(MM^\dagger) = \langle\langle\, W, \overline{W}\,|\Delta_L \otimes \Theta^{\otimes(N-k)}\,|\, V, \overline{V}\,\rangle\rangle \tag{74}$$

$$\mathscr{D}_N(MM^\dagger) = \langle\langle\, W, \overline{W}\,|\Theta^{\otimes(N-1)} \otimes \Delta_R\,|\, V, \overline{V}\,\rangle\rangle \tag{75}$$

with

$$\Delta_B = \sum_\beta \sum_{\alpha\alpha'} \mathscr{D}^B\left(E^{\alpha\alpha'}\right) \otimes \Omega_1^{\alpha\beta} \bar{\Omega}_2^{\alpha'\beta}, \quad B \in \{L, R\}. \tag{76}$$

Using (47) one finds after relabeling of indices

$$\Delta_B = \sum_{\alpha\alpha'} E^{\alpha\alpha'} \otimes \Delta_B^{\alpha\alpha'} \tag{77}$$

with $\Delta_B^{\alpha\alpha'}$ defined by (52). Thus

$$\mathscr{L}(MM^\dagger) = \langle\langle\, W, \overline{W}\,|\,(\Delta_L + \Gamma_L + \Lambda) \otimes_p \Theta^{\otimes_p(N-1)}\,|\, V, \overline{V}\,\rangle\rangle + \langle\langle\, W, \overline{W}\,|\Theta^{\otimes_p(N-1)} \otimes_p (\Delta_R + \Gamma_R - \Lambda)\,|\, V, \overline{V}\,\rangle\rangle = 0 \tag{78}$$

by the Lindblad boundary matching condition (53). □

4 The Heisenberg Ferromagnet

We have skirted the issue of existence of representations of the quadratic algebra arising from the local divergence condition and vectors satisfying the Lindblad boundary matching condition. In order to demonstrate that the matrix product construction of the previous section is not only non-empty but also allows for concrete non-trivial results we review the application to the isotropic Heisenberg ferromagnet [16, 17, 25]. Important other models where the matrix product construction has been employed include the one-dimensional Hubbard model [28] and the spin-1 Lai-Sutherland chain [14].

4.1 Definitions and Notation

It is expedient to introduce the Levi-Civita symbol $\varepsilon_{\alpha\beta\gamma}$ (defined $\alpha, \beta, \gamma \in \{1, 2, 3\}$) by $\varepsilon_{123} = 1$ and $\varepsilon_{\alpha\beta\gamma} = (-1)^\pi \varepsilon_{\pi(\alpha\beta\gamma)}$ for any permutation $\pi(\cdot)$. We also define $\zeta_{0\alpha\beta} = \zeta_{\alpha 0\beta} = \zeta_{\alpha\beta 0} = \delta_{\alpha,\beta}$ for $\alpha, \beta \in \{0, 1, 2, 3\}$ and $\zeta_{\alpha\beta\gamma} = i\varepsilon_{\alpha\beta\gamma}$ for $\alpha, \beta, \gamma \in$

$\{1, 2, 3\}$ and introduce the two-dimensional unit matrix and the Pauli matrices

$$\sigma^0 \equiv \mathbb{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma^1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma^2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma^3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{79}$$

which form a complete basis of $\mathfrak{End}(\mathbb{C}^2)$. They satisfy

$$\sigma^\alpha \sigma^\beta = \sum_{\gamma=0}^{3} \zeta_{\alpha\beta\gamma} \sigma^\gamma. \tag{80}$$

For $\alpha \in \{1, 2, 3\}$ the matrices σ_k^α are related by a unitary transformation U with the property

$$U \sigma_k^\alpha U^\dagger = \sigma_k^{\alpha+1}, \quad \forall k \in \{1, \ldots, N\}, \quad \alpha \bmod 3. \tag{81}$$

Straightforward computation shows that this transformation is realized by the tensor product

$$U = u^{\otimes N} \tag{82}$$

with

$$u = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tag{83}$$

which is unique up to a non-zero factor.

We shall also use the notation $\hat{n} \equiv E^{00} = (1 + \sigma^z)/2, \sigma^+ \equiv E^{01} = (\sigma^x + i\sigma^y)/2$, $\sigma^- \equiv E^{10} = (\sigma^x - i\sigma^y)/2$, $\hat{v} \equiv E^{11} = (1 - \sigma^z)/2$ and the representation of the local basis vectors as column vectors as

$$|0) := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1) := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{84}$$

For later use we also introduce the notation $\sigma^x \equiv \sigma^1$, $\sigma^y \equiv \sigma^2$, $\sigma^z \equiv \sigma^3$ and the three-vectors $\boldsymbol{\sigma} = (\sigma^1, \sigma^2, \sigma^3)$ with the dot product $\mathbf{A} \cdot \mathbf{B} := \sum_{i=1}^3 A^i B^i$. Here the A_i and B_i can be real numbers or Pauli matrices. The reason for introducing this definition is the interpretation of the upper indices of the Pauli matrices as the components of the (quantum) angular momentum vector of an atom in the coordinate directions x, y, z of $\mathbb{R}^3$. If $\mathbf{A} \in \mathbb{R}^3$ is a vector of (Euclidean) length $\mathbf{A} \cdot \mathbf{A} = 1$, then the quantum expectation $\langle\, \mathbf{A} \cdot \boldsymbol{\sigma} \,\rangle$ is the mean of the projection of the angular momentum vector in the direction defined by the vector $\mathbf{A}$.

The Lie algebra $\mathfrak{gl}_2(\mathbb{C})$ with generators X^α, $\alpha \in \{0, 1, 2, 3\}$ is defined by Lie brackets

$$\left[X^0, X^\alpha\right] = 0 \tag{85}$$

$$\left[X^\alpha, X^\beta\right] = 2i \sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma} X^\gamma, \quad \alpha, \beta \in \{1, 2, 3\}. \tag{86}$$

The two-dimensional unit matrix $\mathbb{1}$ and Pauli matrices σ^α (79) are representation matrices for $\mathfrak{gl}_2(\mathbb{C})$ with the Lie-bracket represented by the commutator. Since $\sigma_k^\alpha \sigma_l^\alpha = \sigma_l^\alpha \sigma_k^\alpha$ for $l \neq k$ it follows that also $\mathbf{1} \in \mathfrak{P}$ together with

$$S^\alpha = \sum_{k=1}^{N} \sigma_k^\alpha \in \mathfrak{P} \tag{87}$$

are representation matrices of $\mathfrak{gl}_2(\mathbb{C})$. We say that an endomorphism G on $\mathfrak{P}$ is $SU(2)$-symmetric if its representation matrix satisfies $[G, S^\alpha] = 0$ for $\alpha \in \{1, 2, 3\}$.

We also define the generators

$$X^\pm := \frac{1}{2}\left(X^1 \pm i X^2\right), \quad X^z := \frac{1}{2} X^3. \tag{88}$$

In terms of these generators the defining relations (85), (86) of $\mathfrak{gl}_2(\mathbb{C})$ read

$$\left[X^0, X^{\pm,z}\right] = 0 \tag{89}$$

$$\left[X^+, X^-\right] = 2X^z, \quad \left[X^z, X^\pm\right] = \pm X^\pm. \tag{90}$$

An infinite-dimensional family of representations $X^0 \mapsto I$, $X^{\pm,z} \mapsto S^{\pm,z}$ is given by matrices I, $S^{\pm,z}$ with matrix elements

$$I_{kl} = \delta_{k,l}, \quad S^+_{kl} = l\delta_{k+1,l}, \quad S^-_{kl} = (2p - l)\delta_{k,l+1}, \quad S^z_{kl} = (p - l)\delta_{k,l} \tag{91}$$

for the non-negative integers $k, l \in \mathbb{N}_0$ and parameter $p \in \mathbb{C}$.

4.2 Boundary-Driven Lindblad–Heisenberg Chain

We consider an open chain of $N \geq 2$ quantum spins in contact with boundary reservoirs for which we wish to construct the stationary density matrix defined by (20). For the unitary part of the time evolution we consider the isotropic spin-1/2 Heisenberg Hamiltonian [3, 13] defined with the dot-product by

$$H = \sum_{k=1}^{N-1} \boldsymbol{\sigma}_k \cdot \boldsymbol{\sigma}_{k+1} \tag{92}$$

for N quantum spins at positions k along the chain.

Before defining the boundary dissipators we point out that H is manifestly rotation invariant in $\mathbb{R}^3$ which due to the quantum nature of the spin is equivalent to the symmetry $[H, S^\alpha] = 0$ under the Lie-algebra $SU(2)$ with representation matrices (87). Thus the spin components are locally conserved with associated locally conserved currents j_k^α defined by (8) with $F = \sigma_k^\alpha$. For $1 < k < N$ the action of the adjoint generator (8) yields

$$\mathscr{L}^\dagger(\sigma_k^\alpha) = j_{k-1}^\alpha - j_k^\alpha \tag{93}$$

with

$$j_k^\alpha = 2\sum_{\beta=1}^{3}\sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma}\sigma_k^\beta\sigma_{k+1}^\gamma, \quad 1 \le k < N. \tag{94}$$

In the steady state the current expectations $j^\alpha := \langle j_k^\alpha \rangle$ are position-independent.

We choose two boundary Lindblad operators $D^{L,R}$ to favour a relaxation of the boundary spins towards target states given by density matrices ρ_L, ρ_R satisfying $\mathscr{D}_1(\rho_L) = \mathscr{D}_N(\rho_R) = 0$. As target states we choose fully polarized states of one boundary spin

$$\rho_L = \frac{1}{2}(\mathbb{1} + \mathbf{n}_L \cdot \boldsymbol{\sigma}) \otimes \tilde{\rho}, \quad \rho_R = \tilde{\rho} \otimes \frac{1}{2}(\mathbb{1} + \mathbf{n}_R \cdot \boldsymbol{\sigma}) \tag{95}$$

where $|\mathbf{n}_L| = |\mathbf{n}_R| = 1$ and $\tilde{\rho}$ is an arbitrary reduced density matrix for the remaining $N-1$ spins. The reduced single-site boundary density matrix $\rho_B^{(1)} = \frac{1}{2}(\mathbb{1} + \mathbf{n}_B \cdot \boldsymbol{\sigma})$ is a pure state since for a projection direction given by

$$\mathbf{n}_B = (\sin(\phi_B)\cos(\theta_B), \sin(\phi_B)\sin(\theta_B), \cos(\phi_B)). \tag{96}$$

One has $\rho_B^{(1)} = |\psi_B) \otimes (\psi_B|$ with

$$|\psi_B) = \mathrm{e}^{i\alpha_B}\begin{pmatrix} \cos(\phi_B/2)\mathrm{e}^{-i\theta_B/2} \\ \sin(\phi_B/2)\mathrm{e}^{i\theta_B/2} \end{pmatrix} \tag{97}$$

and arbitrary phase $\alpha_B \in [0, 2\pi)$. The notion "full polarization" means that the expectation of the spin projection $\mathbf{n}_B \cdot \boldsymbol{\sigma}$ in the space-direction defined by $\mathbf{n}_B$ is given by $\langle\, \mathbf{n}_B \cdot \boldsymbol{\sigma}\, \rangle = 1$.

Due to the rotational symmetry (87) of H only the angle between the two boundary polarization vectors plays a role. Therefore we may, without loss of generality, choose $\phi_L = \phi_R = \pi/2$ and fix the coordinate frame in $\mathbb{R}^3$ such that the X–axis points in the $\mathbf{n}_L$ direction (corresponding to $\theta_L = 0$) and to let the XY-plane be spanned by the family vectors $\mathbf{n}_R(\theta)$, i.e.,

$$\mathbf{n}_L = (1, 0, 0), \quad \mathbf{n}_R = (\cos\theta, \sin\theta, 0), \quad 0 \le \theta \le \pi \tag{98}$$

corresponding to $\theta_R = \theta$.

It is easy to verify that there are two families of local Lindblad operators satisfying $\mathscr{D}_1[\rho_L] = 0$, viz. $D_1^R = a(\sigma_1^2 + i\sigma_1^3) + b(\mathbf{1} - \sigma_1^1)$ and $D_1^{R'} = a'\mathbf{1} + b'\sigma_1^1$. Following [16, 17] we choose D^R with $b = 0$ (so that $\mathrm{Tr}(D_1^R) = 0$) and coupling strength $a = \sqrt{\Gamma}$. Similarly, we choose for the right boundary site N the rotated projection to arrive at

$$D_1^R = \sqrt{\Gamma}(\sigma_1^2 + i\sigma_1^3), \quad D_N^L = \sqrt{\Gamma}(\sigma_N^2 \cos\theta - \sigma_N^1 \sin\theta + i\sigma_N^3). \tag{99}$$

Then in absence of the unitary term in (20) the boundary spins relax with characteristic times $\propto \Gamma^{-1}$ to approach ρ_L, ρ_R: Writing $\rho_L(t) = 1/2(\sigma_1^0 + x(t)\sigma_1^1 + y(t)\sigma_1^2 + z(t)\sigma_1^3) \otimes \tilde{\rho}$ one has $x(t) = 1 + (x(0) - 1)\exp(-4\Gamma t)$, $y(t) = y(0)\exp(-2\Gamma t)$, $z(t) = z(0)\exp(-2\Gamma t)$, and similarly for $\rho_R(t)$.

Remark 4 In the untwisted case $\mathbf{n}_L = \mathbf{n}_R := \mathbf{n}$ corresponding to $\theta = 0$ the Lindblad equation (20) for the stationary density matrix is trivially solved by [27]

$$\rho_N(\Gamma, 0) = \left(\frac{\mathbb{1} + \mathbf{n}\cdot\boldsymbol{\sigma}}{2}\right)^{\otimes N}. \tag{100}$$

This is a pure state of the form $\rho_N(\Gamma, 0) = |\,\Psi\,\rangle\langle\,\Psi\,|$ where $|\,\Psi\,\rangle = |\,\psi\,\rangle^{\otimes N}$ and

$$|\,\psi\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}. \tag{101}$$

This pure state is not of the form $\exp(-\beta H)/Z$ for any β and therefore not an equilibrium state.

4.3 Matrix Product Solution

From now on we exclude $\theta = 0$ so that the boundary coupling introduces a twist in the XY-plane, which drives the system perpetually out of equilibrium.

Theorem 2 *Let $I, S^{\pm,z} \in \mathfrak{End}(\mathfrak{A})$ be the infinite dimensional representation (91) of $\mathfrak{gl}_2(\mathbb{C})$ with representation parameter*

$$p = i\Gamma^{-1} \tag{102}$$

and let $\Omega^{00} = -\Omega^{11} = iS^z$, $\Omega^{01} = iS^+$, $\Omega^{10} = iS^-$. Furthermore, let

$$\langle\,\overline{W}, W\,| = \langle\,0\,|\otimes\langle\,0\,|, \quad |\,V, \overline{V}\,\rangle = \sum_{m,n=0}^{\infty}\left(-\cot\frac{\theta}{2}\right)^{m+n}\binom{2p}{m}\binom{2\bar{p}}{n}|\,m\,\rangle\otimes|\,n\,\rangle. \tag{103}$$

Then for ρ in matrix product from (33) and U defined by (81) the density matrix

$$\rho_N(\Gamma,\theta) = U\rho U^\dagger \tag{104}$$

is the unique solution of the quantum master equation (20) for the Heisenberg ferromagnet (92) with boundary dissipators (99).

Uniqueness is guaranteed by the structure of the Lindblad dissipators, see [29]. The proof of (104) follows from verifying the local divergence condition (49) with $\Xi^{00} = \Xi^{11} = iI$, $\Xi^{01} = \Xi^{10} = 0$ and the Lindblad boundary matching condition (53) by (somewhat lengthy but straightforward) explicit computation [16, 17]. Proposition 1 then yields for the non-equilibrium partition function (22)

$$Z_N(\Gamma,\theta) = \langle\, \overline{W}, W\,|\Theta_0^N|\, V, \overline{V}\,\rangle \tag{105}$$

with $\Theta_0 = 2S_1^zS_2^z + S_1^+S_2^+ + S_1^-S_2^-$ defined by (39).

We summarize the main conclusions of [16, 17, 25] drawn from Theorem 2 and the underlying quadratic algebra and Lindblad boundary matching property for $0 < \theta < \pi$.

(1) Dropping the arguments Γ, θ, the stationary magnetization currents are given by

$$j_N^x = -8ip\frac{Z_{N-1}}{Z_N}, \quad j_N^y = -\cot\frac{\theta}{2}j_N^x, \quad j_N^z = -4\frac{\frac{\mathrm{d}}{\mathrm{d}\theta}Z_{N-1}}{Z_N}. \tag{106}$$

Based on numerically exact computation up to $N = 100$ we conjectured that for any fixed coupling strength Γ one has [17]

$$\lim_{N\to\infty} N^2\frac{Z_{N-1}(\Gamma,\theta)}{Z_N(\Gamma,\theta)} = \frac{1}{4}\theta^2. \tag{107}$$

For the currents this result implies

$$\lim_{N\to\infty} N^2 j_N^x(\Gamma,\theta) = \frac{2\theta^2}{\Gamma}, \quad \lim_{N\to\infty} N j_N^z(\Gamma,\theta) = 2\theta. \tag{108}$$

Some rigorous results have been obtained for the Zeno limit $\Gamma \to \infty$ [25]. Rescaling the normalization factor yields a finite limit

$$\tilde{Z}_N(\theta) := \frac{1}{4}\lim_{\Gamma\to\infty}\Gamma^2 Z_N(\Gamma,\theta) \tag{109}$$

which was computed explicitly. Then for small twist angle $\theta = o(1/N)$ the conjecture (107) can be proved rigorously. For the currents one therefore finds

Theorem 3 *Let $\tilde{j}_N^\alpha(\theta) := \lim_{\Gamma\to\infty} j_N^\alpha(\Gamma,\theta)$ be the stationary currents of the Heisenberg chain in the Zeno limit. Then for any N one has*

$$\tilde{j}_N^x(\theta) = \tilde{j}_N^y(\theta) = 0 \quad \forall \theta \in [0, \pi[\tag{110}$$

and for any real $\varepsilon > 0$ *and real* $\theta_0 > 0$

$$\lim_{N\to\infty} N^{2+\varepsilon}\, \tilde{j}_N^z \left(\frac{\theta_0}{N^{1+\varepsilon}} \right) = 2\theta_0. \tag{111}$$

The first statement is a trivial consequence of the explicit expressions (106) and the result that $\tilde{Z}_N(\theta)$ is finite and non-zero. The second statement follows from the explicit form of $\tilde{Z}_N(\theta)$ given in [25].
(2) In terms of

$$B^x := \Theta^{01} + \Theta^{10}, \quad B^y := i\left(\Theta^{01} - \Theta^{10}\right), \quad B^z = \Theta^{00} - \Theta^{11} \tag{112}$$

the multiplication property (80) yields

$$\langle\, \sigma_k^\alpha \,\rangle_N = \frac{S_{k,N}^\alpha(\Gamma, \theta)}{Z_N(\Gamma, \theta)} \tag{113}$$

with

$$S_{k,N}^\alpha(\Gamma, \theta) = \langle\, \overline{W},\, W\, |\Theta_0^{k-1} B^\alpha \Theta_0^{N-k}|\, V, \overline{V}\,\rangle. \tag{114}$$

(3) The quadratic algebra implies that the operators Θ_0 and B^α satisfy the remarkable *cubic* relation

$$[\Theta_0, [\Theta_0, B^\alpha]] + 2\{\Theta_0, B^\alpha\} - 8p^2 B^\alpha = 0 \tag{115}$$

which was found earlier for a specific representation by using computer algebra [27]. This relation induces recursion relations for the unnormalized correlation functions $Z_N \langle\, \sigma_{k_1}^{\alpha_1} \ldots \sigma_{k_n}^{\alpha_n} \,\rangle$. In particular, with

$$\begin{aligned} B^x &= S_1^+ S_2^+ - S_1^- S_2^-, \quad B^y = S_1^z(S_2^- - S_2^+) + (S_1^- - S_1^+)S_2^z, \\ B^z &= i\, S_1^z(S_2^- + S_2^+) - i(S_1^- + S_1^+)S_2^z. \end{aligned} \tag{116}$$

one finds for the unnormalized one-point function (dropping the arguments)

$$S_{k+2,N+1}^\alpha + S_{k,N+1}^\alpha - 2S_{k,N}^\alpha + 2(S_{k,N}^\alpha + S_{k+1,N}^\alpha) - 8p^2 S_{k,N-1}^\alpha = 0. \tag{117}$$

By setting $r = k/N$ and taking the continuum limit $k, N \to \infty$ such that the macroscopic coordinate r remains fixed this recursion together with (107) yields the simple ordinary differential equation $m''(r) + \theta^2 m(r) = 0$ for the large-scale magnetization profile $m^\alpha(r) := \lim_{k,N\to\infty} \langle\, \sigma_k^\alpha \,\rangle_N$. The boundary conditions are given by the microscopic complete polarizations so that

$$m^x(r) = \cos(\theta r), \quad m^y(r) = \sin(\theta r), \quad m^z(r) = 0, \quad 0 \le r \le 1 \tag{118}$$

corresponding to a spin helix state [25]. This implies a strongly sub-diffusive current $0 = \lim_{N\to\infty} N j_N^x = \lim_{N\to\infty} N j_N^y$ inside the twist-plane.

For maximal twist $\theta = \pi$ one has by symmetry $j_N^y(\Gamma, \pi) = j_N^z(\Gamma, \pi) = 0$ for all N, Γ [23]. Magnetization profiles and correlation functions were computed in [8, 27] using the cubic relation (115) and the resulting continuum approximation, see also the review [29]. Remarkably, the correlations along the twist axis are of a form reminiscent of what was obtained for the symmetric simple exclusion process with open boundaries, using the fluctuating hydrodynamics approach [31]. This similarity suggest that also the boundary driven quantum problem may be understood in terms of fluctuating hydrodynamics.

Acknowledgements VP and GMS thank T. Prosen for useful discussions and DFG for financial support.

References

1. Alcaraz, F.C., Dasmahapatra, S., Rittenberg, V.: N-species stochastic models with boundaries and quadratic algebras. Phys. A **257**, 1 (1998)
2. Attal, S., Joye, A., Pillet, C.-A. (eds.): Open Quantum Systems II. The Markovian Approach. Springer, Berlin (2006)
3. Baxter, R.J.: Exactly Solved Models in Statistical Mechanics. Academic, New York (1982)
4. Bertini, L., De Sole, A., Gabrielli, D., Jona Lasinio, G., Landim, C.: Macroscopic fluctuation theory for stationary non-equilibrium states. J. Stat. Phys. **107**, 635–675 (2002)
5. Bertini, L., De Sole, A., Gabrielli, D., Jona Lasinio, G., Landim, C.: Macroscopic fluctuation theory. Rev. Mod. Phys. **87**, 593–636 (2015)
6. Blythe, R.A., Evans, M.R.: Nonequilibrium steady states of matrix-product form: a solver's guide. J. Phys. A: Math. Theor. **40**, R333–R441 (2007)
7. Breuer, H.-P., Petruccione, F.: The Theory of Open Quantum Systems. Oxford University Press, Oxford (2002)
8. Buča, B., Prosen, T.: Connected correlations, fluctuations and current of magnetization in the steady state of boundary driven XXZ spin chains. J. Stat. Mech. 023102 (2016)
9. Derrida, B.: An exactly soluble non-equilibrium system: the asymmetric simple exclusion process. Phys. Rep. **301**, 65–83 (1998)
10. Derrida, B., Lebowitz, J.L., Speer, E.R.: Large deviation of the density profile in the steady state of the open symmetric simple exclusion process. J. Stat. Phys. **107**, 599–634 (2002)
11. Frigerio, A., Spohn, H.: Stationary states of quantum dynamical semigroups and applications. In: Accardi, L., Gorini, V., Paravicini, G. (eds.) Proceedings of Mathematical Problems in the Theory of Quantum Irreversible Processes, Laboratoria di Cibernetica del CNR, pp. 115–135 (1978)
12. Gorini, V., Kossakowski, A., Sudarshan, E.C.G.: Completely positive dynamical semigroups of N-level systems. J. Math. Phys. **17**(5), 821–825 (1976)
13. Heisenberg, W.: Zur Theorie des Ferromagnetismus. Z. Phys. **49**, 619–636 (1928)
14. Ilievski, E., Prosen, T.: Exact steady state manifold of a boundary driven spin-1 Lai-Sutherland chain. Nucl. Phys. B **882**, 485 (2014)
15. Ising, E.: Beitrag zur Theorie des Ferromagnetismus. Z. Phys. **31**, 253–258 (1925)
16. Karevski, D., Popkov, V., Schütz, G.M.: Exact matrix product solution for the boundary-driven Lindblad XXZ chain. Phys. Rev. Lett. **110**, 047201 (2013)
17. Karevski, D., Popkov, V., Schütz, G.M.: Driven isotropic Heisenberg spin chain with arbitrary boundary twisting angle: exact results. Phys. Rev. E **88**, 062118 (2013)

18. Kipnis, C., Landim, C.: Scaling Limits of Interacting Particle Systems. Springer, Berlin (1999)
19. Kossakowski, A.: On quantum statistical mechanics of non-Hamiltonian systems. Rep. Math. Phys. **3**(4), 247–274 (1972)
20. Liggett, T.M.: Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Springer, Berlin (1999)
21. Lindblad, G.: On the generators of quantum dynamical semigroups. Commun. Math. Phys. **48**(2), 119–130 (1976)
22. Polishchuk, A., Positselski, L.: Quadratic Algebras. University Lecture Series, vol. 37. American Mathematical Society, Providence (2005)
23. Popkov, V., Livi, R.: Manipulating energy and spin currents in non-equilibrium systems of interacting qubits. New J. Phys. **15**, 023030 (2013)
24. Popkov, V., Prosen, T.: Infinitely dimensional Lax structure for one-dimensional Hubbard model. Phys. Rev. Lett. **114**, 127201 (2015)
25. Popkov, V., Schütz, G.M.: Solution of the Lindblad equation for spin helix states. Phys. Rev. E **95**, 042128 (2017)
26. Prosen, T.: Open XXZ spin chain: nonequilibrium steady state and a strict bound on Ballistic transport. Phys. Rev. Lett. **106**, 217206 (2011)
27. Prosen, T.: Exact nonequilibrium steady state of a strongly driven open XXZ chain. Phys. Rev. Lett. **107**, 137201 (2011)
28. Prosen, T.: Exact nonequilibrium steady state of an open Hubbard chain. Phys. Rev. Lett. **112**, 030603 (2014)
29. Prosen, T.: Matrix product solutions of boundary driven quantum chains. J. Phys. A: Math. Theor. **48**, 373001 (2015)
30. Schütz, G.M.: Exactly solvable models for many-body systems far from equilibrium. In: Domb, C., Lebowitz, J. (eds.) Phase Transitions and Critical Phenomena, vol. 19, pp. 1–251. Academic Press, London (2001)
31. Spohn, H.: Long-range correlations for stochastic lattice gases in a non-equilibrium steady state. J. Phys. A: Math. Gen. **16**, 4275–4291 (1983)
32. Spohn, H.: Nonlinear fluctuating hydrodynamics for anharmonic chains. J. Stat. Phys. **154**, 1191–1227 (2014)

Ultradistribution Spaces: Superprocesses and Nonlinear Differential Problems

R. Vilela Mendes

Abstract From branching particle systems one obtains, in the scaling limit, measure-value processes called superprocesses. In addition to providing models for evolving populations, superprocesses provide probabilistic representations of the solutions of nonlinear partial differential equations (PDE's). However, the class of PDE's that can be handled by measure-valued superprocesses is rather limited. This suggests an extension of the configuration space of superprocesses to ultradistribution-valued processes which have a wider range of applications in the solution of PDE's. The relevance of the superprocess representation of PDE's to deal with nonlinear singular problems is also discussed.

Keywords Ultradistributions · Superprocesses · Nonlinear PDE's

1 Distributions and Ultradistributions

One of the motivations to develop distribution theory arose from the need to deal with non-smooth entities in differential equations. In particular the generalization of the notion of derivative led to the spaces of distributions ($\mathscr{D}'$) and tempered distributions ($\mathscr{S}'$). However, the theory of distributions is not just $\mathscr{D}'$ and $\mathscr{S}'$. There are many other interesting spaces of "*generalized functions*". In the Figs. 1 and 2 (adapted from [1]) are displayed some other test function spaces, their dense embeddings and Fourier maps (Fig. 1) as well as their corresponding duals (distribution spaces) (Fig. 2).

As $\mathscr{D}'$ and $\mathscr{S}'$ are a tool of choice to deal with linear differential and partial differential equations, some of the other spaces might be more appropriate to deal with other types of mathematical problems. The main properties of the spaces listed in the figures are summarized here:

Test Function Spaces

R. Vilela Mendes (✉)
Centro de Matemática e Aplicações Fundamentais, University of Lisbon, Lisbon, Portugal
e-mail: rvmendes@fc.ul.pt; rvilela.mendes@gmail.com

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_12

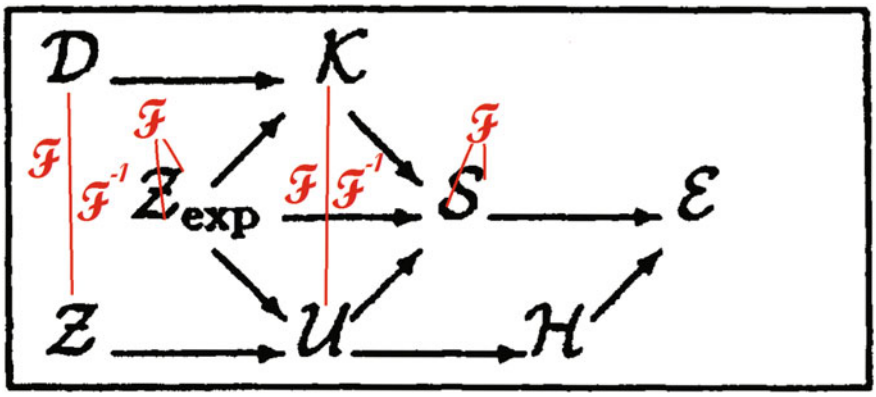


Fig. 1 Test function spaces (adapted from [1])

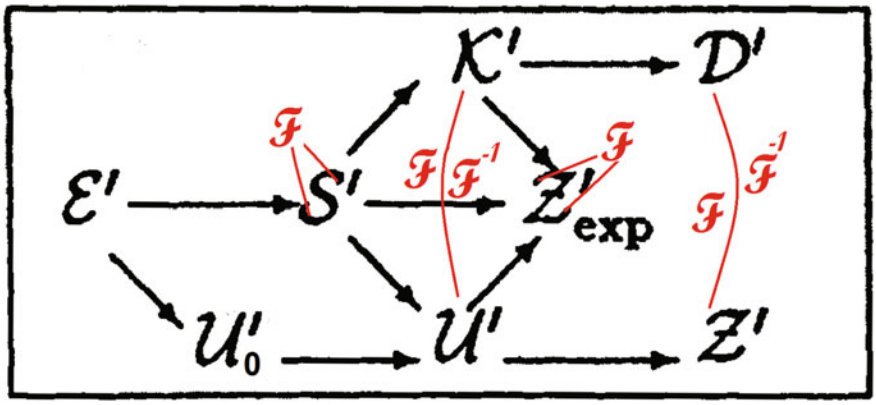


Fig. 2 Distribution spaces (adapted from [1])

$\#\ \mathscr{D} = \cup_K \{\mathscr{D}_K : \varphi \in C^\infty, \text{supp}\,(\varphi) \subset K\}\,;\ \|\varphi\|_{(p,K)} = \max_{0\le r\le p}\left\{\sup\left|\varphi^{(r)}\right|\right\}$
$\#\ \mathscr{K} = \cap_{p=0}^{\infty}\mathscr{K}_p;\ \mathscr{K}_p = \text{completion of } \mathscr{D} \text{ for the norm } \|\varphi\| = \max_{0\le q\le p}\left\{\sup\left|e^{p|x|}\varphi^{(q)}\right|\right\}$
$\#\ \mathscr{S} = \cap\mathscr{S}_{p,r} = \left\{\varphi \in C^\infty : \|\varphi\|_{p,r} = \sup\left|x^p\varphi^{(r)}\right|\right\}$
$\#\ \mathscr{E} = \varphi \in C^\infty$ with uniform convergence on compacts
$\#\ \mathscr{Z} = \varphi : \mathscr{F}\{\varphi\} \in \mathscr{D}, \varphi(z)$ entire: $\left|z^k\varphi(z)\right| \le C_k e^{a|Im(z)|}$
$\#\ \mathscr{U} = \cap_{p=0}^{\infty}\mathscr{U}_p;\ \mathscr{U}_p = \left\{\varphi : \mathscr{F}\{\varphi\} \in \mathscr{K}_p\right\};\ \|\varphi\|_p = \sup_{z\in\Lambda_p}\left\{(1+|z|^p)\,|\varphi(z)|\right\}$
$\#\ \mathscr{H} =$ Entire functions with topology of uniform convergence on compacts of $\mathbb{C}$
$\#\ \mathscr{Z}_{\exp} = \cap_{j=1}^{\infty}\mathscr{Z}_{\exp,j};\ \mathscr{Z}_{\exp,j} = \left\{\varphi : \|\varphi\|_{\exp,j} = \max_{k\le j}\left\{e^{j|Re(z)|}\left|\varphi^{(k)}(z)\right|\right\}\right\}$

Distribution Spaces

$\#\ \mathscr{D}' =$ Schwartz distributions; locally $\mu(x) = D^k(f(x))$
$\#\ \mathscr{K}' =$ Distributions of exponential type, $\mu(x) = D^k\left(e^{a|x|}f\right)$
$\#\ \mathscr{S}' =$ Tempered distributions
$\#\ \mathscr{E}' =$ Subspace of $\mathscr{D}'$ of distributions of compact support

$\mathscr{Z}'$ = Ultradistributions, $\mathscr{D}' \overset{\mathscr{F}}{\to} \mathscr{Z}'$; $\mathscr{Z}' \overset{\mathscr{F}^{-1}}{\to} \mathscr{D}'$
$\mathscr{U}'$ = Tempered ultradistributions
$\mathscr{U}_0'$ = Dual of $\mathscr{H}$, ultradistributions of compact support
$\mathscr{Z}_{\exp}'$ = Topological dual of $\mathscr{Z}_{\exp}$, contains $\mathscr{U}'$ and $\mathscr{K}'$ as proper subspaces

Of particular relevance is the relation of the upper and lower lines in the figures through the Fourier transform. In particular the fact that the Fourier transform of $\mathscr{Z}$ has compact support endows the ultradistribution space $\mathscr{Z}'$ with a rich analytical structure. Ultradistributions not only have derivatives of all orders, like the distributions, but also have Taylor series expansions. This fact, among other things, makes them more convenient than distributions in some application problems.

These generalized function spaces emphasize the role of the Fourier transform. Other techniques have been used to define test function spaces smaller than $\mathscr{D}$ and therefore distribution spaces larger than $\mathscr{D}'$. Among them are the test function spaces defined through weight functions or weight sequences. Given an integrable function of compact support φ, the condition for infinite differentiability ($\varphi \in \mathscr{D}$) is expressed through the Fourier transform by

$$\int \mathscr{F}\{\varphi\}(x)\, e^{n\log(1+|x|)} dx < \infty \qquad \forall n.$$

Replacing $\log(1+|x|)$ by a larger function $\omega(|x|)$

$$\int \mathscr{F}\{\varphi\}(x)\, e^{\lambda\omega(|x|)} dx < \infty$$

one obtains a smaller test function space. With different definitions of projective limit on λ and inductive limit on the compacts, one obtains ultradistribution spaces of class ω of Beurling [2] or Roumieu [3] type.

Another trend defines ultradifferentiable functions [3–6] of class (M_p) (or $\{M_p\}$) by $\forall h > 0\ \exists C > 0$ (or $\exists h > 0$ and $C > 0$) such that

$$\sup \left|\varphi^{(n)}(x)\right| \leq C h^n M_n \qquad \forall n \in \mathbb{N}$$

on compacts, M_p being a sequence of positive numbers.

In this paper I will be mainly concerned with the space $\mathscr{U}'$ of Silva tempered ultradistributions [7, 8] and the space $\mathscr{U}_0'$ of ultradistributions of compact support.

1.1 Silva Tempered Ultradistributions

$\mathscr{U} \subset \mathscr{S}$, is the space of functions in $\mathscr{S}$ that may be extended into the complex plane as entire functions of rapid decrease on strips. Namely $\mathscr{U}(\mathbb{C})$ consists of all entire functions φ for which

$$\|\varphi\|_p = \sup_{|Imz|<p} \left\{\left(1+|z|^p\right)|\varphi(z)|\right\} < \infty \qquad \forall p \in \mathbb{N} \tag{1}$$

$\mathscr{U}$ topologized by the norms $\|\varphi\|_p$ is a Fréchet space and for each $\varphi(z) \in \mathscr{U}$, $\varphi(z)|_{\mathbb{R}} \in \mathscr{S}(\mathbb{R})$.

$\mathscr{U}'$, the dual of $\mathscr{U}$ is Silva space of tempered ultradistributions [7, 8]. It may also be characterized as the space of all Fourier transforms of distributions of exponential type, that is

$$\mathscr{U}' = \mathscr{F}\left\{\mathscr{K}'\right\} \tag{2}$$

$\mathscr{K}'$ being the space of finite order derivatives of some exponentially bounded continuous function, i.e. for each $f \in \mathscr{K}'$ there is $b \geq 0$, $m \in N$ and a bounded, continuous function g such that

$$f(t) = (e^{b|t|}g(t))^{(m)}$$

However, the representation of tempered ultradistributions by analytical functions is the most convenient one for practical calculations. Define B_η as the complement in $\mathbb{C}$ of the strip $\Lambda_\eta = \{z : \mathrm{Im}(z) \leq \eta\}$

$$B_\eta = \{z : \mathrm{Im}(z) > \eta\} \tag{3}$$

and H_η as the set of functions which are holomorphic and of polynomial growth in B_η

$$\varphi(z) \in H_\eta \Longrightarrow \exists M, \alpha : |\varphi(z)| < M|z|^\alpha, \forall z \in B_\eta. \tag{4}$$

Let H_ω be the union of all such spaces

$$H_\omega = \underset{\eta\geq 0}{\cup} H_\eta \tag{5}$$

and in H_ω define the equivalence relation Ξ by

$$\varphi \overset{\Xi}{\simeq} \psi \text{ if } \varphi - \psi \text{ is a polynomial.}$$

Then, the space of tempered ultradistribution is

$$\mathscr{U}' = H_\omega/\Xi \tag{6}$$

and $[\phi(z)]$ will denote the equivalence class. The vectorial operations as well as derivation and multiplication by polynomials, defined on H_ω, are compatible with the equivalence relation and $\mathscr{U}'$ becomes a vector space with these operations.

The Schwartz space $\mathscr{S}'$ of tempered distributions may be identified with a subspace of $\mathscr{U}'$ by the Stieltges transform, that is, a linear mapping of $\mathscr{S}'$ on a subspace $\mathscr{U}'^*$ of $\mathscr{U}'$. Namely, given $\nu(x) \in \mathscr{S}'$

$$\varphi(z) = \frac{p(z)}{2\pi i}\int \frac{\nu(x)}{p(x)(x-z)}dx + P(z) \tag{7}$$

$[\varphi(z)] \in \mathscr{U}'$. Here $p(z)$ is a polynomial such that $\nu/p \sim O\left(x^{-1}\right)$ in the Silva-Cesàro sense and $P(z)$ is an arbitrary polynomial [4, 9, 10].

Operations on tempered ultradistributions $f \in \mathscr{U}'$ are performed using their analytical images $\varphi(z)$. For example f is integrable in $\mathbb{R}$ if there is an $y_0 \in \mathbb{R}$ and a $\varphi(z)$ in $[\varphi(z)] \in \mathscr{U}'$ such that $\varphi(x+iy_0) - \varphi(x-iy_0)$ is integrable in $\mathbb{R}$ in the sense of distributions. Then

$$\langle\varphi|g\rangle = \oint_{\Gamma_{y_0}} \varphi(z)\, g(z)\, dz \tag{8}$$

$\varphi \in \mathscr{U}'$, $g \in \mathscr{U}$ and the integral runs around the boundaries of the strip $\mathrm{Im}(z) \leq y_0$.

An ultradistribution vanishes in an open set $A \in \mathbb{R}$ if $\varphi(x+iy) - \varphi(x-iy) \to 0$ for $x \in A$ when $y \to 0$ or, equivalently, if there is an analytical extension of φ to the vertical strip $\mathrm{Re}\, z \in A$, being of at most polynomial growth there. The support of ν is the complement in $\mathbb{R}$ of the largest open set where ν vanishes.

All these notions are easily generalized to $\mathbb{R}^n$ [8, 11] by considering products of semiplans as in (3) and the corresponding polynomial bounds. For the equivalence relation Ξ one uses pseudopolynomials, that is, functions of the form

$$\sum_{j,k} \rho\left(z_1, \cdots, \hat{z_j}, \cdots, z_n\right) z_j^k,$$

$\hat{z_j}$ meaning that this variable is absent from the arguments of ρ.

An ultradistribution ν in $\mathbb{R}^n$ has compact support if there is a disk D such that any φ in $[\varphi(z)] \in \mathscr{U}'$ has an analytic extension to $(\mathbb{C}/D)^n$, being of at most polynomial growth there. Then the integral in (8) is around a closed contour containing the support of the ultradistribution. In particular for a tempered ultradistribution of compact support there is a unique representative function $\varphi^0(z)$ vanishing at ∞. Then from its Laurent expansion it follows

$$\left[\mu^0\right] = \left[\sum_{i=1}^{\infty} c_n \frac{1}{(z-a)^n}\right] \to -\sum_{i=1}^{\infty} (-1)^n \frac{2\pi i}{n!} c_n \delta^{(n)}(z-a) \tag{9}$$

showing that any ultradistribution of compact support has a representation as a series of multipoles [8]. The space of *tempered ultradistributions of compact support* will be denoted $\mathscr{U}_0'$. $\mathscr{U}_0'$ may be identified with $\mathscr{H}'$, the space of analytic functionals, dual of the space $\mathscr{H}$ of entire functions with the topology of uniform convergence on compacts of $\mathbb{C}$.

For all practical purposes an analytic representative $\varphi(z)$ in $\mathbb{C}\backslash\Lambda_\eta$ of a tempered ultradistribution corresponds (up to a common polynomial) to a pair of

functions $\left[\varphi_+(z), \varphi_-(z)\right]$ which are holomorphic and of polynomial growth respectively above and below some strip $\Lambda_\eta = \{z : \mathrm{Im}(z) \leq \eta\}$.

For comparison, it is perhaps useful to recall the corresponding analytic representation of distributions and tempered distributions. The Cauchy representation of a distribution of compact support $f \in \mathscr{D}'(\Omega)$

$$C(f)(z) = \frac{1}{2\pi i}\left\langle f(t), \frac{1}{t-z}\right\rangle \tag{10}$$

is analytic in $\mathbb{C}\backslash\mathbb{R}$ and f is recovered by

$$f(x) = \lim_{y\to 0^+}\{C(f)(x+iy) - C(f)(x-iy)\} \tag{11}$$

For a general distribution on an open interval I, the Cauchy representation cannot be defined as above. Nevertheless, for $\Lambda_I = \{I + i\mathbb{R}\}$, there is a function F analytic in $\Lambda_I\backslash I$ such that the jump operator as in (11) recovers the distribution. One has the isomorphism

$$\mathscr{D}'(I) \cong \mathscr{H}_{\mathscr{D}'}(\Lambda_I\backslash I)/\mathscr{H}_{\mathscr{D}'}(\Lambda_I)$$

$\mathscr{H}_{\mathscr{D}'}(\Lambda_I\backslash I)$ being the space of functions analytic in $\Lambda_I\backslash I$ and for which $\exists N$ such that

$$\sup |Imz|^N F(z) < \infty$$

on vertical strips contained on $\Lambda_I\backslash I$. The kernel of the homomorphism is the space of such functions which are analytic in the whole of Λ_I.

For $\mathscr{S}'(\mathbb{R})$ one has

$$\mathscr{S}'(\mathbb{R}) \cong \mathscr{H}_{\mathscr{S}'}(\mathbb{C}\backslash\mathbb{R}) / (\mathscr{H}_{\mathscr{S}'}(\mathbb{C}\backslash\mathbb{R}) \cap \mathscr{H}_{\mathscr{S}'}(\mathbb{C}))$$

$\mathscr{H}_{\mathscr{S}'}(\mathbb{C}\backslash\mathbb{R})$ being the functions analytic in $\mathbb{C}\backslash\mathbb{R}$ such that $\forall R > 0\ \exists m, N \in \mathbb{N}$

$$\sup_{z\in\Lambda_R\backslash\mathbb{R}} \frac{|Imz|^N |F(z)|}{(1+|Rez|)^m} < \infty$$

and $\mathscr{H}_{\mathscr{S}'}(\mathbb{C})$ all such functions which are analytic in the whole of $\mathbb{C}$.

2 Superprocesses on Ultradistributions

A superprocess describes the evolution of a population, without a fixed number of units, that evolves according to the laws of chance. It involves both propagation and branching of paths. They have been extensively used to model population dynamics and, more recently, as a tool for the construction of solutions of nonlinear partial differential equations [12–14].

Given a countable dense subset Q of $[0, \infty)$ and a countable dense subset F of a separable metric space E, the countable set

$$M_1 = \left\{ \sum_{i=1}^{n} \alpha_i \delta_{x_i} : x_1 \cdots x_n \in F; \alpha_1 \cdots \alpha_n \in Q; n \geq 1 \right\} \tag{12}$$

is dense (in the topology of weak convergence) on the space $M(E)$ of finite Borel measures on E [14]. This is at the basis of the interpretation of the limits of evolving particle systems as measure-valued superprocesses. The representation of an evolving measure as a collection of measures with point support is useful for the construction of solutions of nonlinear partial differential equations as rescaling limits of measure-valued superprocesses.

However, as far as representations of solutions of nonlinear PDE's, superprocesses constructed in the space $M(E)$ of finite measures have serious limitations. The set of nonlinear terms that can be handled is limited (essentially to powers $u^{\alpha}(x)$ with $\alpha \leq 2$) and derivative interactions cannot be included as well. The first obvious generalization would be to construct superprocesses on distributions of point support, because any such distribution is a finite sum of deltas and their derivatives [15]. However, because in a general branching process the number of branches is not bounded, one really needs a framework that can handle arbitrary sums of deltas and their derivatives. This requirement leads naturally to the space of ultradistributions of compact support $\mathscr{U}_0'$, by virtue of the multipole expansion property (9) mentioned before.

The limitations of superprocesses on measures and the generalization to superprocesses on ultradistributions are described in detail in Refs. [16–18]. Here the main results will be summarized.

Let the underlying space of the superprocess be $\mathbb{R}^n$ and denote by $\left(X_t, P_{0,\nu}\right)$ a branching stochastic process with values in $\mathscr{U}_0'$ and transition probability $P_{0,\nu}$ starting from time 0, $x \in \mathbb{R}^n$ and $\nu \in \mathscr{U}_0'$. The process is assumed to satisfy the *branching property,* that is, given $\nu = \nu_1 + \nu_2$

$$P_{0,\nu} = P_{0,\nu_1} * P_{0,\nu_2}. \tag{13}$$

After the branching $\left(X_t^1, P_{0,\nu_1}\right)$ and $\left(X_t^2, P_{0,\nu_2}\right)$ are independent and $X_t^1 + X_t^2$ has the same law as $\left(X_t, P_{0,\nu}\right)$. In terms of the *transition operator* V_t operating on functions on $\mathscr{U}$ this would be

$$\langle V_t f, \nu_1 + \nu_2 \rangle = \langle V_t f, \nu_1 \rangle + \langle V_t f, \nu_2 \rangle \tag{14}$$

with V_t defined by $e^{-\langle V_t f, \nu \rangle} = P_{0,\nu} e^{-\langle f, X_t \rangle}$ or

$$\langle V_t f, \nu \rangle = -\log P_{0,\nu} e^{-\langle f, X_t \rangle} \tag{15}$$

$f \in \mathscr{U}$, $\nu \in \mathscr{U}_0'$.

In $M = [0, \infty) \times \mathbb{R}^n$ consider an open regular set $Q \subset M$ and the associated exit process $\xi = \left(\xi_t, \Pi_{0,x}\right)$ with parameter k defining the lifetime. The process starts from $x \in \mathbb{R}^n$ carrying along an ultradistribution in $\mathscr{U}_0'$ indexed by the path coordinate. At each branching point (ruled by $\Pi_{0,x}$) of the ξ_t-process there is a transition ruled by a P probability in $\mathscr{U}_0'$ leading to one or more elements in $\mathscr{U}_0'$. These $\mathscr{U}_0'$ elements are then carried along by the new paths of the ξ_t-process. By construction, in each path, the process never leaves $\mathscr{U}_0'$. The whole process stops at the boundary ∂Q, finally defining an exit process $\left(X_Q, P_{0,\nu}\right)$ on $\mathscr{U}_0'$. If the initial ν is δ_x and $f \in \mathscr{U}$ a function on ∂Q one writes

$$u(x) = \left\langle V_Q f, \delta_x \right\rangle = -\log P_{0,x} e^{-\left\langle f, X_Q \right\rangle} \tag{16}$$

$\left\langle f, X_Q \right\rangle$ being computed on the (space-time) boundary with the exit ultradistribution generated by the process.

The connection with nonlinear PDE's is established by defining the whole process to be a $(\xi, \psi)-$*superprocess* if $u(x)$ satisfies the equation

$$u + G_Q \psi(u) = K_Q f \tag{17}$$

where G_Q is the Green operator,

$$G_Q f(0, x) = \Pi_{0,x} \int_0^\tau f(s, \xi_s)\, ds \tag{18}$$

and K_Q the Poisson operator

$$K_Q f(x) = \Pi_{0,x} 1_{\tau < \infty} f(\xi_\tau) \tag{19}$$

$\psi(u)$ means $\psi(0, x; u(0, x))$ and τ is the first exit time from Q, Eq. (17) being recognized as the integral version of a nonlinear partial differential equation with the Green operator determined by the linear part of the equation and $\psi(u)$ by the nonlinear terms. If the equation does not possess a natural Poisson clock for the branching one has to introduce an artificial lifetime for the particles in the process (e^{-k}), which in the end must vanish $(k \to \infty)$ through a rescaling method.

The superprocess is constructed as follows: Let $\varphi(s, x; z)$ be the branching function at time s and point x. Then denoting $P_{0,x} e^{-\left\langle f, X_Q \right\rangle}$ as $e^{-w(0,x)}$ one has

$$e^{-w(0,x)} = \Pi_{0,x} \left[e^{-f(\tau, \xi_\tau)} + k \int_0^\tau ds \left[\varphi\left(s, \xi_s; e^{-w(\tau - s, \xi_s)}\right) - e^{-w(\tau - s, \xi_s)} \right] \right] \tag{20}$$

where τ is the first exit time from Q and $f(\tau, \xi_\tau) = \left\langle f, X_Q \right\rangle$ is computed with the exit boundary ultradistribution. Existence of $\left\langle f, X_Q \right\rangle$ and hence of $e^{-w(0,x)}$ is insured if $f \in \mathscr{U}$ and the branching function is such that the exit $X_Q \in \mathscr{U}_0'$.

Equation (17) is then obtained by a limiting process. Let in (20) replace $w(0, x)$ by $\beta w_\beta(0, x)$ and f by βf. In a branching particle interpretation of the superprocess β may be interpreted as the mass of the particles and when the $\mathscr{U}_0'$-valued process

$X_Q \to \beta X_Q$ then $P_\mu \to P_{\frac{\mu}{\beta}}$.

$$e^{-\beta w(0,x)} = \Pi_{0,x}\left[e^{-\beta f(\tau,\xi_\tau)} + k_\beta \int_0^\tau ds\left[\varphi_\beta\left(s,\xi_s; e^{-\beta w(\tau-s,\xi_s)}\right) - e^{-\beta w(\tau-s,\xi_s)}\right]\right] \tag{21}$$

With

$$u_\beta^{(1)} = \left(1 - e^{-\beta w_\beta}\right)/\beta \quad ; \quad f_\beta^{(1)} = \left(1 - e^{-\beta f}\right)/\beta$$

or

$$u_\beta^{(2)} = \frac{1}{2\beta}\left(e^{\beta w_\beta} - e^{-\beta w_\beta}\right) \quad ; \quad f_\beta^{(2)} = \frac{1}{2\beta}\left(e^{\beta f} - e^{-\beta f}\right)$$

and

$$\psi_\beta^{(i)}\left(0,x;u_\beta^{(i)}\right) = \frac{k_\beta}{\beta}\left(\varphi\left(0,x;1-\beta u_\beta^{(i)}\right) - 1 + \beta u_\beta^{(i)}\right)$$

$u_\beta^{(i)} \to w_\beta$ and $f_\beta^{(i)} \to f$ when $\beta \to 0$, Eq. (17) being obtained in this limit.

Let $z = e^{-\beta w(\tau-s,\xi_s)} = P_{0,x}e^{-\langle \beta f, X\rangle}$. For the branching function $\varphi\,(s,x;z)$, in contrast with the measure-valued case, in addition to branchings of deltas into other deltas one also has:

(1) A change of sign in the point support ultradistribution

$$e^{\langle \beta f, \delta_x\rangle} = e^{\beta f(x)} \to e^{\langle \beta f, -\delta_x\rangle} = e^{-\beta f(x)}$$

which corresponds to

$$z \to \frac{1}{z}$$

and

(2) A change from $\delta^{(n)}$ to $\pm\delta^{(n+1)}$, for example

$$e^{\langle \beta f, \delta_x\rangle} = e^{\beta f(x)} \to e^{\left\langle \beta f, \pm\delta_x'\right\rangle} = e^{\mp\beta f'(x)}$$

which corresponds to

$$z \to e^{\mp\partial_x \log z}.$$

Case (1) corresponds to an extension of superprocesses on measures to superprocesses on signed measures and the second to superprocesses in $\mathscr{U}_0'$.

Existence of the superprocess is existence of a unique solution for the Eq. (21) and its rescaling limit. It will depend on the appropriate choice of the branching function $\varphi\,(s,y;z)$. Suppose that such a ultradistribution branching is specified. Associated to the ultradistribution superprocess Γ with branching function φ there is an *enveloping measure superprocess* $\widetilde{\Gamma}$ with branching function $\widetilde{\varphi}$ that has the same branching topology as Γ but without any derivative change in the original delta measure nor on its sign. General existence conditions for measure-valued superprocesses have been

found in the past [19–21]. Namely $\widetilde{\varphi}$ should have the form

$$\widetilde{\varphi}(s, y : z) = -b(s, y)\, z - c(s, y)\, z^2 + \int_0^\infty \left(e^{-\lambda z} + \lambda z - 1\right) n(s, y; d\lambda)\,. \quad (22)$$

Suppose that the branching $\widetilde{\varphi}$ for the process $\widetilde{\Gamma}$ is of the form (22). This insures almost sure existence of $e^{-\langle f, \widetilde{X}\rangle}$, $\widetilde{X}$ being the exit measure generated by the $\widetilde{\Gamma}$ process. Then, for the corresponding ultradistribution Γ superprocess one has

$$\langle f, X\rangle \leq M \int_{\partial Q} \sum_{n=0} \left|f^{(n)}\right|.$$

and the following result is obtained [18]:

Proposition 1 *A $\mathscr{U}_0'$ ultradistribution-valued exit superprocess Γ exists if the branching function $\widetilde{\varphi}$ of the associated enveloping exit measure process $\widetilde{\Gamma}$ is as in Eq. (22) and the boundary function f is such that the integral over the exit boundary of $\Sigma_n \left|f^{(n)}\right|$ is finite.*

3 Nonlinear Differential Problems, Distributions and Superprocesses

3.1 Nonlinear Theories of Generalized Functions

The general treatment of derivatives in distribution theory provides a powerful symbolic calculus for linear differential problems. However, when modeling natural phenomena, the most interesting problems are very often nonlinear. Of course, if the solutions of a nonlinear problem are known to be smooth, there is no problem because many nice algebras can be found in the domain of smooth functions. However in cases where the solutions are singular or the sources are concentrated (point, line or sheet charges, for example) the application of distribution theory becomes problematic. The problem was identified long ago by Schwartz in his impossibility result [22] which implies that $\mathscr{D}'$ cannot be linearly embedded into a differential algebra with the unit function as unit, with a derivation D satisfying Leibnitz rule, $D|_{C^1(\mathbb{R})}$ being the usual derivative and containing $C(\mathbb{R})$ as a subalgebra.

The need to deal with systems of PDE's without classical solutions, shock waves and singular sources led to many attempts to embed the space of distributions into an algebra, of course violating in a minimal way, one or more of the conditions of Schwartz result. They may be broadly classified into sequential and complex analysis methods and the type of products that are defined are either *intrinsic*, in the sense that the product of two distributions is still a distribution, or products which may lead to generalized functions different from distributions. For reviews see [23] and

[1]. Colombeau's algebra of generalized functions, which is of the second type, has found a widespread use in the applications. It has been applied to singular shock problems [24, 25], to symmetric hyperbolic systems with discontinuous coefficients [26], to equations with distributions as initial conditions [27], to generalized stochastic processes [28, 29], to general relativity [30, 31], etc.

Colombeau's approach [30, 32–34], in line with the sequential approach to distribution theory [35], considers a distribution as an equivalence class of weakly converging sequences of smooth functions and choosing the appropriate quotient constructs a differential algebra $\mathscr{A}$ with a product $\circ$ that satisfies the desired conditions except that instead of having $C(\mathbb{R})$ as a subalgebra it is $\circ|_{C^\infty \times C^\infty}$ that corresponds to the usual pointwise product of smooth functions.

Given some $\Phi \in \mathscr{D}(\mathbb{R}^n)$ with integral one, a family of functions

$$\Phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\Phi\left(\frac{x}{\varepsilon}\right) \tag{23}$$

has the property $\Phi_\varepsilon \to \delta$ in $\mathscr{D}'$ as $\varepsilon \to 0$. Φ is called a *molifier*. Convolution of $f \in C^\infty(\mathbb{R}^n)$ with Φ_ε yields a family

$$f_\varepsilon(x) = \frac{1}{\varepsilon^n}\int f(y)\Phi\left(\frac{y-x}{\varepsilon}\right) d^n y. \tag{24}$$

of smooth functions that converge to f in $\mathscr{D}'$ as ε tends to zero. Using a Taylor series expansion to compare the difference between two such sequences they are said to be equivalent if they differ by a *negligible function*. To define a differential algebra $\mathscr{G}$ as the quotient by negligible functions one needs to restrict the set of functions to *moderate functions*. The canonical choice introduces a grading on the space of molifiers

$$\begin{aligned}
\mathscr{A}_0 &:= \{\Phi \in \mathscr{D}(\mathbb{R}^n) \,:\, \int \Phi(x)\,dx = 1\} \\
\mathscr{A}_q &:= \{\Phi \in \mathscr{A}_0 \,:\, \int \Phi(x)x^\alpha\,dx = 0\,,\ 1 \le |\alpha| \le q\} \quad (q \in \mathbb{N})
\end{aligned} \tag{25}$$

taking the basic function space to be

$$\mathscr{E}^e := \{f :\ \mathscr{A}_0 \times \mathbb{R}^n \to \mathbb{R}^n,\ f \text{ smooth }\} \tag{26}$$

and defines as *moderate functions*

$$\begin{aligned}
\mathscr{E}^e_M(\mathbb{R}^n) := \ \{f \in \mathscr{E}^e : \forall K \subset\subset \mathbb{R}^n\ \forall \alpha \in \mathbb{N}_0^n\ \exists p \in \mathbb{N}_0\ \forall \Phi \in \mathscr{A}_p :\\
\sup_{x\in K} |D^\alpha f(\Phi_\varepsilon, x)| = O(\varepsilon^{-p}) \text{ as } \varepsilon \to 0\}
\end{aligned} \tag{27}$$

and *negligible functions*

$$\mathcal{N}^e(\mathbb{R}^n) := \{f \in \mathcal{E}^e(\Omega) : \forall K \subset\subset \mathbb{R}^n \, \forall \alpha \in \mathbb{N}_0^n \, \forall p \in \mathbb{N}_0 \, \exists q \, \forall \Phi \in \mathcal{A}_q : \sup_{x \in K} |D^\alpha f(\Phi_\varepsilon, x)| = O(\varepsilon^p) \text{ as } \varepsilon \to 0\} \tag{28}$$

Then the algebra is

$$\mathcal{G}^e(\mathbb{R}^n) := \mathcal{E}_M^e(\mathbb{R}^n) \,/\, \mathcal{N}^e(\mathbb{R}^n) \,. \tag{29}$$

with the distributions being embedded into $\mathcal{G}^e$ by convolution with the molifiers

$$\iota(T) = [T * \Phi] \tag{30}$$

$\mathcal{G}^e$ is a commutative differential algebra where $\mathcal{D}'$ is embedded as a linear space but not as an algebra. The results of multiplication in this algebra may frequently be interpreted in terms of distributions by using the concept of *association*. A generalized function f is said to be associated to a distribution $T \in \mathcal{D}'$ if for one (hence any) representative $\{f_\varepsilon\}$ we have

$$\forall \phi \in \mathcal{D}, \quad \lim_{\varepsilon \to 0} \int f_\varepsilon(x) \phi(x) d^n x = \langle T, \phi \rangle \tag{31}$$

Not all elements of $\mathcal{G}^e$ are associated to distributions. Association is an equivalence relation which respects addition and differentiation. It also respects multiplication by smooth functions but by the Schwartz impossibility result cannot respect multiplication in general.

For hyperbolic systems with rapidly growing nonlinear terms or with time-dependent nonregular coefficient the need arises of going beyond the nonlinear theory of generalized functions to a nonlinear theory of ultradistributions. After several other attempts, in a recent paper, Debrouwere, Vernaeve and Vindas have achieved an embedding of ultradistributions into differential algebras [36] of the same type as Colombeau algebra.

When applying these differential algebras to physical problems and in addition to the fact that many potentially interesting products in the Colombeau algebra cannot be associated to distributions, there is a problem of interpretation of the results because of the departure from the notion of pointwise multiplication that is behind the derivation of the physical equations. Instead the multiplication is a kind of tensor product multiplication. For example, if a weak solution of an equation behaves locally like a delta and if the equation has a term u^2 one has locally a δ^2, but in Colombeau-type algebras deltas are multiplied as

$$(\delta, \varphi)\,(\delta, \varphi)$$

which, even in a sequential approach, is very far from a pointwise multiplication and leads to a nonlinear functional

$$\left(\delta^2, \varphi\right) = \varphi^2(0)$$

therefore not a distribution.

3.2 Superprocesses and Nonlinear Differential Equations

The extension of superprocess from measures to ultradistributions, discussed in Sect. 2 allows to deal with a large class of nonlinear differential equations. When solutions are constructed by superprocesses, nonlinear terms do not raise any special problem. Let u be the solution of some equation. Existence of $n-$powers of u in the equation means that there is a splitting of the stochastic path into n paths, a derivative means a transition from $\delta^{(n)}$ to $\delta^{(n+1)}$, etc. It all boils down to the choice of the appropriate branching function and rescaling limit. No nonlinear operations on distributions are required. In the end the boundary process, the ultradistribution X_Q, need not have smooth properties. The only limitation to insure existence of $\langle f, X_Q \rangle$ is a sufficiently smooth boundary condition.

How superprocesses on ultradistributions provide solutions to nonlinear differential equations, which cannot be obtained by superprocesses on measures and avoid any explicit use of the nonlinear theory of generalized functions, is illustrated by the following results, proved in [18].

Proposition 2 *The superprocess with branching function*

$$\varphi(0, x; z) = p_1 e^{\partial_x \log z} + p_2 e^{-\partial_x \log z} + p_3 z^2$$

provides a solution to the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - 2u^2 - \frac{1}{2}(\partial_x u)^2$$

whenever the boundary function $u|_{\partial Q}$ *satisfies the condition of Proposition 1.*

Proposition 3 *The superprocess associated to the branching function*

$$\varphi(0, x; z) = p_1 z^2 + p_2 \frac{1}{z}$$

provides a solution to the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + u^3$$

whenever the boundary function $u|_{\partial Q}$ *satisfies the condition of Proposition 1.*

Here the construction of solutions of nonlinear differential equations by stochastic processes (*stochastic solutions*) has been discussed using superprocesses. There is another method, which has been called the McKean method, which may also deal with nonlinear terms and derivative interactions [37, 38]. Refer to [16] for a comparison of the two methods. In general the superprocess method seems more appropriate for Dirichlet boundary conditions and McKean's for Cauchy conditions. A limitation of this last method, in some cases, is that only finite time solutions may be constructed.

References

1. Hoskins, R.F., Sousa Pinto, J.: Theories of Generalized Functions: Distributions, Ultradistributions and Other Generalized Functions, 2nd edn. Woodhead Publ, Philadelphia (2011)
2. Björck, G.: Linear partial differential operators and generalized distributions. Ark. Mat. **6**, 351–407 (1966)
3. Roumieu, C.: Sur quelques extensions de la notion de distribution. Ann. Sci. École Norm. Sup. Sér. **3**(77), 41–121 (1960)
4. Pilipović, S.: Beurling-Gevrey tempered ultradistributions as boundary values. Portugaliae Math. **48**, 483–504 (1991)
5. Komatsu, H.: Ultradistributions I. Structure theorems and a characterization. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **20**, 25–105 (1973)
6. Debrouwere, A., Vindas, J.: On the non-triviality of certain spaces of analytic functions. Hyperfunctions and ultrahyperfunctions of fast growth. arXiv:1608.07859
7. Sebastião e Silva, J.: Les fonctions analytiques comme ultra-distributions dans le calcul opérationnel. Math. Annalen **136**, 58–96 (1958)
8. Sebastião e Silva, J.: Les séries de multip ôles des physiciens et la théorie des ultradistributions. Math. Annalen **174**, 109–142 (1967)
9. Sebastião e Silva, J.: Integrals and order of distributions. Inst. Gulbekian de Ciencia, Lisboa **35** (1964)
10. Estrada, R.: The Cesàro behavior of distributions. Proc. R. Soc. Lond. A **454**, 2425–2443 (1998)
11. Hasumi, M.: Note on the n-dimensional tempered ultradistributions. Tôhoku Math. J. **13**, 94–104 (1960)
12. Dynkin, E.B.: Diffusions, Superdiffusions and Partial Differential Equations. AMS Colloquium Pubs, Providence (2002)
13. Dynkin, E.B.: Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations. AMS, Providence (2004)
14. Li, Z.: Measure-Valued Branching Markov Processes. Springer, Berlin (2011)
15. Stein, E.M., Shakarchi, R.: Princeton Lectures in Analysis IV, Functional Analysis: Introduction to Further Topics in Analysis. Princeton Universiyt Press, Princeton (2011)
16. Vilela Mendes, R.: Stochastic solutions of nonlinear PDE's: McKean versus superprocesses. In: Leoncini, X., Leonetti, M. (eds.) Chaos, Complexity and Transport, pp. 143–154. World Scientific (2012)
17. Vilela Mendes, R.: Stochastic solutions of nonlinear pde's and an extension of superprocesses. In: Stochastic and Infinite dimensional Analysis, Trends in Math., pp. 243–261. Birkhäuser (2016)
18. Vilela Mendes, R.: Superprocesses on ultradistributions. Stochastics (2016). doi:10.1080/17442508.2016.1269768
19. Dynkin, E.B.: Branching exit Markov systems snd superprocesses. Ann. Probab. **29**, 1833–1858 (2001)

20. Fitzsimmons, P.J.: Construction and regularity of measure-valued Markov branching processes. Isr. J. Math. **64**, 337–361 (1988)
21. Fitzsimmons, P.J.: On the Martingale problem for measure-valued Markov branching processes. In: Seminar on Stochastic Processes 1991, Progress in Probability, vol. 29, pp. 39–51. Springer, Boston (1992)
22. Schwartz, L.: Sur l'impossibilité de la multiplication des distributions. C.R. Acad. Sci. Paris **239**, 847–848 (1954)
23. Li, C.K.: A review on the products of distributions. In: Kenan Tas et al. (ed.) Mathematical Methods in Engineering. Springer (2007)
24. Keyfitz, B.L., Kranzer, H.C.: A strictly hyperbolic system of conservation laws admitting singular shocks. In: Nonlinear Evolution Equations that Change Type, pp. 107–125. Springer (1990)
25. Le Floch, Ph.: Shock waves for nonlinear hyperbolic systems in nonconservative form. Inst. for Math. and Applications - Univ. Minnesotta, preprint , vol. 593 (1989)
26. Lafon, F., Oberguggenberger, M.: Generalized solutions to symmetric hyperbolic systems with discontinuous coefficients: the multidimensional case. J. Math. Anal. Appl. **160**, 93–106 (1991)
27. Langlais, M.: Generalized functions solutions of monotone and semilinear parabolic equations. Monatshefte Math. **110**, 117–136 (1990)
28. Mirkov, R., Pilipovic, S., Selesi, D.: Generalized stochastic processes in algebras of generalized functions. J. Math. Anal. Appl. **353**, 260–270 (2009)
29. Oberguggenberger, M.: Generalized functions and stochastic processes. Prog. Probab. **36**, 215–229 (1995)
30. Grosser, M., Kunzinger, M., Oberguggenberger, M., Steinbauer, R.: Geometric Theory of Generalized Functions Mathematics and its Applications, vol. 537. Kluwer (2001)
31. Steinbauer, R., Vickers, J.A.: The use of generalized functions and distributions in general relativity. Class. Q. Grav. **23**, R91–R114 (2006)
32. Colombeau, J.F.: New Generalized Functions and Multiplication of Distributions. North-Holland-Elsevier (1984)
33. Colombeau, J.F.: Nonlinear generalized functions: their origin, some developments and recent advances. São Paulo J. Math. Sci. **7**, 201–239 (2013)
34. Aragona, J., Juriaans, S.O., Colombeau, J.F.: Locally convex topological algebras of generalized functions: compactness and nuclearity in a nonlinear context. Trans. Am. Math. Soc. **367**, 5399–5414 (2015)
35. Mikusinski, J.: Sur la méthode de gé neralisation de M. Laurent Schwartz sur la convergence faible. Fundam. Math. **35**, 235–239 (1948)
36. Debrouwere, A., Vernaeve, H., Vindas, J.: Optimal embeddings of ultradistributions into differential algebras. arXiv:1601.03512
37. Vilela Mendes, R.: Stochastic solutions of some nonlinear partial differential equations. Stochastics **81**, 279–297 (2009)
38. Vilela Mendes, R.: Poisson-Vlasov in a strong magnetic field: A stochastic solution approach. J. Math. Phys. **51**, 043101 (2010)

A New Mathematical Model for Environmental Monitoring and Assessment

Roberto Monaco and Ana Jacinta Soares

Abstract In this paper we are concerned with a quantitative method of Landscape Ecology. More in details we consider an environmental system distributed in landscape units (ecological sectors) and we propose a new mathematical model in order to implement a method for the evaluation of the ecological state of the system under investigation. After having performed a stability analysis of the model, we apply the proposed procedure first by considering separately each landscape unit and then extending our investigation to the system as a whole, by taking into account the connections between all the landscape units themselves. Our investigation includes some numerical computations that were performed for a Northern district of the Turin Province, using an approximation procedure that should avoid stiffness problems.

Keywords Landscape ecology · Mathematical modeling
Environmental dynamics · Stability analysis · Numerical computations

1 Introduction

For a quantitative evaluation of the ecological state of an environmental system an important contribution is given in the book [10]. Other important contributions to this matter can be found as well in the paper [19] and in the book [5] where for the first time the use of the so-called *ecological graph* is proposed. For a quantitative description of a territory relevant indicators have been also proposed in the references [15, 16] where ideas coming from the mathematical theory of communication [17] and from conceptual models [1] have been transposed to Landscape Ecology.

R. Monaco
DIST, Politecnico di Torino, Viale Mattioli 39, 10125 Torino, Italy
e-mail: roberto.monaco@polito.it

A.J. Soares (✉)
CMAT, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
e-mail: ajsoares@math.uminho.pt

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_13

In the quoted bibliography, as we shall see, the environment is distributed in different *landscape units* (often mentioned in what follows as *ecological sectors*) which can be more or less connected by flows of material and biological energy. In this sense a relevant variable to describe the territory is certainly the last quantity which can be coupled to the consistency of green areas of high ecological quality.

In the past years some mathematical models useful for a quantitative monitoring of environments have been proposed in a couple of papers [7, 8]. In particular in the former a mathematical model, with only two state equations that consider the environment as a whole, has been deduced, assuming as state variables the bio-energy and the percentage of green areas in the territory under investigation.

In the latter such a model has been studied in terms of its stability analysis, finding also bifurcations and discussing on its qualitative properties. Moreover in paper [8], as a future perspective, a new model including state variables at the level of landscape units has been suggested. These models, on the mathematical ground, are represented by an autonomous system of evolution ordinary differential equations, whose equilibrium solutions express the future scenarios of the environment itself.

Starting from this last idea in the present paper we present a model where each landscape unit is represented by two time-dependent variables, namely the extent of green areas of high ecological quality and a suitable function depending on the biological energy per year. Such a new version of the model is capable to identify the characters of the territory at a more detailed level, so that the territory itself results to be more readable.

For a first moment, in Sect. 2, we are concerned with the model stability analysis, considering separately the landscape units. We then apply, in Sect. 3, such an analysis to a Northern district of the Turin Province. This district has been previously studied by De Palma in her master degree thesis [4], deriving all the relevant indicators of this environmental system. Moreover, in Sects. 4 and 5, we complete the environment analysis by coupling all together the landscape units, showing that, in the present case study, connectivity plays a crucial role. For this last analysis the model becomes rather cumbersome and may present some instabilities since it includes 48 coupled ordinary differential equations. For this reason in Sect. 4 we propose an explicit approximation procedure, discussed in paper [9] starting from the methods presented in the book [11], and bibliography therein cited. Finally, conclusions and some ideas for future work are presented in Sect. 6.

2 The Mathematical Modeling for Landscape Units: Equilibrium Solutions and Stability

An environmental system is an isolated system that may be distributed in n landscape units (LU) divided by natural or anthropological barriers (roads, motorways, railways, buildings, industrial infrastructures, rivers, hill ridges and so on). At the same time a LU is formed by several biotopes which are patches characterized by

an uniform land cover. Such a representation of an environment can be given by the Geographic Information System (GIS) [13] and its ecological indicators [3] can be deduced by the ecological graph (see its construction in the book [5] and in the paper [7]).

The nature of each biotope is identified by the value of its bio-potential territorial capacity (BTC) index [10]. In what follows, such an index, measuring the biological energy per year and per square meter, produced by the vegetation inside the biotope, will be indicated by B_{ji}, $j = 1, \ldots, q_i$, being i a subscript which indicates that the j-th biotope belongs to the i-th LU, $i = 1, \ldots, n$. The variable B_{ji} assumes values [7, 10] in the range $[0, B_{max}]$ where $B_{max} = 6.5\ Mcal/(m^2 \cdot year)$ at the European latitudes and corresponds to oak woods.

Moreover the BTC index identifies five ecological classes distributed as follows

$$\mathscr{C}_1 = [0, 0.4),\ \mathscr{C}_2 = [0.4, 1.2],\ \mathscr{C}_3 = (1.2, 2.4],\ \mathscr{C}_4 = (2.4, 4.0],\ \mathscr{C}_5 = (4.0, 6.5].$$

The total value of BTC of each LU, in $Mcal/year$, is given by

$$B_i(t) = \sum_{j=1}^{q_i} B_{ji} s_{ji}, \tag{1}$$

where s_{ji} is the area of the biotope j. Accordingly the area of the corresponding LU is $S_i = \sum_{j=1}^{q_i} s_{ji}$.

To the BTC is often associated another quantity that can be interpreted as a generalized function of BTC (GBTC) defined by the following formula [5]

$$M_i(t) = (1 + K_i) B_i(t), \tag{2}$$

where the constant parameters $K_i \in [0, 1]$ depend upon the physical and morphological features of the LU in such a way that the BTC itself is incremented. Such an increment takes into account the capacity of the LU to transmit energy to the neighboring LUs. In this paper K_i will be expressed in terms of three other parameters

$$K_i = (K_i^{sh} + K_i^{ec} + K_i^{pe})/3\ ,$$

where $K_i^{sh}, K_i^{ec}, K_i^{pe} \in [0, 1]$ are, respectively, the Shannon landscape diversity parameter (evenness [15, 16]), the ecotonal parameter [3] and the LU-border permeability parameter [7]. They are defined as follows:

- the Shannon parameter

$$K_i^{sh} = \left(\sum_{\ell=1}^{5} \frac{n_{\ell i}}{5} \log \frac{n_{\ell i}}{5} \right) / \log \frac{1}{5}\ ,$$

where $n_{\ell i}$ is the number of biotopes belonging to the ecological class $\mathscr{C}_\ell$, $\ell = 1, \ldots, 5$, assuming that $n_{\ell i} \log n_{\ell i} = 0$ if $n_{\ell i} = 0$;
• the ecotonal parameter (i.e. the length of borders between biotopes)

$$K_i^{ec} = 1 - P_i \Big/ \sum_{j=1}^{q_i} P_{ji},$$

where P_i is the perimeter of the i-th LU and P_{ji} are the perimeters of all the biotopes except those belonging to the ecological class $\mathscr{C}_1$;
• the LU-border permeability parameter

$$K_i^{pe} = \frac{1}{P_i} \sum_{r=1}^{s} L_i^r p^r,$$

where L_i^r is the length of the portion r of the border, divided in s parts, and $p^r \in [0, 1]$ is the permeability index, with $p = 0$ for an impermeable barrier, and $p = 1$ for a complete permeable barrier (see [7, 10]).

According to previous versions of the mathematical model already indicated in the Introduction [7, 8], the state variables for each LU of the present model are two, precisely the area $V_i(t)$ of the biotopes belonging to the ecological classes $\mathscr{C}_4$ and $\mathscr{C}_5$ (high ecological quality of green) and the GBTC $M_i(t)$. The right-hand-side of the evolution equations on $V_i(t)$ and $M_i(t)$ consists in a gain term of logistic type and a loss term accounting for environmental impact. The equations have the following form

$$V_i'(t) = b_i(M_i) \left(1 - \frac{V_i(t)}{S_i}\right) V_i(t) - h_i U_i V_i(t), \tag{3}$$

$$M_i'(t) = c_i \left(1 - \frac{M_i(t)}{M_i^{max}}\right) M_i(t) - r_i \left(1 - \frac{V_i(t)}{S_i}\right) M_i(t). \tag{4}$$

In Eq. (3) the coefficient b_i will be expressed, as we shall see, in terms of the GBTC, whereas the other parameters h_i and U_i are, respectively, given by the ratio between the sum of the perimeters of the built-up areas and the total perimeter of the LU, and by the ratio between the sum of the built-up areas and the total area S_i of the LU. Therefore, these parameters, assumed as constant, can be considered a measure, respectively, of the dispersion and of the intensity of edification inside the LU. According to its definition the parameter h_i can assume values greater than one (values greater than one mean that edification dispersion in the LU is significantly remarkable); on the other hand U_i is ranging in [0, 1].

In Eq. (4) the coefficient c_i is the connectivity index between the i-th LU and its neighbors, whereas the other parameter r_i is defined as the ratio between the surface area of the impermeable barriers present in the LU and the total area S_i of the LU itself. Both parameters are supposed to range in [0, 1] and for this preliminary

analysis the indexes c_i are assumed to be constant. Let us now deal with the definition of the parameter $b_i(M_i)$, i.e.

$$b_i := \frac{B_i}{B_i^{max}}, \qquad B_i^{max} = B_{max} S_i.$$

Thus b_i expresses the production percentage of BTC with respect to the maximum value that each LU can produce, assuming that all its biotopes have BTC index equal to B_{max}. If now we define the normalized GBTC

$$m_i := \frac{M_i}{M_i^{max}}, \qquad M_i^{max} = 2B_i^{max},$$

then we have

$$m_i = \frac{(1 + K_i)B_i}{2B_i^{max}} = \frac{1 + K_i}{2} b_i$$

and therefore

$$b_i = \frac{2}{1 + K_i} m_i := a_i m_i, \qquad a_i \in [1, 2]. \tag{5}$$

If now we normalize V_i as well, defining $v_i := V_i / S_i$, both state variables v_i and m_i range in [0, 1]. Thus if we divide Eqs. (3) and (4), respectively, by S_i and M_i^{max}, the model equations assume the following final form

$$v_i'(t) = a_i m_i(t)[1 - v_i(t)]v_i(t) - h_i U_i v_i(t), \tag{6}$$

$$m_i'(t) = c_i[1 - m_i(t)]m_i(t) - r_i[1 - v_i(t)]m_i(t). \tag{7}$$

Moreover to these equations we join the initial data ($t = 0$)

$$v_i(0) = v_{i0}, \qquad m_i(0) = m_{i0}, \tag{8}$$

which must be determined directly from the GIS maps of the environment under investigation.

Next step consists in finding the equilibrium solutions [12] of the system (6) and (7). Solving the algebraic equations

$$v_i[a_i m_i(1 - v_i) - h_i U_i] = 0, \qquad m_i[c_i(1 - m_i) - r_i(1 - v_i)] = 0, \tag{9}$$

one obtains the following equilibria.

• The first is given by

$$\left(v_i^{(1)},\ m_i^{(1)}\right) = (0,\ 0)\,, \tag{10}$$

which corresponds to a scenario where the environment tends to lose substantially its ecological quality since it is characterized by a strong fragmentation;

• the second is expressed by

$$\left(v_i^{(2)},\ m_i^{(2)}\right) = \left(0,\ \frac{c_i - r_i}{c_i}\right), \tag{11}$$

corresponding to a scenario of weak ecological quality characterized by a moderate level of bio-energy (such a scenario is typical of a territory with a predominant agricultural production); the admissibility condition of such solution results to be $c_i > r_i$;
• the third and fourth equilibria are those of coexistence, showing a good level of bio-energy production together with a high ecological quality of green areas; they are given by

$$\left(v_i^{(3)},\ m_i^{(3)}\right) = \left(\frac{2r_i - c_i - D_i}{2r_i}\ ,\ \frac{2h_i U_i r_i}{a_i(c_i + D_i)}\right) \tag{12}$$

$$\left(v_i^{(4)},\ m_i^{(4)}\right) = \left(\frac{2r_i - c_i + D_i}{2r_i}\ ,\ \frac{2h_i U_i r_i}{a_i(c_i - D_i)}\right) \tag{13}$$

where

$$D_i = \sqrt{c_i(a_i c_i - 4h_i U_i r_i)/a_i}\ .$$

The existence of these solutions requires that

$$c_i > \frac{4h_i U_i r_i}{a_i}\ . \tag{14}$$

Moreover the third solution exists if

$$\frac{1}{2}(c_i + D_i) < r_i < \frac{1}{2}(c_i + D_i)\frac{a_i}{h_i U_i}\ , \tag{15}$$

whereas the fourth requires that

$$c_i > D_i\ , \qquad \frac{1}{2}(c_i - D_i) < r_i < \frac{1}{2}(c_i - D_i)\frac{a_i}{h_i U_i}\ . \tag{16}$$

We are now concerned with asymptotic stability [12] of the previous equilibria. The Jacobian matrix joined to the system (6) and (7) is given by

$$J(v_i, m_i) = \begin{pmatrix} a_i m_i - 2a_i v_i m_i - h_i U_i & a_i(1 - v_i)v_i \\ r_i m_i & c_i - 2c_i m_i - r_i(1 - v_i) \end{pmatrix}. \tag{17}$$

• We have for the first equilibrium

$$J(v_i^{(1)}, m_i^{(1)}) = \begin{pmatrix} -h_i U_i & 0 \\ 0 & c_i - r_i \end{pmatrix}. \tag{18}$$

The eigenvalues of the matrix are $\lambda_1 = -h_i U_i < 0$ and $\lambda_2 = c_i - r_i$. Thus the equilibrium solution $(v_i^{(1)}, m_i^{(1)})$ is asymptotically stable if $c_i < r_i$; otherwise it is a saddle point. Moreover the stability of this solution implies that the equilibrium $(v_i^{(2)}, m_i^{(2)})$ does not exist (see the admissibility condition of the second equilibrium).

• For the second equilibrium we get

$$J(v_i^{(2)}, m_i^{(2)}) = \begin{pmatrix} \dfrac{a_i(c_i - r_i) - h_i U_i c_i}{c_i} & 0 \\ \dfrac{r_i(c_i - r_i)}{c_i} & -c_i + r_i \end{pmatrix}, \tag{19}$$

whose eigenvalues are $\lambda_1 = \dfrac{a_i(c_i - r_i) - h_i U_i c_i}{c_i}$ and $\lambda_2 = -c_i + r_i$. Existence of this equilibrium implies that the eigenvalue λ_2 is always negative. Thus the asymptotic stability of $(v_i^{(2)}, m_i^{(2)})$ requires that $a_i c_i < a_i r_i + h_i U_i c_i$.

• Moreover for the third equilibrium we have

$$J(v_i^{(3)}, m_i^{(3)}) = \begin{pmatrix} \dfrac{h_i U_i (c_i + D_i - 2r_i)}{c_i + D_i} & \dfrac{a_i(c_i + D_i)(2r_i - c_i - D_i)}{4r_i^2} \\ \dfrac{2h_i U_i r_i^2}{a_i(c_i + D_i)} & \dfrac{a_i(c_i^2 - D_i^2) - 8h_i U_i c_i r_i}{2a_i(c_i + D_i)} \end{pmatrix}. \tag{20}$$

If we write the characteristic equation of the Jacobian in the form $\lambda^2 + A_1\lambda + B_1 = 0$, after simple computations we get

$$A_1 = -\frac{2a_i h_i U_i (c_i + D_i - 2r_i) + a_i(c_i^2 - D_i^2) - 8h_i U_i c_i r_i}{2a_i(c_i + D_i)} \tag{21}$$

$$B_1 = \frac{h_i U_i (c_i + D_i - 2r_i)[a_i c_i (c_i + D_i) - 4h_i U_i c_i r_i]}{a_i (c_i + D_i)^2}. \tag{22}$$

Therefore the stability condition for the third equilibrium requires that $A_1 > 0$ and $B_1 > 0$. Conversely, if $A_1 > 0$ and $B_1 < 0$ or if $A_1 < 0$ and $B_1 < 0$, we have that such an equilibrium is a saddle point, whereas if $A_1 < 0$ and $B_1 > 0$ we have instability. In addition, in the case of stability, if $A_1^2 - 4B_1 > 0$ the equilibrium is a node, whereas $A_1^2 - 4B_1 < 0$ corresponds to a focus. In cases that $A_1 = 0$ or $B_1 = 0$, the asymptotic stability is not assured. Nevertheless, in practice, since A_1 and B_1 depend on territorial indexes of different nature, the vanishing of these quantities is strongly unlikely and not consistent with the real state of the environment.

• Finally for the last equilibrium $(v_i^{(4)}, m_i^{(4)})$ the Jacobian has the form

$$J(v_i^{(4)}, m_i^{(4)}) = \begin{pmatrix} \dfrac{h_i U_i (c_i - D_i - 2r_i)}{c_i - D_i} & \dfrac{a_i (c_i - D_i)(2r_i - c_i + D_i)}{4r_i^2} \\ \dfrac{2h_i U_i r_i^2}{a_i (c_i - D_i)} & \dfrac{a_i (c_i^2 - D_i^2) - 8h_i U_i c_i r_i}{2a_i (c_i - D_i)} \end{pmatrix}, \quad (23)$$

and the coefficients of its characteristic equation $\lambda^2 + A_2\lambda + B_2 = 0$ are given by

$$A_2 = -\frac{2a_i h_i U_i (c_i - D_i - 2r_i) + a_i (c_i^2 - D_i^2) - 8h_i U_i c_i r_i}{2a_i (c_i - D_i)} \quad (24)$$

$$B_2 = \frac{h_i U_i (c_i - D_i - 2r_i)[a_i c_i (c_i - D_i) - 4h_i U_i c_i r_i]}{a_i (c_i - D_i)^2}. \quad (25)$$

The stability or instability discussion on this equilibrium is just like that of the third, substituting only A_1 and B_1 with A_2 and B_2.

We underline that system (6) and (7) is cooperative of Lotka–Volterra type, and the solution (v_i,m_i) lies in the square $[0, 1] \times [0, 1]$. Thus the system possesses one stable equilibrium at least (see [18]).

In Sect. 3 we will apply the model studied in this section to an environmental system situated in the Northern district of the Province of Turin (Italy). According to the analysis carried on in this section it is evident that the various scenarios admitted by the model and the asymptotic trend to a stable equilibrium solution crucially depends on the values of the model parameters. Thus, it is necessary an accurate determination of these parameters in the environment under investigation through the GIS data. This will be carried out in Sect. 3 where we show such an accurate determination of parameters and initial data of the state of each LU. In particular in that section we propose as well some phase diagrams (Figs. 1, 2, 3, 4, 5, 6) of the system variables, where the sensitivity of the trend from the initial state to the final one can be observed.

3 Stability Analysis for Each LU of the Environment

In order to check the theoretical analysis presented in the previous section, we consider the afore-mentioned environment of the Turin Province that has been studied in the thesis [4]. Such a system has been divided into 24 LUs, corresponding to an area of several municipalities placed around the city of Cirié.

First of all let us mention that the stability analysis carried on all the 24 LUs show that the third equilibrium, $(v_i^{(3)}, m_i^{(3)})$, never exists since conditions (15) are not satisfied.

Moreover the analysis shows that 13 LUs have only the stable equilibrium $(v_i^{(1)}, m_i^{(1)})$, see LUs 2, 3, 4, 11, 12, 13, 14, 15, 16, 19, 21, 22, 23 in Table 1. This means that these LUs have a bad ecological state and present a strong fragmentation

Table 1 Data of the environmental system and indication of the stable equilibria for each LU

LU	v_{i0}	m_{i0}	a_i	h_i	U_i	c_i	r_i	S_i	P_i	nodes
1	0.30	0.19	1.13	1.83	0.09	0.12	0.03	19156648	25248	(4)
2	0.13	0.05	1.36	1.39	0.60	0.07	0.20	2260943	11109	(1)
3	0.20	0.04	1.43	0.81	0.41	0.04	0.14	764299	3791	(1)
4	0.25	0.11	1.27	1.45	0.25	0.07	0.08	2285714	7271	(1)
5	0.26	0.13	1.20	2.51	0.23	0.08	0.07	17004435	26335	(2)
6	0.28	0.27	1.09	1.36	0.17	0.20	0.06	13162369	22717	(4)
7	0.32	0.53	1.06	1.17	0.03	0.26	0.01	29896496	33085	(4)
8	0.26	0.20	1.12	1.72	0.23	0.30	0.08	10139270	20838	(4)
9	0.30	0.20	1.21	0.60	0.11	0.16	0.04	6584277	24162	(4)
10	0.21	0.09	1.43	1.20	0.36	0.17	0.12	837269	5113	(2), (4)
11	0.18	0.06	1.44	1.48	0.47	0.04	0.16	991018	5007	(1)
12	0.23	0.07	1.69	1.39	0.32	0.09	0.11	1395933	5704	(1)
13	0.26	0.10	1.36	1.81	0.23	0.07	0.08	2431435	7903	(1)
14	0.20	0.08	1.39	1.05	0.40	0.04	0.13	1452734	5483	(1)
15	0.11	0.04	1.62	1.42	0.66	0.04	0.22	1072430	5355	(1)
16	0.17	0.06	1.41	1.09	0.49	0.05	0.16	3416393	10922	(1)
17	0.26	0.13	1.22	1.58	0.21	0.14	0.07	6369795	12599	(4)
18	0.31	0.34	1.06	1.97	0.07	0.17	0.02	69754645	60482	(4)
19	0.13	0.17	1.24	1.88	0.61	0.17	0.20	4589299	18604	(1)
20	0.30	0.45	1.00	2.20	0.09	0.19	0.03	42953048	38826	(4)
21	0.09	0.02	2.00	1.06	0.74	0.06	0.25	459102	2887	(1)
22	0.05	0.01	2.00	1.29	0.86	0.05	0.29	302009	2792	(1)
23	0.15	0.05	1.52	1.27	0.54	0.05	0.18	1059134	4720	(1)
24	0.26	0.11	1.24	1.91	0.22	0.08	0.07	16437048	18458	(2)

due to the presence of a significant edification sprawl. Therefore both bio-energy and extension of areas with a BTC index in classes $\mathscr{C}_4$ and $\mathscr{C}_5$ present a decrement of their values which asymptotically tend to zero. Such a result is somehow in accordance with the analysis carried out, with other methods, in the thesis mentioned above. On the other hand 11 LUs exhibit stable equilibria different from $(v_i^{(1)}, m_i^{(1)})$. For these LUs the first equilibrium is always a saddle point. In Table 1, deduced by the GIS map, we report the relevant data of the model (the area and the perimeter of the LUs are indicated in m^2 and m, respectively). In the last column, for each LU, we indicate which of the four equilibria results to be a stable node.

Let us note that the connectivity indexes c_i, which in the following Sect. 4 will be defined as functions of $M_i(t)$ and $M_k(t)$, here are assumed constant with their values recovered by the GIS data. They are computed here by formulas (26)–(28) of Sect. 4 setting $M_i = m_{i0}M_i^{max}$ and $M_k = m_{k0}M_k^{max}$, for all $i, k = 1, \ldots, 24$.

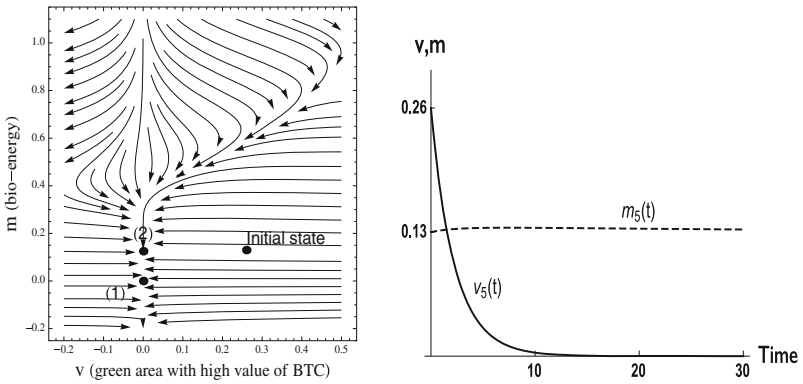


Fig. 1 Representation of LU 5 (scenario typical of agricultural areas). *Left* – Phase diagram: state $(v_5^{(1)}, m_5^{(1)})$ is a saddle point and state $(v_5^{(2)}, m_5^{(2)})$ is a stable node. *Right* – Time evolution of v_5 and m_5 towards the equilibrium $(v_5^{(2)}, m_5^{(2)})$

From the stability analysis it results that LUs 1, 6, 7, 8, 9, 17, 18, 20 admit the equilibrium $(v_i^{(2)}, m_i^{(2)})$ as a saddle point and the other $(v_i^{(4)}, m_i^{(4)})$ as a stable node. The diagrams in the phase plane show that the state of all these LUs, starting from the initial data v_{i0} and m_{i0} reported in Table 1, converges towards the fourth equilibrium which corresponds to a scenario of high ecological quality (see the examples reported below).

Conversely LUs 5 and 24 present as a unique stable node the point $(v_i^{(2)}, m_i^{(2)})$, since $(v_i^{(4)}, m_i^{(4)})$ does not satisfy the existence conditions of Eq. (16). Therefore the LUs show a trend towards a scenario typical of agricultural areas.

Finally LU 10 admits as stable nodes both equilibria $(v_i^{(2)}, m_i^{(2)})$ and $(v_i^{(4)}, m_i^{(4)})$. In this case the asymptotic behavior depends crucially on initial data. In particular with those of Table 1 the actual attractor results to be the second equilibrium, i.e. the one of agricultural scenario.

In the following Figs. 1, 2, 3, 4, 5 and 6, we show some representative behaviors through phase plane diagrams (left plots) and graphics of v_i and m_i versus time at an arbitrary scale (right plots). The simulations have been performed with the software *Mathematica*, version number 10.0.0.0.

In particular, Fig. 1 is referred to LU 5 that, as discussed previously, admits a unique stable node $(v_i^{(2)}, m_i^{(2)})$, as it can be seen in the phase plane diagram, with v_i tending to zero and m_i almost constant as indicated by the time-dependent plot.

Figure 2 refers to LU 8 and shows a trend to a scenario of high ecological quality since, after a transient time, the vegetation area v_i and the biological energy function m_i show an increasing behavior towards the equilibrium $(v_i^{(4)}, m_i^{(4)})$. This trend does not depend on the initial data v_{i0} and m_{i0}, in the sense that $(v_i^{(4)}, m_i^{(4)})$ represents the unique attractor of the LU, as shown in the phase diagram on the left plot, and the system asymptotically converges to such an attractor.

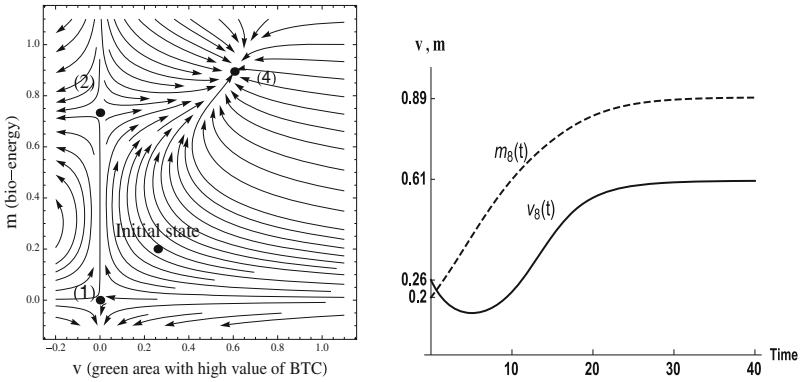


Fig. 2 Representation of LU 8 (scenario of high ecological quality). *Left* – Phase diagram: states $(v_8^{(1)}, m_8^{(1)})$ and $(v_8^{(2)}, m_8^{(2)})$ are saddle points and state $(v_8^{(4)}, m_8^{(4)})$ is a stable node. *Right* – Time evolution of v_8 and m_8 towards the equilibrium $(v_8^{(4)}, m_8^{(4)})$

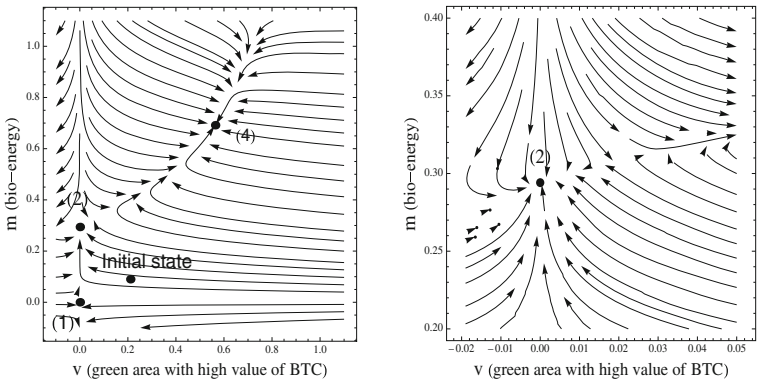


Fig. 3 Representation of LU 10. *Left* – Phase diagram: state $(v_{10}^{(1)}, m_{10}^{(1)})$ is a saddle point, states $(v_{10}^{(2)}, m_{10}^{(2)})$ and $(v_{10}^{(4)}, m_{10}^{(4)})$ are stable nodes. *Right* – Zoom of the phase diagram around the stable node $(v_{10}^{(2)}, m_{10}^{(2)})$ (disregard the negative part of the picture due to the choice of the scales determined by the software Mathematica)

Figures 3 and 4 refer to LU 10 and, as indicated in Table 1, such sector shows two stable equilibria. Accordingly, the phase diagram on the left plot of Fig. 3 shows the existence of two stable nodes $(v_i^{(2)}, m_i^{(2)})$ and $(v_i^{(4)}, m_i^{(4)})$. The fact that the equilibrium $(v_i^{(2)}, m_i^{(2)})$ is a stable node can be visualized with a zoom around such a point, as shown in the right frame of Fig. 3. For the initial data reported in Table 1, the left plot of Fig. 4 shows the time evolution of v_i and m_i towards the equilibrium $(v_i^{(2)}, m_i^{(2)})$, which is typical of agricultural areas. The basin of attraction of the equilibrium $(v_i^{(2)}, m_i^{(2)})$ is represented in the right plot of Fig. 4 by the grey region, showing that the initial state considered in Table 1 belongs to such basin. The case of LU 10 is very interesting from the dynamical point of view, since it represents a bistable situation.

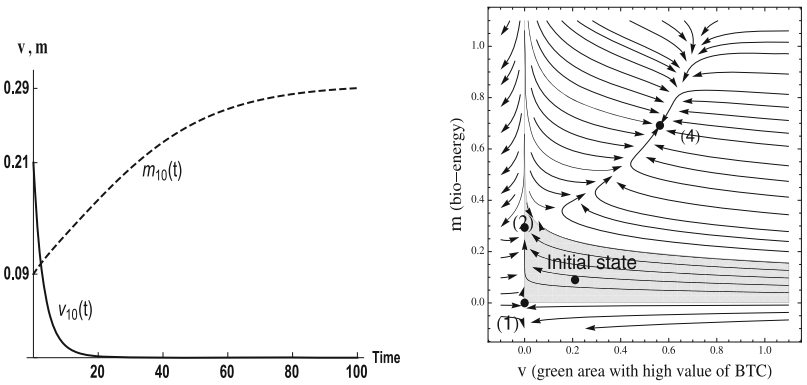


Fig. 4 Representation of LU 10. *Left* – Time evolution of v_{10} and m_{10} towards the equilibrium $(v_{10}^{(2)}, m_{10}^{(2)})$ with the initial data of Table 1. *Right* – Basin of attraction for the node $(v_{10}^{(2)}, m_{10}^{(2)})$ represented by the *grey* region

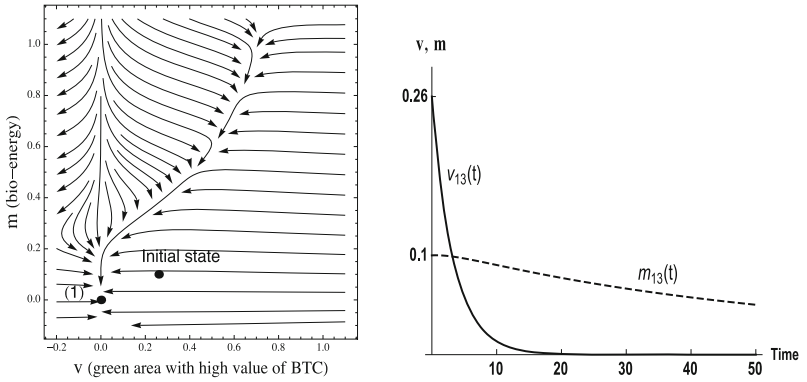


Fig. 5 Representation of LU 13 (scenario of ecological quality loss). *Left* – Phase diagram: state $(v_{13}^{(1)}, m_{13}^{(1)})$ is the unique equilibrium and it is given by a stable node. *Right* – Time evolution of v_{13} and m_{13} towards the equilibrium $(v_{13}^{(1)}, m_{13}^{(1)})$

Moreover, in Sect. 5 when the system will be treated as a whole, the behavior of LU 10 will change significantly.

Figure 5 shows the behavior of LU 13 that, as discussed before, is one of the ecological sectors presenting a strong fragmentation. Therefore the picture in the phase plane shows the existence only of the unique stable node $(v_i^{(1)}, m_i^{(1)})$ in accordance with the time-dependent plot.

Finally, Fig. 6 corresponds to LU 20 and shows a scenario of high ecological quality, similar to the one of LU 8 represented in Fig. 2, but with a stronger trend to an equilibrium of high ecological quality. In fact, for the initial data v_{i0} and m_{i0}, reported in Table 1 for this LU, the variable v_i presents a monotonic increase from

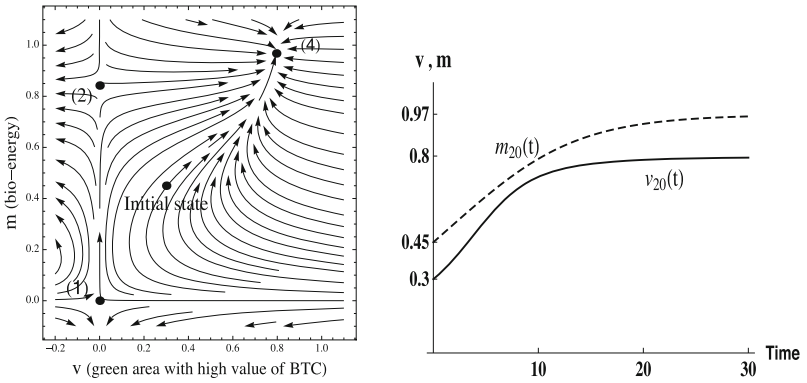


Fig. 6 Representation of LU 20 (scenario of high ecological quality). *Left* – Phase diagram: states $(v_{20}^{(1)}, m_{20}^{(1)})$ and $(v_{20}^{(2)}, m_{20}^{(2)})$ are saddle points and $(v_{20}^{(4)}, m_{20}^{(4)})$ is a stable node. *Right* – Time evolution of v_{20} and m_{20} towards the equilibrium $(v_{20}^{(4)}, m_{20}^{(4)})$

the initial state to the equilibrium $(v_i^{(4)}, m_i^{(4)})$, conversely to that non monotonic of LU 8.

Let us finally comment that we have also considered in our simulations the case of LU 24 which however presents exactly the same behavior as LU 5. Thus, we do not include here the plots of LU 24. Nevertheless, when studying in Sect. 5 the behavior of these LUs in the whole environmental system, our numerical simulations will take into account the correct connectivity to the neighboring sectors, founding that LU 5 and LU 24 exhibit a significantly different dynamics.

4 The Mathematical Model Extended to the Whole Environmental System

In this section we extend the mathematical model to the whole territory under investigation by coupling the equations of each LU with those of their neighbors. Such a coupling is determined by the connectivity indexes c_i which can be computed through the GBTC fluxes Φ_i between the i LUs and all their k neighbors. We get [8]

$$\Phi_i(t) = \sum_{k \in I_i} \frac{M_i(t) + M_k(t)}{2(P_i + P_k)} H_{ik}, \qquad H_{ik} = \sum_{r=1}^{s} L_{ik}^r p^r, \tag{26}$$

where L_{ik}^r is the length of the portion r of the border with a permeability p^r. Moreover P_i and P_k are, as already defined, the perimeters of the two LUs and where the sum is extended to the set I_i including all the neighbors of the i-th LU. In addition it results

$$\sum_{r=1}^{s} L_{ik}^{r} = L_{ik},$$

L_{ik} being the length of the border.

Taking into account that complete permeability implies $p_r = 1$, the corresponding maximum value of Φ_i is given by

$$\Phi_i^{max} = \sum_{k \in I_i} \frac{M_i^{max} + M_k^{max}}{2(P_i + P_k)} L_{ik}. \tag{27}$$

Finally the connectivity index of the i-th LU is defined by [9]

$$c_i(t) = \Phi_i(t)/\Phi_i^{max}, \tag{28}$$

so that $c_i(t) \in [0, 1]$ for all t ($c_i = 0$ corresponds evidently to no connectivity, and $c_i = 1$ to total connectivity).

Thus the equations of the system (6) and (7) are coupled through the coefficients $c_i(t)$ which depend on $M_k(t)$ and definitively on $m_k(t)$.

Because of the great number of LUs that an environment can possess, solving (6) and (7) through a numerical integrator may be costly and stiffness problems may arise. Morever in [9], the numerical integration of a system similar to that of Eqs. (6) and (7) has shown the presence of some instability due to the presence of such a large amount of equations. Thus, an approximation method has been there proposed in order to transform the system of ODEs in an algebraic closed hierarchy, evaluating as well its accuracy (see also [11]). The problem of solving Eqs. (6) and (7) through an algebraic hierarchy instead of a numerical integration allows the use of the model also by persons not acquainted with ODE integrators.

For this reason in the present paper we adopt such a method and we derive as follows such an algebraic hierarchy. If one assumes for a moment that the Eqs. (6) and (7) are completely uncoupled, meaning that the quantities m_i and v_i are constant, respectively, in Eq. (6) and in Eq. (7), and that the coefficient c_i is constant as well, then the system itself, starting from the initial data $v_i(t = 0) = v_{i0}$ and $m_i(t = 0) = m_{i0}$, has the following explicit solution, thanks to its classical logistic structure [12]

$$v_i(t) = \frac{a_i m_i - h_i U_i}{D_i^v \exp[-(a_i m_i - h_i U_i)t] + a_i m_i} \tag{29}$$

$$m_i(t) = \frac{c_i - r_i(1 - v_i)}{D_i^m \exp\left[- \Big(c_i - r_i(1 - v_i)\Big)t \right] + c_i} \tag{30}$$

where

$$D_i^v = \frac{a_i m_i (1 - v_{i0}) - h_i U_i}{v_{i0}}, \qquad D_i^m = \frac{c_i (1 - m_{i0}) - r_i (1 - v_i)}{m_{i0}}.$$

Such a solution can be used by discretizing from $t_0 = 0$, the time axis in intervals $\Delta t = t_s - t_{s-1}$ sufficiently small, so that the quantities c_i, v_i and m_i can be assumed constant in the time interval Δt. Then the solution of Eqs. (6) and (7), by extending the formulae (29) and (30), can be approximated by the hierarchy (see the discussion in paper [9] according to the book [11])

$$v_i(t_s) = \frac{a_i m_i(t_{s-1}) - h_i U_i}{D_i^v(t_{s-1}) \exp\Big[- \Big(a_i m_i(t_{s-1}) - h_i U_i\Big)\Delta t\Big] + a_i m_i(t_{s-1})} \tag{31}$$

$$m_i(t_s) = \frac{c_i(t_{s-1}) - r_i\Big(1 - v_i(t_{s-1})\Big)}{D_i^m(t_{s-1}) \exp\Big[- \Big(c_i(t_{s-1}) - r_i\big(1 - v_i(t_{s-1})\big)\Big)\Delta t\Big] + c_i(t_{s-1})} \tag{32}$$

where

$$D_i^v(t_{s-1}) = \frac{a_i m_i(t_{s-1})[1 - v_i(t_{s-1})] - h_i U_i}{v_i(t_{s-1})} ,$$

$$D_i^m(t_{s-1}) = \frac{c_i(t_{s-1})[1 - m_i(t_{s-1})] - r_i[1 - v_i(t_{s-1})]}{m_i(t_{s-1})} .$$

Of course the determination of such a hierarchy at time t_s must take into account that for any solution m_i at time t_{s-1} it is necessary to compute, for all the LUs, the values of the GBTC $M_i = m_i M_i^{max}$, of the fluxes Φ_i and then of the connectivity indexes c_i. On the other hand the values of h_i, U_i, r_i, M_i^{max} and H_{ik} are constant and consequently can be computed before the generation of the hierarchy itself.

Finally, it is worthwhile to note that the hypothesis that territorial quantities are almost constant during the time interval Δt is justified by the fact the relaxation time of an environmental system is sufficiently long.

5 Evaluation Analysis of the Whole Environmental System

In this section we examine the dynamics of the whole environmental system, providing the solution obtained with the iterative scheme explained in the previous section, using Eqs. (31) and (32). The computations have been performed using the software *Mathematica*, version number 10.0.0.0. We give in the Table 2 the values of the quantities H_{ik} and L_{ik}. Note that in the first column of the table the couple of the neighboring LUs is indicated only for those LUs that present permeable borders.

Figures 7, 8 and 9 show the time evolution of v_i and m_i (for $i = 5, 8, 10, 13, 20, 24$), at an arbitrary scale, when the corresponding LUs are connected in the whole environmental system, through time dependent connectivity indexes $c_i(t)$.

In particular Fig. 7 presents the time evolution of LU 5 and LU 24. In Sect. 3, when studying separately the landscape units, we have mentioned that these two LUs

Table 2 Values of H_{ik} and L_{ik} for the neighboring LUs presenting permeable borders

LUs	H_{ik}	L_{ik}
1__2	391	782
1__17	3017	7543
1__18	3664	7327
1__24	2896	5792
2__3	1098	1569
2__4	1292	1845
4__5	964	1377
4__15	1062	2654
5__2	262	374
5__6	3837	7675
5__9	2804	7011
5__14	553	1383
5__15	633	1583
5__21	267	381
5__22	122	175
6__7	5416	9027
6__8	533	927
7__8	1460	2919
8__9	4231	8991
9__11	241	602
9__12	1931	2758
9__13	1072	1531
9__23	771	1543
10__9	425	850
10__20	305	763
11__12	855	2137
11__19	228	1138
12__13	984	2461
13__19	821	1172
13__23	680	3401
14__15	634	1585
14__16	548	1096
14__22	109	156
16__17	1785	3570
16__21	311	778
16__22	281	702
18__17	5429	10857

(continued)

Table 2 (continued)

LUs	H_{ik}	L_{ik}
18__19	3352	6705
18__20	5390	10780
19__20	1553	3106
21__22	184	459
21__23	197	493
24__2	2557	5114

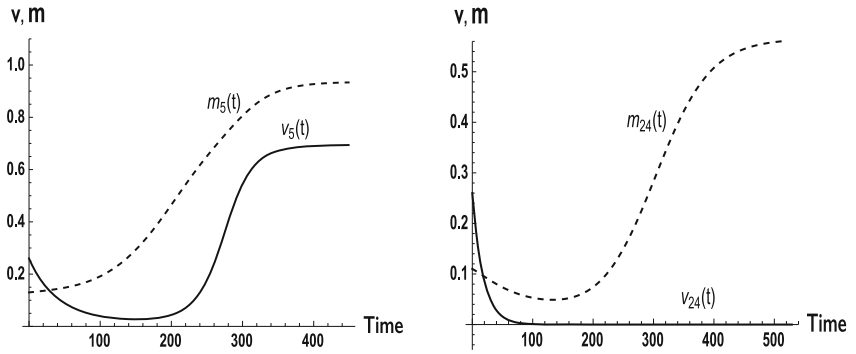


Fig. 7 Representation of the LUs in the whole environmental system. *Left* – LU 5. Evolution of v_5 and m_5 versus time. *Right* – LU 24. Evolution of v_{24} and m_{24} versus time

exhibit the unique stable node $(v_i^{(2)}, m_i^{(2)})$, see Table 1, and have a similar behavior, so that we have shown only the plots of Fig. 1 concerning LU 5. The behavior of these LUs is completely different when the environmental system is considered as a whole. In fact, a new stable node of high ecological quality appears for LU 5, say $(v_5^{(4)}, m_5^{(4)})$, thanks to the monotonic increasing of the GBTC variable m_5. Despite the fact that for a long initial transient the variable v_5 assumes values close to zero, LU 5 tends to the node $(v_5^{(4)}, m_5^{(4)})$. Conversely LU 24, because of a significant decreasing of m_{24} during the transient behavior, does not reach an equilibrium of high ecological quality and leads to the one typical of agricultural areas.

Moreover Fig. 8 describes the dynamics of LU 8 and LU 20 which, when isolated, exhibit a trend to the stable nodes $(v_i^{(4)}, m_i^{(4)})$ of high ecological quality, as shown in Figs. 2 and 6 of Sect. 3. In the whole system simulation, LU 8 still reaches the equilibrium $(v_8^{(4)}, m_8^{(4)})$, whereas LU 20 presents a loss of ecological quality and evolves towards the equilibrium $(v_{20}^{(2)}, m_{20}^{(2)})$, since for a long time interval it exhibits a strong decrease of the GBTC variable m_{20} which causes the decay to zero of the variable v_{20}.

Finally, Fig. 9 considers LU 10 and LU 13. The stability analysis carried out in Sect. 3 for LU 10 shows, as visualized by Fig. 3–*left*, that this sector admits the two stable attractors $(v_{10}^{(2)}, m_{10}^{(2)})$ and $(v_{10}^{(4)}, m_{10}^{(4)})$, and consequently the dynamics of such

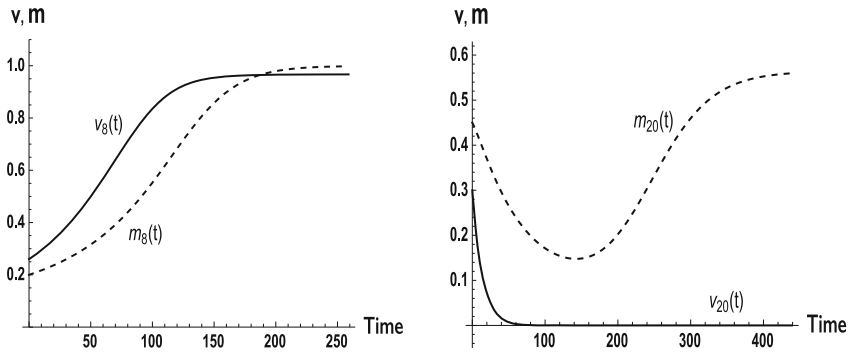


Fig. 8 Representation of the LUs in the whole environmental system. *Left* – LU 8. Evolution of v_8 and m_8 versus time. *Right* – LU 20. Evolution of v_{20} and m_{20} versus time

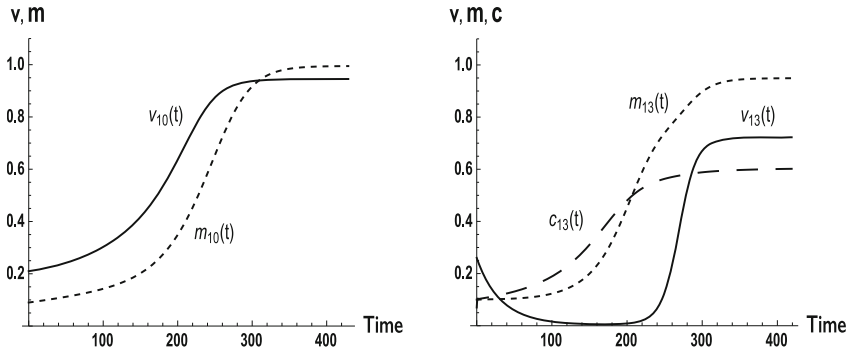


Fig. 9 Representation of the LUs in the whole environmental system. *Left* – LU 10. Evolution of v_{10} and m_{10} versus time. *Right* – LU 13. Evolution of v_{13}, m_{13} and c_{13} versus time

LU depends strongly on the initial data (in particular for those of Table 1 the equilibrium solution reached was $(v_{10}^{(2)}, m_{10}^{(2)})$). When the whole system is considered, LU 10 improves its environmental properties and tends to a scenario of high ecological quality reaching the equilibrium $(v_{10}^{(4)}, m_{10}^{(4)})$, since evidently the basin of attraction of node $(v_{10}^{(2)}, m_{10}^{(2)})$ is now different. Analogously, also LU 13 presents such an improvement when well connected to its neighbors, as it can be seen by Fig. 9–*right*: after a long initial transient where the variable v_{13} is close to zero, then there is a strong growth of v_{13} itself due to the monotonic increasing of both the GBTC variable m_{13} and the connectivity index c_{13} (dashed line). Thus, LU 13 changes from a scenario showing a complete loss of ecological quality, defined by the stable node $(v_{13}^{(1)}, m_{13}^{(1)})$ when it is isolated, to an opposite scenario of high ecological quality stated by $(v_{13}^{(4)}, m_{13}^{(4)})$, when it is considered connected to the whole environmental system. Such examples evidence how a LU can commute to a scenario of high ecological quality when it is well connected to its neighbors, even when a bad trend is found in the stability analysis developed individually for each LU. Such a behavior influences also the

time scaling of the system. Let us note in fact that the time scales of Figs. 7, 8 and 9 are much greater than those provided when the LUs were considered separately. This peculiarity seems reasonable since, when the environmental system is treated as a whole, then the relaxation times increase and fluctuations of territorial quantities get slower.

6 Conclusions and Future Perspectives

In this paper we have proposed a new mathematical model for the evaluation of the ecological state of an environmental system distributed in landscape units (LUs). More specifically, starting from previous ideas advanced in paper [8], our model acts at the level of each LU, instead at that of the whole system, and introduces as state variables the extent of green area of high ecological quality and a generalized biological energy of each landscape unit. The model is then capable to describe the territory at a more detailed level, so that its properties are better apprehended. From the mathematical point of view, the model is represented by an autonomous system of ordinary differential equations of cooperative Lotka–Volterra type [14]. The stability analysis developed in Sect. 2 for each LU determines the equilibrium solutions of the equations, whose qualitative trend indicates the future possible scenarios of the LU itself. The analysis is then applied in Sect. 3 to a Northern district of the Turin Province consisting of 24 different ecological sectors linked through a constant connectivity index recovered from the GIS. It allows to identify the LUs with high ecological quality, showing a great potential to evolve to a favorable scenario, and, conversely, those presenting a bad ecological state with a tendency to a scenario of ecological quality loss. It also allows to identify the LUs showing a different asymptotic equilibrium, in particular that typical of agricultural areas.

For a comprehensive description of the whole environmental system, the dynamics of the landscape units is then investigated in Sect. 4, considering all LUs connected to their neighboring sectors. The resulting model incorporates the connectivity issues among the neighboring sectors showing that the connectivity index plays an important role since now it has become time dependent through the state variables. The analytical treatment of such a system is rather complicated, since it is represented by 48 coupled ordinary differential equations. Our strategy was then to use an approximation procedure based on an algebraic hierarchy and, following the ideas discussed in paper [9], we propose an explicit algorithm presented in Sect. 5. The simulations show how a LU can commute from a certain scenario to a completely different situation, due to the influence of its neighbors, even if a different trend is predicted in the stability analysis developed individually for each LU.

Therefore, the study developed in this paper indicates that the connectivity among the neighboring sectors has a significant impact in the dynamics of the LUs when they are considered as parts of a whole. Even a rather complete analysis of the individual LUs is not enough to describe the whole system and a model taking into account the connectivity issues is an appropriate tool.

In our opinion, the model proposed here offers promising results and motivates future perspectives in terms of networking systems accounting for neighboring sectors. In fact, we think that it is possible to propose a model similar to the one presented here for what concerns the state variables, but different for the LUs coupling, borrowing some ideas from electrical synapses linking neurons [2] and exploiting the analysis of the landscape connectivity [6]. Moreover, another development could take into account a model with more state variables, namely considering variables v's for each ecological class $\mathscr{C}_2, \ldots, \mathscr{C}_5$ with non-null BTC indexes. In such a way, the model would also include the effects due to the presence of landcover areas with weak production of biological energy. These developments can be introduced in a forthcoming paper.

Acknowledgements The present research has been partially supported by Portuguese Funds FCT Project UID/MAT/00013/2013, and by the National Group GNFM of INdAM. One of the Authors (AJS) thanks this institution for the financial support given in her visiting professor program in Italy.

References

1. Argent, R.M., Sojda, R.S., Guipponi, C., McIntosh, B., Voinov, A.A., Maier, H.R.: Best practices for conceptual modelling in environmental planning and management. Environ. Model. Softw. **80**, 113–121 (2016)
2. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: structure and dynamics. Phys. Rep. **424**, 175–308 (2006)
3. Cassatella, C., Peano, A. (eds.): Landscape Indicators: Assessing and Monitoring Landscape Equality. Springer, Dordrecht (2011)
4. De Palma I.: Methodologies of landscape ecology and mathematical models for environments evaluation and ecoservice systems estimate, Master Degree Thesis, Politecnico di Torino (2014)
5. Fabbri, P.: Principi Ecologici per la Progettazione del Paesaggio (Ecological Principles of Landscape Planning). Franco Angeli, Milano (2007)
6. Foltête, J.-C., Clauzel, C., Vuidel, G.: A software tool dedicated to the modelling of landscape networks. Environ. Model. Softw. **38**, 316–327 (2012)
7. Gobattoni, F., Lauro, G., Leone, A., Monaco, R., Pelorosso, R.: A procedure for the mathematical analysis of landscape evolution and scenarios assessment. Landsc. Urban Plan. **103**, 289–302 (2011)
8. Gobattoni, F., Lauro, G., Monaco, R., Pelorosso, R.: Mathematical models in landscape ecology: stability analysis and numerical tests. Acta Appl. Math. **125**, 173–192 (2013)
9. Gobattoni, F., Groppi, M., Monaco, R., Pelorosso, R.: New developments and results for mathematical models in environment evaluations. Acta Appl. Math. **132**, 321–331 (2014)
10. Ingegnoli, V.: Landscape Ecology: A Widening Foundation. Springer, Berlin (2002)
11. Ixaru, L.G.: Numerical Methods for Differential Equations and Applications. Reidel, Dordrecht (1984)
12. Jordan, D.W., Smith, P.: Nonlinear Ordinary Differential Equations. Clarendon Press, Oxford (1977)
13. Murray, J.D.: Mathematical Biology I. An Introduction. Springer, Berlin (2002)
14. National Geographic website (visited in November 2016). http://nationalgeographic.org/encyclopedia/geographic-information-system-gis/
15. O'Neill, R.V., et al.: Indices of landscape pattern. Landsc. Ecol. **3**, 153–162 (1988)
16. Pielou, E.C.: Ecological Diversity. Wiley, New York (1975)

17. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana (1949)
18. Smith, H.L.: Monotone Dynamical Systems: An Introduction to The Theory of Competitive and Cooperative Systems, vol. 41. American Mathematical Society, Providence (2008)
19. Turner, M.G.: Landscape ecology: the effect of pattern on process. Annu. Rev. Ecol. Syst. **20**, 171–197 (1989)

Derivation of Models for Thin Sprays from a Multiphase Boltzmann Model

Valeria Ricci

Abstract We shall review the validation of a class of models for thin sprays where a Vlasov type equation is coupled to an hydrodynamic equation of Navier–Stokes or Stokes type. We present a formal derivation of these models from a multiphase Boltzmann system for a binary mixture: under suitable assumptions on the collision kernels and in appropriate asymptotics (resp. for the two different limit models), we prove the convergence of solutions to the multiphase Boltzmann model to distributional solutions to the Vlasov–Navier–Stokes or Vlasov–Stokes system. The proofs are based on the procedure followed in Bardos et al. (J Stat Phys 63:323–344 (1991), [2]) and explicit evaluations of the coupling terms due to the interaction between the two components of the mixture. The results reviewed in this article are proved in detail in Bernard et al. (A derivation of the Vlasov-Navier-Stokes model for aerosol flows from kinetic theory (2016), [4], A derivation of the Vlasov-Stokes system for aerosol flows from the kinetic theory of binary gas mixtures (2016), [5]).

Keywords Two-component systems · Vlasov–Navier–Stokes system Boltzmann equation · Hydrodynamic limit · Aerosols · Sprays · Gas mixture

MSC 35Q20 · 35B25 · 82C40 · 76T15 · 76D05 · 76D07

1 Introduction

In this paper we shall deal with the modelling of two phase systems where a thin component, given by a solid or condensed phase, is dispersed in a fluid phase. These systems (called *aerosols* or *sprays*) are sometimes modelled, at a macroscopic level, through a system of partial differential equations consisting of a hydrodynamic equation of Navier–Stokes or Stokes type coupled to a kinetic equation of Vlasov type;

V. Ricci (✉)
Dipartimento di Matematica e Informatica, Università di Palermo,
Via Archirafi 34, I90123 Palermo, Italy
e-mail: valeria.ricci@unipa.it

© Springer International Publishing AG 2017

P. Gonçalves and A.J. Soares (eds.), *From Particle Systems to Partial Differential Equations*, Springer Proceedings in Mathematics & Statistics 209,
DOI 10.1007/978-3-319-66839-0_14

the coupling is given by the drag force exerted by the fluid phase on the dispersed phase.

We want to illustrate how to obtain the derivation of such a kind of models from a model on a smaller scale, which could not be the scale corresponding to the elementary components, i.e. particles.

More specifically, our smaller scale system will be a mesoscopic one, given by a system of coupled Boltzmann equations, one for each species, and we shall study suitable asymptotics for the parameters which lead to the Vlasov–(Navier–)Stokes system in the limit.

The paper reviews results obtained in [4, 5].

The class of models we want to analyse is given by the (incompressible) Vlasov–Navier–Stokes system:

$$\begin{cases} \partial_t F + v \cdot \nabla_x F - \dfrac{\kappa}{m_p} \mathrm{div}_v((v-u)nF) = 0\,, \\ \rho_g(\partial_t u + u \cdot \nabla_x u) + \nabla_x p = \rho_g \nu \Delta_x u + \kappa \displaystyle\int_{\mathbf{R}^3} (v-u)F \,\mathrm{d}v\,, \\ \mathrm{div}_x u = 0\,, \end{cases} \tag{1}$$

or Vlasov–Stokes system

$$\begin{cases} \partial_t F + v \cdot \nabla_x F - \dfrac{\kappa}{m_p} \mathrm{div}_v((v-u)nF) = 0\,, \\ -\rho_g \nu \Delta_x u + \nabla_x p = \kappa \displaystyle\int_{\mathbf{R}^3} (v-u)F \,\mathrm{d}v\,, \\ \mathrm{div}_x u = 0\,, \end{cases} \tag{2}$$

which describes the dispersed phase through its number density of particles or droplets (we shall denote these entities as *particles* in what follows) with velocity v located at the position x at time t, $F \equiv F(t,x,v) \geq 0$, and the gas through its velocity field $u \equiv u(t,x) \in \mathbf{R}^3$. The parameters $\kappa, m_p, \rho_g, \nu > 0$ represent resp. the friction coefficient of the gas on the dispersed phase, the mass of a particle, the gas density and the kinematic viscosity of the gas.

The derivation of (1) or (2) from a genuine particle system, i.e. a microscopic dynamics such that both the dispersed and the fluid phase are described by particles interacting through deterministic laws, is presently out of reach. The next level of validation from a smaller scale description would be to consider the derivation of (1) or (2) as the mean-field limit of a system where a large number of particles (describing the dispersed phase) are immersed in a viscous fluid. Unfortunately, this validation presents major difficulties. Derivations of the Stokes and Navier–Stokes equation with a force term including the drag force exerted by the particles on the fluid can be obtained [1, 9], but all these derivations rely on methods of homogenization of elliptic operators on perforated domains where the holes have finite capacity ([7, 12], for a reformulation of [9] in the frame of potential theory see [13]). These methods

require an uniform control on the distance among particles, which can be assured when the particles can be considered as fixed, as in [1, 9], but which can not be preserved by the dynamics when they are allowed to move.

This leads to consider a third class of systems at a smaller scale which can be used to validate (1) or (2), i.e. the class where both species are described at a mesoscopic level: we consider then a system of two coupled Boltzmann equations, one for each species, and we consider its hydrodynamic limit, defining a scaling such that we obtain (1) or (2) in the correspondent asymptotics.

Although this validation does not start from first principles, it presents many advantages, first of all the absence of constraints on minimal distances among particles; moreover, using a multiphase Boltzmann system as a starting point allows to consider a large class of interactions among the dispersed phase and the gas.

We shall describe in what follows the formal derivation, in two different asymptotics, of (1) and (2) from a multiphase Boltzmann system (see (3) below).

2 The Multiphase Boltzmann Model

Let consider as a starting point the mesoscopic system

$$\begin{aligned} \partial_t F + v \cdot \nabla_x F &= \mathscr{D}(F, f) + \mathscr{B}(F)\,, \\ \partial_t f + w \cdot \nabla_x f &= \mathscr{R}(f, F) + \mathscr{C}(f)\,, \end{aligned} \tag{3}$$

where $F(t, x, v) \geq 0$ denotes the distribution function of particles and $f(t, x, w) \geq 0$ the distribution function of gas molecules. $\mathscr{B}(F)$ and $\mathscr{C}(f)$ are Boltzmann collision integrals, resp., for pairs of particles and for pairs of gas molecules, and $\mathscr{D}(F, f)$ and $\mathscr{R}(f, F)$ are Boltzmann type collision integrals describing, resp., the deflection of particles colliding with gas molecules, and the slowing down of gas molecules by collisions with particles.

The collision integrals verify, for F and f defined a.e. on $\mathbf{R}^3$ and rapidly decaying at infinity, the conservation laws:

$$\int_{\mathbf{R}^3} \mathscr{C}(f)(w) \begin{pmatrix} 1 \\ w \\ |w|^2 \end{pmatrix} \mathrm{d}w = 0\,, \tag{4}$$

$$\int_{\mathbf{R}^3} \mathscr{B}(F)(v) \begin{pmatrix} 1 \\ v \end{pmatrix} \mathrm{d}v = 0\,, \tag{5}$$

$$\int_{\mathbf{R}^3} \mathscr{D}(F, f)(v)\,\mathrm{d}v = \int_{\mathbf{R}^3} \mathscr{R}(f, F)(w)\,\mathrm{d}w = 0\,, \tag{6}$$

and satisfy the local balance of momentum in the aerosol

$$m_p \int_{\mathbf{R}^3} \mathscr{D}(F, f)(v) v \, \mathrm{d}v + m_g \int_{\mathbf{R}^3} \mathscr{R}(f, F)(w) w \, \mathrm{d}w = 0 \,, \tag{7}$$

where m_g is the mass of gas molecules and m_p the mass of particles. If the collisions between gas molecules and particles are elastic, they satisfy in addition the local balance of energy.

We consider then typical physical parameters, length L, number densities, $\mathscr{N}_p$ and $\mathscr{N}_g$, and thermal speeds, V_p and V_g, resp. for the particles and the gas molecules, average cross sections for the particle-particle (S_{pp}), particle-molecule (S_{pg}) and molecule-molecule (S_{gg}) collisions, and the three parameters $\eta = m_g/m_p$, $\varepsilon = V_p/V_g$ and $\mu = (m_g \mathscr{N}_g)/(m_p \mathscr{N}_p)$, and we rewrite (3) in dimensionless form (assuming as typical time L/V_p and measuring the speed of each species in terms of its thermal speed), obtaining [4, 5]

$$\begin{cases} \partial_t F + v \cdot \nabla_x F = \mathscr{N}_g S_{pg} L \dfrac{V_g}{V_p} \mathscr{D}(F, f) + \mathscr{N}_p S_{pp} L \mathscr{B}(F) \,, \\ \partial_t f + \dfrac{V_g}{V_p} w \cdot \nabla_x f = \mathscr{N}_p S_{pg} L \dfrac{V_g}{V_p} \mathscr{R}(f, F) + \mathscr{N}_g S_{gg} L \dfrac{V_g}{V_p} \mathscr{C}(f) \,. \end{cases} \tag{8}$$

In what follows, we assume $\mathscr{N}_p S_{pp} L \ll 1$, so that the collisions between particles can be neglected (this hypothesis can be formally justified on the basis of the chosen scaling, we shall nevertheless assume it here without further explanation).

Since the thermal speed of the particles is in general smaller than the thermal speed of the molecules, and the mass of the particles is much bigger than the mass of the gas molecules, we assume $\varepsilon \ll 1$ and $\eta \ll 1$ (notice that there is no need in general to make assumptions concerning the size or the relative size of the molecules to the particles, if not the ones allowing to consider (8) as a valid starting system).

We assume moreover

$$\frac{\eta}{\mu} = \frac{\mathscr{N}_p}{\mathscr{N}_g} \in [0, 1] \,, \quad \mathscr{N}_p \, S_{pg} \, L = \frac{\varepsilon}{\mu} \,, \quad \text{and} \quad \mathscr{N}_g \, S_{gg} \, L = \frac{\mu}{\varepsilon} \,.$$

The relative size of μ with respect to ε will differ for the two different scalings.

With this assumptions, the dimensionless scaled system becomes

$$\begin{cases} \partial_t F + v \cdot \nabla_x F = \dfrac{1}{\eta} \mathscr{D}(F, f) \,, \\ \partial_t f + \dfrac{1}{\varepsilon} w \cdot \nabla_x f = \dfrac{1}{\mu} \mathscr{R}(f, F) + \dfrac{\mu}{\varepsilon^2} \mathscr{C}(f) \,. \end{cases} \tag{9}$$

We shall then consider two different asymptotics.

First,we assume $\mu = 1$: when $\mu = 1$, the mass density of the gas and the mass density of particles have the same order of magnitude, we have only two small parameters, η and ε, s.t. $\eta \ll \varepsilon^2$, and the scaled Boltzmann system, whose asymptotics leads to the Vlasov–Navier–Stokes system, reads

$$\begin{cases} \partial_t F + v \cdot \nabla_x F = \dfrac{1}{\eta} \mathscr{D}(F, f)\,, \\ \partial_t f + \dfrac{1}{\varepsilon} w \cdot \nabla_x f = \mathscr{R}(f, F) + \dfrac{1}{\varepsilon^2} \mathscr{C}(f)\,. \end{cases} \tag{10}$$

Then, we shall consider $\mu \to 0$: here the mass density of the gas is much smaller than the mass density of the particles, the limit of (9) is the Vlasov–Stokes system and the three small parameters involved are such that $\eta \ll \varepsilon^2$ and $\varepsilon \ll \mu^2 \ll 1$.

As a last remark, we recall that systems similar to (1), (2) are used in applications. As an example, the Vlasov–Stokes system is used in the modelling of medical aerosols in the trachea and of drug deposition in the upper part of the lungs and the values of the parameters involved fit the asymptotics analysed in [5].

3 The Collision Kernels

In (9) we define the (ε- and η-dependent) dimensionless collision integrals as

$$\mathscr{C}(f)(w) = \int_{\mathbf{R}^3 \times \mathbf{S}^2} (f(w') f(w'_*) - f(w) f(w_*)) c(w - w_*, \omega)\, \mathrm{d}w_* \mathrm{d}\omega\,, \tag{11}$$

$$\begin{aligned} \mathscr{D}(F, f)(v) = &\int_{\mathbf{R}^3 \times \mathbf{R}^3} F(V) f(W) \Pi_{pg}(v, \mathrm{d}V\, \mathrm{d}W) \\ &- F(v) \int_{\mathbf{R}^3} f(w)\, |\varepsilon v - w|\, \Sigma_{pg}\left(|\varepsilon v - w|\right)\, \mathrm{d}w\,, \end{aligned} \tag{12}$$

$$\begin{aligned} \mathscr{R}(f, F)(w) = &\int_{\mathbf{R}^3 \times \mathbf{R}^3} F(V) f(W) \Pi_{gp}(w, \mathrm{d}V\, \mathrm{d}W) \\ &- f(w) \int_{\mathbf{R}^3} F(v)\, |\varepsilon v - w|\, \Sigma_{pg}\left(|\varepsilon v - w|\right)\, \mathrm{d}v\,, \end{aligned} \tag{13}$$

where the measure–valued functions Π_* and the functions Σ_* (both nonnegative) are related by:

$$\begin{aligned} \int_{\mathbf{R}^3_v} \mathrm{d}v\, \Pi_{pg}(v, \mathrm{d}v'\, \mathrm{d}w') &= |\varepsilon v' - w'| \Sigma_{pg}(|\varepsilon v' - w'|)\, \mathrm{d}v'\, \mathrm{d}w'\,, \\ \int_{\mathbf{R}^3_w} \mathrm{d}w\, \Pi_{gp}(w, \mathrm{d}v'\, \mathrm{d}w') &= |\varepsilon v' - w'| \Sigma_{pg}(|\varepsilon v' - w'|)\, \mathrm{d}v'\, \mathrm{d}w'\,. \end{aligned} \tag{14}$$

The collisions between molecules are elastic, so that

$$\begin{aligned} w' &= w - (w - w_*) \cdot \omega \omega\,, \\ w'_* &= w_* + (w - w_*) \cdot \omega \omega\,, \end{aligned} \tag{15}$$

and the collision kernel c, which is associated to a hard potential satisfying Grad's cut-off assumptions, is of the form

$$c(w - w_*, \omega) = |w - w_*|\sigma_{gg}(|w - w_*|, |\cos(\widehat{w - w_*}, \omega)|). \tag{16}$$

As for the collisions between molecules, some more technical assumptions will be needed in order to pass to the limit. We define the linearised collision integral around the Maxwellian

$$M(w) := \tfrac{1}{(2\pi)^{3/2}} e^{-|w|^2/2} \tag{17}$$

as

$$\mathscr{L}\phi := -M^{-1} D\mathscr{C}(M) \cdot (M\phi)\,, \tag{18}$$

where D is the functional derivative. $\mathscr{L}$ is a Fredholm, self-adjoint operator in $L^2(Mdw)$ s. t. $\mathrm{Im}\,\mathscr{L} = \mathrm{Ker}\,\mathscr{L}^{\perp}$ (e.g. Theorem 2.1 in [4]).

Defining

$$A(w) := w \otimes w - \tfrac{1}{3}|w|^2 I \tag{19}$$

we have $A \perp \mathrm{Ker}\,\mathscr{L}$ in $L^2(M\mathrm{d}v)$ and $\exists$ a unique $\tilde{A} \in \mathrm{Dom}\,\mathscr{L}$ such that

$$\mathscr{L}\tilde{A} = A, \qquad \tilde{A} \perp \mathrm{Ker}\,\mathscr{L}\,. \tag{20}$$

By the symmetry properties of the collision integral and the rotation invariance of (17), $\tilde{A}$ can be shown to be of the form

$$\tilde{A}(w) = \alpha(|w|)A(w)\,, \tag{21}$$

where α is a measurable function such that $\int_{\mathbf{R}^3} \alpha^2 |w|^4 (\bar{c} \star M)^2 M\mathrm{d}w < \infty$.

We shall assume $\alpha \in L^\infty(\mathbf{R}_+)$, but in the Vlasov–Stokes asymptotics we shall need to add the following condition

(A). $\nabla\tilde{A} \le C(1 + |w|^2)$.

As for the class of collision integrals $\mathscr{D}$ and $\mathscr{R}$ which will be analysed, they shall verify the following properties:

(H1). (Conservation of mass) $\exists$ a measurable function $q \equiv q(r)$ s.t. $0 \le q(r) \le C(1 + r)$ for some $C > 0$ and Π_{pg} and Π_{gp} satisfy

$$\int_{\mathbf{R}^3} \Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\mathrm{d}v = \int_{\mathbf{R}^3} \Pi_{gp}(w, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}w = q(|\varepsilon V - W|)\mathrm{d}V\mathrm{d}W\,.$$

(H2). (Conservation of momentum) $\exists\, Q \equiv Q(r) \in C(\mathbf{R}_+^*)$, with $Q \ge 0$ and $Q(r) + |Q'(r)| \le C(1 + r)$ for some $C > 0$, such that Π_{pg} and Π_{gp} satisfy

$$\varepsilon \int_{\mathbf{R}^3} \mathrm{d}v\,(v-V)\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W) = -\eta \int_{\mathbf{R}^3} \mathrm{d}w\,(w-W)\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)$$
$$= -\frac{\eta}{1+\eta}(\varepsilon V - W)Q(|\varepsilon V - W|)\mathrm{d}V\mathrm{d}W\,.$$

(H3). $\exists$ a constant $C > 0$ such that Π_{pg} satisfies

$$\int_{\mathbf{R}^3} \mathrm{d}v \left|\varepsilon v - \frac{\varepsilon V + \eta W}{1+\eta}\right|^2 \Pi_{pg}(v,\mathrm{d}V\mathrm{d}W) \le C\,\eta^2\,(1+|\varepsilon V - W|^2)q(|\varepsilon V - W|)\mathrm{d}V\mathrm{d}W,$$

(q is the function in (H1)).

The next hypothesis will have different form for the Navier–Stokes and the Stokes asymptotics. After the common part in its statement, we split it then into two parts, the first one referring to the Vlasov–Navier–Stokes model and labelled as (H4)VNS and the second one referring to the Vlasov–Stokes model labelled as (H4)VS:

(H4). The limiting measure $\Pi^{0,0}_{gp}$ satisfies[1] $\mathscr{T}_R\#\Pi^{0,0}_{gp} = \Pi^{0,0}_{gp} \quad \forall\, R \in O_3(\mathbf{R})$, where $\mathscr{T}_R : (w,V,W) \mapsto (Rw,V,RW)$, and either

(H4)VNS $\quad \forall\, \Phi := \Phi(w,W) \quad$ s. t. $\quad |\Phi(w,W)| \le C(1+|w|^2+|W|^2)M(W)$, $\exists\, p > 3$,

$$\int_{\mathbf{R}^3} (1+|V|^2)^{-p} \left|\int_{\mathbf{R}^3\times\mathbf{R}^3} \Phi(w,W)(\Pi^{\varepsilon,\eta}_{gp}(w,\mathrm{d}V\mathrm{d}W) - \Pi^{0,0}_{gp}(w,\mathrm{d}V\mathrm{d}W))\,\mathrm{d}w\right| \overset{\varepsilon,\eta\to 0}{\to} 0.$$

Moreover,

$$\int_{\mathbf{R}^3\times\mathbf{R}^3} \mathrm{d}w\,(1+|w|^2+|W|^2)M(W)\Pi^{0,0}_{gp}(w,dVdW) \in L^1((1+V^2)^{-3}dV)\,;$$

or

(H4)VS $\quad \forall \Phi \in C^1(\mathbb{R}^3\times\mathbb{R}^3)$ s. t. $|\Phi(w,W)| + |\nabla_w\Phi(w,W)| \le C(1+|w|^2+|W|^2)M(W)$, $\forall p > 3$,

$$\int_{\mathbf{R}^3} (1+|V|^2)^{-p} \left|\int_{\mathbf{R}^3\times\mathbf{R}^3} \Phi(w,W)(\Pi^{\varepsilon,\eta}_{gp}(w,\mathrm{d}V\mathrm{d}W) - \Pi^{0,0}_{gp}(w,\mathrm{d}V\mathrm{d}W))\,\mathrm{d}w\right| = O(\varepsilon+\eta)$$

when $\varepsilon, \eta \to 0$.

The last hypothesis, valid for both asymptotics, is:

(H5). $\forall h \in L^2(M(w)\mathrm{d}w)$ and $\eta, \varepsilon \approx 0$, $\exists\, C > 0$ independent of η and ε such that:

$$\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} \frac{(1+|W|^2)}{(1+|V|^2)^p}(1+|w|^2)M(W)|h(W)|\Pi^{\varepsilon,\eta}_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w \le C||h||_{L^2(M(w)\mathrm{d}w)}.$$

[1] The notation $\mathscr{T}\#m$ designates the push-forward of the measure m by the transformation $\mathscr{T}$.

Hypotheses (H1) to (H5) are satisfied in some relevant cases for the collision between molecules and particles, such as elastic collisions associated to a hard potential, and inelastic and diffuse as in [6].

4 The Main Results

We now state a theorem, collecting the results proved in [4, 5] and give the sketch of its proof: details can be found in [4, 5].

In what follows, for each $\phi \in L^1(M\,\mathrm{d}v)$, we set $\langle\phi\rangle := \int_{\mathbf{R}^3} \phi(w)M(w)\,\mathrm{d}w$.

Theorem 1 *Let $g_n \equiv g_n(t,x,w) \geq 0$ and $F_n \equiv F_n(t,x,v) \geq 0$ be sequences of smooth (at least C^1) functions, and*

$$f_n(t,x,w) := M(w)(1+\varepsilon_n g_n(t,x,w)), \tag{22}$$

where M is the Maxwellian (17). Assume that (F_n, f_n) is a solution to (9) (with $\mathscr{C}, \mathscr{D}, \mathscr{R}$ defined in Sect. 3), where $\Pi_{pg}^{\varepsilon_n,\eta_n}$ and $\Pi_{gp}^{\varepsilon_n,\eta_n}$ satisfy (H1)–(H3), the first part of (H4), (H5), and the molecular collision kernel c (such that (15), (16) hold) is associated to a hard potential satisfying Grad's cut-off assumptions.

Let α in (21) be such that $\alpha \in L^\infty(\mathbf{R}_+)$, and

$$\varepsilon_n \to 0\,, \quad \eta_n/\varepsilon_n^2 \to 0\,, \quad \varepsilon_n/\mu_n^2 \to 0.$$

Assume $F_n \rightharpoonup F$ in L^∞_{loc} weak-, $g_n \rightharpoonup g$ in $L^2_{loc}(\mathbf{R}_+^* \times \mathbf{R}^3 \times \mathbf{R}^3)$ weak and that (a) $\exists p > 3$ such that, $\forall R > 0$,*

$$\sup_{n\geq 1} \sup_{(t,x,v)\in[0,R]\times[-R,R]^3\times\mathbf{R}^3} (1+|v|^2)^p F_n(t,x,v) \leq C_R < \infty$$

(b) the sequence $\langle g_n^2\rangle$ is bounded in $L^1_{loc}(\mathbf{R}_+^* \times \mathbf{R}^3)$,
(c)

$$\langle g_n\phi(w)\rangle \to \langle g\phi(w)\rangle \tag{23}$$

strongly in $L^2_{loc}(\mathbf{R}_+^* \times \mathbf{R}^3)$ for each $\phi \in C_c(\mathbf{R}^3)$.

Then $\exists\, L^\infty$ functions $\rho \equiv \rho(t,x), \theta \equiv \theta(t,x) \in \mathbf{R}$, and a L^∞ vector field $u \equiv u(t,x) \in \mathbf{R}^3$ such that, for a.e. $(t,x,w) \in \mathbf{R}_+ \times \mathbf{R}^3 \times \mathbf{R}^3$,

$$g(t,x,w) = \rho(t,x) + u(t,x)\cdot w + \theta(t,x)\tfrac{1}{2}(|w|^2 - 3)\,, \tag{24}$$

and the pair (F,u) satisfies, in the sense of distributions,

1. *[4] the Vlasov–Navier–Stokes system*

$$\begin{cases} \partial_t F + v \cdot \nabla_x F = \kappa \mathrm{div}_v((v-u)F), \\ \mathrm{div}_x u = 0, \\ \partial_t u + \mathrm{div}_x(u \otimes u) = \nu \Delta_x u - \nabla_x p + \kappa \displaystyle\int (v-u)F \, \mathrm{d}v, \end{cases} \tag{25}$$

if $\mu_n = 1$ *and* $\Pi_{pg}^{\varepsilon_n,\eta_n}$ *and* $\Pi_{gp}^{\varepsilon_n,\eta_n}$ *verify (H4)VNS;*

2. *[5] the Vlasov–Stokes system*

$$\begin{cases} \partial_t F + v \cdot \nabla_x F = \kappa \mathrm{div}_v((v-u)F), \\ \mathrm{div}_x u = 0, \\ -\nu \Delta_x u - \nabla_x p = \kappa \displaystyle\int (v-u)F \, \mathrm{d}v, \end{cases} \tag{26}$$

if $\mu_n \to 0$ *and* $\Pi_{pg}^{\varepsilon_n,\eta_n}$, $\Pi_{gp}^{\varepsilon_n,\eta_n}$ *verify (H4)VS, and (A) and the additional assumption that, for some* $q > 1$ *and* $\forall R > 0$,

$$\sup_{n \geq 1} \sup_{(t,x) \in [0,R] \times [-R,R]^3} \langle (1+w^2)^p g_n^2 M(w) \rangle < \infty \tag{27}$$

are valid.

In both cases

$$\nu := \tfrac{1}{10} \int \tilde{A} : \mathscr{L} \tilde{A} M(w) \, \mathrm{d}w > 0 \,, \quad \kappa := \tfrac{1}{3} \int Q(|w|)|w|^2 M(w) \, \mathrm{d}w > 0, \tag{28}$$

where Q *is defined in assumption (H2) and* $\tilde{A}$, $\mathscr{L}$ *are defined by (20)–(18).*

The proof of Theorem 1 is based on the formal derivation of the incompressible fluid dynamic limit of the Boltzmann equation formulated in [2] and can be split in the following steps.

4.1 Asymptotic Form of the Molecular Distribution Function

Proposition 1 *Under the assumptions of Theorem 1,* $\exists\, L^\infty$ *functions* $\rho \equiv \rho(t,x)$, $\theta \equiv \theta(t,x) \in \mathbf{R}$, *and a* L^∞ *vector field* $u \equiv u(t,x) \in \mathbf{R}^3$ *s.t. (24) holds. The field* u *satisfies in* $\mathscr{D}'(\mathbf{R}_+^* \times \mathbf{R}^3)$ *the incompressibility condition*

$$\mathrm{div}_x u = 0 \tag{29}$$

and $\tilde{A}$, *defined in (20), is such that*

$$\langle \tilde{A}(w) w \cdot \nabla_x g \rangle = \nu(\nabla_x u + (\nabla_x u)^T) \,,$$

with ν defined in (28).

Proof We have first:

$$\begin{aligned}\mathscr{C}(M(1+\varepsilon_n g_n)) =& \mathscr{C}(M) + \varepsilon_n D\mathscr{C}(M)\cdot(Mg_n) + \varepsilon_n^2 \mathscr{C}(Mg_n)\\ =& -\varepsilon_n M\mathscr{L} g_n + \varepsilon_n^2 M\mathscr{Q}(g_n),\end{aligned}$$

where $\mathscr{L}\phi$ is defined by (18) and

$$\mathscr{Q}(\phi) := M^{-1}\mathscr{C}(M\phi). \tag{30}$$

The second line of Eq. (9) can be recast as:

$$\partial_t g_n + \frac{1}{\varepsilon_n} w\cdot\nabla_x g_n + \frac{\mu_n}{\varepsilon_n^2}\mathscr{L} g_n = \frac{1}{\mu_n\varepsilon_n} M^{-1}\mathscr{R}(M(1+\varepsilon g_n), F_n) + \frac{\mu_n}{\varepsilon_n}\mathscr{Q}(g_n)\,, \tag{31}$$

from which we obtain (by multiplying each side of the equation by ε_n^2/μ_n)

$$\mathscr{L} g_n = (\frac{\varepsilon_n}{\mu_n^2} M^{-1}\mathscr{R}(M(1+\varepsilon g_n), F_n) + \varepsilon_n\mathscr{Q}(g_n)) - \frac{\varepsilon_n^2}{\mu_n}\partial_t g_n - \frac{\varepsilon_n}{\mu_n} w\cdot\nabla_x g_n. \tag{32}$$

The two last terms of (32) converge to 0 in the sense of distributions since $g_n \rightharpoonup g$ weak in L^2_{loc}.

We deal now with the first term. $\forall\phi\in C_c(\mathbf{R}^3)$,

$$\begin{aligned}\int_{\mathbf{R}^3}\mathscr{Q}(g_n)(w)\phi(w)\,\mathrm{d}w = \int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{S}^2} &\left(M^{-1}(w')\phi(w') - M^{-1}(w)\phi(w)\right)\\ &\times M(w_*)g_n(w_*)M(w)g_n(w)c(w-w_*,\omega)\,\mathrm{d}\omega\mathrm{d}w_*\mathrm{d}w\,,\end{aligned}$$

where w', w'_* are given by (15). By the Cauchy–Schwarz inequality,

$$\begin{aligned}\left|\int_{\mathbf{R}^3}\mathscr{Q}(g_n)\phi(w)\,\mathrm{d}w\right| &\le C\int_{\mathbf{R}^3\times\mathbf{R}^3} M(w_*)g_n(w_*)M(w)g_n(w)(1+|w|+|w_*|)\,\mathrm{d}w_*\mathrm{d}w\\ &\le C\langle g_n^2\rangle\int_{\mathbf{R}^3} M(w)(1+|w|)^2\,\mathrm{d}w\,,\end{aligned}$$

so that $\int_{\mathbf{R}^3}\mathscr{Q}(g_n)\phi(w)\,\mathrm{d}w$ is bounded in $L^1_{loc}(\mathbf{R}^*_+\times\mathbf{R}^3)\forall\phi\in C_c(\mathbf{R}^3)$, and $\varepsilon_n\mathscr{Q}(g_n)\to 0$ in the sense of distributions.

Likewise, for each $\phi\in C_c(\mathbf{R}^3)$, we deduce from (H1) that

$$\int_{\mathbf{R}^3}\mathscr{R}(f_n,F_n)M^{-1}(w)\phi(w)\,\mathrm{d}w =$$

$$\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3}(M^{-1}(w)\phi(w) - M^{-1}(W)\phi(W))f_n(W)F_n(V)\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w,$$

so that, according to (H1) and assumption (a),

$$\left|\int_{\mathbf{R}^3} \mathscr{R}(f_n, F_n)M^{-1}(w)\phi(w)\,\mathrm{d}w\right| \le C\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V)f_n(W)q(|\varepsilon_n V - W|)\,\mathrm{d}V\mathrm{d}W$$
$$\le C\langle(1+\varepsilon_n g_n)(1+|W|)\rangle\,,$$

which is bounded in $L^2_{loc}(\mathbf{R}^*_+\times\mathbf{R}^3)$ thanks to (b) in Theorem 1. Therefore

$$\frac{\varepsilon_n}{\mu_n^2}\mathscr{R}(f_n, F_n)M^{-1}(w)\to 0$$

in the sense of distributions.

Hence $\mathscr{L}g_n\to\mathscr{L}g=0$ in the sense of distributions and g is of the form (24).

We come then to the incompressibility condition. Multiplying (31) by $\varepsilon_n M(w)$ and integrating in w gives (according to (4))

$$\varepsilon_n\partial_t\langle g_n\rangle + \mathrm{div}_x\langle wg_n\rangle = 0\,.$$

Since $g_n\rightharpoonup g$ in $L^2(\mathbf{R}^*_+\times\mathbf{R}^3\times\mathbf{R}^3)$ weak and thanks to assumption (b) in Theorem 1, $\langle g_n\rangle\to\langle g\rangle$ and $\langle wg_n\rangle\to\langle wg\rangle$ in $L^2_{loc}(\mathbf{R}^*_+\times\mathbf{R}^3)$ weak, so that $\mathrm{div}_x\langle wg_n\rangle = -\varepsilon_n\partial_t\langle g_n\rangle\to 0$ in $\mathscr{D}'(\mathbf{R}^*_+\times\mathbf{R}^3)$ and $\mathrm{div}_x\langle wg\rangle = 0$. Since $\langle wg\rangle = u$, (29) holds.

Finally, since the tensor field $w\mapsto A(w)w$ is odd, we have

$$\langle\tilde{A}(w)w\cdot\nabla_x g\rangle = \langle\tilde{A}(w)\otimes A(w)\rangle : \nabla_x u.$$

By Lemma 4.4 in [3] (see formulas (4.13a) and (4.10), which are based on elementary symmetry arguments, in particular $A(Rw) = RA(w)R^T\ \forall R\in O_3(\mathbf{R})$), we have $\langle\tilde{A}_{ij}A_{kl}\rangle = \nu(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}-\frac{1}{3}\delta_{ij}\delta_{kl})$, with

$$\nu := \tfrac{1}{10}\langle\tilde{A} : \mathscr{L}\tilde{A}\rangle > 0.$$

Therefore, since $\mathrm{div}_x u = 0$,

$$\langle\tilde{A}(w)w\cdot\nabla_x g\rangle = \nu(\nabla_x u + (\nabla_x u)^T - \tfrac{2}{3}(\mathrm{div}_x u)I) = \nu(\nabla_x u + (\nabla_x u)^T\,.$$

4.2 Asymptotic Deflection and Friction Terms

The computation of the collision integral $\mathscr{D}(F, f)$ is the central point in the proof. Because $\eta\ll 1$, the heavy particles are only slightly deflected in the collision with the light molecules, explaining how the collision integral in the kinetic equation for the distribution function of the dispersed phase converges to the term which appears in the Vlasov equation.

Proposition 2 *Under the assumptions of Theorem 1,*

$$\frac{1}{\eta_n}\mathscr{D}(F_n, f_n) \to \kappa \operatorname{div}((v-u)F) \quad \text{in } \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3 \times \mathbf{R}^3)\,, \tag{33}$$

$$\frac{1}{\varepsilon_n}\int_{\mathbf{R}^3} w\mathscr{R}(f_n, F_n)\,\mathrm{d}w \to \kappa \int_{\mathbf{R}^3} (v-u)F\,\mathrm{d}v \quad \text{in } \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3), \tag{34}$$

with κ defined in (28).

Proof We have

$$\begin{aligned}
&\frac{1}{\eta_n}\int_{\mathbf{R}^3} \mathscr{D}(F_n, f_n)(v)\phi(v)\,\mathrm{d}v \\
&= \frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W) \int_{\mathbf{R}^3} (\phi(v)-\phi(V))\Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}v \\
&= \frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\nabla\phi(V)\cdot\int (v-V)\Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}v \\
&+\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W) \int H(v,V) : (v-V)^{\otimes 2}\Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}v,
\end{aligned} \tag{35}$$

where $H(v,V) := \int_0^1 (1-t)\nabla^2\phi((1-t)V+tv)\,\mathrm{d}t$.

According to (H2), the first term in (35) is given by

$$\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\nabla\phi(V)\cdot\int (v-V)\Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}v =$$

$$-\int_{\mathbf{R}^3} F_n(V)\nabla\phi(V)\cdot\frac{\frac{1}{\varepsilon_n}\int_{\mathbf{R}^3} f_n(W)(\varepsilon_n V - W)Q(|\varepsilon_n V - W|)\,\mathrm{d}W}{1+\eta_n}\,\mathrm{d}V\,.$$

Hence

$$\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\nabla\phi(V)\cdot\int (v-V)\Pi_{pg}(v, \mathrm{d}V\mathrm{d}W)\,\mathrm{d}v =$$

$$I_n^1 + I_n^2 + I_n^3 + I_n^4 + I_n^5\,,$$

with

$$I_n^1 = -\varepsilon_n \int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)g_n(W)V\,Q(|\varepsilon_n V - W|)\,\mathrm{d}W\mathrm{d}V\,,$$

$$I_n^2 = \int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)g_n(W)W(Q(|\varepsilon_n V - W|) - Q(|W|)\,\mathrm{d}W\,,$$

$$I_n^3 = \int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)g_n(W)WQ(|W|)\,\mathrm{d}W\mathrm{d}V\,,$$
$$I_n^4 = -\int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)VQ(|\varepsilon_n V - W|)\,\mathrm{d}W\mathrm{d}V\,,$$
$$I_n^5 = \frac{1}{\varepsilon_n}\int F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int M(W)W(Q(|\varepsilon_n V - W|) - Q(|W|))\,\mathrm{d}W\mathrm{d}V\,.$$

Since by the Cauchy–Schwarz inequality

$$\begin{aligned}&\left|\int_{\mathbf{R}^3} M(W)g_n(W)VQ(|\varepsilon_n V - W|)\,\mathrm{d}W\right|\\ &\le C\int_{\mathbf{R}^3}(1+\varepsilon_n|V|+|W|)M(W)|g_n(W)||V|\,\mathrm{d}W\\ &\le C|V|(1+|V|)\sqrt{\langle g_n^2\rangle},\end{aligned}$$

we have $I_n^1 \to 0$ in $L^2_{loc}(\mathbf{R}_*^+\times\mathbf{R}^3)$. Then,

$$\begin{aligned}&\left|\int_{\mathbf{R}^3} M(W)g_n(W)W(Q(|\varepsilon_n V - W|) - Q(|W|))\,\mathrm{d}W\right|\\ &\le C\varepsilon_n\int_{\mathbf{R}^3} M(W)|g_n(W)||W||V|(1+\varepsilon_n|V|+|W|)\,\mathrm{d}W\\ &\le C\varepsilon_n(1+|V|^2)\sqrt{\langle g_n^2\rangle}\,,\end{aligned}$$

so that $I_n^2 \to 0$ in $L^2_{loc}(\mathbf{R}_*^+\times\mathbf{R}^3)$. By assumption (c) in Theorem 1

$$\begin{aligned}\int_{\mathbf{R}^3} M(W)g_n(W)WQ(|W|)\,\mathrm{d}W &\to \int_{\mathbf{R}^3} M(W)g(W)WQ(|W|)\,\mathrm{d}W\\ &= \tfrac{1}{3}u\int_{\mathbf{R}^3} M(W)|W|^2Q(|W|)\,\mathrm{d}W = \kappa u\end{aligned}$$

in $L^1_{loc}(\mathbf{R}_*^+\times\mathbf{R}^3)$, and therefore $I_n^3 \to \kappa u\cdot\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\,\mathrm{d}V$ in $\mathscr{D}'(\mathbf{R}_*^+\times\mathbf{R}^3)$.

Then

$$\begin{aligned}&\left|\int_{\mathbf{R}^3} M(W)V(Q(|\varepsilon_n V - W|) - Q(|W|))\,\mathrm{d}W\right|\\ &\le C\varepsilon_n\int_{\mathbf{R}^3} M(W)|V|^2(1+\varepsilon_n|V|+|W|)\,\mathrm{d}W\\ &\le C\varepsilon_n|V|^2(1+|V|)\,,\end{aligned}$$

so that

$$\int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)V(Q(|\varepsilon_n V-W|)-Q(|W|))\,\mathrm{d}W\mathrm{d}V\to 0$$

locally uniformly on $\mathbf{R}^+_*\times\mathbf{R}^3$, and

$$I^4_n\to-\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot V\int_{\mathbf{R}^3} M(W)Q(|W|)\,\mathrm{d}W\mathrm{d}V \text{ in } \mathscr{D}'(\mathbf{R}^+_*\times\mathbf{R}^3)\,.$$

Finally

$$\begin{aligned}\left|\int_{\mathbf{R}^3} M(W)W\left(\frac{Q(|\varepsilon_n V-W|)-Q(|W|)}{\varepsilon_n}+\frac{W}{|W|}\cdot VQ'(|W|)\right)\mathrm{d}W\right|&\\ \le\int_{\mathbf{R}^3} M(W)|W||V|\int_0^1|Q'(|\theta\varepsilon_n V-W|)-Q'(|W|)|\,\mathrm{d}\theta\mathrm{d}W&\\ \le C|V|(1+|V|)&\end{aligned}$$

and, for all $V\in\mathbf{R}^3$,

$$\int_{\mathbf{R}^3} M(W)|W||V|\int_0^1|Q'(|\theta\varepsilon_n V-W|)-Q'(|W|)|\,\mathrm{d}\theta\mathrm{d}W\to 0$$

by dominated convergence. With assumption (a) in Theorem 1, we get

$$I^5_n+\int_{\mathbf{R}^3} F_n(V)\frac{\nabla\phi(V)}{1+\eta_n}\cdot\int_{\mathbf{R}^3} M(W)\frac{W}{|W|}W\cdot VQ'(|W|)\,\mathrm{d}W\mathrm{d}V\to 0$$

locally uniformly on $\mathbf{R}^+_*\times\mathbf{R}^3$, and therefore

$$I^5_n\to-\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot\int_{\mathbf{R}^3} M(W)\frac{W}{|W|}W\cdot VQ'(|W|)\,\mathrm{d}W\mathrm{d}V \text{ in } \mathscr{D}'(\mathbf{R}^+_*\times\mathbf{R}^3)\,.$$

By isotropy, we have $\int_{\mathbf{R}^3} M(W)\frac{W}{|W|}W\cdot VQ'(|W|)\,\mathrm{d}W=\frac{1}{3}V\int_{\mathbf{R}^3} M(W)|W|Q'(|W|)\,\mathrm{d}W$, so that

$$I^4_n+I^5_n\to-\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot V\int_{\mathbf{R}^3} M(W)(Q(|W|)+\tfrac{1}{3}|W|Q'(|W|))\,\mathrm{d}W\mathrm{d}V$$

in $\mathscr{D}'(\mathbf{R}^+_*\times\mathbf{R}^3)$. Moreover, since $W\cdot\nabla M(W)=-|W|^2M(W)$,

$$\int_{\mathbf{R}^3} M(W)|W|^2 Q(|W|)\,\mathrm{d}W = -\int_{\mathbf{R}^3} W\cdot\nabla M(W) Q(|W|)\,\mathrm{d}W$$
$$= \int_{\mathbf{R}^3} M(W)\mathrm{div}(W Q(|W|))\,\mathrm{d}W$$
$$= \int_{\mathbf{R}^3} M(W)(3Q + |W|Q')(|W|)\,\mathrm{d}W\,.$$

Hence

$$I_n^4 + I_n^5 \to -\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot V \int_{\mathbf{R}^3} \tfrac{1}{3}|W|^2 M(W) Q(|W|)\,\mathrm{d}W\mathrm{d}V$$
$$= -\kappa \int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot V\,\mathrm{d}V$$

in $\mathscr{D}'(\mathbf{R}_*^+\times\mathbf{R}^3)$, and finally

$$\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\nabla\phi(V)\cdot\int (v-V)\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v \to$$

$$-\kappa\int_{\mathbf{R}^3} F(V)\nabla\phi(V)\cdot(V-u)\,\mathrm{d}V \quad \text{in } \mathscr{D}'(\mathbf{R}_*^+\times\mathbf{R}^3)\,,$$

with

$$\kappa = \tfrac{1}{3}\int_{\mathbf{R}^3} M(W)Q(|W|)|W|^2\,\mathrm{d}W > 0\,.$$

Next we treat the second term in (35). We have

$$\left|\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\int H(v,V) : (v-V)^{\otimes 2}\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v\right|$$

$$\le \frac{1}{2\eta_n}\|\nabla^2\phi\|_{L^\infty}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V) f_n(W)\int_{\mathbf{R}^3}|v-V|^2\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v$$

and, with $U = \frac{\varepsilon_n V+\eta_n W}{1+\eta_n}$,

$$\varepsilon_n^2|v-V|^2 \le 2|\varepsilon_n v - U|^2 + 2|U-\varepsilon_n V|^2 = 2|\varepsilon_n v - U|^2 + \frac{2\eta_n^2}{(1+\eta_n)^2}|\varepsilon_n V - W|^2\,.$$

According to assumption (H3),

$$\int_{\mathbf{R}^3}|v-V|^2\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v \le 2C\frac{\eta_n^2}{\varepsilon_n^2}(1+|\varepsilon_n V - W|^2)q(|\varepsilon_n V - W|)\,,$$

so that

$$\left|\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V)f_n(W)\int H(v,V):(v-V)^{\otimes 2}\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v\right|$$

$$\leq \frac{\eta_n}{\varepsilon_n^2}|C|\nabla^2\phi||_{L^\infty}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V)f_n(W)(1+|\varepsilon_n V-W|^2)q(|\varepsilon_n V-W|)\,\mathrm{d}V\mathrm{d}W\,.$$

By (H1) and assumption (a) in Theorem 1,

$$\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V)f_n(W)(1+|\varepsilon_n V-W|)^2 q(|\varepsilon_n V-W|)\,\mathrm{d}V\mathrm{d}W$$
$$\leq CC_R\int_{\mathbf{R}^3\times\mathbf{R}^3}\frac{(1+\varepsilon_n|V|+|W|)^3}{(1+|V|)^p}M(W)(1+\varepsilon_n g_n)(W)\,\mathrm{d}V\mathrm{d}W$$

for $(t,x)\in[0,R]\times[-R,R]^3$. By assumption (b) in Theorem 1 and the Cauchy–Schwarz inequality, the right hand side is bounded in $L^2_{loc}(\mathbf{R}^+_*\times\mathbf{R}^3)$. Hence

$$\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3} F_n(V)f_n(W)\int H(v,V):(v-V)^{\otimes 2}\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v\to 0\quad\text{in } L^2_{loc}(\mathbf{R}^+_*\times\mathbf{R}^3)$$

since $\eta_n/\varepsilon_n^2\to 0$, which leads to (33).

Then, by assumptions (H1)–(H2),

$$\frac{1}{\varepsilon_n}\int_{\mathbf{R}^3} w\mathscr{R}(f_n,F_n)\,\mathrm{d}w=\frac{1}{\varepsilon_n}\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3}(w-W)f_n(W)F_n(V)\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w$$
$$=-\frac{1}{\eta_n}\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3}(v-V)F_n(V)f_n(W)\Pi_{pg}(v,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}v=-\frac{1}{\eta_n}\int_{\mathbf{R}^3} v\mathscr{D}(F_n,f_n)\,\mathrm{d}v\,,$$

so that

$$-\frac{1}{\eta_n}\int_{\mathbf{R}^3}\phi(v)\mathscr{D}(F_n,f_n)\,\mathrm{d}v\to\kappa\int_{\mathbf{R}^3}F(V)\nabla\phi(V)\cdot(V-u)\,\mathrm{d}V \tag{36}$$

in $\mathscr{D}'(\mathbf{R}^*_+\times\mathbf{R}^3)$ for each test function $\phi\in C^2(\mathbf{R}^3)$ and $\nabla\phi,\nabla^2\phi\in L^\infty(\mathbf{R}^3)$. Setting $\phi(v)=v$ in (36) leads to (34).

Notice that, in the previous steps, neither $\mu_n\to 0$ nor any specific hypothesis in Part 1 or 2 of Theorem 1 is needed, so that the proofs of the previous propositions are valid for both the Vlasov–Navier–Stokes and the Vlasov–Stokes asymptotics.

4.3 Convection Term

This proposition is needed only to deal with the asymptotics in Theorem 1 Part (1) (it can be skipped when dealing with Part 2).

Proposition 3 *Under the assumptions of Theorem 1,*

$$\langle \tilde{A}(w)\mathscr{Q}(g)\rangle = A(u)$$

where $\tilde{A}$ is defined in (20) and $\mathscr{Q}$ is defined in (30).

Proof By Proposition 1, $g(t,x,\cdot) \in \operatorname{Ker}\mathscr{L}$ for a.e. $(t,x) \in \mathbf{R}_+^* \times \mathbf{R}^3$ and (see formula (60) in [2]) $\mathscr{Q}(g(t,x,\cdot)) = \frac{1}{2}\mathscr{L}(g(t,x,\cdot)^2)$, for a.e. $(t,x) \in \mathbf{R}_+^* \times \mathbf{R}^3$. Since $\mathscr{L}$ is self-adjoint on $L^2(M\,\mathrm{d}w)$ (Theorem 2.1 in [4]) and $g^2 \in \operatorname{Dom}\mathscr{L}$, we get

$$\langle \tilde{A}(w)\mathscr{Q}(g)\rangle = \langle \tilde{A}(w)\tfrac{1}{2}\mathscr{L}(g^2)\rangle = \tfrac{1}{2}\langle(\mathscr{L}\tilde{A})g^2\rangle = \tfrac{1}{2}\langle Ag^2\rangle\,.$$

Eliminating the odd component of g^2 ($w \mapsto A(w)$ is even) we get

$$\langle Ag^2\rangle = \langle A\otimes w\otimes w\rangle : (u\otimes u) + \left\langle A\left(\rho + \theta\tfrac{1}{2}(|w|^2-3)\right)^2\right\rangle.$$

But $A(Rw) = RA(w)A^T$ and $\operatorname{trace}(A) = 0$ (Lemma 4.2 in [11]), so:

$$\left\langle A\left(\rho + \theta\tfrac{1}{2}(|w|^2-3)\right)^2\right\rangle = \tfrac{1}{3}\left\langle \operatorname{trace}(A)\left(\rho + \theta\tfrac{1}{2}(|w|^2-3)\right)^2\right\rangle I = 0.$$

Moreover (Lemma 4.2 in [11]) $\langle A\otimes w\otimes w\rangle_{ijkl} = \langle A_{ij}A_{kl}\rangle = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \frac{2}{3}\delta_{ij}\delta_{kl}$, so that finally

$$\langle A\otimes w\otimes w\rangle : (u\otimes u) = 2u\otimes u - \tfrac{2}{3}|u|^2 I$$

and the proposition is proved.

4.4 Asymptotic Friction Flux

Proposition 4 *Under the assumptions of Theorem 1,*

$$\frac{1}{\mu_n}\int \tilde{A}(w)\mathscr{R}(f_n,F_n)(w)dw \to 0 \quad \text{in } \mathscr{D}'(\mathbf{R}_+^*\times\mathbf{R}^3)\,.$$

Proof First, we write

$$\begin{aligned}\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(f_n,F_n)(w)\,\mathrm{d}w = \\ \int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w + \varepsilon_n\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(Mg_n,F_n)(w)\,\mathrm{d}w\,.\end{aligned} \tag{37}$$

The first term on the right-hand side of (37) is then given by:

$$\int_{\mathbf{R}^3} \tilde{A}(w)\mathscr{R}(M, F_n)(w)\,\mathrm{d}w =$$

$$\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F_n(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi_{gp}^{\varepsilon_n,\eta_n}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\,.$$

Now we have to distinguish the two different asymptotics, $\mu_n = 1$ and $\mu_n \to 0$.

When $\mu_n = 1$ (Vlasov–Navier–Stokes asymptotics, Theorem 1 part 1), we use the bound

$$\begin{aligned}
&\left|\int_{\mathbf{R}^3}\left(\tilde{A}\mathscr{R}(M,F_n)-\int_{\mathbf{R}^3\times\mathbf{R}^3} F(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\right)\mathrm{d}w\right| \\
&\quad\le \left|\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F_n(V)M(W)(\tilde{A}(w)-\tilde{A}(W))(\Pi_{gp}^{\varepsilon_n,\eta_n}-\Pi_{gp}^{0,0})(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\right| \\
&\quad+\left|\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} (F_n(V)-F(V))M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\right|.
\end{aligned}$$

The first term on the right hand side vanishes in $\mathscr{D}'(\mathbf{R}_+^*\times\mathbf{R}^3)$ because of (H4)VNS and $\alpha\in L^\infty(\mathbf{R}_+)$. The second term on the right hand side also vanishes in $\mathscr{D}'(\mathbf{R}_+^*\times\mathbf{R}^3)$ because of (H4)VNS.

According to the first part of assumption (H4), $\forall R\in O_3(\mathbf{R})$

$$\begin{aligned}
&\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w \\
&=\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\mathscr{T}_R\#\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w \\
&=\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F(V)M(W)(\tilde{A}(Rw)-\tilde{A}(RW))\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\,.
\end{aligned}$$

Because of (21), $\tilde{A}(Rw) = R\tilde{A}(w)R^T$, $\forall\, R\in O_3(\mathbf{R})$, and, $\forall R\in O_3(\mathbf{R})$,

$$\mathscr{A} := \int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} F(V)M(W)(\tilde{A}(Rw)-\tilde{A}(RW))\Pi_{gp}^{0,0}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w = R\mathscr{A}R^T$$

a.e. on $\mathbf{R}_+^*\times\mathbf{R}^3$, so that $\mathscr{A}(t,x) = \frac{1}{3}\,\mathrm{trace}(\mathscr{A}(t,x))I = 0$ (e.g. Lemma 4.2 in [4]).

Hence, the first term in (37) is s.t. $\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w\to 0$ in $\mathscr{D}(\mathbf{R}_+^*\times\mathbf{R}^3)$.

Next, we deduce from (H1) that, for all $(t,x)\in[0,K]\times[-K,K]^3$, by (H5) and assumptions (b) in Theorem 1

$$\left|\int_{\mathbf{R}^3} \mathscr{R}(Mg_n, F_n)\tilde{A}(w)\,\mathrm{d}w\right| =$$

$$\left|\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} (\tilde{A}(w)-\tilde{A}(W))M(W)g_n(W)F_n(V)\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\right|$$

$$\leq C_K \int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3} (|w|^2+|W|^2)M(W)|g_n(W)|(1+|V|^2)^{-p}\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w$$

$$\leq CC_K\|g\|_{L^2(M\,\mathrm{d}w)}. \tag{38}$$

From assumption (b) in Theorem 1, (38) is bounded in $L^2_{loc}(\mathbf{R}^*_+\times\mathbf{R}^3)$, so that the second term in (37) vanishes.

When $\mu_n\to 0$ (Vlasov–Stokes asymptotics, Theorem 1 Part 2), we consider the additional hypotheses in Theorem 1 part 2 and we get a simplified procedure.

We use as a test function $\Phi(w,W)=M(W)(\tilde{A}(w)-\tilde{A}(W))$ and we write

$$\left|\int \tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w-\int F_n(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi^{0,0}_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w\right|$$

$$\leq C\int(1+|V|^2)^{-p}\left|\int\Phi(w,W)(\Pi^{\varepsilon_n,\eta_n}_{gp}(w,dVdW)-\Pi^{0,0}_{gp}(w,dVdW))dw\right|.$$

By assumption (A), $\Phi\in C^1(\mathbf{R}^3\times\mathbf{R}^3)$ and $|\Phi(w,W)|+|\nabla_w\Phi(w,W)|\leq C(1+|w|^2+|W|^2)M(W)$. Therefore, (H4)VS implies that

$$\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w=$$

$$\int F_n(V)M(W)(\tilde{A}(w)-\tilde{A}(W))\Pi^{0,0}_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w+O(\varepsilon_n+\eta_n).$$

By using the symmetry assumption as in the Vlasov–Navier–Stokes computation, we get for all $R>0$ the bound

$$\sup_{t+|x|<R}\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w=O(\varepsilon_n+\eta_n), \tag{39}$$

so that $\frac{1}{\mu_n}\int_{\mathbf{R}^3}\tilde{A}(w)\mathscr{R}(M,F_n)(w)\,\mathrm{d}w$ vanishes.

Then, for some $p>3$,

$$\left|\int_{\mathbf{R}^3}\mathscr{R}(Mg_n,F_n)\tilde{A}(w)\,\mathrm{d}w\right|\leq$$

$$C_p\int_{\mathbf{R}^3\times\mathbf{R}^3\times\mathbf{R}^3}\frac{|w|^2+|W|^2}{(1+|V|^2)^p}M(W)|g_n(W)|\Pi_{gp}(w,\mathrm{d}V\mathrm{d}W)\,\mathrm{d}w,$$

where $C_p \equiv C_p(t,x) \in L^\infty_{loc}(\mathbf{R}^*_+ \times \mathbf{R}^3)$. The integral on the right hand side of this inequality is bounded in $L^\infty_{loc}(\mathbf{R}^*_+ \times \mathbf{R}^3)$ by (H5) and assumption (27) in Theorem 1 part 2. Since $\frac{\varepsilon_n}{\mu_n} \to 0$, we conclude the proof by obtaining

$$\frac{\varepsilon_n}{\mu_n} \int_{\mathbf{R}^3} \tilde{A}(w) \mathscr{R}(Mg_n, F_n)\, \mathrm{d}w \to 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)\,.$$

4.5 *End of the Proof of Theorem 1*

We shall now give the final result, recalling that $\lim_n \mu_n = 1$ in the Vlasov–Navier–Stokes asymptotics (Part 1 of Theorem 1) and $\lim_n \mu_n = 0$ in the Vlasov–Stokes asymptotics (Part 2 in Theorem 1).

Since $\mathscr{L}$ is self-adjoint in $L^2(M\,\mathrm{d}w)$,

$$\frac{\mu_n}{\varepsilon_n}\langle A(w) g_n\rangle = \frac{\mu_n}{\varepsilon_n}\langle (\mathscr{L}\tilde{A})(w) g_n\rangle = \left\langle \tilde{A}(w) \frac{\mu_n}{\varepsilon_n} \mathscr{L} g_n \right\rangle.$$

Following the same procedure as in [2], we use the Boltzmann equation for g_n in the form (32) to express the term $\frac{1}{\varepsilon_n}\mathscr{L} g_n$:

$$\frac{\mu_n}{\varepsilon_n}\langle A(w) g_n\rangle = \mu_n \langle \tilde{A}(w) \mathscr{Q}(g_n)\rangle - \langle \tilde{A}(w)(\varepsilon_n \partial_t + w\cdot\nabla_x) g_n\rangle + \frac{1}{\mu_n}\langle \tilde{A}(w) M^{-1} \mathscr{R}(f_n, F_n)\rangle. \tag{40}$$

We first pass to the limit in the term $\langle \tilde{A}(w)(\varepsilon_n\partial_t + w\cdot\nabla_x) g_n\rangle$ in $\mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)$. Since $\alpha \in L^\infty$, we have $\langle (1+|w|)^2 |\tilde{A}(w)|^2\rangle < \infty$. By assumption (b) in Theorem 1 and the Cauchy–Schwarz inequality,

$$\langle \tilde{A} g_n\rangle \to \langle \tilde{A} g\rangle \text{ and } \langle w\tilde{A} g_n\rangle \to \langle w\tilde{A} g\rangle \text{ in } L^2_{loc}(\mathbf{R}^*_+ \times \mathbf{R}^3) \text{ weak}.$$

Hence $\langle \tilde{A}(w)(\varepsilon_n\partial_t + w\cdot\nabla_x) g_n\rangle = \varepsilon_n \partial_t \langle \tilde{A} g_n\rangle + \mathrm{div}_x \langle w\tilde{A} g_n\rangle \to \mathrm{div}_x \langle w\tilde{A} g\rangle$ in $\mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)$, so that, by Proposition 1,

$$\langle \tilde{A}(w)(\varepsilon_n\partial_t + w\cdot\nabla_x) g_n\rangle \to \nu(\nabla_x u + (\nabla_x u)^T) \text{ in } \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)\,. \tag{41}$$

Next, we evaluate the term $\mu_n \langle \tilde{A}(w) \mathscr{Q}(g_n)\rangle$. We use the identity

$$\langle \tilde{A} \mathscr{Q}(g_n)\rangle = \int_{\mathbf{R}^3\times\mathbf{R}^3} P(w, w_*) M(w_*) g_n(w_*) M(w) g_n(w) \mathrm{d}w \mathrm{d}w_*$$

where $P(w,w_*) := \int_{\mathbf{S}^2} (\tilde{A}(w') - \tilde{A}(w)) c(w - w_*, \omega)\, \mathrm{d}\omega$. Since c satisfies Grad's cut-off assumption and $\alpha \in L^\infty$, $|P(w,w_*)| \le C(1 + |w|^3 + |w_*|^3)$.

The next step is needed only in the Vlasov–Navier–Stokes asymptotics ($\mu_n = 1$).

We define

$$h_n(t,x,w) := \int_{\mathbf{R}^3} P(w,w_*)M(w_*)g_n(t,x,w_*)\,\mathrm{d}w_*\,.$$

Assumption (b) in Theorem 1 implies that $\sup_{n\geq 1}\int_{[0,R]\times[-R,R]^3}\langle g_n\rangle^2\,\mathrm{d}x\mathrm{d}t<\infty$ so that, by the Cauchy–Schwarz inequality, uniformly in $n\geq 1$ as $R\to\infty$,

$$\int_{|w_*|>R}|P(w,w_*)||g_n(t,x,w_*)|M(w_*)\,\mathrm{d}w_*\to 0 \text{ in } L^2_{loc}(\mathbf{R}_+^*\times\mathbf{R}^3\times\mathbf{R}^3).$$

We deduce then from assumption (c) in Theorem 1 that, in $L^2_{loc}(\mathbf{R}_+\times\mathbf{R}^3\times\mathbf{R}^3)$,

$$h_n(t,x,w)\to\int_{\mathbf{R}^3}P(w,w_*)M(w_*)g(t,x,w_*)\,\mathrm{d}w_*=:h(t,x,w).$$

In particular, by weak-strong continuity of the pointwise product,

$$\int_{|w|\leq K}h_n(t,x,w)M(w)g_n(t,x,w)\mathrm{d}w\to\int_{|w|\leq K}h(t,x,w)M(w)g(t,x,w)\mathrm{d}w$$

in $\mathscr{D}'(\mathbf{R}_+\times\mathbf{R}^3)$ for all $K>0$. Since $M(w)h_n(t,x,w)^2\leq C(1+|w|^3)^2M(w)\langle g_n^2\rangle$, by the Cauchy–Schwarz inequality,

$$\int_{|w|>K}h_n(t,x,w)M(w)g_n(t,x,w)\mathrm{d}w\leq\sqrt{C}\left(\int_{|w|>K}(1+|w|^3)^2M(w)\,\mathrm{d}w\right)^{1/2}\langle g_n^2\rangle\to 0$$

in $L^2_{loc}(\mathbf{R}_+^*\times\mathbf{R}^3)$ as $K\to+\infty$ uniformly in $n\geq 1$, according to assumption (b) in Theorem 1. Hence, from Proposition 3

$$\langle\tilde{A}\mathscr{Q}(g_n)\rangle=\langle h_ng_n\rangle\to\langle hg\rangle=\langle\tilde{A}\mathscr{Q}(g)\rangle=A(u)\quad\text{in}\quad\mathscr{D}'(\mathbf{R}_+\times\mathbf{R}^3)\,.\tag{42}$$

Notice that, when considering the Vlasov–Stokes asymptotics, $\mu_n\to 0$ and it is sufficient to consider the bound

$$|\langle\tilde{A}(w)\mathscr{Q}(g_n)\rangle|\leq\left(\langle g_n^2\rangle\langle g_n^2\rangle\right)^{1/2}\left(\int_{\mathbf{R}^3\times\mathbf{R}^3}M(w_*)M(w)(1+|w|^3+|w_*|^3)^2\,\mathrm{d}w_*\mathrm{d}w\right)^{1/2},$$

so that $\mu_n\langle\tilde{A}(w)\mathscr{Q}(g_n)\rangle\to 0$ in $\mathscr{D}'(\mathbf{R}_+^*\times\mathbf{R}^3)$.

Since the last term on the right hand side of (40) vanishes by Proposition 4, collecting all the estimates we get

$$\frac{\mu_n}{\varepsilon_n}\langle A(w)g_n\rangle\to(\lim_n\mu_n)A(u)-\nu\left((\nabla_x u)+(\nabla_x u)^T\right)\quad\text{in }\mathscr{D}'(\mathbf{R}_+\times\mathbf{R}^3)\,.$$

In particular, by the divergence-free condition in Proposition 1,

$$\operatorname{div}_x \frac{\mu_n}{\varepsilon_n}\langle A(w)g_n\rangle \to (\lim_n \mu_n)\left(\operatorname{div}_x(u\otimes u) - \tfrac{1}{3}\nabla_x|u|^2\right) - \nu\Delta_x u - \nu\nabla_x \operatorname{div}_x u$$
$$= (\lim_n \mu_n)\left(\operatorname{div}_x(u\otimes u) - \tfrac{1}{3}\nabla_x|u|^2\right) - \nu\Delta_x u$$

in $\mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)$. Hence, for each compactly supported, smooth vector field $\xi \equiv \xi(x) \in \mathbf{R}^3$ s. t. $\operatorname{div}_x \xi = 0$ we have in $\mathscr{D}'(\mathbf{R}^*_+)$

$$\int_{\mathbf{R}^3} \frac{\mu_n}{\varepsilon_n}\langle w\otimes w g_n\rangle(t,x) : \nabla\xi(x)\,\mathrm{d}x = \int_{\mathbf{R}^3} \frac{\mu_n}{\varepsilon_n}\langle A(w)g_n\rangle(t,x) : \nabla\xi(x)\,\mathrm{d}x$$
$$\to \int_{\mathbf{R}^3} ((\lim_n \mu_n)u\otimes u - \nu\nabla_x u)(t,x) : \nabla\xi(x)\,\mathrm{d}x.$$

We recall now the momentum balance law for the Boltzmann equation for gas molecules

$$\partial_t\langle w g_n\rangle + \frac{1}{\varepsilon_n}\operatorname{div}_x\langle w^{\otimes 2} g_n\rangle = \frac{1}{\mu_n\varepsilon_n}\langle wM^{-1}\mathscr{R}(f_n, F_n)\rangle\,. \tag{43}$$

By Proposition 1, $\langle w g_n\rangle \to \langle wg\rangle = u$ in $L^2(\mathbf{R}^*_+ \times \mathbf{R}^3)$ weak, and $\frac{1}{\varepsilon_n}\langle wM^{-1}\mathscr{R}(f_n, F_n)\rangle \to \kappa\int(v-u)F\,dv$ in $\mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)$.

Thus, for each divergence-free, compactly supported, smooth vector field $\xi \equiv \xi(x) \in \mathbf{R}^3$, passing to the limit in the weak formulation of (43), we get

$$(\lim_n \mu_n)\partial_t \int_{\mathbf{R}^3} u(t,x)\cdot\xi(x)\,\mathrm{d}x = \int_{\mathbf{R}^3} ((\lim_n \mu_n)u\otimes u - \nu\nabla_x u)(t,x) : \nabla\xi(x)\,\mathrm{d}x$$
$$+ \kappa\int_{\mathbf{R}^3\times\mathbf{R}^3} \xi(x)\cdot(v-u(t,x))F(t,x,v)\,\mathrm{d}v\mathrm{d}x\,.$$

By Theorem 17' in [8], there exists $p \in \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3)$ such that

$$(\lim_n \mu_n)\partial_t u + \operatorname{div}_x((\lim_n \mu_n)u\otimes u - \nu\nabla_x u) - \kappa\int_{\mathbf{R}^3}(v-u)F\,\mathrm{d}v = -\nabla_x p\,.$$

This gives the third equation in (25) when $\mu_n = 1$ or in (26) when $\lim_n \mu_n = 0$.

Finally, the equation for the distribution function of the dispersed phase is (first line of (9)):

$$\partial_t F_n + v\cdot\nabla_x F_n = \frac{1}{\eta_n}\mathscr{D}(F_n, f_n)\,.$$

The assumptions on the convergence of F_n in Theorem 1 imply

$$\partial_t F_n + v\cdot\nabla_x F_n \to \partial_t F + v\cdot\nabla_x F \quad \text{in } \mathscr{D}'(\mathbf{R}^*_+ \times \mathbf{R}^3 \times \mathbf{R}^3)$$

leading, together with Proposition 2, to

$$\partial_t F + v \cdot \nabla_x F = \kappa \mathrm{div}_v((v-u)F),$$

concluding the proof of Theorem 1.

5 Conclusion

In this paper we gave a review of the results obtained in two previous papers [4, 5] about the asymptotics leading from a multiphase Boltzmann system to models for sprays where an hydrodynamic equation of Stokes or Navier–Stokes type is coupled to a Vlasov equation.

The validation is formal, in the sense that some convergence properties of the sequences of solutions to the systems of partial differential equations involved in the asymptotics are assumed instead to be proved.

The theorems proved in the articles [4, 5] are presented under the form of a unique theorem, so to make easier the comparison and to highlight the analogies and the differences in the two proofs, leading to the limit systems of Vlasov–Navier–Stokes and Vlasov–Stokes in the two different asymptotics.

References

1. Allaire, G.: Homogenization of the Navier-Stokes equations in open sets perforated with tiny holes. Arch. Ration. Mech. Anal. **113**, 209–259 (1991)
2. Bardos, C., Golse, F., Levermore, C.D.: Fluid dynamic limits of kinetic equations. I. Formal Derivations. J. Stat. Phys. **63**, 323–344 (1991)
3. Bardos, C., Golse, F., Levermore, C.D.: Fluid dynamic limits of kinetic equations. I. Convergence proofs for the Boltzmann equation. Commun. Pure Appl. Math. **46**, 667–753 (1993)
4. Bernard, E., Desvillettes, L., Golse, F., Ricci, V.: A derivation of the Vlasov-Navier-Stokes model for aerosol flows from kinetic theory. Commun. Math. Sci. **15**, 1703–1741 (2017)
5. Bernard, E., Desvillettes, L., Golse, F., Ricci, V.: A derivation of the Vlasov-Stokes system for aerosol flows from the kinetic theory of binary gas mixtures. Kinet. Relat. Models **11**, 43–69 (2018)
6. Charles, F.: Modélisation mathématique et étude numérique d'un aérosol dans un gaz raréfié. Application à la simulation du transport de particules de poussière en cas d'accident de perte de vide dans ITER. Ph.D. thesis, ENS Cachan (2009)
7. Cioranescu, D., Murat, F.: Un terme étrange venu d'ailleurs. Nonlinear Partial Differential Equations and their Applications, Collège de France Seminar vol 2. (Paris 1979–1980). Research Notes in Mathematics, vol. 60, pp. 98–138. Pitman, Boston (1982)
8. de Rham, G.: Differentiable Manifolds: Forms, Currents, Harmonic Forms. Springer, Berlin (1984)
9. Desvillettes, L., Golse, F., Ricci, V.: The mean-field limit for solid particles in a Navier-Stokes flow. J. Stat. Phys. **131**, 941–967 (2008)
10. Desvillettes, L., Golse, F., Ricci, V.: Derivation of a homogenized two-temperature model from the heat equation. M2AN **48**, 1583–1613 (2014)
11. Golse, F.: Fluid dynamic limits of the kinetic theory of gases. From Particle Systems to Partial Differential Equations. Springer Proceedings in Mathematics and Statistics, vol. 75, pp. 3–91 (2013)

12. Marchenko, V.A., Khruslov, E.Ya.: Homogenization of Partial Differential Equations. Birkhäuser, Boston (2006)
13. Ricci, V.: Modelling of systems with a dispersed phase: measuring small sets in the presence of elliptic operators. From Particle Systems to Partial Differential Equations III. Springer Proceedings in Mathematics and Statistics, vol. 162, pp. 285–300 (2016)
14. Ricci, V.: Large number asymptotics for two–component systems with self–consistent coupling. From Particle Systems to Partial Differential Equations. Springer Proceedings in Mathematics and Statistics, vol. 75, pp. 285–300 (2014)